



Philippe Fuchs, *Les casques de réalité virtuelle et de jeux vidéo*, Paris : Presses des MINES, collection Mathématiques et informatique, 2016.

© Presses des MINES - TRANSVALOR, 2016  
60, boulevard Saint-Michel - 75272 Paris Cedex 06 - France  
[presses@mines-paristech.fr](mailto:presses@mines-paristech.fr)  
[www.pressesdesmines.com](http://www.pressesdesmines.com)

© Illustration de couverture : Corentin Echivard

ISBN : 978-2-35671-396-4

Dépôt légal : 2016  
Achevé d'imprimer en 2016 (Paris)

Tous droits de reproduction, de traduction, d'adaptation et d'exécution réservés pour tous les pays.

# Les casques de réalité virtuelle et de jeux vidéo

Collection Mathématiques et informatique

Dans la même collection :

Georges Matheron,  
*Estimating and choosing*

M. Schmitt, J. Mattioli  
*Morphologie mathématique*

J.-M. Martins da Cruz  
*Spam*

X. Deshen, P. Montesinos  
*Proceedings of the third International Workshop  
on Image Analysis*

S. Desprès, M. Crampes  
*IC 2010*  
*21<sup>es</sup> Journées Francophones d'Ingénierie  
des Connaissances*

C. Lurgeau  
*Le siècle de la voiture intelligente*

Dir. P. Fuchs, Coord. générale G. Moreau,  
Coord. du vol., S. Donikian,  
*Le traité de la réalité virtuelle vol.5*  
*Les humains virtuels*

Dir. P. Fuchs,  
Coord. générale G. Moreau, coord. du  
vol. B. Arnaldi et P. Guitton  
*Le traité de la réalité virtuelle vol.4*  
*Les applications de la réalité virtuelle*

Dir. P. Fuchs,  
Coord. générale G. Moreau, coord. du vol.  
G. Moreau et J. Tisseau  
*Le traité de la réalité virtuelle vol.3*  
*Outils et modèles informatiques*

Dir. P. Fuchs,  
Coord. générale G. Moreau, coord. du  
vol. J.-M. Burkhardt et S. Coquillart  
*Le traité de la réalité virtuelle vol.2*  
*L'interfaçage, l'immersion et l'interaction en  
environnement virtuel*

Dir. P. Fuchs,  
Coord. générale G. Moreau, coord. du vol.  
A. Berthoz et J.L. Vercher  
*Le traité de la réalité virtuelle vol.1*  
*L'homme et l'environnement virtuel*

John Cagnol & Jean-Paul Zolesio (eds)  
*Information Processing*  
*Recent Mathematical Advances in Optimization  
and Control*

F. Goulette  
*Modélisation 3D automatique*  
*Outils de géométrie différentielle*

# Les casques de réalité virtuelle et de jeux vidéo

Philippe Fuchs



---

## Préface

---

Depuis le début de l'année 2014 se succèdent un grand nombre de communiqués émanant de constructeurs de matériel informatique annonçant de nouveaux produits dans le domaine de la réalité virtuelle. Oculus Rift, Hololens, Kinect, Leap Motion, Magic Leap... autant de noms nouveaux qui offrent de nouvelles opportunités. Quel que soit leur nature (casques, capteurs, lunettes, etc.), ces nouveaux matériels possèdent un dénominateur commun : leur coût bien inférieur aux équipements utilisés jusque-là, permettant d'envisager des applications destinées au grand public.

Ces annonces ont été relayées par une grande quantité d'articles dans les media mais, malheureusement, au-delà des effets d'annonces, la plupart d'entre eux proposent des analyses incomplètes, voire fausses, à cause d'une méconnaissance de la «réalité virtuelle». Notamment parce que beaucoup de ces chroniqueurs découvrent un domaine loin d'être émergent, qui existe depuis plus d'un quart de siècle avec ses usages, ses entreprises et ses laboratoires de recherche. De façon non exhaustive, citons les applications utilisées depuis longtemps dans les secteurs de la construction automobile, de l'aménagement de bâtiments, de la conception et de l'apprentissage de gestes chirurgicaux, de reconstitutions archéologiques, de la formation, etc.

Au premier rang des erreurs les plus courantes, figure la confusion entre «réalité virtuelle» et «immersion», proposant à l'utilisateur de «plonger» dans un environnement virtuel en fournissant à son cerveau des stimuli sensoriels artificiels. Si l'immersion est nécessaire à la réalité virtuelle, elle n'est pas suffisante. Le second pilier indispensable pour pouvoir parler de réalité virtuelle est l'interaction de l'utilisateur pour qu'il puisse réaliser une tâche dans un environnement virtuel. Ainsi, on ne peut pas parler de réalité virtuelle pour une exploration d'un site via une visualisation à 360° à l'aide d'un visiocasque ou pour une séance de cinéma, même si l'image est projetée en relief sur un très grand écran, accompagnée d'un son spatialisé, parce que le spectateur n'y est que passif.

Beaucoup plus grave que cette confusion terminologique, les professionnels, tant dans le monde de l'entreprise que dans celui de la recherche, savent depuis de nombreuses années qu'une immersion en réalité virtuelle pose des problèmes qu'il ne faut pas négliger. Allant d'un sentiment d'inconfort à des malaises, la manifestation la plus connue a été baptisée mal de l'espace virtuel (cybersickness pour les anglophones). Elle est étudiée depuis longtemps pour tenter d'en comprendre les causes et d'en atténuer les effets. Or, curieusement, nulle trace sur cette question dans ces articles et encore moins d'étude proposée par les fabricants de visiocasques qui se bornent à faire accepter une licence les excluant de toute responsabilité en cas de problème. Illustration : dans la notice fournie avec un visiocasque désormais célèbre, le constructeur déconseille de prendre le volant d'un véhicule juste après une utilisation de son produit ! La volonté de se prémunir de tout risque juridique ne suffit pas à

justifier une telle recommandation. En pratique, les professionnels limitent la durée de l'usage des visiocasques pour les raisons évoquées précédemment. Que se passera-t-il si nos enfants jouent plusieurs heures quotidiennement pour progresser dans un jeu immersif? A l'âge où leur système visuel est encore en développement, quelles conséquences sont à craindre? Il est vraiment très surprenant qu'à une époque où le principe de précaution s'impose dans beaucoup de domaines, cette interrogation ne soulève pas plus d'intérêt, ni, à notre connaissance, d'études de longue durée impliquant des experts de différentes spécialités.

Il est donc indispensable de faire preuve de pédagogie pour présenter au plus grand nombre la réalité virtuelle, expliquer les principes de fonctionnement de ces nouveaux casques et en étudier les usages pour comprendre tant les opportunités que les risques. Malheureusement, il y avait vraiment un chaînon manquant entre ces articles destinés au grand public et les communications réservées aux experts et publiées dans des revues spécialisées, qu'elles soient industrielles ou académiques.

C'est pour combler ce manque que Philippe Fuchs a décidé de rédiger cet ouvrage. Il y décrit de façon complète les concepts, les modes de fonctionnement, les usages et les solutions pour éviter inconfort et malaises éventuels. Je fais le pari que ce livre rencontrera le succès et ce pour plusieurs raisons : tout d'abord, Philippe Fuchs est un pionnier de la réalité virtuelle dans notre pays et il a su s'entourer d'experts pour l'aider à rédiger ce livre. Ensuite, parce qu'il a déjà démontré les qualités nécessaires à une telle réussite en étant à l'initiative et aux commandes du «Traité de la réalité virtuelle», encyclopédie à succès qui s'est traduit par trois versions françaises et une anglophone. Rappelons qu'il n'existe aucun équivalent dans le monde de cet ouvrage dont la dernière édition forte de plus de 2000 pages a réuni 101 auteurs !

Mais finalement rien de surprenant pour un homme qui aime relever les défis les plus fous, et pas uniquement académiques, comme courir pendant cinq mois soixante-dix kilomètres par jour, six jours par semaine, pour relier Paris à Pékin pour l'ouverture des J.O. en 2008 ! La question que je me pose maintenant est de savoir quel sera son prochain défi ?

Pascal Guitton

Professeur d'Informatique à l'Université de Bordeaux

Membre fondateur et Président de l'Association Française de Réalité Virtuelle de 2009 à 2011

Directeur de la Recherche d'Inria de 2010 à 2014

---

# Avant-propos

---

Après les journées de l'AFRV (Association Française de la Réalité Virtuelle)<sup>1</sup> en novembre 2015 à Bordeaux, où nous avons débattu de la future introduction des visiocasques auprès du grand public, j'ai pris la décision d'écrire ce livre, mon dernier... Un des objectifs de l'ouvrage est de faire connaître aux personnes voulant exploiter cette nouvelle interface visuelle, que des recherches et des développements en réalité virtuelle ont déjà été entrepris depuis une vingtaine d'années par la communauté RV, en France et dans le Monde. Il serait inopportun d'ignorer ces travaux. C'est le fruit de tout cette R&D amont qui m'a permis de proposer dans cet ouvrage ma démarche théorique et pragmatique pour l'exploitation des casques de réalité virtuelle.

Désirant la parution du livre pour mars 2016, lors de «Laval Virtual», l'événement incontournable en RV, j'ai dû rédiger en peu de temps cet ouvrage. J'espère que ce ne sera pas au détriment de la qualité du livre que j'aurais voulu améliorer sur certains chapitres. Cette rédaction n'aurait pas pu se réaliser sans l'importante aide de mon collègue, Olivier Hugues, chercheur à Mines ParisTech. Il a accepté de consacrer du temps à la longue rédaction du chapitre sur les produits commercialisés et, surtout, à une relecture attentive et pointilleuse de tous mes chapitres. Je tiens à le remercier vivement ainsi que tous les autres co-auteurs qui m'ont proposé, en si peu de temps, un chapitre sur différents domaines applicatifs des visiocasques, donnant ainsi une plus grande dimension à ce livre : Andras Kemeny, Daniel Mestre, Judith Guez, Jean-François Jego et Stéphane Faudeux.

Je remercie Silvia Dekorsy, directrice des Presses des Mines, et sa collaboratrice, Sandra Rodrigues, qui m'ont aidé à la publication de ce livre et qui, avec beaucoup de patience, ont pris le temps nécessaire à corriger et améliorer cet ouvrage. Ma gratitude va également à Corentin Echivard pour la réalisation de la couverture.

Mes remerciements s'adressent à Pascal Guitton, ancien directeur de la recherche de l'INRIA et président de l'AFRV, qui a accepté de préfacer mon livre. Il m'avait déjà amicalement aidé pour la publication du «Traité de la réalité virtuelle», en prenant en charge la coordination d'un volume, et pour l'édition anglaise du traité «Virtual Reality : Concepts and Technologies».

Mes remerciements s'adressent aussi à tous les chercheurs et industriels avec qui j'ai beaucoup échangé et appris sur le développement de la réalité virtuelle.

Philippe Fuchs

---

<sup>1</sup> [www.af-rv.fr](http://www.af-rv.fr)



---

# Introduction et enjeux

---

Après un quart de siècle de recherche et développement pour des applications principalement professionnelles, la réalité virtuelle semble envisageable aujourd'hui pour le grand public. Les récentes évolutions ont cependant eu une influence sur le nom même du domaine, dont le terme est aujourd'hui très souvent détourné et son usage fortement restreint. En effet, la «réalité virtuelle» est simplement réduite à l'utilisation d'un matériel, appelés casques immersifs ou casques de réalité virtuelle. Cette dernière appellation, trop vite interprétée, insinuerait que l'utilisation d'un tel matériel est nécessaire et suffisante pour mettre en œuvre une application de réalité virtuelle. Il n'en est rien et, pour éviter toute confusion, nous lui préférons le terme neutre de visiocasques<sup>1</sup>, car il s'agit avant tout d'un nouveau dispositif d'affichage pour **voir** des images. L'usage détourné de l'expression «réalité virtuelle» est porté par certains pour s'accaparer l'image de marque de cette innovation. Celle-ci est bien présente : l'utilisateur ne verra plus son monde virtuel sur un simple écran mais à travers un visiocasque, offrant une nouvelle possibilité d'immersion visuelle, avec ses avantages et ses contraintes, qui sont justement les objets principaux de cet ouvrage.

Il est illusoire de vouloir modifier cet usage restrictif du terme «réalité virtuelle», les médias n'écoulant guère les spécialistes du domaine pour définir correctement les innovations. Nous l'avons déjà constaté pour les applications exploitant les images stéréoscopiques<sup>2</sup> : depuis plus d'un siècle, nous connaissons les films stéréoscopiques ou les films en relief. Subitement, l'industrie cinématographique imposa, il y a quinze ans environ, le terme «film 3D» et «télévision 3D», créant la confusion auprès du public qui avait connu les jeux vidéo «3D» dans les années 1990, mais ces derniers n'avaient pas d'images stéréoscopiques ! Nous expliciterons au chapitre 2 les notions de vision binoculaire et d'images stéréoscopiques pour augmenter la perception du relief. Ces sociétés, ayant un poids économique immense, relèguent le domaine des applications professionnelles de la réalité virtuelle au second plan. Elles imposeront peut-être l'usage du terme «réalité virtuelle» pour toute application utilisant un visiocasque. Ce constat ne doit pas conduire obligatoirement les entreprises voulant utiliser les techniques RV à choisir un visiocasque quand d'autres interfaces visuelles seraient plus adaptées. Nous ferons le point sur cette question.

Cette confusion de terme me rappelle ma propre expérience, il y a vingt ans environ, quand je discutais de l'introduction des techniques RV chez PSA Peugeot-Citroën avec les spécialistes de simulation de transport. À l'époque déjà, les techniques de la «réalité virtuelle» impliquaient obligatoirement, pour eux, l'usage d'un visiocasque

---

1 Le terme «visiocasque» a été la première traduction française du terme anglais *Head Mounted Display* (HMD) début 1990, que nous avons repris dans les livres du domaine dont le «Traité de la Réalité Virtuelle».

2 Images stéréoscopiques : couple de deux images différentes, chacune étant vue par un des deux yeux de l'observateur.

plutôt que celui des grands écrans semi-cylindriques issus de leur simulateur de conduite. J'ai dû les convaincre qu'un domaine scientifique et technique ne se définit pas avec le matériel exploité et que leurs futures applications exigeraient différents types d'interfaces visuelles suivant leurs besoins. L'histoire me donna raison car, après les avoir accompagnés dans l'évaluation d'un visiocasque de l'époque, le DataVisor80, ils décidèrent d'utiliser principalement des CAVEs<sup>3</sup>. Ce visiocasque, malgré son coût conséquent, avait des caractéristiques trop limitées dues aux technologies de l'époque, dont un champ de vision trop restreint, ce qui ne pouvait répondre aux attentes de simulation de l'intérieur d'un véhicule.

Bien que le terme «réalité virtuelle» n'ait été utilisé jusqu'à ce jour que pour les applications professionnelles basées sur des environnements interactifs, les chercheurs et les industriels en RV savent que leur domaine fait appel aux mêmes concepts et aux mêmes technologies de base, matérielles et logicielles, que celles des jeux vidéo. Les exemples sont nombreux, comme les logiciels de RV : Virtools et Unity qui ont été principalement développés pour les jeux vidéo et ont été (pour Virtools) ou sont encore (pour Unity) utilisés pour les applications RV professionnelles. En ce qui concerne le matériel, la Kinect conçue par Microsoft est une interface destinée à la console Xbox 360, permettant de contrôler des jeux vidéo. Mais elle fut très vite détournée de son usage premier par les chercheurs, pour être aussi très utilisée dans les applications RV non ludiques. Bien sûr, l'évolution de ces technologies n'est due qu'à l'importance économique des entreprises des jeux vidéo qui peuvent se permettre d'investir des sommes gigantesques dans le développement de technologies innovantes. Il est aisé de comprendre que les enjeux sont de grande taille en terme économique, sans besoin de rappeler l'engouement suscité par l'achat d'Oculus par Facebook.

Il y a déjà vingt ans, des concepteurs de jeux vidéo voulaient introduire des visiocasques auprès du public, mais les technologies n'étaient pas assez performantes. Il y eu bien des échecs, comme la console de jeux Virtual Boy de Nintendo, accompagnée d'un «visiocasque stéréo» très peu vendu et abandonnée en 1996. À cette époque, avant les fêtes de fin d'année, il y eu plusieurs annonces de visiocasques pour les jeux vidéo, qui prêtaient à sourire, compte tenu des faibles caractéristiques de ces dispositifs dédiés au «grand public». Un peu gênés par une très faible définition des écrans, les constructeurs communiquaient un nombre de pixel multiplié par trois (pour les 3 couleurs RGB)!

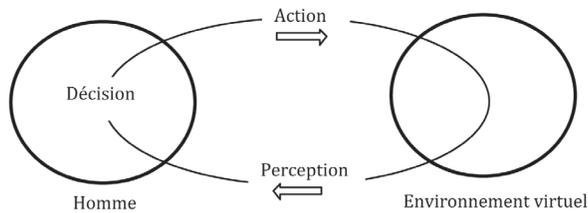
Après avoir redéfini le domaine de la réalité virtuelle dans cet ouvrage, il est ensuite nécessaire d'aborder certaines notions du fonctionnement sensorimoteur de l'être humain, en particulier celui de la vision, dans le monde réel comme dans un environnement virtuel. Car c'est à l'utilisateur que s'adresse cette innovation, *invasive au niveau visuel*, et cela a des impacts sur ses autres sens et sur ses actions motrices. La méconnaissance du fonctionnement sensorimoteur de l'être humain pour la plupart d'entre nous, dont moi au début de ma carrière, impose de faire dans cet ouvrage un rappel des connaissances de base sur les sens... qui sont plus nombreux que cinq! Cette

---

3 Un CAVE est un ensemble d'écrans formant quelques faces (de 4 à 6) d'un cube de quelques mètres de côté entourant l'utilisateur.

réalité aura une incidence importante pour la compréhension des problèmes et des solutions d'exploitation des visiocasques. Et la vision humaine doit être suffisamment comprise pour exploiter correctement un visiocasque.

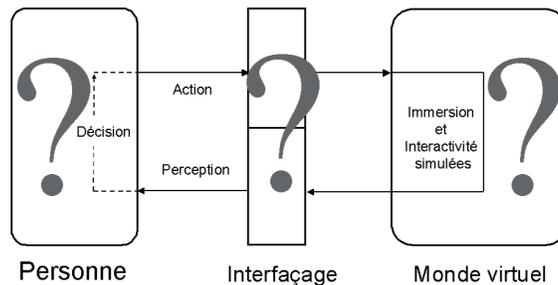
L'homme est au centre de ce nouvel espace créé virtuellement, dans lequel il est immergé et peut (inter)agir. Face à cette problématique générale en réalité virtuelle, qui fait appel aux sciences de l'ingénieur et aux sciences cognitives, j'ai pris l'initiative avec d'autres chercheurs en RV, il y a une vingtaine d'années, de réunir ces deux communautés pour travailler sur les différentes problématiques interdisciplinaires de la réalité virtuelle. Dans toute application de réalité virtuelle, la personne est en immersion et en interaction dans un environnement virtuel. Elle perçoit, décide et agit dans cet environnement, processus schématisé par la classique boucle «perception, décision, action», qui doit être réalisée malgré les contraintes technologiques, physiologiques et cognitives (Figure 1).



**Figure 1 :** la boucle classique «perception, décision, action»

De ce schéma découlent trois problématiques fondamentales de la réalité virtuelle (Figure 2) :

- La problématique de l'analyse et de la modélisation de l'activité humaine en environnement virtuel ;
- La problématique de l'analyse et de la réalisation de l'interfaçage du sujet pour son immersion et son interaction dans un environnement virtuel ;
- La problématique de la modélisation et de la réalisation de l'environnement virtuel.



**Figure 2 :** les trois problématiques fondamentales de la réalité virtuelle

Ces trois problématiques, différentes mais fortement imbriquées, sont développées en détail dans les trois premiers volumes du «*Traité de la Réalité Virtuelle*»<sup>4</sup>, que tout lecteur voulant compléter ses connaissances peut consulter pour l'aider à aborder les questions suivantes :

- Quel est le comportement de l'homme, équipé d'un visiocasque, confronté à des latences et des incohérences sensorimotrices dans un monde virtuel ?
- Quelles sont les interfaces et les techniques d'interaction qui génèrent ces latences et ces incohérences ?
- Quels sont les outils et les algorithmes informatiques qui limitent la nuisance de ces artefacts ?

Une démarche basée sur le modèle 3I<sup>2</sup>, développée à l'école Mines ParisTech, présentera cette problématique générale d'immersion et d'interaction pour toute application RV, indépendamment des dispositifs utilisés, pour se focaliser ensuite sur l'immersion visuelle via un visiocasque.

Avant d'aborder la conception des visiocasques, une présentation de tous les types d'interfaces visuelles sera brièvement proposée pour comparaison. Les interfaces associées aux visiocasques seront brièvement présentées dans cet ouvrage, plus sous les aspects fonctionnels que techniques. Un chapitre suivant présentera les différents visiocasques commercialisés avec leurs principales caractéristiques. L'usage des visiocasques peut poser des problèmes de confort et de santé dans certains types d'applications, dus principalement aux incohérences sensorimotrices induites et aux temps de latence trop élevés. Un chapitre y sera consacré et le suivant donnera une méthode et des solutions pour pallier à ce risque important de rejet de la part des futurs utilisateurs.

Tous les développeurs ont-ils bien conscience des difficultés et des règles à respecter pour que leurs applications soient efficaces ? Les utilisateurs auront-ils la capacité d'adaptation suffisante à cette immersion visuelle ? L'objectif principal de cet ouvrage est de faire le point sur cette question ouverte. Ayons bien à l'esprit que ces technologies «*perturbent*» le fonctionnement physiologique et sensorimoteur de l'utilisateur. On peut faire le rapprochement avec le cas de la vision stéréoscopique qui crée une incohérence sensorimotrice entre l'accommodation et la vergence des yeux. Mais les spécialistes de ce dernier domaine, les stéréographes, connaissent depuis forts longtemps les règles à respecter pour exploiter les images stéréoscopiques (3Ds). La première règle, vieille de plus d'un siècle et demi, a été déterminée par David Brewster en 1856, et beaucoup d'études expérimentales ont ensuite analysé précisément ce que le public pouvait tolérer. Avec le visiocasque, cette nouvelle interface visuelle intrusive, l'expérience de son usage est très limitée. Il est souhaitable que les spécialistes étudient cette question qui a été très peu abordée à ce jour.

---

<sup>4</sup> Fuchs P., Moreau G., *et al.*, *Le Traité de la réalité virtuelle*, troisième édition, 5 volumes : «*L'Homme et l'environnement virtuel*», «*Interfaçage, immersion et interaction en environnement virtuel*», «*Les outils et les modèles informatiques des environnements virtuels*», «*Les applications de la réalité virtuelle*», «*Les humains virtuels*», Presses des Mines, 2009. [www.pressesdesmines.com](http://www.pressesdesmines.com)

D'ailleurs, certains concepteurs de visiocasques, en partie pour se protéger d'une absence d'études approfondies, alertent sur les risques éventuels liés à l'utilisation de leur produit et en interdisent l'usage en dessous d'un certain âge. Cette problématique semble plus complexe que celle concernant la vision stéréoscopique... Mais ceci n'est qu'intuitif. Les chercheurs sont en manque de données issues d'études sur un grand nombre de personnes, puisque les visiocasques ne sont pas encore utilisés par le grand public. Quel sera le niveau d'adaptation de l'être humain à cette immersion visuelle dans un monde virtuel? Certains seront plus sensibles que d'autres mais sans trop savoir profondément pourquoi. Une autre question peut être posée: est-ce que le risque d'addiction aux jeux vidéo pour certaines personnes peut-être plus élevé si le joueur utilise un visiocasque à la place d'un simple écran? Cette question, adressée aux psychologues et aux psychiatres, ne sera pas débattue dans cet ouvrage. Je suppose qu'elle n'a pas encore été étudiée.

Ce nouvel artefact pose des problèmes de confort, de santé mais ouvre de nouveaux horizons puisque le virtuel permet d'aller au-delà du réel. Comment peut-on se percevoir, se sentir présent dans un espace où l'on ne voit pas son corps ou seulement partiellement (par exemple, seules les mains virtuelles de son propre avatar sont visibles)? Et que dire si l'on se situe visuellement dans un autre corps, en dehors de son corps? Les chercheurs en Sciences Cognitives peuvent entreprendre des études instructives sur le comportement humain. Il faudra aussi connaître les capacités d'adaptation de l'être humain à des environnements virtuels qui peuvent être très éloignés de notre environnement réel. J'ai eu la chance de pouvoir longuement méditer sur mon adaptation sensorimotrice lors de mes randonnées en courant à travers l'Europe et l'Asie, ayant découvert, vers la quarantaine, les capacités insoupçonnées de l'endurance du corps humain. On connaît très peu son fonctionnement et ses propres capacités physiques et psychiques, qui permettent, entre autres, d'aller de Paris à Pékin en courant :

*« Comment est-ce possible qu'un corps s'adapte à une si longue activité sportive? C'est une question que l'on me pose sans cesse au sujet de mes courses peu ordinaires. L'incompréhension vient du fait que nous avons oublié que l'homme a vécu sans aide technique durant des millénaires et qu'il a intrinsèquement une grande capacité d'adaptation, en particulier à des efforts prolongés en courant. Cette aptitude d'adaptation est fondamentale pour tout être vivant et nous en profitons dans de multiples occasions, même lors de l'introduction de nouveaux dispositifs techniques : les outils, les moyens de transport, les machines, les ordinateurs... jusqu'aux environnements virtuels, mon domaine de recherche »<sup>5</sup>.*

L'ouvrage se terminera par un tour d'horizon des domaines d'applications exploitant judicieusement ces nouvelles interfaces visuelles: comment mettre à profit les avantages de cette nouvelle interface visuelle, tout en minimisant ses inconvénients? L'exploitation raisonnée et efficace des visiocasques est un défi auquel cet ouvrage veut apporter sa contribution. Je crois qu'il est indispensable d'éclairer le professionnel et le public sur cette innovation qui est loin d'être anodine.

---

<sup>5</sup> Extrait du livre P. Fuchs, *Adaptation et conditionnement - Réflexions en courant de Paris à Pékin*, 2015, [www.edilivre.com](http://www.edilivre.com)



# Partie I



---

## Concepts de base en réalité virtuelle

---

### 1. DÉFINITIONS DE LA RÉALITÉ VIRTUELLE

Définir la réalité virtuelle est une tâche indispensable. On trouve encore dans la littérature et dans les médias des définitions qui mélangent malencontreusement la finalité de la réalité virtuelle, ses fonctions, ses applications et les techniques sur lesquelles elle repose, telle que le visiocasque. Il faut rejeter ces approches, à la fois parce qu'elles sont trop centrées sur une technologie particulière et, à la fois, parce qu'elles sont trop fortement restrictive quant aux enjeux scientifiques et commerciaux. Nous proposons des définitions à plusieurs niveaux pour clarifier le domaine de la réalité virtuelle.

#### 1.1. Finalité de la réalité virtuelle

Avant de se focaliser sur des fonctions ou des techniques, il semble judicieux de déterminer d'abord la finalité de la réalité virtuelle que partagent tous les acteurs travaillant dans ce domaine depuis vingt ans. Après avoir étudié l'objectif commun de chacune de ses applications, nous pouvons affirmer que [Fuchs P., 1996]:

*La finalité de la réalité virtuelle est de permettre à une personne (ou à plusieurs) une activité sensorimotrice et cognitive dans un monde artificiel, créé numériquement, qui peut être imaginaire, symbolique ou une simulation de certains aspects du monde réel.*

**Une simulation de certains aspects du monde réel**: ceux-ci sont à déterminer lors de la conception de l'application. Nous verrons que cette phase initiale de conception est fondamentale et doit être analysée explicitement. L'erreur, souvent rencontrée, est celle du concepteur qui recherche le plus grand « degré de réalisme ». Cette approche erronée est entreprise sans se soucier de savoir précisément quels sont les aspects de la réalité que l'application exige. L'idée naïve de vouloir, si possible, le comportement du monde virtuel entièrement identique à celui du monde réel est absurde. Si on veut faire une réalité « virtuelle », c'est bien dans le but de modifier des aspects de la « vraie » réalité ; par exemple, réaliser une formation sans danger réel pour le formé dans un environnement virtuel (Figure 1.1).



**Figure 1.1 :** Simulation de certains aspects du monde réel

**Un monde symbolique :** on peut aussi exploiter des représentations symboliques pour améliorer la compréhension du monde simulé. La réalité virtuelle est alors exploitée soit pour représenter un phénomène, (la structure de molécules, l'écoulement de fluide, etc.), soit pour ajouter au monde réel simulé des concepts symboliques. Ceux-ci permettent à l'utilisateur de se faire une meilleure représentation mentale de son environnement. Par exemple : nous pouvons afficher des informations schématiques pour l'utilisateur, lui permettant de mieux saisir la structure d'un mécanisme ou la planification d'une tâche à accomplir (Figure 1.2).



**Figure 1.2 :** Simulation d'un monde symbolique

**Un monde imaginaire :** la virtualité est employée pour créer un monde irréel, sorti de l'imagination de l'artiste ou de l'auteur de science-fiction. Dans ce cas, le monde créé n'a pas l'obligation d'être une simulation du monde réel (Figure 1.3).



**Figure 1.3 :** Simulation d'un monde imaginaire

En conclusion, il y a bien des façons variées d'exploiter les potentialités de la réalité virtuelle, les trois cas pouvant évidemment être associés dans une même application.

Avec une telle finalité de la réalité virtuelle, les jeux vidéo en font partie, même si l'activité sensorimotrice (physique) du joueur est faible jusqu'à présent et va devenir plus importante avec l'introduction des visiocasques et des dispositifs associés. Cette évolution explique pourquoi l'industrie des jeux vidéo vient de s'accaparer l'expression «**réalité virtuelle**» si porteur auprès du grand public. Ce dernier ignore que depuis un quart de siècle des chercheurs et des entreprises<sup>1</sup> exploitent la réalité virtuelle, comme ce fut notre cas à l'école des Mines ParisTech. Mais parler de réalité virtuelle lorsque l'on propose des vidéos à 360 degrés d'environnement réels filmés, même observées avec un visiocasque, est exagéré. Mais on ne pourra guère empêcher les entreprises de ce nouveau secteur d'employer un terme aussi porteur commercialement.

## 1.2. Définition fonctionnelle

En 1995, j'ai proposé une taxonomie basée sur les fonctionnalités «théoriques» de la réalité virtuelle. Face à sa propre perception de la réalité, l'homme a conceptualisé les notions de temps et d'espace sur lequel il ne peut interagir que suivant des lois physiques immuables. À l'opposé :

*La réalité virtuelle va lui permettre de s'extraire de la réalité physique pour changer virtuellement de **temps**, de **lieu** et(ou) de type d'**interaction** : interaction avec un environnement simulant la réalité ou interaction avec un monde imaginaire ou symbolique.*

Cette définition fait allusion à l'exigence inverse des auteurs des tragédies littéraires du XVII<sup>e</sup> siècle prônant la règle des trois unités de temps, de lieu et d'action. Une réflexion sur cette approche peut être consultée dans l'article [Nannipieri O., Fuchs P., 2009].

## 1.3. Définition technique

Une définition plus technique de la réalité virtuelle va s'attacher à caractériser le domaine. Deux mots sont la clef de voûte de la réalité virtuelle : l'immersion et l'interaction. La définition technique de la réalité virtuelle est [Arnaldi B., Fuchs P., 2003] :

*La réalité virtuelle est un domaine scientifique et technique exploitant l'informatique et des **interfaces comportementales** en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en **interaction en temps réel** entre elles et avec une personne en **immersion pseudo-naturelle**.*

Cette définition introduit une terminologie nécessitant quelques explications permettant de la situer par rapport aux arguments développés dans l'introduction.

---

<sup>1</sup> Les centres de recherches et les entreprises du domaine de la réalité virtuelle sont regroupés dans l'Association Française de la Réalité Virtuelle (AFRV), <http://www.af-rv.fr>

Nous exploitons des interfaces matérielles de la réalité virtuelle, que nous appelons « **interfaces comportementales** » car exploitant des comportements humains. Elles sont composées « d'interfaces sensorielles », « d'interfaces motrices » et « d'interfaces sensorimotrices ». Les interfaces sensorielles (visuelles, tactiles, sonores, etc.) informent l'utilisateur par ses sens de l'évolution du monde virtuel. Une interface visuelle est toujours évidemment exploitée : un visiocasque, un CAVE ou un simple écran, etc. Les interfaces motrices informent l'ordinateur des actions motrices de l'homme sur le monde virtuel (manettes, gants de données, tapis de marche, etc.). Les interfaces sensorimotrices informent dans les deux sens (les interfaces à retour d'effort<sup>2</sup>). Le nombre et le choix de ces interfaces dépendent de l'objectif poursuivi de l'application. L'**interaction en temps réel** est obtenue si l'utilisateur ne perçoit pas de décalage temporel, appelé « temps de latence », entre son action, via les interfaces motrices, sur l'environnement virtuel et la réponse sensorielle de ce dernier via les interfaces sensorielles. Cette contrainte est difficile à satisfaire et nous en parlerons.

L'utilisateur doit être en « **immersion pseudo-naturelle** » la plus efficace possible dans le monde virtuel. Cette sensation est une notion en partie subjective qui dépend de l'application et du dispositif utilisé (interfaces, logiciels, etc.). Nous parlerons en détail des concepts d'immersion et d'interaction qui doivent être bien définis et analysés à plusieurs niveaux.

Il résulte de cette analyse globale un principe de la réalité virtuelle. Celui-ci est contenu dans la boucle de la figure 1.4. L'utilisateur agit sur l'environnement virtuel grâce à l'usage d'interfaces motrices qui captent ses actions (gestes, déplacements, voix, etc.). Ces activités sont transmises à l'ordinateur qui l'interprète comme une demande de modification de l'environnement. Conformément à cette sollicitation de modification, l'ordinateur évalue les transformations à apporter à l'environnement virtuel et les restitutions sensorielles (images, son, efforts, etc.) à transmettre aux interfaces sensorielles. Cette boucle en environnement virtuel interactif n'est que la transposition de la boucle « perception, décision, action » du comportement de l'homme dans un mode réel. Mais deux contraintes majeures, inhérentes aux techniques, vont perturber la boucle « perception, décision, action » et en conséquence le comportement du sujet : les **incohérences sensorimotrices**, dont le cas spécifique des temps de latence. Le temps de latence est un fait technique inhérent aux environnements virtuels interactifs et va influencer la qualité de toute application de réalité virtuelle. Quel que soit le nombre de canaux sensoriels exploités dans une application, quel que soit le nombre d'interactions à la disposition du sujet, il y a presque toujours des incohérences sensorimotrices par rapport au comportement sensorimoteur du sujet dans le monde réel. Ces incohérences et les temps de latence perturberont-ils le comportement de l'utilisateur ? Nous traiterons cette problématique au chapitre 7.

---

<sup>2</sup> Les interfaces d'effort appliquent des forces et des couples sur la main de l'utilisateur qui peut en retour manipuler l'objet virtuel. Elles sont décrites dans le volume deux du *Traité de la Réalité Virtuelle*, éditeur Presses des Mines, [www.pressesdesmines.com](http://www.pressesdesmines.com)

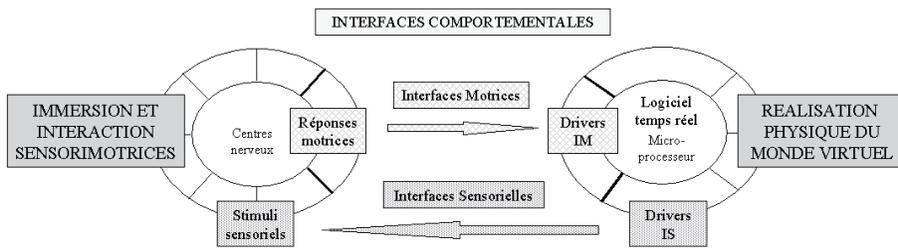


Figure 1.4 : niveau sensorimoteur de l'immersion et de l'interaction

## 2. MODÈLE DE RÉFÉRENCE POUR L'IMMERSION ET L'INTERACTION

### 2.1. Les trois niveaux d'immersion et d'interaction

Dans tout dispositif de réalité virtuelle, l'homme est au centre du système car l'application virtuelle lui est adressée. Partant de ce constat, il est souhaitable d'adopter une approche anthropocentrique (dirigée vers l'homme) et non technocentrique (dirigée vers l'ordinateur), symbolisée par le **schéma anthropocentrique** (Figure 1.5) de l'immersion de l'homme dans le monde virtuel, tel que celui-ci doit le percevoir en tant qu'utilisateur.

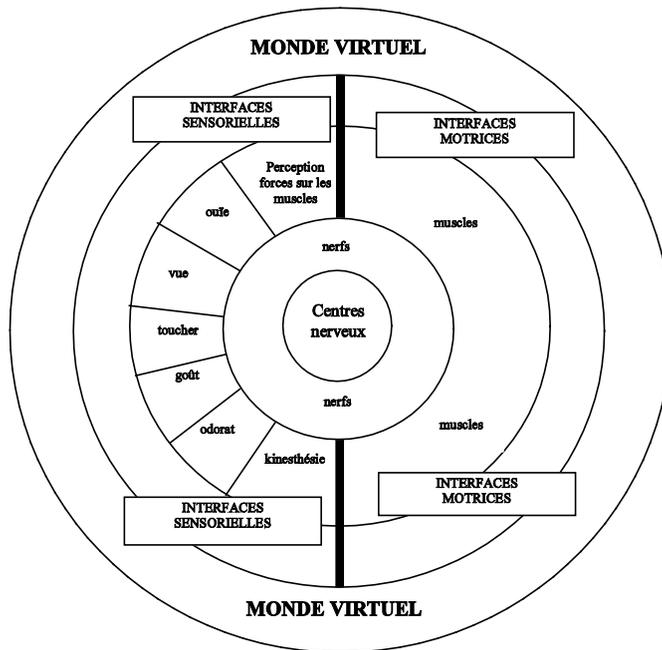


Figure 1.5 : schéma anthropocentrique de l'immersion de l'homme dans le monde virtuel

Mais si le concepteur doit se référer à ce schéma anthropocentrique, objectif à réaliser pour l'utilisateur, il doit aussi le compléter car il est nécessaire d'analyser finement le processus d'interfaçage et les dispositifs à concevoir. Nous avons fait le choix fondamental d'analyser ce processus à trois niveaux. Au premier niveau, on peut schématiser l'interfaçage entre l'homme et le monde virtuel *au niveau physique*. Nous parlons dans ce cas **d'immersion et d'interaction sensorimotrices**, puisque physiquement l'ordinateur est connecté au corps de l'utilisateur au niveau de ses sens et de ses réponses motrices. Pourquoi employons-nous pour les trois niveaux les termes « Immersion ET Interaction » ? Fondamentalement, comme nous l'avons indiqué dans notre définition de la réalité virtuelle, celle-ci est basée sur l'Immersion d'une personne dans un monde virtuel. Mais nous y associons le terme Interaction, car parler seulement d'immersion serait omettre l'innovation de la réalité virtuelle, qui est de permettre au sujet d'agir dans un monde virtuel. D'où, au niveau physique d'immersion et d'interaction sensorimotrices, nous obtenons le schéma technocentrique de la figure 1.4.

Le schéma précédent est restrictif car il ne représente que partiellement les problèmes et les solutions que tout concepteur doit étudier. À un deuxième niveau, nous devons analyser les processus cognitifs que le sujet va mettre en œuvre et qu'il faudra assister dans son interfaçage avec l'environnement virtuel : c'est le niveau de **l'immersion et de l'interaction cognitives**. Notre approche est basée sur l'activité du sujet. Il s'agit de comprendre à partir de quels processus mentaux la personne va penser et agir, comme l'a décrit Rabardel [Rabardel P., 1995] dans son approche instrumentale : l'interface est un médiateur d'activité qui repose matériellement sur des artefacts (ou instruments) et va être utilisée suivant des processus cognitifs de l'utilisateur.

Quelles vont être les catégories de processus cognitifs qui seront impliqués pour permettre une interaction et une immersion efficaces et, si possible, relativement naturelles, proposant alors une immersion et une interactivité pseudo-naturelles ? Concrètement, tout concepteur peut exploiter deux catégories de processus cognitifs, que nous avons proposé de bien différencier : les **schèmes comportementaux importés** et les **métaphores**, avec ou sans substitution sensorielle.

L'utilisateur peut faire appel dans ses activités sensorimotrice et cognitive à certains automatismes qu'il a assimilés dans le monde réel en les mettant en œuvre dans le contexte de l'activité à réaliser dans l'environnement virtuel. Ces automatismes sont à associer à la notion de schème proposée par le psychologue Piaget [Piaget J. *et al.*, 1979] : un schème est l'organisation mentale des actions telles qu'elles se transfèrent ou se généralisent lors de la répétition de cette action en des circonstances analogiques. D'après Piaget, les schèmes constituent les moyens du sujet à l'aide desquels il peut assimiler les situations et les objets auxquels il est confronté, comme c'est le cas avec les nouvelles expérimentations dans un environnement virtuel. Pour donner un exemple : quand on veut proposer à un utilisateur d'observer un lieu dans toutes les directions, il est préférable d'exploiter un visiocasque ou un CAVE, plutôt qu'un simple écran, car l'utilisateur exploite naturellement le schème de changement de point de vue par rotation de sa tête. Mais si l'application RV exige une grande

qualité d'images à faible coût ou pour d'autres raisons, on peut employer un simple écran. Il faut alors permettre le changement de point de vue, via une métaphore : c'est, par exemple, manuellement avec une manette que l'utilisateur peut changer le point de vue de la scène observée. Il est donc utile de réfléchir, lors de la conception d'une application RV, s'il est possible d'exploiter un schème acquis par toute personne dans le monde réel et de l'importer en environnement virtuel. Nous parlons de Schème Comportemental Importé (SCI). Par exemple, pour la conception du déplacement dans un magasin virtuel, nous nous sommes basés sur le schème du «déplacement en poussant un chariot», schème bien maîtrisé par l'ensemble de la population. La commande de déplacement passe par l'intermédiaire d'un chariot pour se rapprocher des actions effectuées dans un magasin réel. L'utilisateur doit pousser le chariot en avant ou en arrière et l'orienter à gauche ou à droite devant l'écran, comme avec tout joystick. Mais pour les personnes n'ayant pas l'habitude d'employer un joystick, il était nettement plus aisé et compréhensible de «pousser son chariot», comme nous avons pu le vérifier après expérimentation (Figure 1.6).



**Figure 1.6 :** déplacement virtuel grâce à un chariot, contrôlé comme un joystick

Notons pour l'instant, comme Piaget l'a énoncé, que l'intelligence sensorimotrice parvient à résoudre un ensemble de problèmes d'action (atteindre des objets, se déplacer, etc.) en construisant un système complexe de schèmes d'assimilation et à organiser le réel selon un ensemble de règles spatiotemporelles et causales. D'où notre postulat fondamental en réalité virtuelle :

*Dans un environnement virtuel interactif, la personne exploite la même démarche que dans un monde réel, pour organiser le virtuel selon un ensemble de règles spatiotemporelles et causales.*

Cette organisation doit être perçue cohérente pour le cerveau de l'utilisateur. Il y a des risques d'éventuelles difficultés si des incohérences sensorimotrices sont créées par l'application. Nous en reparlerons au chapitre 7.

Dans le cas de difficultés techniques, économiques ou théoriques ne permettant pas d'exploiter un Schème Comportemental Importé, nous pouvons contourner ces difficultés en employant une **métaphore**. Au lieu d'exploiter un comportement

sensorimoteur et acquis de la personne, nous lui proposons, visuellement en général, une image symbolique de l'action ou de la perception souhaitée. Par exemple dans un magasin virtuel, on peut proposer au consommateur de valider l'achat d'un produit en cliquant simplement sur son image et ensuite sur une icône représentant une caisse. Cette action devient symbolique et elle n'est plus représentative de l'action sensorimotrice dans un magasin réel, l'immersion et l'interaction y sont moins pseudo-naturelles. Un autre exemple de métaphore : pour éviter de trop fortes incohérences sensorimotrices lors d'une navigation en Environnement Virtuel (EV)<sup>3</sup> avec un visiocasque, il est préférable de se déplacer par téléportation d'un point à un autre. Le choix du lieu d'arrivée est choisi par indication manuelle via un pointeur, action métaphorique, l'utilisateur passant instantanément du lieu de départ au lieu d'arrivée.

La métaphore peut aussi exploiter un sens spécifique, qui soit différent du sens stimulé dans l'action réelle : **métaphore avec substitution sensorielle**, par exemple la détection entre deux objets est rendue par leur changement de couleur. Nous verrons que la plupart des visiocasques occultent la vision périphérique de l'observateur. Il peut donc être judicieux d'activer une source sonore spatialement positionnée dans la zone de vision périphérique pour faire comprendre à l'observateur de tourner la tête dans la direction de la source sonore.

Nous avons présenté deux niveaux d'immersion et d'interaction : le niveau sensorimoteur et le niveau cognitif. Un troisième niveau est nécessaire pour compléter notre démarche. Ce troisième niveau concerne l'application de réalité virtuelle, dont l'objectif est de s'attacher à réaliser une immersion de l'homme pour une activité donnée et non pour une simple immersion cognitive de l'homme dans un monde artificiel. Nous parlons dans ce cas **d'immersion et d'interaction fonctionnelles**<sup>4</sup>. C'est le niveau fondamental que doit étudier en premier tout concepteur d'application : quelles sont les activités (les fonctions) que doit réaliser le sujet ?

Au niveau I<sup>2</sup>fonctionnelles, par rapport à l'application et ses objectifs, il faut se poser la question suivante : quelles sont les activités que l'utilisateur doit exécuter ? Si le lecteur veut prendre le temps d'y réfléchir, il verra rapidement que dans toutes les applications RV, les activités du sujet sont toujours décomposables en quelques comportements de base que nous appelons les « Primitives Comportementales Virtuelles » (PCV). Il faut donc au niveau des I<sup>2</sup>fonctionnelles bien définir les PCV et leurs spécificités. Quelle que soit l'application, ces dernières peuvent être regroupées en quatre catégories :

- observer le monde virtuel ;
- se déplacer dans le monde virtuel ;

---

3 On utilise par la suite la notation EV pour Environnement Virtuel et ER pour Environnement Réel.

4 On utilise pour la suite la notation I<sup>2</sup> pour l'immersion et l'interaction.

- agir sur le monde virtuel, principalement manipuler des objets ;
- communiquer avec autrui ou avec l'application RV<sup>5</sup>.

Les visiocasques sont conçus évidemment pour améliorer la première PCV (Observation) en permettant une immersion pseudo-naturelle, l'utilisateur exploitant un schème naturel. Mais il faut noter que la plupart des visiocasques ont un champ de vision limité qui contraint en partie l'observation de l'utilisateur. Le déplacement (la navigation) dans le monde réel avec un visiocasque pose des problèmes, dus en particulier aux incohérences sensorimotrices qui peuvent être créées. L'observation visuelle peut être ou non couplée au déplacement de l'utilisateur dans le monde virtuel et l'association de ces deux PCV peut aussi induire des difficultés complémentaires. La troisième PCV, la manipulation d'objets, avec un visiocasque, pose le problème de la représentation ou non des membres supérieurs, surtout des mains, de l'utilisateur. Ce sont des difficultés que nous détaillerons au chapitre 7. La quatrième PCV, communication avec autrui, ne sera pas développée dans cet ouvrage. On notera juste que le port du visiocasque pose le problème d'être coupé des personnes situées à côté de vous, situation souvent frustrante. Vous pouvez leur parler mais vous ne les voyez pas. C'est un des avantages d'être dans un CAVE plutôt qu'être immergé dans un visiocasque. C'est une contrainte non négligeable si on souhaite faire de la formation en RV avec un visiocasque. Mais il est toujours possible de communiquer avec d'autres personnes qui sont aussi immergées dans le même EV, en les représentant sous forme d'avatars ou sous forme d'incrustation vidéo.

Nous obtenons en final le schéma de référence de la réalité virtuelle (Figure 1.7) qui va nous servir de canevas dans notre démarche en réalité virtuelle [Fuchs P., *et al.*, 1999]. Il est structuré en trois niveaux et avec deux côtés symétriques qui existent l'un grâce à l'autre et se font écho : le monde virtuel créé cognitivement par la personne et le monde virtuel créé matériellement par l'ordinateur<sup>6</sup>. Avec ce schéma interdisciplinaire<sup>7</sup>, nous avons clarifié les notions d'immersion et d'interaction<sup>8</sup>.

---

5 La communication avec l'application RV, dont la configuration du visiocasque, est en général difficile, peu ergonomique car l'utilisateur n'a pas accès à un simple clavier plus un écran. Il doit souvent faire des commandes avec des boutons, sur son casque ou sur ses manettes, sans les voir !

6 Cette relation bijective est expliquée aussi par Marc Boucher dans la revue d'art en ligne « arts médiatiques & cyberculture » (<http://archee.qc.ca/>), qui parle de « l'immersant fait la réalité virtuelle et la réalité virtuelle fait l'immersant ».

7 Ce schéma fait appel à des connaissances en informatique, en physiologie, en psychologie, en ergonomie et en technologie des interfaces (optique, mécanique, etc.)

8 Cette démarche d'immersion et d'interaction 3I<sup>2</sup> est plus détaillée dans le volume 2 « Interfaçage, Immersion et Interaction en environnement virtuel » du *Traité de la Réalité Virtuelle*.

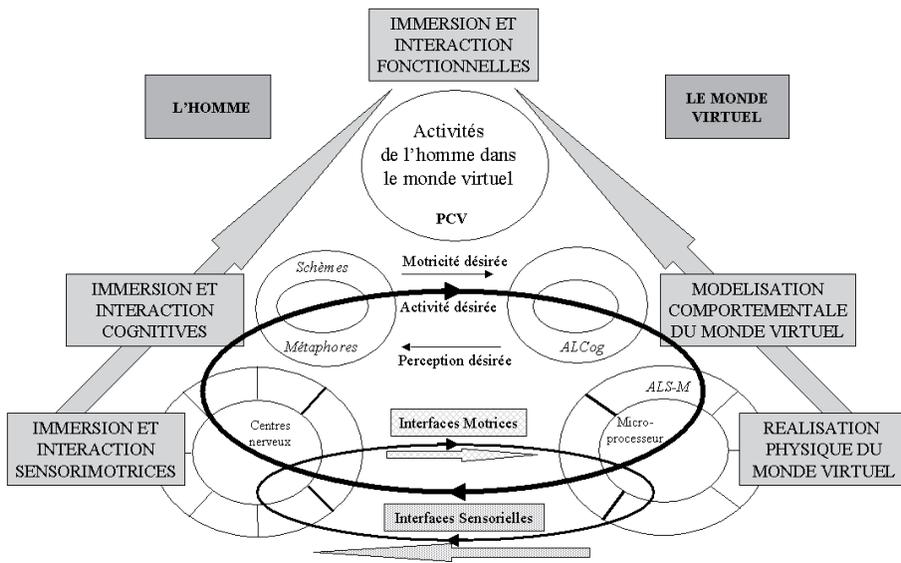


Figure 1.7: schéma général à trois niveaux d'immersion et d'interaction

## 2.2. Les aides logicielles comportementales

En associant un découpage transversal entre la personne et le monde virtuel, nous avons deux niveaux de fonctionnement logiciel pour le monde virtuel. Symétriquement aux  $I^2$  sensorimotrices, l'ordinateur doit gérer la partie logicielle temps réel (noyau temps réel et « drivers » pour les dispositifs matériels des interfaces), permettant une **réalisation physique du monde virtuel**. Cette dernière concerne la simulation basée sur les lois physiques (mécaniques, optiques, biomécaniques, etc.) agissant sur les objets et les êtres animés. Face aux  $I^2$  cognitives, la partie logicielle spécifique de réalité virtuelle doit gérer le **modélisation comportementale du monde virtuel**. Cette partie logicielle doit donc fournir des **Aides Logicielles Comportementales (ALC)**, associées aux processus cognitifs pour faciliter les  $I^2$  cognitives de la personne. Il est ainsi possible, via l'ordinateur, d'aider l'utilisateur à employer efficacement les interfaces. Il y a de multiples aides programmables et exploitables. Le concepteur doit y réfléchir pour proposer des aides efficaces, tout en respectant l'objectif des  $I^2$  fonctionnelles. Nous allons préciser les différentes possibilités d'aider à l'interfaçage comportemental : certaines ALC sont plus orientées sur une aide sensorimotrice (ALSM) et d'autres plus axées sur une aide cognitive (ALC<sub>Cog</sub>). Cette classification en ALSM et ALC<sub>Cog</sub> n'est pas rigoureuse car les comportements sensorimoteurs et cognitifs de l'Homme sont interdépendants. Elle a pour but premier de montrer l'étendue des possibilités d'aides concevables et programmables par le concepteur d'application RV, avec ou sans visiocasque.

### 2.2.1. Les Aides Logicielles SensoriMotrices

Les dispositifs d'interface ne sont pas parfaits dans la transmission des données. De même, le comportement sensorimoteur du sujet imposé en environnement virtuel est imparfait, du moins différent de celui dans le monde réel. Les Aides Logicielles SensoriMotrices servent donc à améliorer l'interfaçage, soit en augmentant la qualité des données transitant entre l'homme et l'ordinateur, soit en proposant une amélioration « psychophysique » des stimuli sensoriels ou des réponses motrices. Dans le premier cas de la qualité d'un signal amélioré, un exemple :

- le filtrage des mesures d'un capteur de localisation est souvent nécessaire pour supprimer les bruits des signaux qui peuvent, par exemple, gêner l'utilisateur dans la manipulation d'un objet virtuel. Si l'utilisateur décide d'arrêter le mouvement de l'objet, il ne faut pas que les bruits des mesures fassent bouger, même légèrement, l'image de l'objet, créant des scintillements de pixels. Mais le filtrage des données impliquent l'augmentation du temps de latence de la boucle « perception, décision, action ». Un optimum est souvent à rechercher.

Dans le second cas pour les ALSM, l'amélioration psychophysique des stimuli sensoriels ou des réponses motrices consiste par logiciel à permettre au sujet d'avoir un comportement sensorimoteur plus aisé dans l'environnement virtuel, par exemple :

- pour un utilisateur ayant un tremblement des mains élevé, il est efficace de filtrer les données des capteurs de localisation pour stabiliser ses actions dans l'environnement virtuel ;
- si le nombre de pixels dans le visiocasque est insuffisant pour lire un texte inscrit dans le monde virtuel, les lettres peuvent être exagérément agrandies ;
- Si le champ de vision d'observation dans le visiocasque est trop étroit pour bien percevoir ce que l'application RV exige, le champ de vision observé dans l'environnement virtuel peut être compressé. Mais cela implique une difficulté de coordination main et vision pour l'utilisateur, donc une incohérence sensorimotrice.

Les ALSM envisageables avec l'emploi d'un visiocasque seront détaillées au chapitre 8.

### 2.2.2. Les Aides Logicielles Cognitives

Les Aides Logicielles Cognitives de motricité et de perception, associées aux PCV, apportent des facilités au sujet lors de la réalisation d'une tâche à accomplir. Les ALCog peuvent être plus axées, soit sur une aide face aux incohérences sensorimotrices, soit sur une aide à l'utilisateur pour réaliser une tâche. Dans le premier cas, on peut citer les exemples suivants :

- on peut rajouter des contraintes spécifiques lors de la manipulation d'un objet par un capteur à 6 degrés de liberté (6ddl) et sans retour d'effort, impliquant des incohérences sensorimotrices lors de la collision de l'objet avec un autre. Pour aider l'utilisateur à obtenir la motricité désirée, par exemple poser un

objet sur une table, les mouvements de l'objet virtuel sont contraints. La face d'appui de l'objet se met parallèlement à la surface de la table quand elles seront proches l'une de l'autre. Ces contraintes irréelles sont appelées guides virtuels (de mouvement) ou effets de magnétisme ;

- pour le déplacement dans les couloirs d'un bâtiment, les trajectoires possibles peuvent être prédéterminées pour simplifier les commandes de déplacement de l'utilisateur. Car il n'est pas toujours nécessaire de laisser à ce dernier la possibilité de se mouvoir librement dans les couloirs virtuels d'un bâtiment. Ce principe de trajectoires prédéterminées peut être évidemment exploité dans d'autres types d'environnement virtuel.

Dans le second cas pour les ALCog, il faut détecter l'intention du sujet en train d'accomplir une tâche, par exemple :

- quand on détecte l'intention de l'utilisateur de se rapprocher d'un objet ou d'une zone de l'environnement virtuel, le déplacement virtuel de l'utilisateur peut être contrôlé pour atteindre automatiquement le lieu désiré sans intervention de l'utilisateur. Ceci peut être employé quand le déplacement virtuel est activé par un simple tapis roulant (1D) et par une manette indiquant les changements d'orientation : l'utilisateur ne peut pas réaliser tous les types de trajectoire (pas de pivotement du corps sur place, etc.) et rater l'endroit à atteindre sans une ALCog adaptée.

Pour résumer notre analyse de l'interfaçage que j'ai dénommé, « interfaçage comportemental » car basé sur un comportement humain.

**L'interfaçage comportemental** est l'ensemble de **l'interface matériel**, du **processus cognitif** mis en jeu et des **ALC** pour aider à son exploitation.

En pratique, les ALCog peuvent être déterminées en fonction des *affordances* des objets ou de l'environnement autour du sujet, comme nous venons de le présenter dans les exemples précédents. D'après la théorie de Gibson [Gibson J.J., 1979], une *affordance* représente les interactions possibles entre l'objet et le sujet. Ces interactions sont perçues par le sujet par rapport à l'idée qu'il se fait de la fonction de l'objet, plus qu'il ne perçoit les caractéristiques physiques et géométriques de l'objet ou de l'environnement. Il est donc plus utile de savoir par avance à quoi un élément de l'environnement va servir que d'avoir des notions précises de sa géométrie ou de ses caractéristiques physiques. Connaissant les *affordances* de tel objet, nous pouvons prévoir des ALCog qui aident le sujet à interagir efficacement avec cet objet. Par exemple, une interaction par reconnaissance gestuelle sur un objet peut être aidée par ses *affordances*. Cet exemple est présenté dans un de nos travaux de recherche [Jego J.F., *et al.*, 2013].

Il est nécessaire de rappeler au lecteur que la principale difficulté de l'immersion et de l'interaction du sujet provient des perturbations de(s) la boucle(s) « perception, décision, action ». Grâce à des choix judicieux des dispositifs d'interfaçage, des processus cognitifs et des ALC, les perturbations peuvent être surmontées pour permettre au sujet d'agir efficacement dans des boucles « perception, décision, action » artificielles, incluant des artefacts (voir la figure 1.7). Nous verrons dans le chapitre 8

tous les solutions susceptibles d'améliorer l'immersion et l'interaction de toute personne dans un EV via un visiocasque.

Pour résumer les objectifs de cette démarche de conception à partir de ce « **modèle 3I<sup>2</sup>** » :

- le niveau sensorimoteur doit être **transparent** pour l'utilisateur. Concrètement, ce dernier ne doit plus se préoccuper du visiocasque et des interfaces associées qui doivent se faire oublier ;
- les choix des processus cognitifs à mettre en œuvre doivent être fait **entre schème et métaphore**, avec ou sans substitution sensorielle. Ces choix doivent être compatibles avec les exigences des PCV déterminées.
- suivant les objectifs de l'application RV, **les PCV et leurs caractéristiques doivent être déterminées**

### 3. IMMERSION ET PRÉSENCE

Sans faire trop de développement sur les deux notions de Présence et d'Immersion, nous donnons quelques points de repères<sup>9</sup>. Si un utilisateur dans un EV fait appel un processus cognitif pour (inter)agir via des interfaces (I<sup>2</sup> cognitives), il a aussi à réfléchir aux actions qu'il doit entreprendre dans son application (I<sup>2</sup> fonctionnelles). Mais indépendamment de ces deux aspects cognitifs, l'utilisateur a le sentiment d'être présent plus ou moins fortement dans un environnement virtuel différent de son environnement quotidien. Etant présent dans cet environnement virtuel, est-ce qu'il s'y plaît, y a peur, ressent d'autres émotions, se sent ailleurs et téléprésent dans un autre lieu réel, mais simulé, etc. ?

Ce sentiment de présence dépend en partie de la notion d'immersion corporelle dans cet espace artificiel. Concrètement, suivant les applications RV, on peut envisager plusieurs degrés d'immersion du corps :

- **présence ou immersion visuelle**, qui correspond à l'observation visuelle dans un visiocasque (sans voir la représentation de ses mains), l'interaction n'étant réalisée que par le mouvement de la tête. C'est ce type de présence qui est proposée dans les vidéo 360 degrés et dans des applications RV où l'on n'a pour objectif que d'observer des lieux ou des produits, comme lors de la revue de projet dans l'industrie. Mais ce dernier cas se fait en général devant un grand écran et rarement via un visiocasque ;
- **présence ou immersion semi-corporelle**, l'utilisateur voit ses mains dans le visiocasque représentées virtuellement. Avec celles-ci, il peut manipuler des objets. Il est un peu plus plongé dans l'EV, comme disent certains. Ce cas est aussi envisageable sans visiocasque. Par exemple, l'observateur voit ses mains réelles à travers un écran semi-transparent placé devant lui, à hauteur de son buste, et des objets virtuels, qui sont affichés en vision stéréoscopique sur l'écran semi-transparent. Les objets sont manipulés par les mains de l'observateur, qui sont traquées ;

<sup>9</sup> La notion de présence se situe au niveau I<sup>2</sup> cognitives dans notre modèle 3I<sup>2</sup>.

- **présence ou immersion corporelle entière**, l'utilisateur se sent totalement immergé via son visiocasque. Il faut dans ce cas que la représentation du corps soit réalisée. L'utilisateur voit tout son corps et peut agir avec tous ses membres. Cela exige qu'en plus de ses mains, ses pieds soient aussi traqués. Ce cas est aussi envisageable sans visiocasque. Par exemple, l'utilisateur est immergé dans un CAVE, ce qui permet naturellement une présence corporelle totale. C'est le principal avantage d'un CAVE par rapport à un visiocasque, le deuxième avantage étant la grande qualité des images que ne peuvent pas rivaliser les visiocasques actuels<sup>10</sup>.

Mais peut-on parler « **d'immersion totale** » comme certains le prétendent dès qu'on exploite un visiocasque ? Il faut au moins que trois conditions soient réunies :

- Il faut que la représentation du corps du sujet soit correctement réalisée en temps réel dans l'environnement virtuel ;
- Il faut que l'immersion visuelle soit totale, donc avec un visiocasque ayant un champ de vision aussi grand que celui du sujet immergé ;
- Il faut que l'EV soit spatialement égal à l'ER pour que dans ce cas le sujet ait les mêmes sensations proprioceptives de son corps (musculaire et kinesthésique) et qu'ainsi tous ses sens créent un environnement cohérent, fusion de l'ER et de l'EV... car le corps du sujet est toujours dans l'ER ! Nous en reparlerons au chapitre 8.

Mais pour bien comprendre ce dernier point, il est souhaitable d'abord d'exposer en détail les sens humains, qui sont plus que cinq comme nous l'avons souligné. C'est l'objectif du chapitre suivant.

Des recherches ont été déjà entreprises et sont à poursuivre sur le sentiment de Présence, sous les aspects philosophiques [Nannipieri O., 2013], psychologiques [Mestre D. and J.L. Vercher, 2011] et techniques.

## 4. BIBLIOGRAPHIE

[Arnaldi B., Fuchs P. 2003] Arnaldi B., Fuchs P. et Tisseau J., *Chapitre 1 du volume 1 du traité de la réalité virtuelle*. Presses des Mines.

[Fuchs P., 1996] Fuchs P. *Les interfaces de la réalité virtuelle*. Presses des Mines.

[Fuchs P., et al., 1999] Fuchs P., Nashashibi F. and Lourdeaux D., "A theoretical approach of the design and evaluation of a virtual reality device". In *Virtual Reality and Prototyping'99*, pp.11-20, Laval, France.

---

<sup>10</sup> La qualité visuelle des visiocasques s'améliora dans l'avenir, mais parallèlement, on peut espérer aussi que le prix d'un CAVE diminue lorsque l'on pourra le concevoir avec des écrans plats et non par vidéoprojection impliquant un grand encombrement spatial.

- [Gibson J.J., 1979] J. J. Gibson. *The ecological approach to visual perception*. Lawrence Erlbaum Associates, Hillsdale, USA.
- [Jego J.F., et al., 2013] Jégo, J. F., Paljic, A., and Fuchs, P. “*User-Defined Gestural Interaction : a Study on Gesture Memorization*”. IEEE Symposium on 3D User Interfaces 3DUI, IEEE Computer Society, Orlando, USA (16, 17, March).
- [Mestre D., J.L. Vercher, 2011] Mestre D. and J.L. Vercher, chap. 5, “*Immersion and presence*”, book “*Virtual Reality: Concepts and Technologies*”, Fuchs P., Moreau G. and Guitton P. editor, 2011, CRC Press, 432 p.
- [Nannipieri O., 2013], Thèse d’Olivier Nannipieri, « Les paradoxes de la présence dans les environnements immersifs : de la réalité à la réalité virtuelle », Thèse de doctorat en Sciences de l’information et de la communication, Université de Toulon.
- [Nannipieri O., Fuchs P., 2009] Nannipieri O. et Fuchs P., « Pour en finir avec la réalité : une approche socio-constructiviste de la réalité virtuelle », *Revue des Interactions Humaines Médiatisées*, 10, 1, pp.83-100.
- [Piaget J. et al., 1979] Piaget J. et Chomsky N. *Théories du langage, théories de l’apprentissage*. Seuil.
- [Rabardel P., 1995] Rabardel P. *Les hommes et les technologies, approche cognitive des instruments contemporains*. Armand Colin.



---

# Quelques notions sur les sens humains

---

### 1. INTRODUCTION

Nous avons vu au chapitre précédent que la conception d'une application de réalité virtuelle se fait à partir de trois niveaux d'Immersion et d'Interaction, dont le premier est le niveau des I<sup>2</sup> sensorimotrices. Les techniques de la réalité virtuelle font intervenir des interfaces comportementales (sensorielles ou(et) motrices) pour coupler physiquement l'ordinateur à l'homme. La conception de telles interfaces a pour objectif de réaliser un dispositif matériel ayant des caractéristiques métrologiques performantes qui doivent, si possible, correspondre aux capacités psychophysiologiques maximales des sens ou(et) des réponses motrices de l'homme<sup>1</sup>. Dans ces conditions, il est indispensable d'avoir une vision claire et précise du comportement sensorimoteur de l'homme dans le monde réel, aussi bien pour analyser efficacement les interfaces comportementales que les I<sup>2</sup> sensorimotrices. C'est pourquoi nous présentons les caractéristiques des organes sensoriels spécifiques (vision, ouïe et toucher) et des récepteurs proprioceptifs, permettant la localisation spatiale, l'équilibre, les mouvements et les déplacements.

L'homme perçoit son environnement par l'intermédiaire de ses sens. La réception d'un stimulus sensoriel est le point de départ du transfert de l'information dans l'organisme par les voies nerveuses. Le message nerveux afférent provenant du récepteur sensoriel est véhiculé vers le système nerveux central (SNC), composé du cerveau, cervelet, tronc cérébral et moelle épinière. Après intégration et traitement de l'information dans ces centres, ceux-ci transmettent le message nerveux efférent aux organes effecteurs (muscles squelettiques, muscles oculaires, muscles des cordes vocales, etc.). Ceux-ci réalisent le comportement moteur, se traduisant par des mouvements, sauf pour le cas particulier des muscles des cordes vocales permettant la parole.

Tout influx sensitif naît au sein de récepteurs spécialisés. Ils peuvent être classés en deux groupes. Les uns, répartis dans tout notre organisme, renseignent sur l'état du corps et sur sa sensibilité générale (somesthésie). D'autres, spécifiques à un sens, sont situés dans l'organe sensoriel correspondant. Ce sont par exemple les

---

<sup>1</sup> Actuellement, seules les interfaces auditives, haut-parleurs et microphones, ont des capacités correspondant aux caractéristiques psychophysiologiques maximales de l'ouïe et de la parole.

photorécepteurs de la rétine pour la vue. Les récepteurs extéroceptifs, sensibles à l'extérieur, nous renseignent sur notre environnement. Les récepteurs proprioceptifs sont sensibles aux actions sur l'organisme : le mouvement du corps, sa position dans l'espace et les forces exercées sur les muscles. Il y a cinq organes sensoriels spécifiques mais plus que cinq sens avec tous les récepteurs proprioceptifs, plus que nécessaires<sup>2</sup>, et que nul ne doit ignorer pour immerger un être humain dans un environnement virtuel. Mais tous ses sens sont-ils exploités bien distinctement par le système nerveux central ?

Des chercheurs [Sherrington, 1906] ont proposés une distinction stricte entre extérocepteurs, propriocepteurs et intérocepteurs. Les extérocepteurs (les yeux, les oreilles, le nez, la bouche et la peau) nous donnent accès aux changements dans le milieu externe et seraient la base de la perception. Les propriocepteurs (extrémités des organes, des muscles, articulations, oreilles internes) nous donnent accès aux sensations relatives à la position et aux mouvements du corps : ils donnent accès à la sensation du mouvement. Les intérocepteurs (nerfs aboutissant dans les viscères) donnent des sensations vagues sur les organes internes. Contre la position défendue par Sherrington, Gibson [Gibson, 1966] suggère que la sensibilité à l'action et la sensibilité au mouvement ne dépendent pas de récepteurs spécialisés. Les yeux, les oreilles ou la peau peuvent enregistrer le comportement d'un individu aussi bien que les événements extérieurs. Les yeux, par exemple, enregistrent les mouvements de la tête (avant, arrière et rotation) au moyen des mouvements de la lumière ambiante, le flux optique (extéroception). De la même manière, des récepteurs dans les articulations ou les oreilles internes (voir plus loin) peuvent enregistrer les mouvements imposés au corps aussi bien que des mouvements initiés par l'individu. Ainsi, la proprioception, considérée comme le moyen d'obtenir des informations sur nos propres mouvements, ne dépend pas nécessairement que des propriocepteurs. Et l'extéroception, considérée comme le moyen d'obtenir des informations sur les événements extérieurs, ne dépend pas nécessairement des extérocepteurs.

La perception a été envisagée, selon une conception ancienne, comme un processus dans lequel le système sensoriel recevait passivement des stimulations, puis traitait ces informations afin d'identifier objets et événements sous la forme d'une représentation interne et, enfin, l'action était envisagée comme une sortie. Mais actuellement, une conception active de la perception est privilégiée. Pour J. Gibson [1966], les sens sont des systèmes perceptifs, essentiellement actifs, qui extraient des informations et sont sources de connaissances sur le monde. Pour la vision, percevoir, c'est extraire, grâce aux mouvements (des yeux, de la tête...), des informations en détectant des propriétés structurelles invariantes des stimuli [M. Auvray, 2006].

---

2 J'y ai pensé bien souvent lors de mes randonnées en courant : d'innombrables stimuli arrivent à mes centres nerveux par différents canaux sensoriels : la vision (1), l'ouïe (2), le toucher (3) dont les pressions plantaires à mes pieds, la proprioception musculaire (4) m'informant des efforts sur mes muscles et mes tendons, le positionnement de mon corps par des récepteurs au niveau de mes articulations (5) et la détermination du mouvement de ma tête par les mesures d'accélération (6) via les systèmes vestibulaires ; tous ces stimuli devant être perçus cohérents.

Des caractéristiques générales de tous les sens peuvent être énoncées. Il existe un seuil minimum en dessous duquel toute stimulation sur un organe sensoriel ne provoque aucun effet. Au-dessus du seuil, la variation minimum perceptible d'un stimulus est proportionnelle à la valeur absolue de l'intensité du stimulus (loi de Weber). Qualitativement, cette loi veut dire dans le cas de la vision : pour une faible intensité lumineuse perçue par l'œil, une petite variation d'intensité peut être détectée tandis que pour une forte intensité lumineuse perçue par l'œil, il faut une grande variation d'intensité pour qu'elle soit détectée. En plus, un stimulus doit avoir une durée minimale pour être perçu. À l'opposé, si une stimulation est maintenue en permanence, il y a généralement disparition ou atténuation de la sensation (phénomène d'adaptation), sauf pour la douleur et certaines fibres nerveuses.

Par ailleurs, quand nous parlerons de la résolution des interfaces, nous verrons l'importance de la densité des récepteurs et de la variation locale de cette dernière dans les organes sensoriels. La précision des sens humains est aussi importante à connaître pour réaliser des interfaces sensorielles adéquates. En général, la précision absolue des sens est faible comparativement à la précision relative, l'homme ayant une grande capacité de comparaison entre deux stimuli, par exemple :

- il est difficile de définir une couleur d'un objet seul, mais il est plus aisé de percevoir une légère différence de deux couleurs observées simultanément ;
- la perception absolue de profondeur est difficile à estimer comparativement à la détection d'une faible différence de profondeur entre deux objets voisins ;
- il est délicat de quantifier la température absolue de l'air ou de l'eau en contact avec notre peau. En revanche, nous percevons facilement une légère différence de température entre deux fluides observés simultanément.

La fréquence maximum  $F_m$  de variation d'un stimulus sensoriel qui soit perceptible par l'homme est aussi une caractéristique importante à connaître. Les interfaces sensorielles doivent avoir, quand cela est techniquement possible, une bande passante de 0 Hz à  $F_m$ . Pour les interfaces visuelles dont les visiocasques, il est souhaitable que les images soient affichées sur les écrans à plus de 30 images par seconde, 60 en stéréoscopie, ce qui ne pose plus trop de difficultés maintenant.

Dans la suite de ce chapitre, nous ne présenterons pas entièrement les connaissances<sup>3</sup> sur tous les sens. Nous nous contentons de détailler certains points sur les principaux sens à connaître, à notre avis, pour exploiter correctement un visiocasque dans une application RV. Le paragraphe sur le système visuel sera évidemment plus développé.

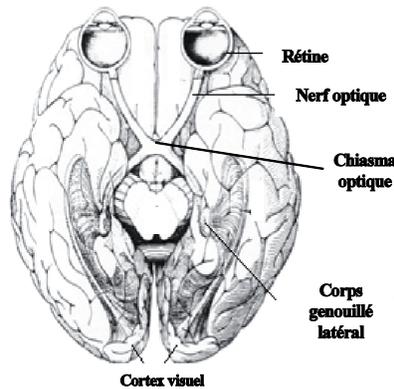
---

3 Il faudrait plus d'un livre entier sur ce sujet.

## 2. LE SYSTÈME VISUEL

### 2.1. L'ensemble du système visuel

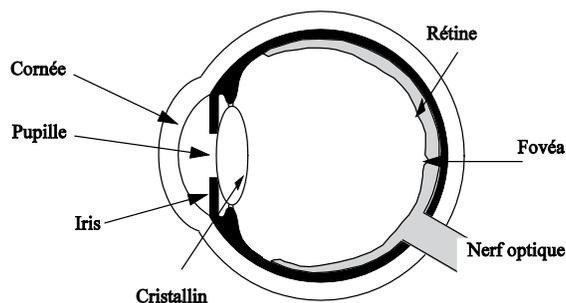
Les deux yeux, récepteurs sensoriels du système visuel, participent à l'observation de l'espace. Les nerfs optiques sortant des yeux se regroupent au niveau du chiasma optique où les fibres temporales gauches rejoignent les fibres nasales droites et inversement. Les fibres du nerf optique sont alors dirigées vers le corps genouillé latéral. De celui-ci, l'information est envoyée au cortex occipital. La Figure 2.1 montre le trajet de l'information visuelle depuis l'œil jusqu'au cortex visuel.



**Figure 2.1 :** le trajet de l'information visuelle depuis l'œil jusqu'au cortex visuel.

### 2.2. L'œil

La fonction des yeux est de canaliser la lumière, de longueurs d'ondes entre 400 et 700 nm, émise ou réfléchie par un objet pour créer une image nette qui s'imprime sur la partie de l'œil couverte de récepteurs sensoriels : la rétine. L'œil est constitué d'une succession de milieux transparents qui jouent le rôle d'une lentille convergente, dont la focale globale peut varier par modification de la courbure du cristallin (Figure 2.2).



**Figure 2.2 :** la structure anatomique de l'œil

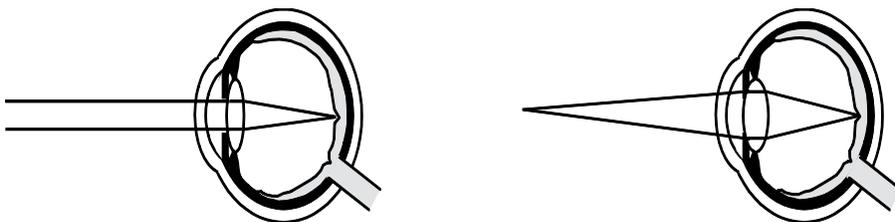
Schématiquement, l'œil est une chambre noire de géométrie sphérique à l'entrée de laquelle se situent plusieurs dioptries et à l'arrière de laquelle se trouve la structure réceptrice :

- la pupille représente le diaphragme du système. Les rayons lumineux sont ensuite projetés sur la rétine qui fait office d'écran sphérique ;
- la cornée est une membrane fibreuse, transparente, qui constitue la lentille principale du système optique. Elle est le pôle antérieur de l'œil et l'axe optique passe en son centre. Son diamètre horizontal est de 11,5 à 12 mm ;
- le cristallin est une lentille biconvexe transparente placée entre l'iris et le corps vitré.

Pour explorer l'espace visuellement, apprécier l'emplacement des objets, naviguer dans l'espace, lire et travailler sur écran, un bon fonctionnement des mouvements de nos deux yeux est indispensable. Plusieurs mouvements oculaires permettent de réaliser ces tâches : les mouvements pour modifier la direction du regard et les mouvements de vergence pour modifier la profondeur de la vision. Les mouvements pour modifier la direction du regard lors de l'observation d'une image sont exécutés par des saccades oculaires qui sont les mouvements les plus rapides du corps humain (leur vitesse peut atteindre  $500^\circ/\text{s}$ ) et qui durent en général moins de 100 ms.

### 2.3. Accommodation et vergence

En regardant un objet à moins de cent mètres environ, l'image réfléchi deviendrait floue sur la rétine si l'œil ne possédait le pouvoir d'accommoder automatiquement. Sous l'action des muscles ciliaires, la puissance du cristallin varie et permet de mettre au point sur la rétine pour voir des objets proches ou éloignés (Figure 2.3) : c'est le phénomène d'accommodation. Le cristallin est le seul dioptré de la chaîne visuelle dont la puissance soit variable. Tous les rayons lumineux sont déviés vers la fovéa qui est le point central de la rétine. Le niveau d'accommodation est réglé pour obtenir une image nette.



**Figure 2.3 :** le phénomène d'accommodation par changement de forme du cristallin (schéma optique simplifié)

Les muscles des globes orbitaux permettent d'orienter les deux yeux en les faisant converger vers le point de l'espace observé. C'est le phénomène de vergence (ou convergence). Comme l'accommodation, la vergence des yeux se fait inconsciemment, sauf si la personne veut loucher. Les mouvements des yeux dans les globes orbitaux,

pour changer le point de vergence ou(et) pour suivre un objet en déplacement, peuvent être très rapides (maximum de 500°/s environ). L'accommodation et la vergence des yeux étant toutes deux liées à la profondeur de l'objet que l'on observe, il en découle une relation entre celles-ci. L'accommodation entraîne la vergence et réciproquement. Si on oblige une personne à accommoder, elle aura tendance à converger et réciproquement. Cette relation est plus ou moins forte suivant les personnes. Pour mesurer le lien entre l'accommodation et la vergence, on oblige la personne à accommoder (facteur A) et on mesure la vergence accommodative (facteur AC) obtenue. Le rapport AC/C de vergence accommodative est compris en moyenne entre 3 et 4.

Cette correspondance n'est pas innée chez l'homme mais est maîtrisée expérimentalement et inconsciemment par l'enfant. Cette relation naturelle sera mise en défaut artificiellement lorsque l'on regarde des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait converger les axes optiques sur lui tandis que les yeux s'accommodent sur l'écran. Cette modification de la relation vergence-accommodation impose une incohérence sensorimotrice et est susceptible de créer des difficultés visuelles et de la fatigue pour l'utilisateur. Ce problème sera étudié dans le paragraphe sur «La restitution visuelle stéréoscopique» dans les chapitres suivants.

## 2.4. La rétine

La rétine est le lieu de la transformation de l'énergie lumineuse en signal électrochimique discret. Ce signal sort de l'œil par le nerf optique. Il existe deux types de photorécepteurs, les cônes et les bâtonnets, de répartition très différente sur la rétine :

- les cônes, concentrés à la fovéa, sont sensibles à la longueur d'onde : ils sont de trois types car respectivement sensibles à des longueurs d'onde centrées sur 560, 530 et 420 nm. L'information de couleur est produite au niveau cortical par la comparaison des informations de plusieurs récepteurs contenant différents photopigments ;
- les bâtonnets sont, à l'inverse, très peu sensibles à la couleur : présents presque partout dans la rétine, ils sont absents au centre de la fovéa, et très denses à environ 20° d'angle visuel.

Même si les bâtonnets sont beaucoup plus nombreux que les cônes (100 millions contre 6 millions), ce sont les cônes qui contribuent en majorité à l'information transmise aux structures profondes du système visuel. Grâce à ses cônes et ses bâtonnets, le système visuel peut s'adapter à l'intensité lumineuse ambiante et couvrir une gamme d'intensités de l'ordre de 7 unités logarithmiques. Sous très faible éclairage, seuls les bâtonnets sont sensibles. Sous faible éclairage, les cônes sont couplés positivement à leurs voisins pour augmenter la surface du détecteur de photons. L'augmentation de sensibilité se fait au détriment de la résolution du système. En revanche, quand l'intensité lumineuse croît, les photorécepteurs sont découplés pour obtenir une résolution maximale.

La rétine est composée de plusieurs couches de neurones qui recouvrent le fond de l'œil. Ces couches forment un tapis de cellules connectées horizontalement au sein d'une même couche, et verticalement d'une couche à l'autre créant une architecture neuronale fortement pyramidale. En effet, si le nombre de photorécepteurs est de l'ordre de 100 millions, le nombre des cellules ganglionnaires, dont les axones constituent le nerf optique, est de l'ordre du million. Les cellules ganglionnaires sont activées en fonction des contrastes entre les stimuli lumineux reçus par les photorécepteurs. Les axones des cellules ganglionnaires sortent toutes de la rétine dans la même région dépourvue de photorécepteurs appelée tâche aveugle ou papille optique.

## 2.5. La perception visuelle de mouvement, la vection

Une scène visuelle en mouvement peut induire pour un observateur immobile, la sensation de mouvement de son propre corps. Ce phénomène est appelé vection. L'exemple le plus connu en environnement réel est le cas où le voyageur dans un train à quai voit par la vitre du wagon un autre train en mouvement lent sur une voie ferrée voisine. Le voyageur aura l'impression que son train démarre dans l'autre sens. Ceci est dû au traitement neuronal des flux optiques sur les rétines qui détectent un mouvement. Cette expérience n'est vécue que si le mouvement de l'autre train est lent, au démarrage par exemple, car les systèmes vestibulaires, voir plus loin, ne détecteront pas l'absence de faible accélération qui aurait lieu si le train du voyageur démarrerait. Ce n'est plus le cas si un TGV passe à grande vitesse une voie ferrée voisine car il n'y a plus d'ambiguïté (l'accélération instantanée du corps serait très importante et automatiquement détectée par les systèmes vestibulaires).

Le phénomène de vection est évidemment largement exploité en environnement virtuel : les mouvements des images font croire à l'observateur qu'il se déplace dans l'environnement virtuel. Nous en reparlerons au chapitre 7.

## 2.6. La perception visuelle de la profondeur

### 2.6.1. Perception cognitive par indices monoculaires

Il faut d'abord bien comprendre que la perception de la profondeur d'un monde tridimensionnel ne découle pas seulement de la vision binoculaire. Avec seulement un œil, l'homme interprète l'image reçue et en déduit des notions de profondeur. Cette interprétation se fait inconsciemment par traitement cognitif grâce à l'expérience apprise dès le plus jeune âge. Au stade précoce du traitement visuel, la décomposition et le traitement de l'information visuelle par différents canaux permettent d'utiliser une série variée d'indices pour la perception de l'espace tridimensionnel, donc de la profondeur. Comme nous l'avons écrit au premier chapitre, il est très aberrant d'employer les termes «film 3D» pour parler des films stéréoscopiques ou des films en relief ainsi que de «télévision 3D», créant la confusion auprès du public... La vision binoculaire et donc les images stéréoscopiques ne font qu'augmenter la perception du

relief. Ceci explique en grande partie l'échec, que nous avons prévu, du faible usage de la télévision dite faussement «3D».

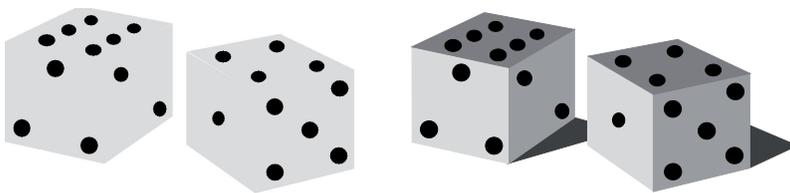
Les indices se décomposent en deux grandes catégories :

- les indices proprioceptifs (dus aux actions des muscles orbitaux et ciliaires), composés de l'**accommodation** et de la **convergence** : ces indices sont fiables à faible distance (quelques mètres) ;
- les indices visuels, composés des **indices binoculaires** et des **indices monoculaires**.

Les indices proprioceptifs, accommodation et convergence, sont ajustés par le système visuel. C'est par la valeur de commande des muscles orbitaux et ciliaires que le cerveau a une perception proprioceptive de la profondeur de l'objet observé. Rappelons que la vision naturelle humaine est **une vision active**. Les yeux sont presque toujours en mouvement pour observer l'environnement et les cristallins se déforment pour permettre l'accommodation. En revanche, un appareil photographique observe en vision passive l'environnement réel. La profondeur de champ de toute image photographique est un indice de profondeur qui correspond partiellement au phénomène d'accommodation. On peut donc utiliser des effets de profondeur de champ sur des images de synthèse pour donner plus de «relief» à l'espace observé.

En vision monoculaire, les indices monoculaires, appris inconsciemment, permettent d'appréhender la troisième dimension avec un œil, même si la perception de la profondeur est mieux quantifiée avec la vision binoculaire. Les indices monoculaires peuvent être classés de la façon suivante :

- Ombres et lumière : Les variations de lumière et les ombres sur les objets permettent d'avoir une meilleure perception de la forme tridimensionnelle de ceux-ci (Figure 2.4).



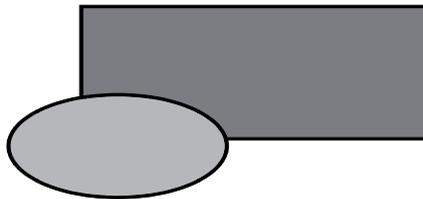
**Figure 2.4 :** variations de lumière et les ombres sur les objets

- Dimensions relatives : Chaque objet envoie au fond de l'œil son image dont les dimensions sont proportionnelles à celle de l'objet et décroissante en fonction de la distance par rapport à l'œil. Le cerveau connaît les dimensions «normales» des objets réels. Il peut donc en déduire leurs distances approximativement. De même l'estimation de profondeur est facilitée pour une série d'objets, réels ou imaginaires, de formes identiques : ils sont *a priori* perçus à taille identique et positionnés à différentes distances dans un espace 3D (Figure 2.5).



**Figure 2.5 :** variations de fleurs de même taille à profondeur différente

- Interposition ou occultation : Un objet peut cacher partiellement un autre objet placé derrière lui, ce qui permet de positionner relativement ces objets en profondeur (Figure 2.6). C'est par un raisonnement cognitif que le cerveau interprète cette image : il perçoit spontanément une ellipse devant un rectangle et non deux formes adjacentes sur un même plan.



**Figure 2.6 :** occultation d'un objet par un autre

- Gradient de la texture d'une surface : La texture d'une surface est perçue plus nettement si la surface est positionnée à faible profondeur. Ce gradient de la texture donne une information complémentaire de profondeur.
- Variation de visibilité dans une scène extérieure : La visibilité d'une scène extérieure est décroissante en fonction de l'épaisseur de la couche atmosphérique, donc de la profondeur.
- Parallaxe due aux mouvements : Quand un observateur se déplace, les images rétiniennes d'objets immobiles ont des mouvements relatifs dépendant des distances de ceux-ci. Il en est de même quand des objets se déplacent relativement entre eux. En pratique, on peut utiliser cette règle pour rendre compte du relief par rotation d'un produit ou de toute une scène 3D visualisés sur un écran monoscopique. On peut aussi modifier les images sur l'écran, monoscopique ou stéréoscopique, en fonction du déplacement de l'observateur devant l'écran. Mais cette solution est plus contraignante en pratique si on se trouve devant un écran car il faut détecter en temps réel la position de la tête de l'observateur. En revanche, c'est un des grands intérêts d'exploiter un visiocasque qui permet en temps réel le changement de point de vue et donc une meilleure perception 3D. Le lecteur doit bien noter que cet indice monoculaire est plus puissant pour percevoir la profondeur (3D) que la vision stéréoscopique, comme nous avons pu le vérifier de nouveau dans nos travaux [Leroy, 2009]. Faites l'expérience de

suite en vision naturelle : tourner autour d'un objet en fermant un œil et observer le sans bouger votre tête avec vos deux yeux ! Mais dans cette expérience, un autre indice intervient : la perspective, paragraphe suivant.

- Perspective : C'est bien sûr la règle la plus efficace pour percevoir le monde en relief sur un écran monoscopique. Celle-ci est employée en peinture depuis l'époque de la Renaissance pour visualiser sur une surface plane l'espace tridimensionnel. À noter de nouveau qu'en passant d'images planes (sans perspective) dans les jeux vidéo aux images avec perspectives sur écran monoscopique, les concepteurs de jeux vidéo parlaient de jeux vidéo «3D». Il existe plusieurs types de perspectives (cavalière, géométrique, photographique, artistique). La perspective cavalière fut utilisée en dessin industriel. Elle est plus éloignée de la vision réelle que la perspective photographique mais tellement plus facile à réaliser à la main. Il n'y a pas de point de fuite, mais des «fuyantes» parallèles, comme sur la Figure 2.7.

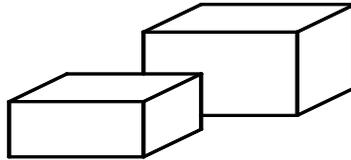


Figure 2.7: perspectives

**Remarque :** Le lecteur peut noter que nous venons d'utiliser dans le livre ces règles, qui sont regroupées sur le graphe de la Figure 2.8. Vous pouvez fermer un œil, la perception de la profondeur est inchangée ! Ces règles sont maintenant couramment employées en images de synthèse pour donner une représentation tridimensionnelle sur un écran d'ordinateur. Dans Open GL par exemple, vous pouvez calculer vos images en perspective cavalière (orthographic projection) avec la commande `gluOrtho2D` ou en perspective photographique<sup>4</sup> (perspective projection) avec la commande `gluPerspective`.

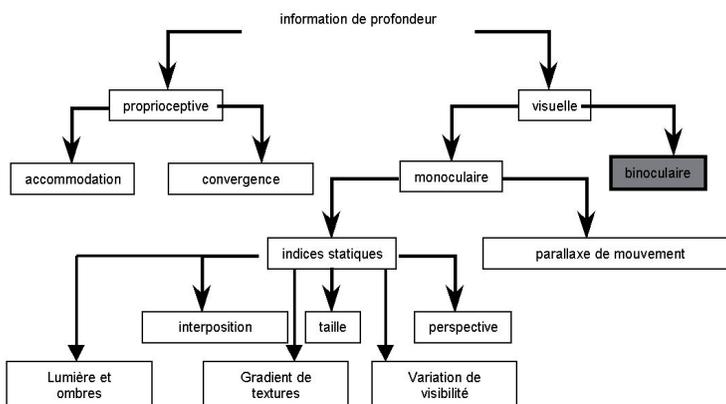
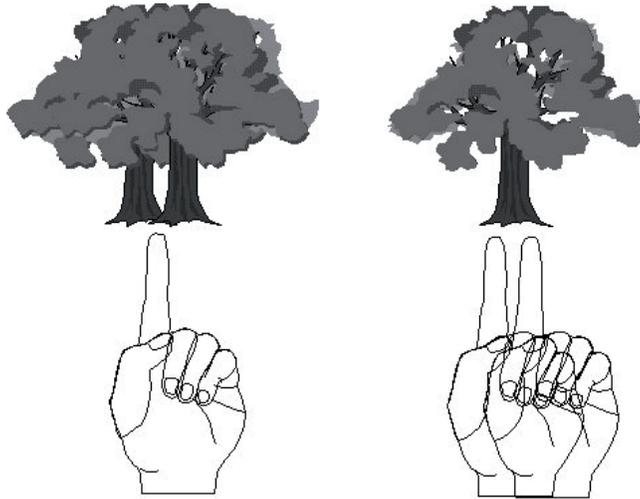


Figure 2.8: tous les indices visuels de perception de la profondeur

4 Plus précisément, avec un modèle sténopé, donc sans faire intervenir les distorsions des optiques.

### 2.6.2. Convergence et disparité rétinienne

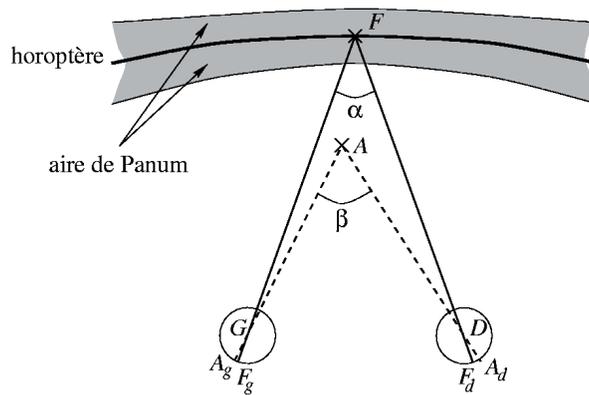
Précisons le phénomène de convergence et de son incidence sur les images rétiniennes : en regardant fixement un de vos doigts placé à 20 cm devant vous, vous voyez une image dédoublée de tout objet situé en arrière-plan. Inversement si vous fixez du regard un objet placé derrière votre doigt, vous voyez ce dernier se dédoubler : le cerveau n'arrive pas à fusionner les deux images différentes de votre doigt, la disparité rétinienne étant trop grande (Figure 2.9).



**Figure 2.9 :** fixation sur le doigt ou fixation sur l'arbre

Du fait du recouvrement (partiel) des champs visuels des deux yeux et de la convergence des axes optiques, nous possédons deux vues légèrement différentes de la même scène permettant d'extraire une information de profondeur. Les axes optiques convergeant sur un objet, le cerveau obtient donc des images dédoublées des autres objets éloignés de ce dernier. Que fait le cerveau de ces images dédoublées ? En pratique, il n'en tient pas compte mentalement. En particulier, quand vous regardez, même au loin, vous voyez toujours deux images floues de votre nez mais vous n'y prêtez pas attention... sauf maintenant... cela va peut-être vous perturber si vous y faites attention ! Dans le cas de la vision naturelle, les yeux n'accommodent pas sur les images dédoublées (les objets correspondants sont trop éloignés ou trop proches). Mais cela n'est pas identique sur un écran où toutes les images des objets sont nettes et à même distance. Nous détaillerons ce problème dans paragraphe sur « La restitution visuelle stéréoscopique » du chapitre 3. Normalement, la convergence des axes visuels est commandée par le cerveau, mais on peut physiquement faire diverger ces axes et voir les images se dédoubler : regardez ce texte et appuyez avec un doigt sur un œil pour le faire tourner, les images se dédoublent, la fusion n'est plus possible.

Mais analysons maintenant plus géométriquement et optiquement ce qui se produit en vision naturelle à partir de la Figure 2.10 où un observateur regarde le point F.



**Figure 2.10 :** perception de la différence de profondeur entre le point F de fixation et le point A à partir de la disparité

La disparité rétinienne horizontale quantifie la notion de différence entre les images en mesurant des écarts. La figure 2.10 présente une illustration de la disparité rétinienne : pour le point de fixation F nécessitant un angle de convergence  $\alpha$ , la disparité  $d$  au point A sous-tendant un angle  $\beta$  est définie par :  $d = \beta - \alpha$  qui correspond au double de l'angle  $DF_dA_d$  quand A et F sont sur la normale à la droite DG. Géométriquement, les disparités ne sont suffisamment différentes que si les objets observés sont proches des yeux. La perception binoculaire de la profondeur n'est donc efficace qu'à petites distances (quelques mètres).

### 2.6.3. Stéréopsie et diplopie

Pour un point de fixation donné, l'ensemble des points de l'espace se projetant en des positions rétiniennes correspondantes forme l'horoptère (sur la figure 2.10), les points perçus à même profondeur que le point de fixation F). L'horoptère théorique est appelé cercle de Vieth–Müller : c'est le cercle qui passe par le point de fixation et le point nodal de chacun des yeux. La détermination de l'horoptère empirique est basée sur différents critères selon la technique de mesure effectivement choisie. Les points situés en avant de l'horoptère sont dits à disparité croisée ou convergente, alors que ceux situés en arrière sont à disparité directe ou divergente.

Si les images rétiniennes sont proches, avec de petites disparités horizontales, chaque couple de stimuli binoculaires est perçu<sup>5</sup> par le cerveau en une seule entité par fusion des points qui se situent dans l'aire de Panum (la zone grisée sur la figure précédente) : c'est la **stéréopsie**. Les objets sont perçus en profondeur en avant ou en arrière du point de fixation par stéréopsie. Au-delà de l'aire de Panum, les stimuli de chaque œil sont perçus séparément. Ils sont dédoublés : c'est la **diplopie**.

<sup>5</sup> Le cerveau traite les deux « images » rétiniennes pour percevoir son environnement mais ne reconstruit pas une pseudo « image tridimensionnelle ».

#### 2.6.4. Mécanismes neurophysiologiques de la perception de la profondeur

Après la mise en correspondance, la profondeur peut être estimée. Ce problème est loin d'être trivial puisque la disparité est un indice ambigu. En effet, la même valeur de disparité peut être associée à des distances différentes, selon la convergence du regard. Les informations de disparité doivent donc être calibrées en fonction de la convergence pour estimer la profondeur véritable : ce processus est nommé constance de profondeur. L'activité des cellules corticales traitant la disparité doit donc être extrêmement spécialisée et modulée par la distance de fixation. Ainsi, Trotter note que lorsqu'on modifie la distance réelle de profondeur, l'activité des neurones change de façon remarquable. Les neurones préfèrent toujours la même disparité rétinienne, par exemple « plus près » ou « plus loin » que le point de fixation. Mais leur niveau d'activité est en grande partie sous la dépendance de la distance de l'objet. Trotter démontre aussi que la perception cohérente de l'espace en trois dimensions est le résultat de la combinaison, au niveau d'une population de neurones dans le cortex visuel primaire, des messages issus de la rétine et d'informations extra-rétiniennes sur la position des yeux. L'information la plus probable est le signal proprioceptif de la musculature extrinsèque des yeux [Trotter, 1995].

Pour terminer ce long paragraphe sur la perception visuelle de la profondeur, nous pouvons faire la remarque synthétique et surprenante<sup>6</sup> sur le phénomène de la vision « 3D » : d'un objet réel 3D, l'observateur recueille deux « images 2D » sur le fond de ses deux rétines. Son système visuel en extrait une perception 3D. En vision artificielle en relief sur un ou deux écrans, nous affichons deux images 2D calculées à partir d'un modèle informatique 3D de l'objet. Les rétines des yeux reçoivent également deux images 2D qui sont perçues par le système visuel en 3D ! Ce n'est donc pas le cinéma soi-disant « 3D » qui réalise techniquement la perception tridimensionnelle mais seulement le cerveau du spectateur avec l'ensemble des indices monoculaires et binoculaires, ces derniers n'étant pas indispensables... mais très attrayants !

#### 2.7. Les caractéristiques psychophysiques de la vision

Nous précisons dans ce paragraphe les caractéristiques utiles pour la compréhension d'une interface visuelle : Celles correspondant aux efforts entrepris actuellement pour améliorer les interfaces visuelles, dont les visiocasques :

- amélioration de la définition de l'image ;
- augmentation du champ de vision ;
- vision stéréoscopique ;
- immersion du regard.

---

6 Pour ceux qui ignorent les principes de la vision humaine.

### 2.7.1. Les acuités visuelles

L'acuité visuelle (monoscopique) n'est pas homogène dans tout le champ de vision. Pour un œil emmétrope (normal), l'acuité monoculaire est très grande pour un cône centré de 2° d'angle, conséquence de la répartition des cônes rétinien limitée à la fovéa de l'œil. La valeur minimum de l'angle sous lequel deux points sont vus séparément est fonction du stimulus observé :

- une fente éclairée sur fond noir: 30 « d'angle » ;
- deux points clairs sur fond noir: 1 ' d'angle ;
- deux points noirs sur fond clair: 2 ' d'angle.

La valeur moyenne (1') est en accord avec la densité des photorécepteurs au niveau de la fovéa. Un angle de 1 minute correspond à la vision de deux points séparés de 0,1 mm pour une distance de 35 cm entre l'image et les yeux. Pour respecter ces caractéristiques, un écran situé à cette distance et large de 25 cm doit afficher horizontalement 2500 pixels. On observe donc que maintenant les écrans d'ordinateur de qualité ont une définition d'image suffisante, ce qui n'était pas le cas il y a encore quelques années. Ces valeurs numériques sont établies pour un écran donnant un champ de vision étroit (40 degrés). Dans un visiocasque où le champ de vision est plus grand, la résolution de (ou des deux) écran(s) n'est pas encore assez satisfaisante au regard des capacités du système visuel humain. La valeur moyenne (1 minute d'angle) de l'acuité visuelle doit être retenue pour juger de la qualité de la résolution d'un écran ou d'un visiocasque en situation dans un dispositif RV. À noter que si la qualité d'un écran, par rapport à l'acuité visuelle humaine, dépend d'abord de sa résolution (en nombre de pixels), elle résulte aussi de la luminosité et du contraste de l'écran. Nous en reparlerons.

L'acuité visuelle binoculaire permet de détecter une différence de profondeur DP entre deux plans. Elle est fonction de la distance P de profondeur du point observé. Une formule approximative<sup>7</sup> est :

$$DP = 0.001 P^2$$

Cela signifie que si un objet est situé à une distance de 10 m de l'observateur, ce dernier ne peut percevoir le relief (la différence de profondeur) que pour deux plans distants d'au moins 0,1 m. Pour un objet à 1 m, le relief n'est perceptible que pour des plans séparés de 1 mm au minimum. En conclusion, la vision en relief naturelle n'est relativement performante que pour de faibles distances puisqu'elle est fonction de l'inverse au carré de la distance de profondeur. Étant donné que cette vision procure des fatigues oculaires, ces considérations devraient inciter les concepteurs à ne pas exploiter la vision stéréoscopique dans toutes les situations, mais seulement pour des cas bien justifiés. La vision stéréoscopique sera surtout utile pour les objets proches de la tête de l'observateur et aussi la représentation virtuelle de ses mains, si cette dernière

<sup>7</sup> Les explications détaillées sur l'acuité stéréoscopiques sont présentées dans le chapitre 3 du volume « L'homme en environnement virtuel » du *Traité de la réalité virtuelle*.

est proposée. On peut aussi exagérer artificiellement la vision stéréoscopique mais dans ce cas, il n'y a plus de correspondance exacte entre la perception de profondeur due aux indices monoculaires et celle due aux indices binoculaires (voir le chapitre 8).

### 2.7.2. Les champs de vision

Les deux yeux participent à l'observation de l'espace. La région de l'espace vue à chaque instant par l'un ou l'autre œil est définie comme le champ visuel. Les points de l'espace qui se situent dans la zone de recouvrement binoculaire sont acquis simultanément par l'œil droit et l'œil gauche (en anglais, «overlap» pour les visiocasques). Les caractéristiques du champ visuel pour un œil immobile sont approximativement :

- Champ Visuel, tête et œil immobiles :
- Un œil, horizontal : axe-tempe  $90^\circ$ , axe-nez  $50^\circ$  à  $60^\circ$
- Deux yeux, horizontal :  $180^\circ$
- vertical : haut  $45^\circ$ , bas  $70^\circ$

Champ de vision binoculaire horizontal :

Avec les deux yeux fixes, il y a une zone de recouvrement correspondant au champ de la vision binoculaire de  $100^\circ$  à  $120^\circ$  en moyenne horizontalement.

Pour obtenir une totale immersion visuelle dans un visiocasque, il faut tenir compte du mouvement des yeux et de la tête. Un œil peut tourner dans son orbite d'environ  $15^\circ$  horizontalement et verticalement. Il faut donc théoriquement au maximum un champ de vision dans le visiocasque de  $210^\circ$  horizontalement et de  $140^\circ$  environ verticalement. En pratique, actuellement, les champs de vision dans les visiocasques sont encore bien faibles, de l'ordre de  $100^\circ$  horizontalement, sauf un seul visiocasque à champ de vision de  $210^\circ$  horizontal. L'impression d'immersion visuelle instantanée y est donc plus ou moins forte.

Le champ de vision total avec mouvement de la tête et des yeux est très grand (horizontalement  $>200^\circ$  côté tempe et  $130^\circ$  côté nez, verticalement :  $140^\circ$  haut et  $170^\circ$  bas). C'est bien sûr l'avantage des visiocasques de pouvoir observer presque tout autour de l'utilisateur, mais en ayant à chaque instant qu'un champ de vision plus étroit que celui de la vision naturelle. Il faut en conséquence bien étudier ce que les I<sup>2</sup> sensorimotrices exigent au niveau visuel.

L'œil tourne dans son orbite à la vitesse maximum de  $600^\circ/s$  environ. La tête peut tourner à la vitesse maximum de  $800^\circ/s$  environ. On comprend aisément que le tracking des yeux est difficile dans un visiocasque pour suivre précisément les mouvements des yeux et obtenir un très faible de temps de latence. De même, on comprend les

difficultés à mesurer les mouvements de la tête, action pourtant indispensable à l'usage des visiocasques. À noter que si on veut connaître l'endroit exact observé par l'utilisateur (la direction des axes optiques des yeux) le temps de latence et la précision dépendront du système de tracking des yeux et de celui du tracking de la tête.

### 2.7.3. La distance interpupillaire

La distance interpupillaire (DIP) est variable pour l'être humain. En moyenne, elle est de 65 mm pour l'homme adulte, légèrement inférieure pour la femme en moyenne, les écarts pouvant être significatifs, de 50 à 75 mm environ, pour la population européenne. Dans les cas extrêmes, la DIP peut s'étendre de 40 à 80 mm. Il est primordial que cette grandeur soit prise en compte pour l'utilisation des visiocasques, les deux optiques devant être bien centrées par rapport aux yeux et ainsi avoir un champ de vision optimum. Pour des réglages fins sur un visiocasque, nous avons mesuré la DIP pour chaque utilisateur, en employant un pupillomètre. Ces mesures permettent de mieux régler la DIP, plutôt que de laisser l'utilisateur faire un réglage approximatif.

### 2.7.4. Pouvoir d'accommodation

Les deux cristallins permettent de voir net le point dans l'espace observé, mais ce pouvoir d'accommodation a ses limites. La limite inférieure est appelée le *punctum proximum* (d'accommodation)<sup>8</sup> qui est le point le plus proche qu'un œil peut voir nettement, en accommodant au maximum. Avec l'âge, le pouvoir d'accommodation diminue avec le vieillissement du cristallin et le point proximum s'éloigne progressivement de l'œil. La personne devient presbyte et doit porter des lunettes. Pour un enfant, le *punctum proximum* est de 7 cm environ et il est de 25 cm environ vers 40 ans.

Concrètement, pour les visiocasques avec des écrans classiques, ceux-ci étant très proches des yeux, il est obligatoire de mettre des lentilles pour que l'observateur puisse accommoder sur les images qui sont perçues à une certaine distance ou à l'infini. Ce système optique à insérer entre les écrans et les yeux posent un problème technique très difficile à concevoir : réaliser l'accommodation sur les écrans tout en ayant un grand champ de vision proche des 180 degrés et non de l'ordre de 100 degrés. À l'inverse du *punctum proximum*, il y a le *punctum remotum* qui est le point le plus éloigné qu'un œil peut voir nettement. Il est situé à l'infini pour une personne ayant une vision emmétrope, c'est-à-dire normale.

### 2.7.5. Fréquence temporelle maximum en vision

La fréquence temporelle des images pour percevoir des mouvements fluides est critique pour des valeurs inférieures à 25 - 30 images par seconde, valeur dépendant

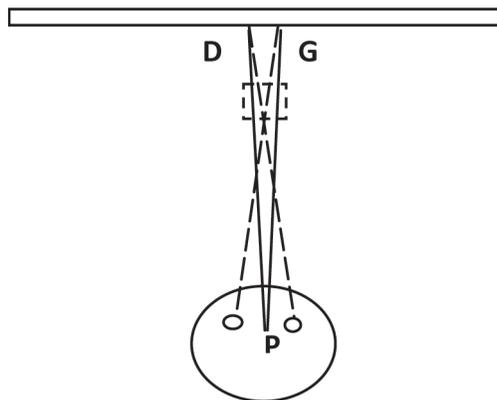
---

<sup>8</sup> Il existe aussi les points limites pour la convergence des yeux.

du type d'images. La fréquence sur des écrans monoscopiques doit être au moins à 50 - 60 Hz (une image est composée de deux trames). Pour les écrans stéréoscopiques, la fréquence de balayage doit être double, au moins 100 Hz. Il faut aussi que la fréquence d'affichage des images ne soit pas trop faible par rapport aux besoins du système visuel pour percevoir des images sans scintillement et avec des mouvements continus des objets en mouvement. Cela ne dépend pas de la rémanence des rétines (ancienne fausse théorie) mais de l'effet « phi » et le mouvement « bêta », mécanismes neurophysiologiques : lorsque le cerveau perçoit, par exemple, deux points lumineux alternativement affichés à une distance angulaire proche l'un de l'autre, le cerveau perçoit le premier point se déplacer à la position du second point, ce qui donne une impression de mouvement à partir d'informations ponctuelles. Plus la fréquence est élevée, plus les mouvements semblent fluides.

### 2.7.6. Les caractéristiques psychophysiques de la vision stéréoscopique

Nous avons explicité que le système visuel arrive à fusionner les deux images rétiniennes pour l'ensemble des points qui se situent dans l'aire de Panum. Dans le cas contraire, le cerveau est incapable de fusionner les deux images ou il le fait avec une certaine fatigue. Pour la vision en relief, l'objectif sera de créer des images stéréoscopiques créant dans les yeux de l'observateur de faibles disparités rétiniennes, tout en gardant un effet de profondeur. Mais quelles sont les limites de fusion acceptables en moyenne pour les êtres humains ?



**Figure 2.11** : la parallaxe horizontale angulaire est définie par l'angle DPG

Nous avons défini la **disparité rétinienne**. Pour des images en relief affichées sur un écran stéréoscopique, la disparité rétinienne horizontale est égale approximativement à la **parallaxe angulaire horizontale**. Celle-ci est définie par l'angle DPG, formé par les deux points homologues des images droite et gauche observés du point de vue de l'observateur (Figure 2.11).

En faisant intervenir la durée de visualisation des images, l'étude de Yei-Yu Yeh examine la limite de la fusion et l'appréciation de la profondeur [Yei-Yu *et al.*

, 1990]. La limite de fusion est analysée en fonction de différents paramètres des stimuli affichés sur un écran cathodique avec lunettes actives : la couleur, le temps d'affichage et la parallaxe horizontale. Les résultats impliquent que la limite de fusion (donnée en angle de parallaxe) est très faible pour un stimulus bref de 0,2 seconde : soit 27 à 24 minutes d'arc comparativement à un stimulus de 2 secondes : 1,5 degré. Il est donc plus difficile d'observer en relief des images en mouvement rapide. En prenant la parallaxe horizontale comme paramètre influençant la stéréopsie, des études expérimentales [Valyus, 1962] ont démontré la difficulté de fusionner deux images planes fournissant des parallaxes horizontales (et donc des disparités rétinienne) d'angle supérieur à 1,5 degrés. Nous avons entrepris des tests en situation de travail dans le cadre de la téléopération avec des images stéréoscopiques réelles. Nos résultats donnent une limite un peu inférieure pour la limite de parallaxe horizontale : 1,2° [Fuchs *et al.*, 1995]. Ces variations sont normales car elles sont variables pour chaque individu qui tolère plus ou moins les images en relief et qui force ou pas son système visuel à fusionner les images. À noter qu'un petit nombre de la population (quelques pourcents) n'arrive pas à fusionner des images en relief sur un écran.

Des études ont montré qu'il est difficile de fusionner des images créant des parallaxes verticales (et donc des disparités verticales). Les yeux, généralement positionnés sur un même plan horizontal, ne sont pas aptes à percevoir des disparités verticales, sauf si elles sont très faibles. Pour les images en relief, les parallaxes verticales doivent être inférieures à un angle de 20' d'arc [Julesz, 1971]. Nous verrons, dans le paragraphe sur la création des images en relief, que cette contrainte est à prendre en compte.

Les études entreprises à l'IRBA (Institut de Recherches Biomédicales des Armées) du Ministère de la Défense donnent des résultats corroborant les précédents, en étant plus précis. La limite de fusion est aussi fonction de la fréquence spatiale<sup>9</sup> des images, ce qui n'est pas étudié dans les travaux cités précédemment. Leurs études permettent de conclure que de plus grandes disparités horizontales sont fusionnables lorsque les stimulations visuelles sont constituées de basses fréquences spatiales. Du point de vue pratique, nous pouvons en conclure que les objets représentés par leurs contours (hautes fréquences spatiales) sont plus difficilement fusionnables. Leurs études ont mis en évidence deux mécanismes impliqués dans la fusion selon le temps du stimulus : une fusion immédiate si la disparité n'est pas trop grande et une fusion non immédiate pour des disparités plus grandes, mettant en jeu des mouvements de vergence réflexes des yeux. Les résultats donnent entre autres : à la fréquence spatiale de 4,8 cycles/degré, fusion immédiate jusqu'à 20' d'arc et fusion maximale à 52' d'arc environ. À la fréquence spatiale de 0,22 cycles/degré, fusion immédiate jusqu'à 80' d'arc et fusion maximale à 176' d'arc environ [Perrin, 1998].

---

<sup>9</sup> La fréquence spatiale d'une image est définie par le nombre de cycles d'une oscillation périodique de luminance ou de couleur dans un degré d'angle visuel. Sa valeur est donc donnée en cycles par degré.

### 2.7.7. Dépendance et indépendance au champ

La dépendance et l'indépendance au champ est une caractéristique de la perception visuelle qui se traduit par la possibilité pour une partie des humains de reconnaître très rapidement des objets dans une scène complexe sans pratiquer d'exploration détaillée (indépendant au champ). Les autres sont obligés d'explorer l'ensemble des objets situés dans le champ visuel pour effectuer la reconnaissance des objets (dépendant au champ). Cette faculté d'indépendance s'acquière à partir de 12 ans et seulement pour un peu moins de 40% des sujets. Environ 30% de la population reste dépendant du champ et le reste est plus ou moins dépendant en fonction des scènes regardées. Ceci a des répercussions sur le sentiment d'immersion dans les systèmes de réalité virtuelle. Les dépendants au champ auraient besoin de champs visuels plus importants pour avoir le sentiment d'immersion visuelle.

## 3. LE SYSTÈME AUDITIF

Nous ne présentons pas le système auditif dans sa totalité, car il est connu. De plus, il ne pose pas de grandes difficultés techniques pour la simulation et la restitution sonores, sauf pour celle d'une source sonore localisée spatialement et qui mérite un développement particulier :

Pour créer un monde virtuel visuel et sonore compatible, il est préférable de pouvoir simuler spatialement des sources sonores qui se superposent aux images des objets correspondants. Mais comment le cerveau humain localise-t-il les sources sonores virtuelles dans n'importe quel point de l'espace ? L'homme perçoit les sons différemment avec ses deux oreilles. Plusieurs différences sont inconsciemment interprétées par le cerveau. Les deux premières sont le décalage temporel et le rapport d'intensité entre les oreilles (Figure 2.12).

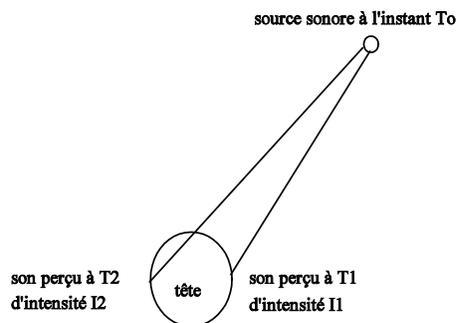


Figure 2.12 : schéma des propagations des sons aux oreilles

Si on considère les deux oreilles comme de simples récepteurs ponctuels, le décalage de temps  $T_2 - T_1$ , de l'ordre de la milliseconde, et la différence d'intensité  $I_2 - I_1$  permettent seulement de localiser la source sonore sur un cône d'axe intra-aural, appelé cône de

confusion. Les valeurs d'intensités des sons permettent d'estimer la distance de la source sonore. Mais dans ce cas de récepteurs ponctuels, il est impossible d'estimer la position avant ou arrière et en haut ou en bas du son reçu. En réalité les oreilles ne sont pas de simples récepteurs ponctuels, mais sont composées de pavillons (oreilles externes) et de conduits internes qui jouent un rôle important dans l'audition. Le son est réfléchi sur le pavillon et il est guidé par celui-ci vers l'oreille interne. Il est aussi réfléchi et en partie absorbé par la tête et les épaules. Ces phénomènes sont évidemment fonction de la position de la source par rapport aux oreilles. Tout ceci implique une troisième règle : la source sonore produit des variations d'intensité pour les deux oreilles qui sont aussi fonction de la fréquence des sons. Cette différence de perception des deux intensités dépendant de la fréquence constitue une caractéristique pour la localisation de la source.

Ces phénomènes de localisation ont été étudiés depuis quelques dizaines d'années et permettent de créer artificiellement des sources sonores virtuelles dans l'espace. Le lecteur ne doit pas confondre cette technique avec celle du son stéréophonique qui ne crée des sons que sur un plan vertical intra-aural. Les études ont montré que chaque personne a sa propre « fonction de transfert auditif de la tête » qui donne la relation entre la différence de perception des intensités  $I_1$  et  $I_2$  et la fréquence du son. Une source sonore peut être localisée, du moins si la personne veut bien y prêter attention et si la source n'est pas trop environnée d'autres sources sonores parasites. Des illusions inter-sensorielles sont toujours possibles entre la vue et l'ouïe. Par habitude, nous positionnons au niveau de la bouche la parole d'une personne. Ainsi le ventriloque peut donner l'illusion que sa marionnette parle à sa place en faisant bouger les lèvres de la marionnette tandis que ses lèvres restent immobiles. De même, par effet Doppler (variation de la fréquence d'un son en fonction de la vitesse de la source sonore), il est possible de suggérer que la source sonore provienne de l'objet en mouvement observé sur l'écran (voiture sur une route virtuelle, par exemple) même si la source sonore est fixe<sup>10</sup>. Dans les applications RV, on peut mettre à profit cette illusion : en fournissant un simple son stéréophonique d'un objet virtuel sonore, l'affichage de celui-ci fait croire à l'observateur que la source sonore est toujours localisée exactement là où se voit l'objet, sans être forcément obligé de fournir un son spatial 3D.

#### **4. LA SENSIBILITÉ CUTANÉE**

Dans la plupart des tissus du corps humain (les muscles, la peau, les articulations, les tendons, etc.), il y a des récepteurs pouvant être activés mécaniquement. Leurs stimuli sensoriels sont séparés en deux parties : la sensibilité cutanée et la proprioception (sensibilité interne).

La sensibilité cutanée est composée du toucher, ainsi que de la sensibilité à la pression, aux vibrations et à la température. L'absence du toucher dans un monde virtuel ôte une

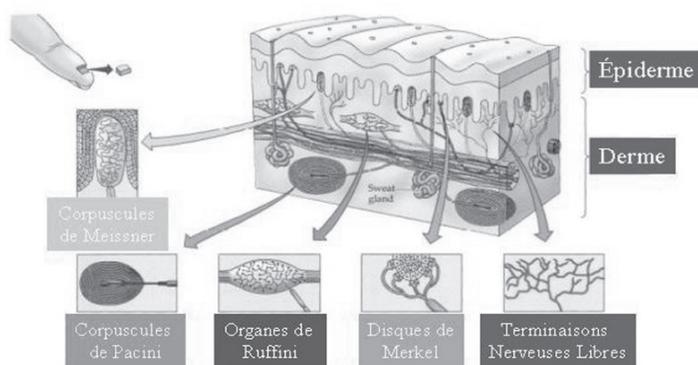
---

<sup>10</sup> Des présentations détaillées du son spatial et des techniques matérielles et logicielles associées sont consultables dans le « Traité de la Réalité Virtuelle ».

source d'informations, ce qui peut être gênant pour certaines simulations virtuelles. Si l'ensemble de la peau participe au sens du toucher, il est important de noter qu'il est surtout efficace au niveau des mains : grâce à celles-ci, on peut déterminer par toucher la géométrie de la surface en contact, sa rugosité et son mouvement relatif. La peau possède une innervation afférente très riche, dont la densité varie beaucoup d'une localité à une autre. À titre d'exemple, les extrémités des doigts sont richement innervées (environ 2500 récepteurs/cm<sup>2</sup>).

Les récepteurs du sens du toucher sont une composante de la peau. Ils nous renseignent sur le poids, la rugosité et la température de l'objet touché. Il y a trois types de récepteurs : les nocicepteurs, thermorécepteurs et mécanorécepteurs. Les nocicepteurs prennent en charge les sensations douloureuses. Les thermorécepteurs sont sensibles à la température qui entre en jeu lors de l'exploration tactile. La sensibilité thermique est due à la présence dans la peau de thermorécepteurs divisés en récepteurs du froid et récepteurs du chaud. La sensibilité thermique est surtout une sensibilité à la variation de température. Les sensations qui apparaissent lors d'un changement de température dépendent de l'étendue de la surface stimulée, de la température initiale et de la vitesse de changement.

Les mécanorécepteurs jouent un rôle très important dans la perception tactile. Ce sont les récepteurs les plus souvent sollicités lorsqu'un contact est établi entre la peau et un objet. La distribution de ces mécanorécepteurs et la combinaison de leurs signaux efférents, en réponse aux stimuli, impliquent une spécialisation de la structure sensorielle. Quatre types de mécanorécepteurs extéroceptifs ont pu être identifiés. Tout d'abord les terminaisons de Meissner qui se situent au niveau des empreintes de la peau de la main ainsi que celles de Merkel. Ces deux premiers types sont caractérisés par une grande résolution spatiale. Les deux types restant sont les corpuscules de Pacini et les terminaisons de Ruffini. On les trouve situés au plus profond de la peau avec des champs de réception plus larges<sup>11</sup>.



**Figure 2.13 :** les différents récepteurs de la sensibilité cutanée

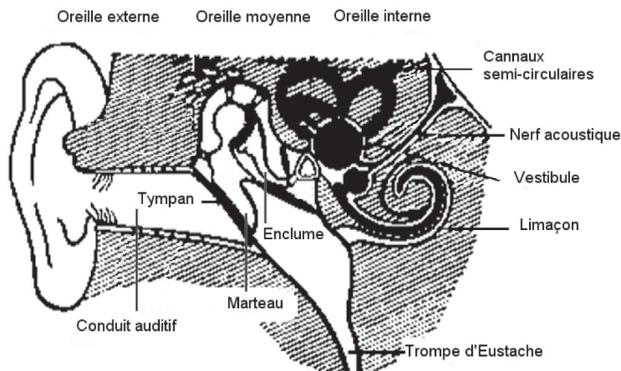
<sup>11</sup> Des explications détaillées sur la sensibilité cutanée sont consultables dans le volume « L'homme et l'environnement virtuel » du *Traité de la réalité virtuelle*.

## 5. LA PROPRICEPTION

Les trois domaines liés à la proprioception sont les sensibilités à la position dans l'espace, au mouvement du corps et aux forces exercées sur les muscles. Les deux premiers domaines correspondent au sens kinesthésique et le dernier correspond à la proprioception musculaire.

La conscience du mouvement est donnée par la position des différentes parties du corps, leur direction, leur taux de mobilité ainsi que par les forces de contractions musculaires nécessaires à la résistance aux mouvements. En plus des organes proprioceptifs des muscles, des tendons et des articulations, les organes qui participent aux sensations de mouvement sont ceux situés au niveau des récepteurs des **deux systèmes vestibulaires** dans les oreilles internes et, pour une part non négligeable, la vision.

C'est pourquoi ce paragraphe comprend une succincte présentation des récepteurs des deux systèmes vestibulaires. Ces derniers informent le cerveau du mouvement et de l'orientation de la tête. Les oreilles internes sont composées de cinq cavités, trois pour les canaux semi-circulaires et deux autres cavités dénommées saccule et utricule. Les récepteurs sensoriels sont composés de canaux semi-circulaires orthogonaux situés dans les trois plans de l'espace, détectant l'inclinaison de la tête et les accélérations angulaires, ainsi que de petites concrétions calcaires, les otolithes, qui détectent les accélérations linéaires de la tête et qui se situent dans le saccule et l'utricule.



**Figure 2.14 :** les récepteurs du système vestibulaire dans l'oreille interne

Les intérieurs des canaux semi-circulaires sont remplis de liquide gélatineux et tapissés de cellules ciliaires stimulées par les mouvements du liquide. Les canaux semi-circulaires sont donc sensibles à l'inclinaison de la tête mais aussi aux accélérations angulaires dans les mouvements de rotation de la tête. Dans le cas d'accélérations linéaires d'un mouvement de translation de la tête, ce sont les otolithes qui imposent des forces sur leurs supports ciliaires dues aux inerties de leur masse et donc informent le cerveau des accélérations linéaires. On peut donc comparer les récepteurs du système vestibulaire à une centrale inertielle.

Si la tête est en mouvement, les messages nerveux renseignent sur l'accélération totale subie par la tête, composée de l'accélération du mouvement et de l'attraction terrestre. Ceci permet de comprendre qu'en inclinant une personne dans une cabine de simulateur, on peut simuler indirectement une accélération de mouvement en présentant visuellement à la personne un déplacement virtuel, que ce soit via un grand écran ou via un visiocasque.

Comme l'explique Alain Bertoz [A. Bertoz, 1997], le problème de coordination entre la position du corps dans l'espace, la posture et les mouvements du corps est complexe. La stabilisation de la tête, grâce entre autres aux systèmes vestibulaires, est indispensable. Par exemple, lorsqu'une personne court, la tête est stabilisée sur une trajectoire malgré les mouvements complexes de ses membres durant la course. Par ailleurs, lorsqu'une personne bouge la tête, elle n'utilise pas seulement les informations de son système visuel pour regarder des objets devant elle. Si l'image du monde bouge de dix degrés par seconde environ sur la rétine, l'image est floue. Or les rotations de la tête se font à plus de cent degrés par seconde. La solution vient du réflexe vestibulo-oculaire qui fait tourner l'œil en sens contraire du mouvement de la tête. Les systèmes vestibulaires, mesurant la rotation de la tête, commandent de produire un mouvement des yeux en sens contraire au mouvement de la tête pour stabiliser l'image rétinienne.

D'autres récepteurs proprioceptifs, répartis dans tout notre organisme, sont sensibles à la position ou au déplacement des différentes parties du corps. Dans les muscles, les récepteurs sont les fuseaux neuromusculaires et les organes tendineux de Golgi. Les récepteurs articulaires répondent à des variations d'étirement et sont donc sensibles au mouvement relatif entre deux parties du corps.

Les informations de tous les récepteurs proprioceptifs sont analysées par le cerveau avec des informations sensorielles transmises par la peau, par les yeux et par l'oreille interne. Ainsi le cerveau en tire une connaissance de la position et du mouvement du corps. La perception des mouvements est donc multimodale et exige la collaboration de récepteurs sensoriels variés pour contribuer à la perception des mouvements. Il est évident que, si dans une application RV avec ou sans visiocasque, certains récepteurs sensoriels du sujet sont incorrectement stimulés, les perceptions de la posture et des mouvements du corps du sujet sont ambiguës. La perception retenue du mouvement sera celle imposée par la vision, grâce à lavection, mais sera source d'incohérences sensorimotrices.

Un autre phénomène à connaître, expliqué par Alain Berthoz [A. Berthoz, 2000], est que le cerveau anticipe et simule les mouvements. Le cerveau planifie une séquence de mouvements et en parallèle, configure les récepteurs sensoriels qui seront utiles. Il anticipe les messages que doivent fournir les récepteurs sensoriels si le mouvement est correctement effectué. En exemple donné par Alain Berthoz, «un skieur n'a pas le temps d'évaluer tous les mouvements de son corps sur la piste. Il simule la trajectoire dans sa tête, et ne fait que vérifier de temps en temps (même si tout se fait en quelques fractions de secondes) l'état de quelques indices sensoriels qui lui sont utiles pour connaître sa position ou la distance au piquet». De plus, le cerveau n'attend jamais que les sens soient activés pour produire quelque chose. Un exemple : pour attraper un

ballon plus ou moins lourd, quelques centaines de millisecondes avant que le ballon ne touche la main, le cerveau produit une contraction des muscles pour bien réceptionner le ballon. Cette contraction n'est possible que grâce à l'existence de modèles internes des actions mécaniques de la gravité sur l'objet. Car autrement, il ne serait pas possible que le cerveau réagisse instantanément aux forces nécessaires lors du contact. Le cerveau a des modèles internes des lois de la mécanique, qui lui permettent d'anticiper sur ces lois. Nous en reparlerons au chapitre 7.

Pour conclure sur la proprioception, la compréhension de ce sens est très importante pour les applications RV. Par exemple, techniquement pour avoir l'illusion d'être plus immergé, il peut être nécessaire de restituer dans le monde virtuel la position exacte du corps du sujet immergé. Il faudra alors disposer de capteurs permettant de transmettre à l'ordinateur générant le monde virtuel les informations concernant la position du corps et de ses membres dans le monde réel où il se trouve. D'autre part, compte tenu de la sensibilité humaine aux déplacements dans les trois directions de l'espace, il peut être aussi nécessaire de restituer au sujet immergé les sensations liées à l'équilibre, à la perception de la verticale et aux mouvements simulés en excitant correctement les systèmes vestibulaires et la proprioception musculaire. Pour ce dernier aspect, nous présenterons au chapitre 5 les « interfaces à simulation de mouvements » du corps du sujet.

Par ailleurs, la proprioception musculaire inclut la sensibilité aux forces exercées sur les muscles qui donne une information globale sur les forces de contact entre la personne et un objet. Il peut alors être envisagé de créer virtuellement des efforts (forces et couples) sur le corps de l'utilisateur, en pratique presque toujours sur ses mains. Dans ce cas, il faut utiliser les « interfaces à retour d'effort » qui ne seront pas présentées dans cet ouvrage<sup>12</sup>.

En conclusion de ce chapitre sur les sens humains, il est important de bien maîtriser leur compréhension qui est plus complexe que le laisse croire notre vie quotidienne où leurs fonctionnements se font presque toujours inconsciemment. En revanche, ce sera une importante problématique de réaliser une simulation virtuelle sans ou avec peu incohérences sensorimotrices pouvant induire de l'inconfort ou des malaises. Nous en reparlerons au chapitre 7.

## 6. BIBLIOGRAPHIE

[Auvray, 2006] M. Auvray, « Théorie de la perception », chap. 7, vol. 1, *Traité de la Réalité Virtuelle*, Presses des Mines.

[Bertoz, 1997] A. Bertoz, *Le sens du mouvement*, O. Jacob.

---

12 Une présentation détaillée sur les interfaces à retour d'effort est consultable dans le volume « Interfaçage, immersion et interaction en environnement virtuel » du *Traité de la réalité virtuelle*.

- [Berthoz, 2000] A. Berthoz, *Le cerveau, le langage et le sens*, l'université des savoirs, O. Jacob.
- [Fuchs *et al.*, 1995] P. Fuchs, D. Ernadotte, D. Maman, C. Laurgeau, and J. Bordas. «Téléprésence virtuelle stéréoscopique». In *Interface des Mondes réels et virtuels* (1995), pages 77–91, Montpellier.
- [Gibson, 1966] J. Gibson. *The senses considered as a perceptual system*. Houghton, Mifflin, Boston, MA.
- [Julesz, 1971] B. Julesz. *Fondation of Cyclopean Perception*. University of Chicago Press, Chicago.
- [Leroy, 2009] Leroy L., Fuchs P., Paljic A., Moreau G., “Some experiments about shape perception in stereoscopic displays”, *IS&T/SPIE Symposium on Electronic Imaging*, Number 7237-45, 18-22 January, San Jose, California USA.
- [Perrin, 1998] J. Perrin, P. Fuchs, C. Roumes, and F. Perret. “Improvement of stereoscopic comfort through control of disparity and spatial frequency content”. In *Visual Information Processing VII*, volume 3387 of Proceedings of SPIE (1998).
- [Sherrington, 1906] C. S. Sherrington. “The Integrative Action of the Nervous System”. Yale University Press (1947 ed.).
- [Trotter, 1995] Y. Trotter. «Bases neuronales de la perception visuelle chez le primate». *Journal français d'orthoptique*, (27).
- [Valyus, 1962] N. Valyus. *Stereoscopy*. Focal Press.
- [Yei-Yu *et al.*, 1990] Y. Yeh and L. Silverstein. “Limits of fusion and depth judgment in stereoscopic color displays”. *The Human Factors Society*, 32:45-60.



---

## Les interfaces visuelles

---

### 1. INTRODUCTION

Avant d'aborder la description des visiocasques, une présentation de tous les types d'interfaces visuelles est brièvement exposée. Ceci permet à tout concepteur d'application RV une comparaison entre les visiocasques et les autres types d'interfaces visuelles pour savoir s'il est bien judicieux d'exploiter un visiocasque. Par notre expérience, nous avons remarqué que certains professionnels voulaient utiliser un visiocasque sans avoir étudié les autres possibilités d'interfaces visuelles. Il n'y a pas d'interface idéale et universelle ! Si elle existait, elle devrait avoir des caractéristiques métrologiques correspondant aux capacités maximales du système visuel humain pour bien exploiter ce canal sensoriel. Cette interface devrait alors offrir quatre capacités de plus que celles d'un écran ordinaire :

- de **grands champs de vision horizontaux et verticaux** correspondant à ceux des yeux ;
- une **vision stéréoscopique dans tout le champ de vision binoculaire** ;
- une **très haute résolution graphique** exploitant toutes les performances des acuités monoscopique et stéréoscopique ;
- une **immersion du regard**<sup>1</sup> dans le monde virtuel.

Ce dernier point n'est possible que si l'interface visuelle est couplée à un capteur localisant l'orientation de la tête de l'opérateur. La puissance informatique doit bien sûr être aussi suffisante pour afficher en temps réel devant les yeux des images gauche et droite à haute résolution, correspondant à la direction de son regard.

### 2. FONCTIONNALITÉS POUR LA PERCEPTION D'UN ESPACE TRIDIMENSIONNEL

Avant d'aborder les technologies des interfaces visuelles, il faut préciser les différentes fonctionnalités que l'on peut employer pour percevoir un espace en trois dimensions. Ces fonctionnalités se réfèrent aux caractéristiques de la vision humaine : la perception d'un espace 3D, de la profondeur, dépend des indices proprioceptifs (accommodation et vergence) et des indices visuels binoculaires ou monoculaires (voir le chapitre

---

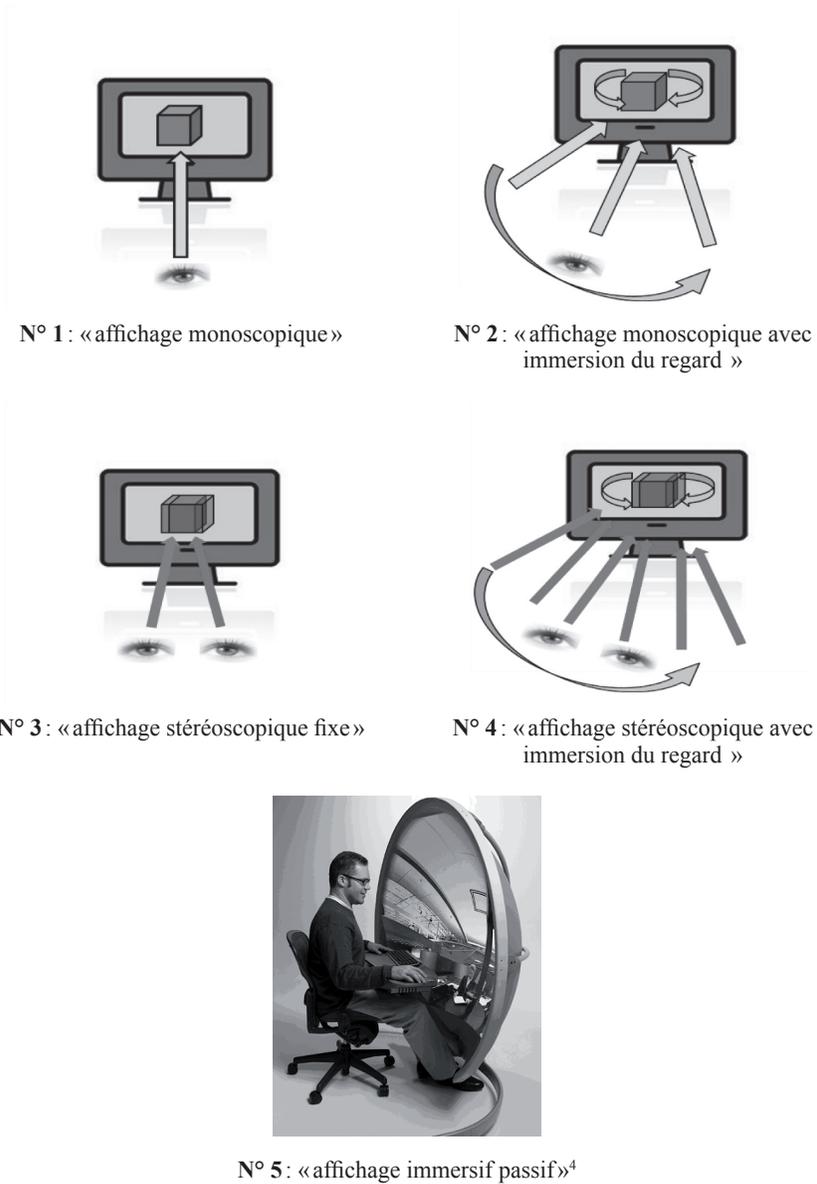
<sup>1</sup> L'immersion du regard : l'utilisateur voit toujours la scène virtuelle, même s'il tourne sa tête (son regard) dans une autre direction.

précédent). Comme nous l'avons déjà indiqué, il faut bien retenir que c'est l'*observateur* qui réalise la perception 3D de l'espace et non la technologie. Différentes fonctionnalités (Figure 3.1) peuvent être employées pour observer et être immergé dans un espace en trois dimensions :

- N° 1, la fonctionnalité la plus simple, « **affichage monoscopique** », est de projeter l'espace 3D sur un plan, donc sur un écran monoscopique, en exploitant des indices monoculaires, dont la perspective. Dans ce cas le champ de vision est limité et le changement de point de vue est commandé manuellement. Pour l'instant, la plupart des jeux vidéo utilisent cette simple fonctionnalité via un écran d'ordinateur ou de console de jeux, couplé à un joystick ;
- N° 2, la fonctionnalité « **affichage monoscopique avec immersion du regard** », consiste toujours à projeter l'espace 3D sur un écran monoscopique, mais avec contrôle de l'image affichée par l'observateur lui-même. Le champ de vision est toujours limité. Cette fonctionnalité, encore peu utilisée, quelques fois sur smartphone, exploite l'indice monoculaire de parallaxe de mouvement pour mieux percevoir l'espace tridimensionnel. Il y a deux sous cas :
  - soit l'écran est fixe et la tête de l'observateur est traquée, et donc en se déplaçant partiellement, l'observateur perçoit la conséquence de son changement de point de vue sur l'affichage ;
  - soit l'écran est mobile (tablette numérique ou smartphone) et l'observateur manipule l'écran, dont la position est mesurée par un capteur d'orientation, pour changer l'orientation du point de vue observée ;
- N° 3, la fonctionnalité « **affichage stéréoscopique** » exploite les indices monoculaires et binoculaires en même temps. C'est le cas classique pour le cinéma<sup>2</sup> et la télévision 3Ds. Le champ de vision est toujours limité. Comme nous l'avons indiqué précédemment, la fonctionnalité N°2 « affichage monoscopique avec immersion du regard » permet une meilleure perception des formes et des distances en profondeur que la fonctionnalité N°3 ;
- N° 4, la fonctionnalité « **affichage stéréoscopique avec immersion du regard** » avec un écran rassemble les fonctionnalités 2 et 3, permettant une meilleure perception de la troisième dimension. Le champ de vision est toujours limité ;
- N° 5, la fonctionnalité « **affichage immersif passif** » où l'affichage des images recouvrent tout le champ de vision de l'observateur sans que celui-ci puisse commander le changement de point de vue. Cette fonctionnalité est proposée, par exemple, dans les géodes<sup>3</sup>, avec ou non la vision stéréoscopique ;
- N° 6, la fonctionnalité « **affichage immersif actif** » où l'affichage des images stéréoscopiques recouvrent tout le champ de vision de l'observateur. Celui-ci commande le changement de point de vue par mouvement de la tête. C'est la fonctionnalité du visiocasque ou du CAVE. Dans un visiocasque, les champs de vision (monoscopique et stéréoscopique) sont, hélas, très rarement aussi grands que ceux de la vision humaine. Nous en reparlerons. Noter aussi que dans un CAVE, les champs de vision sont aussi perturbés par le port des lunettes. Avec les techniques actuelles, nous sommes encore loin de l'interface visuelle idéale...

2 La notation « 3Ds » (pour 3D stéréoscopique) est utilisée pour éviter l'ambiguïté de l'expression « 3D ».

3 Les géodes sont en général de très grandes dimensions pour offrir l'affichage à un grand nombre de spectateurs mais il existe aussi des géodes de petites dimensions pour un seul observateur.



**Figure 3.1 :** présentation des 5 premières fonctionnalités pour voir un espace en 3 dimensions<sup>5</sup>

<sup>4</sup> Visionstation d'Elumens, copyright photo Elumens.

<sup>5</sup> Schémas du rapport « les effets sanitaires potentiels des technologies audiovisuelles en 3D stéréoscopique », Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, [www.anses.fr/fr/system/files/AP2011sa0334Ra.pdf](http://www.anses.fr/fr/system/files/AP2011sa0334Ra.pdf)

La perception avec la fonctionnalité N° 6 est actuellement la plus proche de la perception naturelle humaine. Mais nous pourrions rajouter une future fonctionnalité N° 7 qui sera encore plus proche de la vision humaine puisque l'indice proprioceptif d'accommodation sera similaire en vision artificielle et naturelle : pour les fonctionnalités N° 1 à N° 6, l'accommodation est figée puisque les yeux doivent se focaliser sur le plan de l'écran, ce qui n'est pas naturel et donne un mauvais indice de distance de profondeur. Avec de nouveaux types d'écrans, comme les écrans «Light Field» (voir le chapitre suivant), l'accommodation sera variable et se fera à la bonne distance de profondeur pour chaque entité dans l'environnement virtuel. Ce type d'écran est prévu pour de futurs visiocasques RA (Réalité Augmentée).

Les fonctionnalités avec affichage stéréoscopique (N°3 à N°6) fournissent des informations plus réalistes si les images stéréoscopiques sont projetées en *vision orthostéréoscopique*, c'est-à-dire que les échelles de profondeur sont en concordance pour les indices monoculaires et binoculaires. Ce sera le cas *a priori* dans la plupart des usages des visiocasques. Mais ce n'est pas le cas, par exemple, pour les images stéréoscopiques aériennes prises à haute altitude, avec les deux points de vue séparés de plusieurs centaines de mètres : les bâtiments ne sont pas alors perçus à leur taille réelle mais paraissent très petits, comme dans une maquette de la région observée. Si on traque le mouvement des yeux de l'observateur, prévu sur certains futurs visiocasques, on peut traiter les images stéréoscopiques pour améliorer le confort visuel et la perception de la profondeur, mais cela ne crée pas une nouvelle fonctionnalité (dans la mesure où cela ne fournit aucune information supplémentaire sur la taille ou la profondeur des objets observés).

Cette description de la grande variété de fonctionnalités proposées par les différentes interfaces réalisées techniquement à ce jour, a pour but de montrer que le choix de l'interface visuelle appropriée pour une application RV n'est pas simple. Il ne se réduit pas au simple choix entre un écran classique et un visiocasque. Sans compter que l'environnement virtuel peut être configuré spatialement d'une façon particulière pour permettre d'exploiter une interface visuelle plutôt qu'une autre. En exemple, pour l'immersion visuelle dans un magasin virtuel, nous avons choisi un grand écran à haute résolution devant l'observateur, plutôt qu'un visiocasque, le champ de vision étant suffisant, de 130° environ. Tous les produits étaient sur les mêmes côtés des rayonnages et ainsi l'observateur n'avait donc pas besoin de pivoter sa tête, ce qui aurait exigé l'emploi d'un visiocasque, mais avec une qualité d'image très insuffisante.

### **3. LES INTERFACES À ÉCRANS FIXES**

#### **3.1. Écrans plats monoscopiques ou stéréoscopiques**

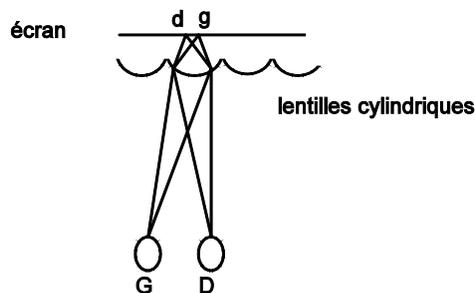
Les écrans plats classiques sont monoscopiques ou stéréoscopiques, cette dernière fonctionnalité est souvent proposée sur les téléviseurs, même si elle est peu utilisée par les téléspectateurs. On peut noter qu'avec un ensemble de plusieurs écrans plats

de grandes dimensions, il est possible de réaliser un CAVE rudimentaire, même s'il est difficile de trouver des écrans plats sans bord ou à bord très fins.

Quelle que soit la technique employée en vision stéréoscopique, l'objectif est toujours le même : fournir à chaque œil une image légèrement différente. Pour ceci, il y a deux principes : soit on place près des yeux deux petits écrans (cas des visiocasques), soit on sépare physiquement les deux images affichées sur un seul écran, celui-ci étant plus ou moins éloigné des yeux. Dans cette dernière catégorie, le principe repose sur un dispositif technique permettant de séparer les images pour ne visualiser sur chaque œil que l'image désirée. La classification qui suit est donc fonction du lieu où se fait cette séparation et du type de dispositif employé.

### 3.1.1. Séparation au niveau de l'écran

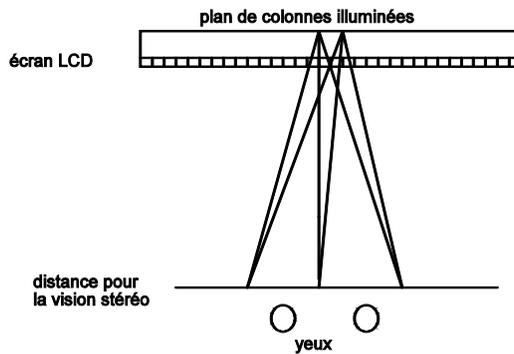
**Les écrans autostéréoscopiques** permettent de faire la séparation du couple d'images stéréoscopiques au niveau de l'écran, en jouant sur le décalage de position des deux yeux de l'observateur. Le principal avantage des écrans autostéréoscopiques est de permettre à l'observateur de ne porter aucune lunette (grand intérêt surtout dans les lieux publics pour une vision immédiate). L'écran autostéréoscopique est recouvert d'un réseau lenticulaire : un plan composé de lentilles demi-cylindriques d'axes parallèles aux colonnes de l'écran. Celles-ci ont des caractéristiques optiques qui impliquent que chaque œil de l'utilisateur voit des colonnes différentes de pixels, correspondant à deux images différentes. Derrière chaque lentille cylindrique, on a pour chaque œil 1, 2 ou 4 colonnes de pixels. S'il y en a plusieurs, en déplaçant légèrement la tête, l'observateur voit la scène affichée en continu sous différents points de vue. Dans ce cas, la perception de la profondeur est augmentée (indice de mouvement de parallaxe). Mais ce changement de point de vue n'est possible que sur une petite portion d'espace (Figure 3.2).



**Figure 3.2 :** principe de la vision avec l'écran autostéréoscopique à réseau lenticulaire

Il existe aussi des écrans autostéréoscopiques à illumination ou à barrières de parallaxe. Pour les écrans à illumination, chaque colonne de lumière est espacée de telle façon qu'un observateur dans une bonne position voit celle-ci à travers la colonne paire de pixels avec son œil gauche, et à travers la colonne impaire de pixels avec son œil droit. Les pixels ne sont visibles dans ce dispositif que s'ils sont éclairés (Figure 3.3). Les *écrans autostéréoscopiques* à barrières de parallaxe sont constitués de deux filtres,

la barrière de parallaxe, qui oriente la lumière vers chacun des deux yeux et chaque colonne de pixels est destinée pour un œil.



**Figure 3.3 :** principe de la vision avec l'écran autostéréoscopique à illumination

Comme nous l'avons déjà indiqué, de nouvelles technologies d'écrans autostéréoscopiques sont en développement dont certaines seront exploitées dans le futur pour les écrans de visiocasques. La technologie la plus prometteuse est celle des écrans «Light Field», que l'on peut traduire par «champ lumineux». Nous détaillerons cette technologie dans le chapitre suivant sur les visiocasques.

### 3.1.2. Séparation par des lunettes

La *séparation* peut se réaliser par différenciation colorimétrique. Le procédé ancien et bien connu est avec des lunettes bicolors (procédé anaglyphe). Mais le procédé a été remis à jour par l'entreprise Barco (procédé Infitec). Chaque œil de l'observateur reçoit une image filtrée différente sur les composantes Rouge, Verte et Bleue (RVB) du spectre des couleurs. Ceci est possible en mettant une paire de lunettes ayant deux filtres colorimétriques différents et en équipant de même les deux vidéoprojecteurs, affichant les images sur un grand écran.

La *séparation* peut se réaliser par lunettes à obturateur électronique (lunettes actives). Ces lunettes sont composées de deux écrans à cristaux liquides qui obturent alternativement 60 (ou 50) fois par seconde chaque œil, pendant que le moniteur affiche les images à la fréquence de 120 ou 100 hertz.

La *séparation* peut aussi se réaliser par lunettes polarisantes (lunettes passives). Devant le moniteur est placé un écran à cristaux liquides permettant une polarisation de la lumière différente dans le temps pour chaque couple d'images stéréoscopiques. Avec des lunettes ayant deux filtres ne permettant que la transmission de l'image désirée pour chaque œil, l'observateur voit les images en relief. La fréquence du moniteur doit être aussi doublée pour la même raison que précédemment et le problème technique majeur est le même que pour les lunettes actives : chevauchement des images droite et gauche entre elles (crosstalk).

## 3.2. Systèmes à projection sur grands écrans

Afin de simuler un grand champ de vision à l'utilisateur, certains dispositifs sont basés sur des vidéoprojecteurs affichant des images stéréoscopiques sur de grands écrans. Ces installations sont exploitées en général pour représenter une scène virtuelle à l'échelle 1. La plupart du temps, la projection se fait via des miroirs pour limiter l'encombrement global du système. Le nombre de ces écrans ainsi que leur disposition et leur forme sont variés. L'un des éléments clés d'un système à base de projection sur grands écrans est le type de vidéoprojecteur utilisé pour afficher les images. Plusieurs technologies sont présentes sur le marché : LCD (Liquid Crystal Display), DLP (Digital Light Processing), LCOS (Liquid Crystal On Silicon), etc.

Il y a deux configurations spatiales. La première est réalisée pour les salles immersives, composées en général d'un grand écran semi-cylindrique. Le champ de vision horizontal pour l'observateur central peut être de 180° ou 210° et l'affichage est réalisé souvent avec un ensemble d'au moins trois vidéoprojecteurs en projection directe, (Figure 3.4).

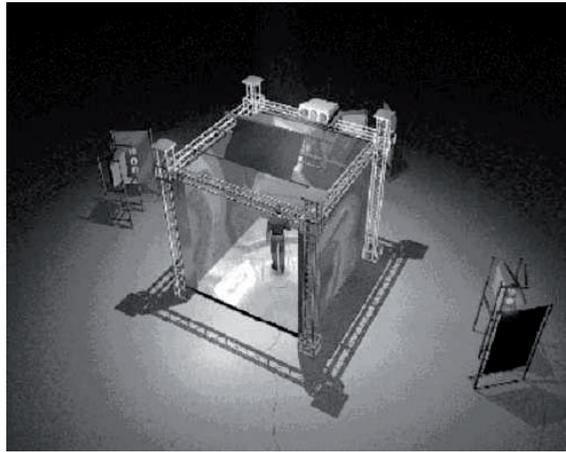


**Figure 3.4 :** salle immersive à trois vidéoprojecteurs

Si le champ de vision verticale est grand aussi, on peut obtenir que les champs de vision de l'observateur soient entièrement couverts, le cas des géodes. Ces salles immersives sont conçues pour un certain nombre de spectateurs et évidemment sans système de tracking du point de vue. L'image est calculée et exacte que pour un seul utilisateur, généralement l'observateur central, les autres observateurs ayant une vision légèrement déformée de la scène virtuelle.

Les CAVEs (Cave Automatic Virtual Environment), se présentent sous la forme d'une pièce cubique d'environ trois mètres d'arête et sont la solution la plus connue à base de projection sur grands écrans. C'est aussi des interfaces visuelles onéreuses. Des images stéréoscopiques du monde virtuel sont projetées sur un nombre de faces allant de quatre (trois murs et le sol) à six pour certaines configurations. Les CAVEs procurent une très bonne sensation d'immersion grâce à la disposition des écrans qui procure à l'utilisateur

un large champ de vision et grâce au changement de point de vue avec capture des mouvements de la tête. En tournant autour d'un objet, l'utilisateur peut ainsi examiner l'objet sous n'importe quel angle. Comparativement à un visiocasque, le CAVE peut avoir une très haute résolution, plusieurs vidéoprojecteurs pouvant même servir à créer l'image d'une face. En exemple, le CAVE installé chez Renault a cinq faces 4K à 120 Hz. Un autre avantage par rapport au visiocasque est que l'utilisateur est très bien « immergé corporellement », sans avoir besoin d'être représenté par son avatar, (Figure 3.5).



**Figure 3.5 :** structure d'un CAVE à 4 faces

## 4. LES INTERFACES PORTABLES

Ces interfaces portables sont principalement les visiocasques et les tablettes numériques. Ces dernières sont rarement utilisées dans des applications RV. Elles peuvent quelquefois servir de dispositifs de commande accessoire pour l'application RV, via des métaphores d'interactions.

Les visiocasques sont présentés en détail dans le chapitre suivant.

## 5. LA RESTITUTION VISUELLE STÉRÉOSCOPIQUE

Beaucoup d'interfaces visuelles proposent la vision stéréoscopique (images en relief). Elle est souvent perçue évidente d'emploi mais demeure une technique délicate à exploiter correctement dans certaines situations. Elle est source de contraintes visuelles pour certaines personnes, allant jusqu'à l'impossibilité de fusionner le couple d'images droite et gauche. Rappelons aussi que la vision stéréoscopique n'est pas indispensable pour la perception de la profondeur. Dans une application RV, tout concepteur doit se poser la question : la vision en relief est-elle utile, indispensable ?

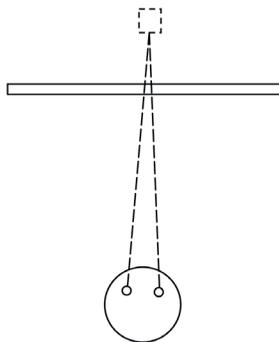
Il s'agit de créer deux images légèrement différentes, l'une correspondant à l'œil droit et l'autre à l'œil gauche. Le couple d'images stéréoscopiques est soit affiché sur deux écrans proches des yeux (le cas des visiocasques), soit sur un même écran, chaque œil ne voyant qu'une des deux images grâce à un dispositif technique.

Dans ce deuxième cas, un point dans l'espace 3D correspond à deux points homologues sur le plan pour le couple d'images stéréoscopiques. La distance entre ces deux points homologues est appelée **parallaxe**<sup>6</sup>. Celle-ci produit la **disparité rétinienne**<sup>7</sup> dans les yeux de l'observateur, explicitée au chapitre précédent. Elle fournit l'effet binoculaire de perception de la profondeur, la stéréopsie.

Dans le premier cas des visiocasques, le raisonnement géométrique est identique si on ne considère pas les écrans réels proches des yeux mais le plan de focalisation de visualisation de deux images « virtuelles<sup>8</sup> ». Les images sont perçues par l'observateur comme affichées à une distance de quelques mètres (1,3 mètres par exemple pour le visiocasque DK2 d'Oculus) grâce au dispositif optique composé de lentilles, voir le chapitre 4.

Il y a plusieurs cas de positionnement géométrique :

- Parallaxe horizontale positive (non croisée) : Quand l'objet observé est situé virtuellement derrière l'écran, la parallaxe horizontale est positive. Les points homologues gauche et droite sont respectivement affichés à gauche et à droite (Figure 3.6).



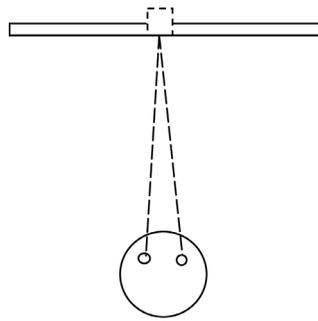
**Figure 3.6 :** parallaxe horizontale positive permettant de visualiser des objets derrière l'écran

- Parallaxe horizontale nulle : Quand la partie d'un objet virtuel se situe au niveau de l'écran, les points homologues ont une parallaxe nulle. Ils sont affichés aux mêmes endroits (Figure 3.7).

<sup>6</sup> On peut parler de parallaxe « absolue » : distance entre les points homologues ou de parallaxe « angulaire » : angle entre les points homologues, perçu depuis le point de vue de l'observateur.

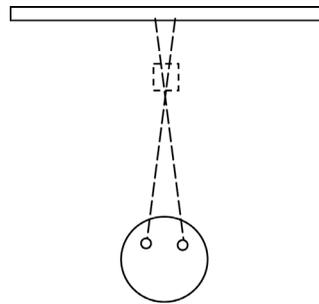
<sup>7</sup> On emploie le terme « parallaxe », au niveau de l'écran, et « disparité rétinienne », au niveau de l'œil. Il y a les parallaxes horizontales et les parallaxes verticales, en considérant la tête non inclinée évidemment, et de même pour les disparités rétinienne.

<sup>8</sup> « Virtuelles » au sens de l'Optique.



**Figure 3.7 :** parallaxe horizontale nulle permettant de visualiser des objets au niveau de l'écran

- Parallaxe horizontale négative (croisée) : Quand l'objet observé est situé virtuellement devant l'écran, la parallaxe est négative. Les points homologues gauche et droite sont respectivement affichés à droite et à gauche (Figure 3.8).



**Figure 3.8 :** parallaxe horizontale négative permettant de visualiser des objets devant l'écran

Nous insistons sur le fait que la vision d'images stéréoscopiques n'est pas identique à la vision spatiale réelle, car les images observées par les yeux ne sont que des images sur un plan, sur lequel la scène tridimensionnelle y est projetée. L'accommodation est figée puisque les yeux doivent se focaliser sur le plan de l'écran, ce qui n'est pas naturel et crée une incohérence entre accommodation et vergence. Bien des personnes pensent qu'une fois le matériel choisi, la création d'images en relief est une question relativement simple à traiter. Nous expliciterons au chapitre 8 que les configurations des algorithmes pour les images de synthèse sont à choisir judicieusement. Pour les images en relief, des compromis doivent être pris. Ils se basent sur des considérations géométriques, psychophysiques et cognitives.

---

# Étude fonctionnelle des visiocasques

---

## 1. INTRODUCTION

Comme nous l'avons précédemment indiqué, il n'y a pas d'interface idéale et universelle... Mais le visiocasque tente de s'en rapprocher pour permettre une grande immersion sensorielle visuelle. Le visiocasque idéal devrait avoir quatre capacités visuelles :

- de grands champs de vision horizontaux et verticaux correspondant à ceux des yeux ;
- une vision stéréoscopique dans tout le champ de vision binoculaire ;
- une très haute résolution graphique exploitant toutes les performances des acuités monoscopique et stéréoscopique ;
- une immersion du regard dans le monde virtuel.

Ce dernier point n'est possible que si l'interface visuelle est couplée à un capteur localisant en temps réel l'orientation de la tête de l'observateur.

Nous pouvons classer les différentes contraintes techniques pour la conception d'un visiocasque :

- à la base, il s'agit de réaliser un dispositif porté par l'utilisateur avec deux petits écrans associés à un **système optique** qui a pour objectif de focaliser deux images virtuelles à une certaine distance des yeux de l'observateur. Cette distance doit être plus grande que le punctum proximum d'accommodation pour que l'observateur puisse voir correctement les images. Associé à ce dispositif, il doit avoir un **capteur, ou un ensemble de capteurs, pour déterminer le mouvement de la tête** de l'utilisateur. Certains capteurs ne mesurent que les rotations de la tête, capteurs à trois degrés de liberté (3 ddl) et d'autres mesurent en plus les trois translations : capteurs à 6 ddl ;
- le système optique doit fournir un **champ de vision le plus grand possible** horizontalement et verticalement, ce qui est la plus difficile contrainte à réaliser ;
- **les mini-écrans** sont évidemment un composant essentiel d'un visiocasque pour fournir des images de qualité, stéréoscopiques non. Leur résolution doit être très grande, de plusieurs millions de pixels, et leur niveau de contraste élevé ;
- **l'ensemble du dispositif doit être ergonomique** pour permettre les bons réglages optiques, en fonction des différentes morphologies des utilisateurs, et être confortable, donc nécessairement léger.

Depuis plus d'un quart de siècle, des entreprises conçoivent et fournissent des visiocasques. Les premiers modèles étaient très rudimentaires, avec des mini-écrans LCD, ou très onéreux, en exploitant des mini-écrans à tube cathodique (CRT), très chers à fabriquer pour un marché de niche. Mais depuis le développement des technologies pour les smartphones avec son très grand marché, il est possible maintenant de fabriquer des visiocasques de meilleure qualité et de coût relativement modeste. Les paragraphes suivants présentent les différentes contraintes et les solutions retenues pour les visiocasques récents : la conception optique, les écrans, les capteurs de mouvements de la tête (le tracking) et la conception ergonomique. Mais avant, un paragraphe est consacré à la présentation des différents types de visiocasques, dont les produits commercialisés correspondants seront présentés au chapitre 6.

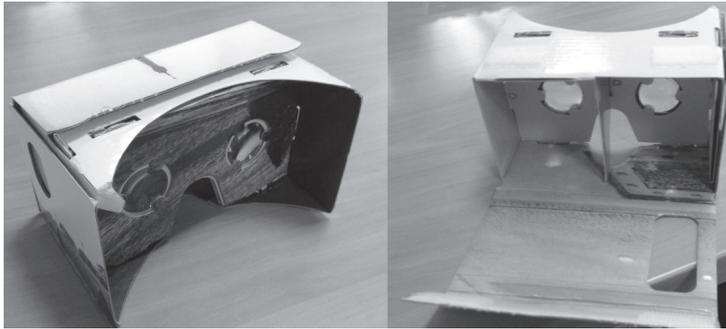
## 2. LES DIFFÉRENTS TYPES DE VISIICASQUES

### 2.1. Les visiocasques réalisés à partir d'un smartphone

L'idée de base semble intéressante : exploiter directement l'écran et les capteurs d'orientation d'un smartphone pour réaliser un visiocasque low-cost, puisqu'il suffit de concevoir et de vendre juste un support. Celui-ci est composé principalement de deux lentilles pour l'observation proche de l'écran et d'un emplacement pour insérer le smartphone. Certains ne proposent même pas une fixation sur la tête.

#### 2.1.1. Les visiocasques en carton

Les visiocasques sans fixation sur la tête sont des produits très bas de gamme, les «cardboards», comme le célèbre cardboard de Google (Figure 4.1). Ils sont d'un prix très modeste permettant même de les offrir dans des campagnes de marketing, avec les innombrables variantes visuelles. Il n'y a pas de réglage possible sur ce type de dispositif. Comme il faut tenir en main le visiocasque, la seule PCV (Primitive Comportementale Virtuelle) permise est l'observation d'un endroit, tel qu'un lieu touristique. Il est possible toutefois de se déplacer, le sens du déplacement étant celui de l'orientation du regard. Mais ce type de dispositif est fait principalement pour un usage ludique et très court, vu que les lentilles sont de qualité optique modeste et que la rigidité du visiocasque est très rudimentaire. La qualité visuelle du dispositif repose en grande partie sur celle du smartphone utilisé. Un petit avantage de ce type d'interface visuelle est d'être aisément transmis à quelqu'un d'autre. Il permet ainsi une utilisation en groupe, l'objet passant de main en main.



**Figure 4.1 :** le cardboard de google

### 2.1.2. Les visiocasques à support rigide

Par rapport à la catégorie précédente, l'usage pourra être plus prolongé car le support est rigide et ergonomique tel que le Gear VR de Samsung (Figure 4.2). Mais comme le prix doit être faible, peu de réglages sont possibles en général. Par exemple, le visiocasque Homido propose un jeu de trois portes lentilles. Il est compatible avec la plupart des smartphones et permet le réglage de la DIP. Quel que soit le soin porté à leur réalisation, la contrainte d'exploiter les structures mécanique et optique d'un smartphone implique de grands compromis pour réaliser un tel type de visiocasque qui n'aura pas une grande qualité. Comme on sait qu'il est très difficile de concevoir et réaliser un visiocasque de qualité, les dispositifs à partir de smartphone ne sont que des solutions secondaires. Leur principal intérêt est de permettre, sans investir dans du matériel onéreux, des essais d'application RV et de domaines connexes, par exemple la visualisation de photo panoramique. Des manettes de jeux peuvent être associées pour se déplacer dans le monde virtuel. Un avantage est l'absence de câble reliant l'ordinateur au visiocasque, la puissance informatique et graphique exploitée étant celle incluse dans le smartphone. Pour les véritables visiocasques, le câble est actuellement encore indispensable mais la puissance informatique est évidemment plus importante.



**Figure 4.2 :** le Gear VR, copyright de Villars VR<sup>1</sup>

<sup>1</sup> Villars VR organise un festival et un forum sur la RV, <http://villars-vr.com>

## 2.2. Les visiocasques conçus intrinsèquement pour la RV

La conception de ce type de visiocasque permet de mieux répondre aux différentes contraintes précédemment énumérées. La conception optique doit permettre de s'adapter davantage aux différentes morphologies des utilisateurs : la DIP est donc réglable et des réglages complémentaires peuvent permettre de rendre l'image la plus nette pour chaque œil indépendamment. Une contrainte souvent respectée est de permettre aux utilisateurs portant des lunettes de les garder en mettant le visiocasque. Les réglages pour chaque utilisateur peuvent se compléter par un calibrage au début de l'exploitation du visiocasque. À l'inverse du cas précédent, il y a en général un câble qui relie l'ordinateur au visiocasque permettant une grande puissance informatique et graphique.

Les différences essentielles entre les visiocasques commercialisés de cette catégorie concernent l'étendue des champs de vision et la résolution des écrans. Actuellement seul un visiocasque recouvre totalement les champs de vision de l'observateur, le StarVR (210° à l'horizontale dont 90° en recouvrement central pour les deux yeux et 130° à la verticale ; pour chaque œil 150° horizontalement), avec une résolution de 2560 x 1440 pixels par œil, ce qui donne une résolution angulaire de  $2560/150 = 17$  pixels par degré (Figure 4.3). Les autres visiocasques ont des champs de vision limités qui imposent de voir la scène virtuelle plus étroitement.

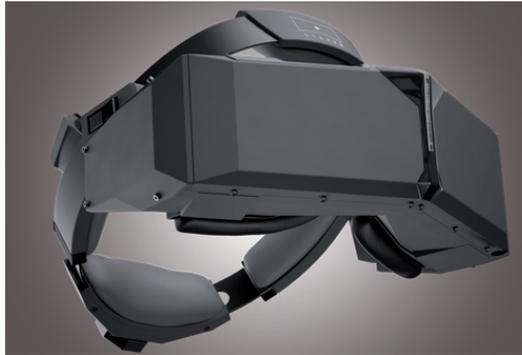


Figure 4.3 : le visiocasque StarVR

Certains produits peuvent incorporer un tracking des yeux qui permet de créer une visualisation plus adaptée à l'observateur comme on l'a déjà signalé, indépendamment de permettre des études ergonomiques sur la vision dynamique de l'utilisateur.

## 2.3. Les visiocasques conçus intrinsèquement pour la RA

Il existe des visiocasques pour la RA (Réalité Augmentée), « see-through HMD » en anglais, qui permettent de superposer des images de synthèse à la vision du monde réel. Ceux-ci ont été conçus au début pour certaines applications spécifiques, comme la vue simultanée d'informations virtuelles et de la vue extérieure pour les pilotes d'avions.

Quelle que soit le visiocasque RA, il repose sur un ou deux écrans semi-réfléchissants sur lesquels sont affichées des images qui se superposent à la vision humaine directe. Cette technique est reprise sur certains visiocasques à destination du public tel que l'Hololens de Microsoft. Dans ce type de visiocasque, le champ de vision total de l'observateur n'est pas occulté, comme dans les visiocasques précédents (visiocasques RV) mais le champ de vision où il y a association d'images virtuelles avec la vision directe du monde réel est en général limité. Le fait d'avoir la vision périphérique du monde réel non occultée sera un avantage, comme nous l'expliciterons au chapitre 8. On peut aussi toujours exploiter un visiocasque conçu pour la RV auquel on rajoute une caméra, ou deux caméras, filmant le devant du visiocasque. Il est alors possible d'associer des images de synthèse avec les images transmises par la caméra. On parle alors de «video see-through» au lieu de «optical see-through». Le champ de vision périphérique est dans ce cas occulté, la grande différence avec les visiocasques RA «optical see-through».

L'incrustation des images virtuelles peut être plus ou moins importante. Pour les Google Glass, l'incrustation était très modeste puisqu'il s'agissait que d'informations textuelles sur une petite zone du champ de vision (un petit écran sur le côté). L'observateur perçoit un petit écran flottant sans lien direct avec sa vision du monde réel. Dans d'autres cas, l'incrustation est plus importante puisqu'il s'agit de vraies images de synthèse et positionnées au centre du champ de vision de l'observateur. Certains dispositifs sont très légers, certains n'exploitent qu'un écran pour un seul œil et sont peu invasifs, voir le chapitre 6.

Dans certains cas, on peut parler de «vidéolunettes» qui sont des interfaces visuelles plus légères, différentes des visiocasques, car elles permettent de voir les images virtuelles avec un champ de vision non entièrement obstrué par le dispositif optique. L'observateur peut voir sur le côté son environnement extérieur. En corollaire, l'utilisateur perçoit alors des images sur un écran virtuel, flottant dans l'air, avec une immersion visuelle moins probante. Certaines vidéolunettes sont employées pour l'affichage d'images vidéo monoscopiques et d'autres avec des images stéréoscopiques. Certaines vidéolunettes sont sans tracking de tête, tandis que d'autres en incorporent.

### **3. LA CONCEPTION OPTIQUE**

Une des grandes difficultés des visiocasques est de permettre d'observer des images sur des écrans très proches des yeux. Pour cela on doit insérer des optiques entre les yeux et les écrans. Il faut parfaitement déterminer la géométrie et l'optique du visiocasque pour pouvoir calculer et afficher correctement les images stéréoscopiques sur les deux écrans. Si nous bougeons la tête, un objet fixe est vu sous différents angles et les projections de ses images sur les deux rétines varient. Mais notre cerveau perçoit sans effort cet objet comme étant fixe, car les «images rétinienne» sont cohérentes avec le mouvement de la tête, détecté par les systèmes vestibulaires. Le visiocasque ne doit donc pas perturber cette perception visuelle des objets immobiles. Les images affichées sur les écrans du visiocasque doivent en conséquence donner les mêmes

sensations visuelles que lorsque l'on regarde le monde réel. Ceci ne dépend pas que des caractéristiques du visiocasque, mais aussi de son correct positionnement par rapport aux yeux, de la qualité du capteur de localisation de la tête et des algorithmes calculant les images à afficher. L'utilisation efficace d'un visiocasque nécessite donc la résolution des problèmes précédemment indiqués.

Les images à afficher dans le visiocasque doivent suivre les règles de la perspective, des dimensions relatives et de la vision stéréoscopique. Concernant ce dernier point, la règle à respecter (limitation de la disparité rétinienne) revient au même raisonnement géométrique si on ne considère pas les écrans réels mais le plan de focalisation de visualisation de deux images « virtuelles », au sens de l'Optique (déjà expliqué dans le chapitre 3). La position du plan de focalisation ne peut être déterminé qu'en connaissant les caractéristiques des systèmes optiques, de leur position par rapport aux deux écrans et aux yeux. Plusieurs sources d'erreurs existent dans les dispositifs optiques. Ceux-ci créent des distorsions (radiales, astigmatisme, etc.). Le constructeur doit en tenir compte dans le modèle optique pour pouvoir corriger les images par un logiciel, qu'il faut obligatoirement utiliser. En stéréoscopie, les deux images virtuelles doivent être affichées en donnant une convergence correcte des yeux en fonction de la profondeur de l'objet observé. La distance interpupillaire étant variable d'une personne à l'autre, il faut prévoir un réglage mécanique. Il y a plusieurs solutions pour les systèmes optiques : soit avec plusieurs lentilles classiques, soit avec des lentilles de Fresnel. Ce dernier type de lentilles est utilisé pour le visiocasque StarVR permettant un grand champ de vision pour l'utilisateur.

## 4. LES ÉCRANS

### 4.1. Les écrans actuels

Il y a plusieurs caractéristiques importantes pour définir la qualité des écrans utilisés dans un visiocasque. La plus importante est la résolution<sup>2</sup> de l'écran en nombre de pixels par pouce. Mais pour comparer entre plusieurs visiocasques et surtout avec l'acuité visuelle humaine, il faut mieux déterminer le nombre de pixels pour un angle de vision donné (unité : nombre de pixels par degré), car dans un visiocasque, l'écran (ou deux écrans) peut recouvrir un champ de vision plus ou moins grand : en général vers 100°, mais à l'extrême 210°.

Comme nous l'avons précédemment indiqué, dans tous les visiocasques actuels la résolution n'est pas encore assez satisfaisante au regard des capacités du système visuel humain. La valeur moyenne (1 minute d'angle) de l'acuité visuelle peut être retenue (mais on pourrait prendre 0,5 minute d'angle) pour calculer le nombre de

---

<sup>2</sup> La résolution d'un écran mesure la finesse des détails d'un écran pour une dimension donnée. Elle s'exprime en pixels par pouce. La définition est le nombre de pixels qu'un écran affiche et s'exprime en nombre de pixels horizontaux et verticaux.

pixels nécessaire pour exploiter la capacité maximale d'acuité visuelle humaine : avec un champ de vision à  $210^\circ$  (horizontal) par  $140^\circ$  (vertical), il faudrait horizontalement au moins  $210 \times 60 = 12600$  pixels et verticalement  $140 \times 60 = 8400$  pixels, total : plus de 100 millions de pixels, et mieux avec le double de pixels... Il faudrait aussi avoir la puissance graphique pour créer de telles images en temps réel ! On est donc encore très loin de l'interface visuelle idéale avec des visiocasques ayant au mieux 5 K pixels.

À noter aussi que si la qualité d'un écran, par rapport à l'acuité visuelle humaine, dépend d'abord de sa résolution (en nombre de pixels par pouce), elle résulte aussi de la luminosité et du contraste de l'écran. La résolution de l'écran doit être satisfaisante : *c'est une condition nécessaire mais pas suffisante*. La luminosité et le rapport de contraste doivent aussi être satisfaisants ! Le rapport de contraste est donc également une caractéristique essentielle d'un écran. Il représente l'écart maximal de luminosité entre la couleur blanche et la couleur noire affichées au même moment sur l'écran. La luminance du noir ne peut pas être nulle car un écran n'est jamais parfait, un très bon écran affiche un noir en dessous de  $0,1 \text{ cd/m}^2$  (candela par mètre carré) et un mauvais écran au-dessus de  $0,3 \text{ cd/m}^2$ . Plus le ratio de contraste statique<sup>3</sup> est élevé, plus la qualité visuelle des images est supérieure pour une même résolution d'écran. Rappelons que le traitement neuronal dans les rétines se fait sur des contrastes et que l'acuité visuelle humaine est établie sur une différence de contraste entre des points noirs et blancs, comme expliqué au chapitre 2. Le rapport de contraste doit être déterminé pour une certaine luminance utile. En exemple, si on peut afficher une couleur blanche à  $200 \text{ cd/m}^2$  et un noir à  $0,25 \text{ cd/m}^2$ , on obtient un rapport de contraste typique de 800:1.

Comme nous l'avons déjà indiqué, les nouveaux visiocasques profitent de l'évolution constante des technologies des écrans pour les smartphones. Ces technologies sont sans cesse en renouvellement pour améliorer les caractéristiques des écrans. Certaines caractéristiques sont moins importantes pour les visiocasques, comme la consommation d'énergie, les commandes tactiles et les angles de vision. La caractéristique la plus importante est évidemment la résolution des écrans. Leurs définitions standard actuelles sont Full HD ( $1940 \times 1080$  pixels) et QHD ( $2560 \times 1440$  pixels).

La technologie de base exploite les écrans à cristaux liquides (LCD pour liquid crystal display). Ils utilisent la polarisation de la lumière grâce à des filtres polarisants et la biréfringence de certains cristaux liquides dont on peut faire varier l'orientation en fonction du champ électrique. L'écran LCD est donc un dispositif passif qui n'émet pas de lumière. Il doit donc disposer d'un éclairage qui sera régulé par la transparence variable des cristaux liquides. Bien des écrans à cristaux liquides couleurs de qualité utilisent la technologie TFT (Thin Film Transistor), dite à matrice active, qui permet d'obtenir des temps de réponse inférieurs à 10 ms. Le contraste reste toutefois limité à environ 300:1. Une deuxième technologie, une amélioration de la technologie TFT, est IPS (In-Place Switching), utilisée sur les smartphones haut de gamme actuels. Une

---

<sup>3</sup> Il existe aussi un rapport de contraste dynamique, qui n'a pas d'intérêt en pratique, l'affichage du blanc et du noir n'étant pas au même moment... Ce rapport sert seulement à fournir de très grande valeur pour impressionner le client!

technologie différente est l'écran AMOLED (Active-Matrix Organic Light-Emitting Diode). C'est un écran associant une technologie OLED avec une matrice active, qui envoie l'information électrique à afficher à **chaque pixel** individuellement. Il n'y a pas besoin de rétro-éclairage et cette technologie offre des temps de réponse très courts, inférieurs à 0,1 ms, ce qui se traduit par une bonne fluidité des images.

## 4.2. De nouveaux types d'écrans

Comme nous l'avons indiqué dans le chapitre précédent, de nouvelles technologies d'écrans sont en développement. Certaines pourront être exploitées pour les écrans de visiocasques. La plus prometteuse est la technologie «Light Field» (notée LF, que l'on peut traduire par «champ lumineux»). Cette technologie a déjà été employée récemment pour les appareils photographiques LF, ou appareils plénoptiques. Le principe de tels appareils repose sur le concept de champ lumineux : c'est une fonction de vecteurs qui décrit la quantité de lumière traversant dans chaque direction de chaque point de l'espace. Le champ lumineux est donné par la fonction 5D plénoptique<sup>4</sup> avec l'intensité de chaque rayon indiquée par sa luminance<sup>5</sup>. En simplifiant, on peut dire qu'il s'agit pour chaque point de l'espace 3D de tous les rayons lumineux traversant le point dans toutes les directions de l'espace. Un appareil photographique classique traite tous ces rayons ensemble et fournit une image 2D (une projection de l'espace 3D sur le plan 2D de l'image). La différence avec les appareils plénoptiques est qu'ils enregistrent le champ lumineux entièrement : la luminance de chaque couleur RGB et aussi la direction de chaque rayon. Ils utilisent un dispositif basé sur une matrice d'un très grand nombre de microlentilles devant le capteur très haute résolution qui capture la luminance pour les trois couleurs et la direction vectorielle des rayons lumineux. On obtient des informations directionnelles que les capteurs classiques ne peuvent pas enregistrer. Cela permet, entre autres, de faire des mises au point en post-traitements<sup>6</sup> à partir du fichier 4D de la photo.

Les futurs écrans LF font de même, mais en inversant le procédé évidemment, en fournissant des champs lumineux à l'observateur. Un écran autostéréoscopique LF exploite un réseau de microlentilles semi-sphériques et non pas un réseau de microlentilles semi-cylindriques, comme pour les écrans autostéréoscopiques classiques, présentés au chapitre précédent. L'écran doit être évidemment à très haute résolution permettant à chaque pixel d'afficher le bon rayon lumineux pour permettre à l'observateur de converger sur un point virtuel de l'espace avec la bonne accommodation, correspondant à la distance de profondeur du point observé. Il n'y a plus dans ce cas d'incohérence entre accommodation et vergence, qui crée des fatigues oculaires, voire pour certaines personnes des impossibilités de fusionner les images perçues par leur deux yeux.

4 Voir [https://en.wikipedia.org/wiki/Light\\_field](https://en.wikipedia.org/wiki/Light_field)

5 La luminance est une grandeur correspondant à la sensation visuelle de luminosité d'une surface. Elle se définit comme le quotient de l'intensité lumineuse de la surface source par l'aire de cette source projetée sur la perpendiculaire à la direction d'observation. Son unité SI est la candela par mètre carré (cd·m<sup>-2</sup>)

6 <https://pictures.lytro.com/>

Une autre technologie d'affichage est en développement et pourrait être exploitée dans de futurs visiocasques. Il s'agit de la technologie VRD « Virtual Retinal Display » (affichage rétinien virtuel) qui consiste à diriger sur l'œil des rayons de lumière sans passer par une réelle image sur un écran. Il n'y a donc plus d'écran réel avec cette technologie. Un faisceau lumineux provenant d'une diode laser, ou autre dispositif, balaie très rapidement la rétine de milliers de pixels qui forment une image virtuelle. La combinaison de rayons bleus, rouges et verts permet d'obtenir théoriquement des représentations en couleurs de haute résolution. Si un grand avantage est de supprimer le système optique de lentilles, lourdes, encombrantes et limitant le champ de vision, une grande difficulté est de pouvoir projeter correctement le faisceau lumineux sur le centre de la rétine, la fovéa. Le principal avantage est de fournir un système d'affichage stéréoscopique avec accommodation adaptée à la profondeur de chaque point de l'espace 3D. Il n'y a plus dans ce cas aussi d'incohérence entre accommodation et vergence. Des explications détaillées peuvent être consultées dans le livre « Displays, Fundamentals & Applications », de Rolf R. Hainich et Oliver Bimber<sup>7</sup>.

## 5. LE TRACKING DE LA TÊTE

### 5.1. Généralités

Pour une immersion du regard, il faut donc avoir un capteur de localisation associé au visiocasque. Celui-ci est positionné sur une position quelconque du visiocasque et il faut, bien sûr, déterminer où se situent les centres optiques des deux yeux par rapport à cette position du capteur. Ce décalage, cet « offset », dépend de la conception mécanique du système optique, de la fixation plus ou moins précise du visiocasque sur la tête de l'observateur et de la DIP de l'observateur. Un calibrage peut être prévu si on souhaite une immersion du regard précise. Les caractéristiques principales du capteur sont le temps de latence (ou latence), sa fréquence de mesure et ses précisions angulaires et en position.

Pour les techniques de réalité virtuelle, les mêmes types de traqueurs sont souvent employés dans les applications RV, non seulement pour la localisation de la tête de l'utilisateur portant un visiocasque mais aussi pour la localisation d'une ou deux mains et éventuellement l'ensemble du corps, la technique « motion capture ou mocap ». Nous présentons dans les paragraphes suivants les principes et les techniques de la localisation spatiale d'un objet ou d'une entité, mais seulement dans une zone restreinte telle que l'espace environnant d'un ordinateur ou d'une pièce (quelques centimètres à quelques décimètres).

Le positionnement d'un repère lié à un objet est parfaitement déterminé si le traqueur permet de mesurer les 6 paramètres (3 angles et 3 distances) liés aux 6 degrés de liberté (ou ddl) du repère. Ces 6 paramètres sont définis par la matrice de transformation

---

7 ISBN 978-1-56881-439-1, éditeur CRC Press.

géométrique entre le repère lié à l'environnement et le repère lié à l'objet. Cette transformation est composée d'une rotation et d'une translation. Les paramètres peuvent être définis de différentes façons, et en général les systèmes de mesure laissent à l'utilisateur le choix entre les différentes représentations. Pour les rotations, il est préférable d'utiliser les quaternions et non les angles d'Euler pour éviter les singularités.

Le capteur de positionnement est soit lié au visiocasque ou une partie du corps ou soit immobile par rapport à l'environnement. Dans ce dernier cas, c'est par l'observation du visiocasque que le capteur détermine les paramètres de localisation. Le traqueur parfait (ou « presque parfait » en ne tenant pas compte du prix) n'existe pas. Tous les traqueurs ont au moins une contrainte d'exploitation ou des performances restrictives. Les techniques de réalité virtuelle sont pourtant très demandeuses de ce type d'interface. Mais concevoir et réaliser un capteur de localisation bon marché qui fournisse six mesures très précises en temps réel (à une fréquence supérieure à 100 Hz et avec un temps de latence de 1 ou 2 ms<sup>8</sup>), sans contrainte d'exploitation, est un challenge technique très difficile. C'est sous l'essor de la réalité virtuelle et des smartphones que sont apparus les capteurs de localisation low-cost les plus évolués. Les principes physiques exploités actuellement sont variés, chacun ayant au moins un avantage par rapport aux autres.

Rappelons pour mémoire qu'il est toujours possible d'envisager l'utilisation d'un capteur basé sur une structure de liaisons mécaniques entre un support fixe de la pièce et le visiocasque. La précision est bien meilleure et le temps de latence bien plus faible que les autres types de capteurs. Les mouvements relatifs entre les pièces mécaniques sont des rotations. Celles-ci sont mesurées grâce à des capteurs optiques incrémentaux. Les mesures sont très rapides (jusqu'à 300 mesures par seconde) et le temps de réponse est très court (2 à 3 ms). Mais la gêne occasionnée par l'utilisation de liaisons mécaniques entre un support fixe de la pièce et le visiocasque fait que ce type n'est plus utilisé, sauf éventuellement pour des expérimentations scientifiques spécifiques (spatiales, etc.)

Il y a deux principaux types de capteurs : les capteurs optiques avec caméras et les centrales inertielles. Ces dernières sont composées de capteurs mécaniques déterminant une orientation, une vitesse ou une accélération, dont nous présentons les principes.

## 5.2. Capteurs basés sur des principes mécaniques

Plusieurs types de traqueurs mécaniques sont regroupés sous cette même partie car ils sont basés sur des principes mécaniques ne fournissant, directement ou indirectement, la mesure de quelques degrés de liberté. Il faut donc les associer entre eux pour réaliser des capteurs de localisation à 3 ou 6 ddl. Ils sont basés sur la technologie MEMS (Microelectromechanical systems) qui est fondée sur un système électromécanique,

---

<sup>8</sup> On rappelle qu'il est nécessaire d'avoir une fréquence élevée pour avoir un temps de latence faible mais cela n'est pas suffisant ! Le temps de latence ne peut être très faible qu'avec une fréquence très élevée.

de dimensions micrométriques, comprenant un ou plusieurs éléments mécaniques et électriques pour réaliser une fonction de capteur. La technologie MEMS, à base de silicium en général, est issue de la technologie microélectronique, permettant une grande production à faible coût. Les différents éléments mécaniques sont sollicités grâce à des forces générées par des transducteurs électromécaniques : électrostatiques, capacitifs, magnétiques, etc. Ils mesurent différentes caractéristiques mécaniques.

Les gyromètres permettent de mesurer des vitesses angulaires, à différencier des gyroscopes, mesurant des angles, dont les premiers réalisés étaient sur le principe mécanique d'un rotor tournant à grande vitesse, dont l'axe garde une direction constante. Mais d'autres phénomènes physiques sont exploités pour mesurer des vitesses angulaires de rotation. Ces capteurs ont le défaut d'avoir une dérive dans le temps.

Les accéléromètres sont bien sûr basés sur la mesure d'une force provenant de l'accélération d'une masse. Les accéléromètres peuvent être en boucle ouverte ou en boucle fermée : pour la mesure en boucle ouverte, la force est mesurée par le faible allongement d'un ressort supportant la masse. Pour ce type d'accéléromètre en boucle ouverte, il y a le problème classique de compromis à faire entre précision et rapidité de réponse. Pour les accéléromètres en boucle fermée, on crée une force opposée à la force inertielle, annulant le déplacement de la masse. Cette force d'opposition est créée par un système moteur dont la commande est fonction de cette force, donc de l'accélération. Ce principe permet d'obtenir des accéléromètres de meilleure précision. La mesure de localisation de l'objet en mouvement est obtenue par double intégration, d'où une augmentation quadratique des erreurs de position. Les dérives dans le temps sont donc le principal inconvénient de ce type de traqueurs.

L'accélération terrestre est à prendre en compte dans les mesures, par exemple avec un inclinomètre. Pour ce type de traqueur, c'est la terre qui est la source émettrice du champ gravitationnel. Sans mouvement de l'inclinomètre, celui-ci indique sa position angulaire avec la verticale du lieu. Mais s'il y a mouvement, il indique la direction de la force extérieure totale sur lui, due à l'accélération du mouvement et à l'attraction terrestre.

La terre crée un grand champ magnétique qui peut être utilisé pour des mesures de localisation. Les compas électroniques miniatures exploitent principalement l'effet Hall. Ce phénomène consiste en une déviation de courant dans une plaque métallique très mince, quand on place celle-ci dans un champ magnétique normal à son plan. Avec deux capteurs à effet Hall perpendiculaires entre eux, on peut déterminer le nord magnétique terrestre.

La fusion des données de tous ces capteurs, que sont les accéléromètres, les gyromètres et les magnétomètres, va permettre de mesurer en temps réel les six degrés de liberté d'un objet, que ce soit un visiocasque ou un smartphone. Pour ce dernier, les capteurs peuvent aussi servir de boussole, de commandes ou d'autres fonctions. Les recherches et développements se poursuivent pour diminuer le temps de réponse et pour pallier aux imprécisions des mesures, même avec la fusion de données. Les avantages et

inconvenients sont à comparer avec l'autre principe général de mesure de localisation à base d'optique, de caméras. Un calibrage des traqueurs est nécessaire et proposé sur les smartphones.

### 5.3. Capteurs basés sur des principes optiques

Il n'y a pas dans cette catégorie de traqueur standard. La technologie des traqueurs optiques a vraiment progressée, notamment celle liée aux capteurs matriciels type CCD ou CMOS, ainsi que les algorithmes utilisés en vision par ordinateur. Les algorithmes de calibrage de caméra et de reconstruction tridimensionnelle sont à présent bien maîtrisés par les industriels.

De manière générale, ces traqueurs fonctionnent en associant des sources lumineuses et des capteurs photosensibles, ponctuels (phototransistor) ou planaires (caméra à technologies CCD ou CMOS). Les sources lumineuses sont en général réalisées à partir de diodes électroluminescentes (LED, Light Emitting Diode) et émettent dans le spectre visible ou proche infrarouge. En réalité virtuelle, les traqueurs utilisent la partie infrarouge justement parce que l'œil humain n'y est pas sensible. Le spectre visible peut cependant être utilisé pour les systèmes portables comme les visiocasques.

Le principe géométrique de la localisation tridimensionnelle de points est généralement basé sur la triangulation. Dans le cas de deux caméras, le problème se pose de la manière suivante : on considère deux caméras dont la position relative des centres optiques est spécifiée par une translation et une rotation. Connaissant les paramètres propres aux caméras et le positionnement relatif des caméras entre elles, il est possible de calculer la position de tout point de l'espace des caméras. Les paramètres propres à la caméra se nomment paramètres intrinsèques. Il s'agit généralement de la distance entre le centre optique et le plan image ( focale), et de la projection du centre optique dans le plan image. Des paramètres de distorsion dus à l'optique peuvent être aussi ajoutés. Tous ces paramètres sont obtenus par une procédure de calibrage. Plusieurs méthodes existent, mais leurs principes reposent tous sur l'acquisition d'images d'un objet dont la structure géométrique est parfaitement connue. En ce qui concerne le positionnement relatif des caméras entre elles, on parle de paramètres extrinsèques. Le problème se généralise à plusieurs caméras. Les systèmes commercialisés font d'ailleurs souvent appel à un nombre de caméras supérieur à 2. Cela permet d'augmenter le volume de travail, mais aussi de pallier le problème majeur des systèmes optiques, à savoir la non-visibilité de la cible.

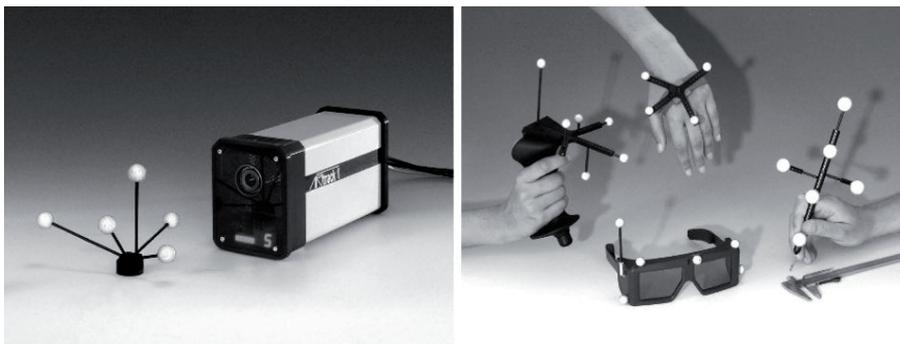
Les traqueurs sont généralement répartis suivant un des deux principes suivants : outside-in ou inside-out. Dans le cas outside-in (de l'extérieur vers l'intérieur), les caméras sont fixes et les marqueurs sont placés sur le visiocasque. La difficulté de ce principe réside dans le fait que les repères ponctuels se déplacent peu relativement pour les rotations du visiocasque, d'où une précision relativement faible en rotation mais très précise en translation, ce qui fait son grand intérêt actuel. Evidemment, concrètement il faut une salle dédiée, équipée de plusieurs caméras fixées sur trépied ou aux murs. Dans le cas inside-out (de l'intérieur vers l'extérieur), les marqueurs sont

fixes, accrochés aux murs de la pièce par exemple, et la (les) caméra(s) est embarquée sur le visiocasque (aisé si le visiocasque est réalisé à partir d'un smartphone). Le principal inconvénient réside dans le fait que les mesures en translation sont moins précises. Actuellement, les deux solutions coexistent pour l'exploitation du tracking de la tête (et autres partie du corps) sur les visiocasques, chacune ayant ses propres avantages et inconvénients.

Il existe des marqueurs passifs ou des marqueurs actifs : les marqueurs passifs sont de petites sphères réfléchissantes, de simples planches de motifs noirs et blancs, etc. L'avantage des marqueurs passifs est d'être libre de tout fil. En revanche, ils nécessitent une source lumineuse. Dans le cas de sphères, la source lumineuse est située proche de l'axe optique des caméras pour que les rayons réfléchis soient perçus par la caméra. Les marqueurs actifs sont généralement constitués de LED. Ils ont l'inconvénient de nécessiter une alimentation. Leur grand avantage réside dans la possibilité de les déclencher les uns après les autres, ce qui rend trivial leur identification dans l'image.

En exemple, l'ensemble logiciel ARToolkit (Augmented Reality Toolkit), proposé en « open source », permet de mettre en œuvre la solution inside-out. Son principe est le suivant : des cibles prédéfinies sont imprimées sur une imprimante standard, puis fixées sur un support rigide. Le programme permet, par analyse d'image, de calculer la localisation et l'orientation de la caméra et donc du casque.

Un autre exemple, le système ARTtrack est commercialisé par la société ART (Advanced Realtime Tracking) (Figure 4.4). Il suit le principe outside-in. Un système traqueur est constitué d'au moins deux caméras (jusqu'à huit), d'un ordinateur et de marqueurs. Le nombre de caméras dépend essentiellement du volume de travail. Les marqueurs se présentent sous la forme de sphères réfléchissantes. Le traqueur renvoie la localisation pour les marqueurs isolés, et donne la localisation et l'orientation pour des structures mécaniquement rigides. Celles-ci sont simplement constituées de plusieurs marqueurs astucieusement disposés. Dans une application typique, les corps rigides se montent sur un visiocasque, sur une paire de lunettes stéréoscopiques, pour traquer le point de vue, ou sur la main pour permettre une interaction avec le monde virtuelle.



**Figure 4.4 :** le système ARTtrack commercialisé par la société ART

Les performances annoncées sont en précision de 0.4 mm en translation et 0.12 degré en rotation, pour une fréquence maximale de 60 Hertz. Le volume de travail est de quelques mètres cube. Les paramètres intrinsèques et de distorsion de chaque caméra sont calculés au sein de la société et stockés dans chaque caméra. Les paramètres extrinsèques liés au traqueur dépendent de la position relative des caméras, et doivent donc être calculés à chaque fois que l'on installe le système sur le site. Cette opération consiste généralement à balayer une structure mécanique connue dans le champ des caméras durant environ une minute. La structure peut être par exemple une poutre d'environ un mètre sur laquelle sont fixées des cibles. Le programme de calibrage calcule les paramètres extrinsèques. Le principal inconvénient des capteurs optiques reste la nécessité de ne pas obstruer la ligne de vue allant des caméras aux marqueurs. Ceci est néanmoins pris en compte dans les systèmes les plus récents, notamment grâce à une redondance des caméras et des marqueurs.

Les technologies des capteurs pour traquer la tête ont beaucoup progressé ces dernières années pour fournir des mesures en temps réel satisfaisantes. Mais il reste encore des recherches et des développements à entreprendre pour améliorer les précisions des mesures et pour diminuer le temps de latence. Plusieurs recherches concernent la fusion de données des capteurs des centrales inertielles avec celles des caméras puisque tout ce type de matériel équipe actuellement les smartphones, comme les travaux de M. Alibay<sup>9</sup>.

## 6. LA CONCEPTION ERGONOMIQUE

L'utilisateur tire profit d'un visiocasque si l'ergonomie de l'interface est suffisante pour qu'elle devienne transparente. Ceci impose que les dispositifs soient très légers, confortables, avec une bonne ventilation et respectant des règles de sécurité. L'utilisation d'un visiocasque est intrusive, en ce sens que le visiocasque est plus ou moins lourd, que ses mouvements peuvent être plus ou moins contraints par des câbles vidéo et que le centre de gravité de la tête est déplacé.

Il peut être nécessaire de chercher à évaluer le biais induit par le simple port d'un visiocasque si, bien sûr, la posture de la tête de l'utilisateur est importante pour l'application (par exemple, l'étude de la visibilité d'une scène routière depuis l'intérieur d'un véhicule). Pour ne pas perturber les mouvements de la tête, il ressort que les visiocasques doivent être symétriques et avoir un centre de gravité vers l'avant de la tête afin de profiter des muscles dorsaux du cou plus puissants et moins sujets à la fatigue.

Pour étudier tous ces problèmes et caractériser la qualité d'un visiocasque comme interface visuelle, il ne faut pas se contenter de demander à l'utilisateur : voyez-vous correctement les images ? Nous avons proposé un certain nombre de tests associés à l'utilisation professionnelle d'un visiocasque. Les résultats ne dépendront pas que des caractéristiques intrinsèques du visiocasque mais aussi de la capacité de l'ordinateur

---

<sup>9</sup> Thèse de M. Alibay, "Extended sensor fusion for embedded video applications", Mines ParisTech, 2015.

en images de synthèse et de votre propre système visuel. Rappelons qu'avant d'étudier l'immersion et l'interaction cognitives et fonctionnelles d'une application RV, il faut valider les I<sup>2</sup> sensorimotrices. Par des tests, il est souhaitable de valider les deux principales fonctionnalités d'un visiocasque, qui peuvent ne pas être correctement réalisées :

- localisation du visiocasque, deux problèmes majeurs :
  - la précision de la localisation de la tête est-elle suffisante pour l'application ?
  - le temps de latence induit par le capteur et par la création des images est-il acceptable ?
- vision, deux problèmes majeurs :
  - la compatibilité des champs de vision dans le visiocasque en environnement virtuel avec les champs de vision exploités en réel (horizontaux, verticaux et zone de recouvrement) ;
  - utilité, validité et contraintes de la vision stéréoscopique pour l'application.

Pour une validation soignée de l'exploitation d'un visiocasque, divers tests techniques et psychophysiques peuvent être entrepris :

- des tests dits préliminaires, qui servent à valider le couple capteur-visiocasque ;
- des tests statiques qui doivent être effectués avant chaque nouvelle utilisation du casque ;
- des tests dynamiques qui sont effectués pour comparer la visibilité en environnement réel et en environnement virtuel, sur des exemples simples ;
- des tests ergonomiques pour étudier le confort et la fatigue du visiocasque.

### Tests préliminaires

Il existe un certain nombre de tests à mener avant toute utilisation généralisée d'un visiocasque :

- répétitivité des réglages de la distance interpupillaire (DIP) ;
- qualité et résolution des images affichées (Pour cela, on peut demander à l'utilisateur de lire un texte à l'écran) ;
- perception des interstices entre pixels, en focalisant ou non votre regard sur les écrans ;
- positionnement vertical des écrans (à vérifier en traçant des lignes horizontales sur les écrans à même hauteur et voir si ces lignes coïncident) ;
- mesures des éventuelles distorsions des images en affichant des grilles, ensemble de droites verticales et horizontales régulièrement espacées. Cela permet de percevoir des formes déformées à la place de carrés, en particulier en vision périphérique ;
- mesure colorimétrique en affichant une grille standard de couleur pour discerner les variations de couleur détectables et tests simples en affichant sur tout l'écran

une seule des trois couleurs R, G ou B pour détecter si la luminosité et la teinte sont bien homogènes ;

- Positionnement des écrans par rapport au champ de vision souhaité. Il ne faut pas oublier que les optiques sont circulaires et que les écrans sont rectangulaires. Comment sont-ils positionnés entre eux et quelle est l'incidence sur les champs de vision horizontaux et verticaux ? Comment se fait précisément la zone de recouvrement binoculaire (overlap) ? Comment sont positionnés visuellement les deux écrans et les deux optiques ?
- Tests pour l'estimation du recouvrement : faire avec de simples lignes verticales sur fond noir pour chaque observateur, qui doit voir la même ligne un œil après l'autre dans la zone de recouvrement

À noter, que généralement le fabricant de visiocasque ne vous fournira pas le modèle optique de sortie de son produit ainsi que ses caractéristiques précises : variation maximale de DIP, tracé précis des champs de vision monoscopiques et stéréoscopiques, positions des centres optiques par rapport à celle du capteur de localisation, etc. Par expérience, les valeurs annoncées et les valeurs mesurées sont des fois différentes...

Si vous souhaitez mesurer les champs de vision monoscopiques et stéréoscopiques, il est possible de le faire en mettant des caches en carton devant les yeux du testeur, obstruant une partie de son champs de vision. Il faut d'abord mesurer ce champ de vision ainsi réduit en vision réelle : en pointant sur un mur les limites du champ de vision et mesures de distance pour calculer les tangentes des angles. Il faut ensuite demander au testeur de mettre le visiocasque tout en gardant le cache. En affichant des marqueurs mobiles sur un des deux écrans (un pixel, une croix, un petit segment horizontal ou vertical, ...), on peut déterminer les positions X et Y en pixel sur l'écran où il y a occultation par le cache en carton. Par répétitions sur tout le pourtour du champ de vision avec plusieurs caches de plus en plus grands, on peut déterminer par calcul le contour du champ de vision effectif pour chaque œil d'un utilisateur donné. On peut procéder à l'inverse : découper un seul cache en carton correspondant au champ de vision d'un œil, en faisant la découpe progressivement par itération en commençant par une petite ouverture. On peut aussi déterminer la répétabilité de mesures en demandant à plusieurs testeurs de mettre plusieurs fois le visiocasque, des légères modifications de position influençant les champs de vision effectifs (voir Figure 4.5).



**Figure 4.5 :** cache en plastique, fait par imprimante 3D et dimensionné progressivement pour encadrer le champ de vision effectif vu par l'observateur

**Tests statiques :**

- positionnement du casque : on fait le réglage de la DIP et on demande à l'utilisateur de voir des rectangles blancs affichés sur les bords des écrans : sont-ils entièrement visibles ?
- acuité visuelle monoscopique : par exemple, séparation d'une ligne noire horizontale d'épaisseur 1 pixel entre deux lignes blanches horizontales d'épaisseur 1 pixel. En vérifiant ainsi l'acuité angulaire effective dans le visiocasque qui dépend de la résolution de l'écran mais aussi de son rapport de contraste et de sa luminosité. Cela permet de relativiser avec l'acuité visuelle naturelle ;
- acuité visuelle stéréoscopique : détermination de la possibilité de perception de profondeur en vision stéréoscopique sans indices monoculaires. Par exemple, en décalant en profondeur un plan à côté d'un autre et en demandant au testeur s'il perçoit la différence en profondeur entre les deux plans (devant ou derrière et en calculant le JND<sup>10</sup>).

**Tests dynamiques :**

- mesure du temps de latence : il est aisé de mesurer le temps de latence de la boucle « action – perception » quand un utilisateur est immergé devant un simple écran, petit ou grand : il suffit avec une caméra à très haute fréquence (1000 Hz pour une précision de 1 ms) de filmer la main traquée du testeur en même temps que les images affichées sur l'écran. Le temps de latence est mesuré ensuite sur la vidéo : c'est l'écart entre le début du mouvement de la main et le début du mouvement de l'animation de l'image due à l'interaction par la main du testeur. Avec un visiocasque, la mesure est un peu plus délicate à effectuer : la caméra doit filmer la manette de jeux en même temps qu'une des deux lentilles de sortie du visiocasque dans laquelle on peut percevoir l'animation de l'image. On peut aussi penser utiliser un écran de contrôle qui affiche les images en parallèle à celles dans le visiocasque mais il faut être sûr qu'elles sont affichées au même temps à la milliseconde près.
- D'autres tests peuvent dépendre de l'application envisagée en fonction des fonctions souhaitées.

**Tests ergonomiques :**

- encombrement et poids du visiocasque ;
- réglages aisés de la DIP, de la netteté et de la fixation de l'appareil sur la tête. Des mesures de répétabilité peuvent être effectuées ;
- fatigue visuelle due à l'immersion et à la stéréoscopie ;
- confort thermique.

Les différentes formes d'installations corporelles pour exploiter un visiocasque sont variées et posent quelques problèmes ergonomiques spécifiques. Nous pouvons lister les différents types d'installation, sans être exhaustif :

---

<sup>10</sup> Le JND (Just Noticeable Difference - seuil différentiel) définit la limite en dessous de laquelle un testeur ne parvient plus à différencier deux stimulations différentes.

- l'utilisateur est assis sur un siège, sans possibilité de rotation de son corps, position classique pour les jeux vidéo. L'immersion du regard ne pourra pas se faire aisément sur 360 degrés;
- pour pallier au cas précédent, l'utilisateur est assis sur un fauteuil pivotant sur 360° et il est ainsi plus immergé visuellement au niveau I<sup>2</sup> sensorimotrices. Le câble éventuel entre le visiocasque et l'ordinateur peut être une gêne;
- l'utilisateur est assis un siège d'une petite plate-forme mobile et ainsi son corps est stimulé par des mouvements de rotation (2 ou 3 rotations : roulis, tangage et lacet) et éventuellement une translation. il est ainsi plus immergé au niveau I<sup>2</sup> sensorimotrices car son système vestibulaire est correctement (ou partiellement)<sup>11</sup> stimulé. L'utilisateur peut être aussi couché et non assis sur une telle plate-forme pour une simulation de vol en EV;
- l'utilisateur est debout sur place et peut tourner sur lui-même, même s'il y a un câble qui peut être gênant. Ce dernier peut être accroché au plafond pour plus de liberté de mouvement de l'utilisateur;
- l'utilisateur est debout sur place, peut tourner sur lui-même et peut marcher sur un tapis roulant 1D ou 2D. Il est ainsi plus immergé au niveau I<sup>2</sup> sensorimotrices car sa proprioception musculaire est correctement ou partiellement<sup>10</sup> stimulée;
- l'utilisateur est debout et peut se déplacer dans toutes directions sur quelques mètres, dans une pièce ou sur plus grande distance dans un lieu aménagé en conséquence. Il est ainsi correctement immergé au niveau I<sup>2</sup> sensorimotrices (pour sa vision, sa kinesthésie et sa proprioception musculaire) lors de ses déplacements, s'ils sont bien sûr cohérents en ER et en EV. Si les déplacements sont dans une pièce, le câble peut être une gêne et il faut prévoir une sécurisation pour que l'utilisateur ne rencontre pas une paroi de la pièce (voir chapitre 8). Si l'espace est plus grand, l'utilisateur doit exploiter un visiocasque avec smartphone ou porté un ordinateur sur soi.

---

<sup>11</sup> Nous expliciterons les divers cas dans le chapitre 7, suivant que les déplacements en EV correspondent totalement ou partiellement aux stimuli sensoriels sur le corps du sujet.

---

## Les interfaces associées au visiocasque

---

### 1. INTRODUCTION

Nous ne présentons pas dans ce chapitre toutes les interfaces de la réalité virtuelle, mais seulement quelques catégories d'interfaces qui peuvent être exploitées avec un visiocasque. Certaines interfaces ont été développées pour tout type de restitution visuelle tandis que d'autres, souvent récentes, ont été développées spécifiquement pour les visiocasques. Plusieurs interfaces de commande manuelle ont été conçues par les fabricants de visiocasques en vue d'être employées avec leur produit. Les interfaces de commande manuelle, les « manettes », sont les plus indispensables pour compléter l'usage des visiocasques dès que l'on veut (inter)agir dans l'EV.

Nous allons présenter les principales interfaces associées aux visiocasques :

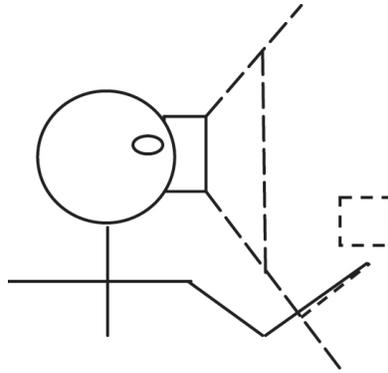
- les interfaces de commande et d'interaction manuelle ;
- les interfaces acoustiques ;
- les interfaces de marche : tapis roulants 1D ou 2D ;
- les interfaces de simulation de mouvement ;
- les interfaces de suivi des yeux.

### 2. INTERFACES DE COMMANDE ET D'INTERACTION

Il est possible de rencontrer en réalité virtuelle toutes les catégories d'interfaces motrices, des plus simples telles que les dispositifs de commande classique au plus complexes, les interfaces à retour tactile et les interfaces sensorimotrices à retour d'effort. Ces dernières, relativement onéreuses, ne seront guère utilisées pour les applications RV grand public. Pour ce type d'application, le retour tactile, moins onéreux, pourra être proposé, guère sur des interfaces spécifiques, mais plus en complément des dispositifs de commande. Ces derniers sont utilisés pour envoyer des ordres au monde virtuel ou au système qui le gère. Parmi ces interfaces de commande, il y a les souris, les joysticks et surtout les manettes. Il existe aussi les pointeurs qui permettent de sélectionner une zone de l'écran. Les pointeurs sont

composés d'un traqueur de localisation en orientation et de plusieurs boutons de commandes binaires. Ils peuvent donc être aussi utilisés dans un visiocasque en dessinant leur axe de pointage tel un faisceau laser. L'utilisation de tous les types d'interfaces classiques de commande est possible mais pour certains l'usage n'est pas pratique quand il faut les utiliser sans les voir. C'est *néanmoins* possible de les manier après apprentissage.

Le plus intéressant dispositif est la manette d'interaction manuelle, en général proposée en double, une pour chaque main, dont les derniers modèles sont en développement chez les constructeurs de visiocasques. Ce sont des interfaces indispensables pour offrir plus de possibilités d'applications RV avec les visiocasques. Quel serait l'intérêt d'être immergé visuellement dans un EV si on ne pouvait pas au minimum manipuler les objets devant soi? La manipulation peut se faire à distance, en « téléopération » ou en colocalisation main-tête (voir le chapitre 8). La colocalisation est la superposition de l'espace visuel avec l'espace de manipulation pour l'utilisateur de l'application RV.



**Figure 5.1 :** Colocalisation en manipulation (ER : traits continus et EV : traits pointillés)

Les constructeurs de visiocasques privilégient avec leurs dernières manettes une manipulation en colocalisation, la main réelle étant perçue par proprioception au même endroit que la main virtuelle vue à travers le visiocasque. Certaines de ces manettes ont été améliorées par des dispositifs à retour tactile, mais ponctuellement, sur un bouton ou un joystick. Elles sont alors équipées de petits vibreurs commandés en intensité et en fréquence pour simuler certains effets. Il n'est guère envisagé d'interface à retour tactile sur toute la main ou seulement sur les bouts des doigts pour des applications RV grand public. Ces dispositifs exigeront l'usage d'un gant, en complément d'une manette. S'il y a eu des essais il y a une vingtaine d'année, ils ne furent pas couronnés de succès et furent vite abandonnés car l'usage de gants pose des contraintes ergonomiques, d'hygiène, de fiabilité et de tailles pour correspondre à toutes les morphologies des mains. Une entreprise développe néanmoins des gants tactiles, les gants Gloveone, qui permettront de ressentir le toucher des objets virtuels manipulés (mais par leur poids et leur rigidité), grâce à l'intégration de dix

mini-capteurs dans les gants créant des petites vibrations à des intensités et des fréquences variables et contrôlables. Mais le prix d'un seul gant est assez élevé (200 € environ en prévision) et il est utile de compléter l'interface avec un système de détection de la position des doigts (comme le capteur de la société « Leap Motion ») pour la reconnaissance des mouvements des doigts (voir plus loin).

Les principes de localisation et du suivi des mains sont similaires à ceux du suivi de la tête, présentés dans le chapitre précédent. Avec les derniers modèles de manettes en développement chez les constructeurs de visiocasques, c'est la même technique qui est mise en œuvre pour le suivi de la tête et des mains, ce qui est compréhensible.

Nous avons vu qu'il y a deux principaux types de capteurs de localisation de la tête : les capteurs optiques avec caméras et les centrales inertielles. Ces dernières peuvent être utilisées pour réaliser des manettes mais ce sont surtout les capteurs optiques qui sont exploités et en développement car ils offrent une bonne précision avec des latences très faibles. Les manettes récentes utilisent des capteurs optiques avec le principe *outside-in* : les caméras sont fixes et des marqueurs sont placés sur les manettes car ce principe permet une très bonne précision en positionnement sur un espace de travail suffisamment vaste. Evidemment, concrètement il faut une salle dédiée, équipée de plusieurs caméras fixées sur trépied ou aux murs. C'est un petit inconvénient mais qui sera rédhibitoire pour tous les usagers qui n'auront pas la possibilité de consacrer une salle, ou au moins quelques mètres carrés, 3 m par 4 m dans l'idéal, à une installation RV. L'inconvénient majeur des capteurs optiques est le risque d'occultation pendant le suivi des mains : pour localiser un visiocasque et ses cibles fixées sur lui, il n'y a pas trop de difficultés car elles sont presque toujours visibles en grande quantité par les caméras placées en hauteur. Mais pour les manettes tenues en main, les cibles incorporées aux manettes, même judicieusement, peuvent être occultées par les mains, les bras ou le corps. La forme des manettes et le positionnement des cibles à suivre en temps réel est un problème ardu pour le concepteur qui doit rechercher la forme optimale des manettes ainsi que la répartition optimale des cibles qui seront relativement nombreuses (pour le visiocasque Vive, il y a 37 cibles). Il est aussi très utile que les manettes soient sans fil, ce qui est prévu pour les manettes récentes.

Les manettes de l'entreprise Oculus, les « Oculus Touch » utilisent deux caméras avec le principe *outside-in*. Les deux manettes sans fil sont en forme de croissants de lune sur lesquels de nombreuses cibles sont positionnées pour qu'il y en ait toujours de visibles par les caméras. Chaque manette intègre un petit joystick manipulable par le pouce et des boutons classiques. Les capteurs permettront de reconnaître certains gestes. L'objectif de ces deux manettes est de permettre aisément et intuitivement de sélectionner des objets et de les manipuler. Les manettes sans fil des entreprises HTC et Valve utilisent aussi des caméras avec le même principe *outside-in*. Des boutons sont disposés sur les manettes ainsi qu'une gâchette et un pavé tactile, qui permet de choisir des options. L'autonomie des manettes est de 4 heures (Figure 5.2).



**Figure 5.2 :** les manettes sans fil des entreprises HTC et Valve, copyright Middle VR<sup>1</sup>

Avec de telles manettes, on se passe de capter les mouvements des doigts, qui ne sont pas indispensables pour des manipulations en EV : on peut profiter d'être en EV pour simplifier la prise d'un objet... ce qui n'est guère possible sans les doigts en ER. Mais si l'on souhaite simuler néanmoins les doigts, que ce soit pour manipuler un objet ou faire toutes autres actions (jouer d'un instrument de musique, etc.), il est possible d'exploiter un capteur de mouvements des doigts. Le plus utilisé à ce jour est l'interface «Leap Motion». Ce dispositif est un capteur qui fournit les positions des extrémités des doigts. Il est capable de les déterminer au millimètre, avec un très faible temps de latence et les mesures sont relevées à la fréquence de 200 Hz. Il se caractérise par un simple petit parallélépipède de 8 x 3 x 1,1 cm et équipé de deux capteurs surmonté de lentilles capables d'offrir un volume de travail d'environ 150° en latéral et 120° en transversal. On le pose à plat devant un écran ou on le colle sur la face avant d'un visiocasque. Les deux caméras filment en stéréoscopie la zone proche dans un volume de 0,2 m<sup>3</sup> et trois LED infra-rouge complètent le système.

### 3. INTERFACES ACOUSTIQUES

Le son est important pour l'immersion de la personne, mais il n'intervient pas de la même façon que la vision : la perception visuelle de l'espace entourant la personne est limitée à chaque instant à son champ de vision, tandis que ce n'est pas le cas pour le son, perçu sans limitation spatiale. Les sources sonores sont localisables dans l'espace, comme expliqué au chapitre 2, mais sans très grande précision. Si le concepteur souhaite fournir une immersion sonore performante, il est souhaitable qu'il crée des sources sonores spatialisées. Au niveau logiciel, il existe plusieurs programmes de création de son 3D qui permettent maintenant de créer, en temps réel, une simulation sonore spatiale de qualité sur des ordinateurs classiques, même s'il y a encore des limites en nombre de sources sonores et en qualité de propagation des sons. La restitution sonore peut se faire de deux manières pour l'utilisateur :

- Couplées au visiocasque, par l'intermédiaire d'interfaces acoustiques le plus fréquemment présentes sous la forme de micro haut-parleurs (écouteurs

<sup>1</sup> La société Middle VR travaille dans la réalité virtuelle, <http://www.middlevr.com>

intra-auriculaires, extra-auriculaires) et plus rarement par conduction osseuse (ostéophonie);

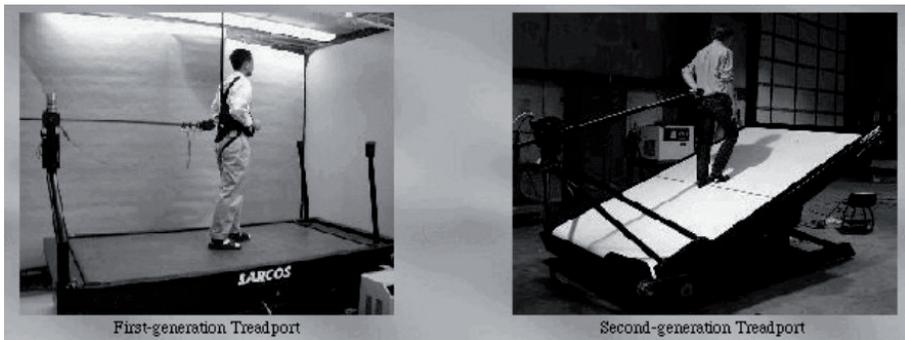
- Il y a aussi les enceintes qui peuvent être réparties en nombre à la périphérie de la pièce.

Au niveau du logiciel, les captures virtuelles des sources sonores de l'EV doivent suivre la position de l'utilisateur pour ensuite les lui restituer. La restitution sonore sera évidemment différente si on utilise des écouteurs ou des enceintes, car ces dernières sont fixes dans l'ER alors que les écouteurs sont en mouvement dans l'ER, mais leurs positions sont connues grâce au tracking de la tête. Le logiciel sonore doit prendre en compte, si possible, les deux modes. Il peut aussi prévoir des réglages individuels car les études ont montré que chaque personne a sa propre « fonction de transfert auditif de la tête ». Des explications sur les interfaces acoustiques et les logiciels pour le son spatialisé peuvent être consultées dans les volumes 2 et 3 du « Traité de la Réalité Virtuelle ».

#### **4. TAPIS ROULANTS 1D ou 2D**

Le principe d'une interface de marche est d'utiliser un tapis roulant qui fait défiler le sol réel pour compenser le mouvement de translation engendré par la marche. C'est une solution efficace pour obtenir l'interface de déplacement la plus naturelle puisqu'elle exploite le schéma de la marche. Cependant, le problème majeur rencontré avec cette solution est l'unique direction imposée par le tapis (tapis 1D), qui doit être complété par une autre interface (volant, manette, boutons...) pour permettre à la personne de changer progressivement de direction de déplacement. Mais dans ce cas, les trajectoires ne sont pas les mêmes qu'en marche naturelle puisque la personne ne peut pas se retourner instantanément. Les trajectoires ont un certain rayon de courbure, comme celles d'une voiture. Face à un tapis asservi, les utilisateurs ne sont pas toujours à l'aise surtout s'ils portent un visiocasque. La visualisation, en surimpression, dans le visiocasque de la zone mobile du tapis peut être une aide pour l'utilisateur. Pour marcher le plus aisément sur un tapis, il suffit de ne pas se préoccuper de son fonctionnement. Mais l'appréhension de la chute en bout de tapis est assez forte pour certains utilisateurs. Dans ce cas, ils essaient d'asservir la vitesse du tapis et s'opposent au contrôleur informatique qui a le même objectif. Ceci crée des instabilités dans le mouvement du tapis. Pour éviter cette appréhension, il faut utiliser un tapis roulant plus long que le strict nécessaire pour le fonctionnement de la marche sur place : une partie à l'avant du tapis ne servira jamais, mais sera rassurante, comme nous avons pu le vérifier.

Le tapis le plus simple permet la marche dans une seule direction sans effort. La présence éventuelle d'un élastique rattaché au marcheur crée une force de rappel pour simuler une impression d'effort. Mais celui-ci n'est pas réglable. La notion de distance dans le monde virtuel s'en trouve accrue. Un tapis peut aussi être incliné pour simuler des efforts dans une montée et réaliser ainsi une simulation plus réaliste (Figure 5.3).



**Figure 5.3 :** interface de locomotion par tapis 1D de Sarcos, copyright photo Sarcos

Pour les tapis 2D, (ou tapis omnidirectionnel), il est bien évident que le problème de fond est de permettre la marche dans deux directions : c'est un problème mécanique complexe si on veut vraiment permettre à la personne de marcher. Certains modèles se conteraient de faire glisser les pieds pour simplifier grandement la réalisation technique. Le glissement des pieds n'est pas naturel mais marcher en restant physiquement sur place (les systèmes vestibulaires sont là pour le confirmer) n'est pas naturel non plus... Comment faire marcher une personne dans toutes les directions tout en la laissant sur place ? Certains ont pensé à exploiter une grande sphère libre en rotation par rapport à son centre, le marcheur étant soit en bas à l'intérieur ou à l'extérieur en haut sur la sphère. Pour que cela fonctionne d'une façon satisfaisante, le rayon de la sphère doit être assez grand. Le marcheur a alors une impression de se déplacer sur une surface plane et non sur une calotte sphérique. Cela implique une interface de très grande dimension et donc d'un coût élevé.

Des solutions de tapis 2D, où l'on peut marcher, ont été réalisées avec des roulements perpendiculaires combinés entre eux mécaniquement et positionnés sur un plan support. L'ensemble des roulements est commandé astucieusement par deux tapis intermédiaires orthogonaux liés à deux moteurs pour faire tourner les billes des roulements dans toutes les directions du plan support. Un autre tapis 2D a utilisé le principe d'un tore composé de 12 tapis simples reliés côte à côte qui peuvent défiler suivant une direction perpendiculaire à leur axe de déroulement dans le tore. Les mouvements des pieds sont mesurés par des capteurs électromagnétiques placés sur les genoux du marcheur. Le sol bouge dans le sens opposé de la marche par asservissement afin que l'utilisateur conserve en permanence sa position au centre du système mécanique (Figure 5.4).

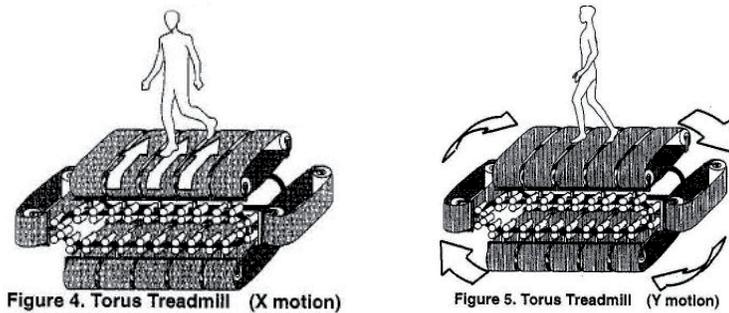


Figure 5.4 : schéma des mouvements sur le tapis Professor Iwata

Il n'est pas toujours très aisé de marcher sur de tels tapis, surtout si la personne ne se tient pas à une barre. Une solution différente a été développée récemment par l'entreprise Virtuix : le tapis Omni dont le principe repose sur le glissement des pieds sur le sol, le «marcheur» se tenant à une barre circulaire qui l'encercle au niveau du bassin (Figure 5.5). Le sol incurvé, et recouvert d'une surface glissante, permet de détecter les pas. Il faut toutefois utiliser une paire de chaussures spéciales, équipées de semelles à faible coefficient de frottement, afin d'assurer au mieux le glissement. La marche, la course et les sauts peuvent être réalisés et détectés avec l'Omni par deux capteurs aux pieds très précis et avec peu de latence. Les chaussures doivent être choisies avec leur taille et plusieurs tailles sont disponibles. Ce dispositif de marche est plus destiné aux salles d'arcade et parc d'attraction qu'à un usage personnel, vu son prix. Un autre tapis de marche, Cyberith Virtualizer<sup>2</sup>, est sur le même principe que le tapis Omni mais avec de petites variantes sur le harnais qui retient le corps au niveau du bassin, permettant à l'utilisateur de glisser facilement sur un sol lisse sous ses pieds.



Figure 5.5 : schéma de principe de tapis Omni : le marcheur se tient debout au milieu, entouré par l'arceau, tout en glissant les pieds sur le socle légèrement concave

2 Voir : <http://cyberith.com/product/>

## 5. INTERFACE DE SIMULATION DE MOUVEMENT

### 5.1. Introduction

Les interfaces à simulation de mouvement ont pour objectif de modifier l'orientation du corps dans l'espace ou de lui faire subir des accélérations. Mais nous devons remarquer que les interfaces de simulation de mouvement imposent en plus des efforts, incontrôlables directement, sur le corps de l'utilisateur, stimulant sa proprioception musculaire, ainsi que des sensations tactiles.

Techniquement, le déplacement du corps d'une personne, ou de plusieurs dans les cabines de simulation, est théoriquement sans problème technique insurmontable. Il suffit de se donner les moyens, mais en pratique cela impose des mécanismes très puissants et par conséquent des prix élevés ou très élevés si l'on veut faire des mouvements de grandes ampleurs. C'est donc bien un problème mécanique à résoudre pour tout concepteur de ce type d'interface. Le poids des personnes étant ce qu'il est, nous ne pouvons jamais espérer avoir des interfaces de simulation de mouvement à prix très modiques. Il faudra toujours fournir une énergie relativement importante, imposant des dispositifs techniques conséquents. À cause d'investissements nécessaires importants, ces interfaces ne sont donc fabriquées et commercialisées que pour des applications professionnelles ou pour des salles de jeux et de spectacles. Si elles pouvaient être peu onéreuses, il est bien évident que le marché des jeux vidéo les aurait exploitées pour des jeux individuels. Le coût moyen d'un siège mobile avec 3 ddl est d'environ dix milles euros.

Nous pouvons classer les interfaces à simulation de mouvement en deux grands domaines, qui ne sont pas strictement disjoints :

- les sièges et autres dispositifs mono utilisateur de simulation, qui changent l'orientation du corps de la personne et(ou) qui parfois appliquent des accélérations alternatives et limitées en intensité ;
- les cabines de simulation, mono ou multi-utilisateurs, modifient l'orientation du corps des personnes et leur imposent de fortes accélérations. Dans ces cabines, le dispositif d'affichage du monde virtuel est en général inclus dans la cabine fermée mais il est aussi possible d'utiliser des visiocasques.

La vision du monde virtuel est donc en général liée au mouvement de la cabine, ce qui n'est pas le cas avec les sièges de simulation, sauf quelques cas particuliers. Dans les interfaces mono utilisateur, la personne n'est pas forcément assise, elle peut être debout ou couchée, si on souhaite la faire voler virtuellement par exemple.

Les accélérations ne peuvent pas être aussi importantes que désirées ou imposées dans la direction voulue. Aussi sont-elles parfois simulées différemment et ne correspondent pas alors exactement au mouvement réel. Dans l'exemple classique de simulation de véhicule, on baisse la cabine vers l'avant, quand il y a un freinage du véhicule virtuel, pour suggérer au cerveau, grâce aux systèmes vestibulaires, une décélération linéaire, plutôt que de réaliser cette décélération. Car cette dernière imposerait un mécanisme

plus cher et surtout plus encombrant. L'utilisateur est abusé sensoriellement car il n'a pas de repère visuel du monde extérieur et parce qu'il voit le véhicule virtuel en train de décélérer dans le monde virtuel. Il n'a donc pas la sensation d'être incliné. Il y a donc dans ce cas précis une limite de décélération ou d'accélération d'intensités maximales inférieure à l'attraction terrestre. Dans tous les cas, il y a des limites physiques et techniques aux simulations de mouvements à accélération importante. Il est impossible actuellement de simuler les accélérations d'une voiture tournant rapidement autour d'un rond-point. Le cas extrême est le simulateur de départ en fusée. Il exigerait d'être conçu à partir d'une vraie fusée pour vaincre l'accélération terrestre d'une façon identique au cas réel ! Mais cela n'empêchent pas les astronautes de s'entraîner à subir les accélérations les plus fortes imposées à des hommes. C'est le cas limite d'interface de simulation de mouvement : la centrifugeuse pour l'entraînement des astronautes...

Un domaine d'exploitation des différents types de simulateur est celui des jeux et des loisirs. Dans ce cas, les dispositifs sont installés dans des salles de jeux ou des parcs d'attraction, permettant un amortissement plus rapide des investissements engagés. Des cabines de simulation sont équipées d'un ensemble de sièges (seize, par exemple). Mais dans ce cas, l'interactivité est faible car les participants sont presque toujours passifs par rapport à la simulation proposée. La visualisation peut être collective ou individuelle par visiocasque.

Un autre domaine est celui des transports, qui soient terrestres, maritimes ou aériens (transport ferroviaire, voitures, engins militaires, avion, bateau, etc.). Certains simulateurs sont plus simples à réaliser que d'autres, s'ils ont moins de degrés de liberté à commander ou s'ils ont des accélérations maximales plus faibles. Les simulateurs de train sont construits avec une simple interface visuelle (un seul écran) et disposent d'actionneurs pour les mouvements. Dans le cas de leur simulation, la cabine est montée sur un système assurant trois déplacements pour subir des accélérations linéaires, latérales et pour imposer de légères inclinaisons. Depuis de nombreuses années, les simulateurs de vol sont conçus pour donner aux pilotes les sensations de vol. C'est un domaine important au niveau économique, dont les solutions proposées sont très satisfaisantes (les pilotes d'avion se forment partiellement et homologuent certaines de leurs compétences sur simulateurs d'avion). Pour les véhicules terrestres, des plates-formes mobiles classiques à 6 axes sont utilisées et aussi à 8 axes (voir plus loin), mais les simulations sont plus éloignées des phénomènes réels que pour les avions de l'aviation civile.

## 5.2. Sièges à simulation de mouvements

Ces interfaces mono utilisateur ont la possibilité de modifier l'orientation du corps de la personne. Celle-ci est soit assise sur un siège mobile, supporté par une base fixe, soit debout sur une plaque mobile, entourée d'une protection circulaire. La personne peut être aussi couchée, si l'application l'exige. Ces dispositifs sont souvent équipés de poignées de commande pour permettre la navigation dans une scène virtuelle. Si des accélérations en translation doivent être subies par la personne, il s'agit en pratique de faibles accélérations dues aux vibrations appliquées au siège. Cette technique est

employée dans des salles de cinéma interactif, où les vibrations et les inclinaisons des sièges sont synchronisées avec les mouvements virtuels projetés à l'écran pour une meilleure sensation d'immersion. Le siège de simulation est aussi employé dans les salles de jeux pour réaliser un simulateur de voiture dans lequel le joueur roule sur un circuit virtuel. Pour une meilleure immersion, les inclinaisons du siège sont synchronisées avec le mouvement du volant. L'affichage peut être un écran ou un visiocasque, quand ils seront disponibles, pour une plus grande immersion visuelle.

Les sièges de simulation peuvent être classés selon le nombre de degrés de liberté qu'ils imposent à l'utilisateur. Nous pouvons avoir des sièges :

- à un degré de liberté, par exemple, des vibrations verticales pour simuler des secousses suivant l'état d'une chaussée et des suspensions du véhicule ;
- à deux degrés de liberté, par exemple, un siège de moto ;
- à trois degrés de liberté, par exemple, un delta plane (figure 5.6) ;
- à six degrés de liberté, par exemple, une planche de surf ou siège sur six vérins (figure 5.7).

Les mouvements des sièges sont commandés par des moteurs électriques, par des vérins pneumatiques ou par des vérins hydrauliques mais plus contraignants techniquement et plus onéreux. Ces derniers actionneurs sont les plus puissants. Ces interfaces ont en général au moins deux inclinaisons, tangage et roulis, 25° à 45° suivant les modèles.



**Figure 5.6 :** interface à simulation de mouvement en vol



**Figure 5.7 :** siège à six vérins

La société SOMNIACS a proposé au cours de l'année 2015 un simulateur de vol d'oiseau nommé Birdly. Le dispositif permet à l'utilisateur de s'allonger sur une plateforme mobile en rotation et en translation et fournit deux supports indépendants destinés à recevoir ses bras. Équipé d'un visiocasque et d'écouteurs, l'utilisateur peut à l'aide de mouvements de ces bras (de bas en haut), simuler les mouvements des ailes d'un oiseau, qu'il incarne dans la simulation 3D, afin d'élever son altitude par exemple. Il peut aussi modifier son altitude en inclinant ces mains par des rotations des poignets, ce qui aura le même effet que les ailerons sur les ailes d'un avion. D'un poids de 132 kg, équipé d'un ventilateur orienté sur l'utilisateur pour simuler le vent relatif, son encombrement est relativement limité (2100 mm x 1400 mm x 1000 mm).

Une nouvelle interface originale est l'utilisation d'un bras robotisé pour y mettre un siège à son extrémité, réalisée par l'entreprise MM-Company et soutenue par l'entreprise Ubisoft : le siège MMOne<sup>3</sup>. Ce siège pivote à 45 degrés sur sa droite et sa gauche et permet une rotation à 360 degrés (Figure 5.8). Le site précise que le MMOne convient à des enfants au-delà de plus de 12 ans. Le siège est doté de 5 ceintures et peut être modifiée en fonction du gabarit du joueur.



**Figure 5.8 :** le siège sur un bras robotisé MMOne

### 5.3. Cabines à simulation de mouvements

Les simulateurs avec mouvement de cabine, nommés souvent simplement « simulateurs », sont employés depuis plusieurs dizaines d'années dans le domaine du transport aérien et terrestre. Ces systèmes permettent des immersions multi-sensorielles de très hautes qualités, indispensables pour la formation et l'entraînement des pilotes et des conducteurs de train. Les actionneurs sont les pièces maîtresses du système. Ils sont pneumatiques, électriques ou hydrauliques pour les charges de plusieurs tonnes. Les actionneurs sont souvent disposés deux par deux sur chaque côté d'un triangle, proposant les six degrés de liberté sur des débattements de 20 à 30 degrés. Ils peuvent atteindre des vitesses d'un mètre par seconde. Les accélérations angulaires de la cabine sont supérieures à  $100^\circ/s^2$ . Ils doivent pouvoir déplacer des masses de plusieurs tonnes et vaincre l'inertie de la cabine. Leurs caractéristiques dynamiques

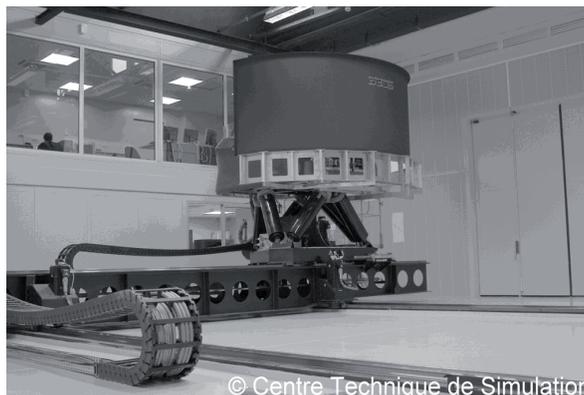
<sup>3</sup> <http://www.realite-virtuelle.com/ubisoft-investit-chaise-realite-virtuelle-mmone>

sont les plus importantes. Les investissements et les coûts de fonctionnement imposent un taux d'utilisation important (Figure 5.9).



**Figure 5.9 :** cabine de simulateur de vol à 6 ddl

Il existe dans le monde quelques systèmes qui sont des simulateurs à 6ddl ayant de grandes courses en translation. Ces dispositifs sont réalisés à partir d'une plate-forme mobile classique montée sur un mécanisme de translation à 2 ou 3 ddl et d'étendue de plusieurs mètres, voire décamètres. Evidemment, la structure mécanique est imposante pour fournir la puissance nécessaire permettant de déplacer une plate-forme mobile standard avec le système d'affichage et l'utilisateur. Le constructeur de véhicules Renault est équipé d'un tel système, l'Ultimate, qui est un simulateur de conduite à hautes performances dynamiques (Figure 5.10). Ce projet a été réalisé par le Centre Technique de Simulation de Renault. ULTIMATE est destiné plus particulièrement aux applications d'ingénierie dans le domaine de la dynamique véhicule, des systèmes d'aide à la conduite, et de l'étude du comportement conducteur en situation dynamique.



**Figure 5.10 :** le simulateur ULTIMATE de Renault

## 6. INTERFACES DE SUIVI DES YEUX

Plusieurs systèmes d'oculomètre ont été développés pour l'ophtalmologie et pour l'aide aux déficients visuels. Les appareils sont basés sur différents principes, dont à base de mesure optique de réflexion sur l'œil. Ces derniers systèmes, les seuls non invasifs, exploitent le traitement d'images actif (un faisceau infrarouge participe à la localisation de l'œil). Ces systèmes sont utilisables à la condition restrictive que la tête ne bouge pas beaucoup face à un écran ou que le dispositif soit porté par la tête, ce qui est le cas avec les quelques visiocasques équipés actuellement.

La technique du reflet cornéen a pour principe d'envoyer au centre de la pupille une lumière infrarouge émise par une diode ou un ensemble de diodes. Le reflet infrarouge renvoyé par la cornée de l'œil est ensuite détecté et ce sont les variations d'intensité de ce reflet qui permettent, après calcul, de repérer le centre de la pupille et de connaître la position de l'œil. Généralement, le système consiste en une caméra, muni ou non d'un filtre infrarouge, et d'un ensemble de LED placé sur le pourtour de l'objectif de la caméra. La position des LED est importante. En effet, le faisceau de lumière émis et celui réfléchi ont la même direction. En conséquence, la lumière émise doit être très proche de l'axe optique de la caméra. Une variante consiste à illuminer la pupille en alternance avec une source lumineuse proche du centre optique, et une source lumineuse éloignée du centre optique. Une image est prise à chaque fois. Dans la première image, la pupille sera en surbrillance. Dans la seconde, elle apparaîtra presque noire. La différence des deux images mettra en valeur la pupille. C'est en fait un mode différentiel que les électroniciens connaissent bien pour être beaucoup plus robuste au bruit. Seul désavantage de ce mode différentiel, l'œil peut avoir bougé entre les deux prises d'image.

Des sociétés spécialisées peuvent adapter leurs produits pour être compatible avec certains visiocasques. La société SMI, par exemple, fournit un service de modification d'un visiocasque pour lui incorporer un capteur de suivi des yeux. Les données de suivi des yeux sont acquises à une fréquence de 60 Hz, pour un champ de vision horizontal de 80° et un champ de vision vertical de 60° et la précision est comprise entre 0,5° et 1° (voir les détails dans le chapitre suivant).



---

# Caractéristiques techniques des visiocasques

---

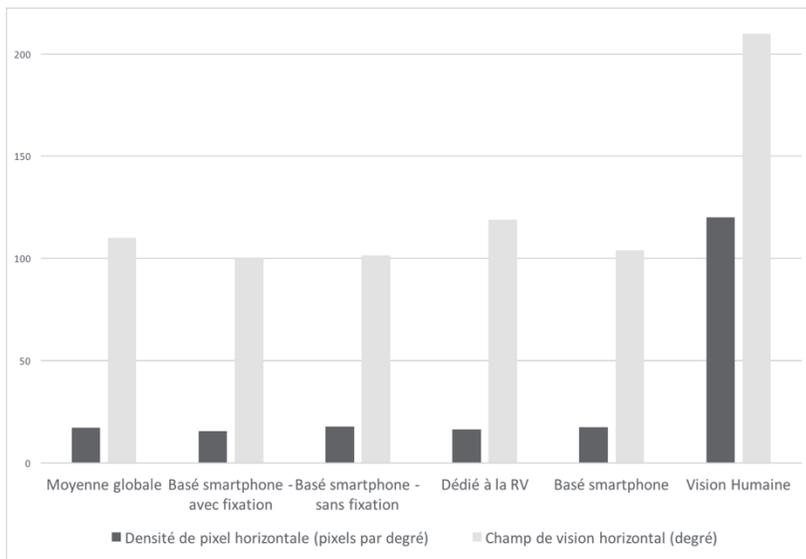
Olivier Hugues

## 1. INTRODUCTION

Dans le chapitre précédent, nous avons présenté de nombreux aspects techniques quant à l'usage des visiocasques. Dans ce chapitre, nous allons faire un tour d'horizon des solutions commerciales existantes ou encore en phase de conception lors de l'écriture de cet ouvrage. Nous allons détailler les caractéristiques principales de ces dispositifs, logicielles et matérielles lorsqu'elles sont disponibles, pour en appréhender leurs avantages et leurs limites. Ainsi, le lecteur pourra choisir le futur dispositif qu'il souhaite utiliser pour répondre favorablement aux besoins de l'application qu'il vise à mettre en œuvre. Les visiocasques peuvent avoir différents usages et il semble donc intéressant de les regrouper selon leur fonctionnalité principale. Pour cela, nous nous sommes basés sur les primitives comportementales virtuelles (PCV) détaillées dans le chapitre 1 en proposant d'abord les dispositifs permettant uniquement d'observer l'environnement virtuel tandis que d'autres permettent de se déplacer et d'agir dans cet environnement virtuel. Ce sont souvent des considérations techniques qui vont limiter ou permettre certains usages avec ces appareils.

## 2. CARACTÉRISTIQUES GÉNÉRALES

Dans les sections suivantes, nous allons présenter quelques fiches techniques regroupant les caractéristiques de nombreux visiocasques disponibles dans le commerce afin d'offrir un moyen de comparer les dispositifs entre eux. Mais il semble d'abord intéressant de comparer leurs caractéristiques principales par rapport aux capacités psychophysiques du système visuel humain.



**Figure 6.1 :** valeurs moyennes de densité de pixel et de champ de vision pour la totalité des 32 modèles analysés comparées à celles des capacités psychophysiques du système visuel humain

Sur 32 modèles répertoriés (voir tableau en fin de chapitre), on peut voir dans la Figure 6.1 que les valeurs moyennes de la densité de pixels et du champ de vision des dispositifs commerciaux sont encore très éloignées des valeurs nécessaires pour atteindre les capacités psychophysiques du système visuel humain. En effet, comme nous l'avons rappelé dans le chapitre 2, le champ de vision horizontal moyen d'une personne n'ayant aucune déficience visuelle est d'environ  $210^\circ$  et  $140^\circ$  verticalement. Ce constat indique que la plupart des visiocasques actuellement commercialisés sont de faible qualité par rapport à ce qu'il faudrait obtenir comparativement à la vision humaine. Il faut espérer que de nouvelles évolutions technologiques permettront d'atteindre des valeurs plus importantes. Cela ne sera pas trop compliqué d'augmenter la densité en pixel mais, en revanche, l'augmentation des champs de vision est un problème optique difficile techniquement sur lequel les constructeurs butent depuis plus de 20 ans.

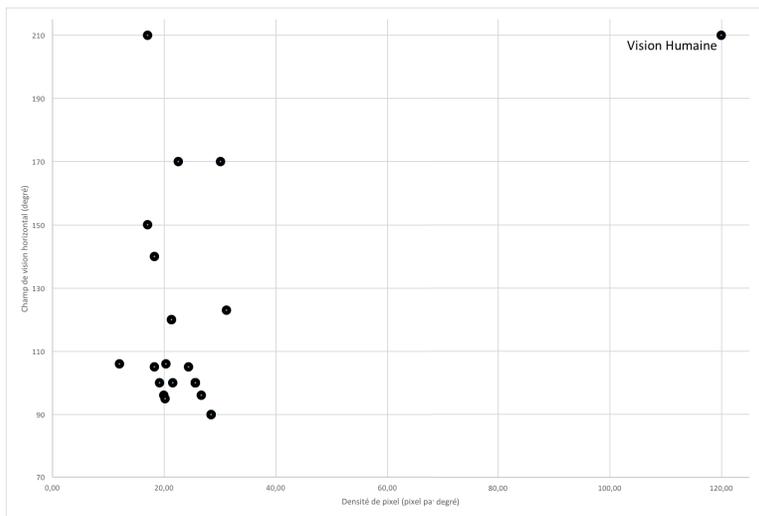
La densité de pixel est un critère très important. Exprimée en nombre de pixel par degré (ppd), elle traduit la résolution du dispositif, indépendamment du champ de vision qu'il offre. Cette densité est à comparer avec l'acuité visuelle de l'utilisateur, dont la moyenne, quant à elle, est d'environ une minute d'angle. Nous pouvons cibler une acuité double de la moyenne, pour atteindre 0,5 minute d'angle afin d'obtenir une qualité perçue acceptable même pour les utilisateurs ayant une acuité au-delà de la moyenne. La densité de pixel horizontale des visiocasques se calcule par la formule ci-dessous avec :

- $Def_H$  : définition horizontale pour un œil, en pixel ;
- $FOV_H$  : champ de vision horizontal total, en degré ;
- $R_H$  : champ de vision de recouvrement, en degré ;

$$D_H = \frac{2 Def_H}{FOV_H + R_H}$$

Il est souvent très difficile d'obtenir la valeur de recouvrement auprès des constructeurs. Lorsque nous ne la connaissons pas, nous avons utilisé la valeur favorable de 100% du champ de vision. Bien que cela soit la norme pour de nombreux modèles, certains fabricants augmentent artificiellement le champ de vision de leur dispositif en concevant un système optique disposant d'un champ de recouvrement limité. Ainsi, si le champ de vision horizontal est plus important, la zone dans laquelle l'utilisateur est capable de percevoir en stéréoscopie devient très limitée, voire nulle. La Figure 6.2 permet de comparer le champ de vision horizontal et la densité de pixels des dispositifs qui ont été sélectionnés dans notre liste par rapport aux caractéristiques optimales du visiocasque exploitant parfaitement l'acuité visuelle et les champs de vision (par simplification, on ne compare que le champ de vision horizontal, le plus essentiel, mais on pourrait aussi le faire aussi avec le champ de vision vertical).

On remarque que la densité de pixels offerte par la plupart des dispositifs inventoriés ne dépasse que légèrement les 20 ppd alors que le système visuel nécessiterait environ 120 ppd. On remarque aussi que la plus grande proportion des dispositifs offre un champ de vision d'environ 100° à 120° (surtout les dispositifs basés sur des téléphones portables ou tablettes).



**Figure 6.2 :** répartition des visiocasques selon leur champ de vision horizontal et leur densité de pixel

### 3. CARACTÉRISTIQUES TECHNIQUES DES VISIICASQUES

#### 3.1. Les visiocasques à partir de smartphone

##### 3.1.1. Généralités

Il est assez difficile d'appréhender globalement ce type de système puisque leurs performances sont extrêmement dépendantes des smartphones utilisés et des applications. En effet, les critères matériels fondamentaux tels que la définition, le champ de vision, la latence, le poids ou encore le contraste dépendent du modèle de smartphone utilisé. Au niveau logiciel, l'application qui tire parti de l'électronique embarquée a aussi un impact sur la performance du dispositif.

De façon générale, ces visiocasques offrent un champ de vision horizontal moyen d'un peu plus de 100°. Si la forme globale des smartphones semble relativement identique quel que soit la marque, ces visiocasques ne sont pas pour autant compatibles avec l'ensemble des smartphones du marché. Les constructeurs font cependant l'effort de concevoir des boîtiers capables d'accueillir les modèles phares des grands constructeurs, mais il est parfois possible d'avoir des modèles incompatibles ou partiellement compatibles avec certains smartphones. Le suivi de la tête de l'utilisateur s'effectue à l'aide des capteurs embarqués du smartphone. Ceux-ci peuvent dériver au cours du temps et provoquer une gêne ou un inconfort (si par exemple l'horizon n'est plus tout-à-fait à l'horizontal par rapport à l'environnement réel). Ils doivent être réinitialisés souvent à l'aide d'une application dédiée.

Dans cette catégorie, on trouve deux types de visiocasque. En effet, il est possible de dissocier les dispositifs qui ne peuvent pas être fixés sur la tête et qui doivent donc être maintenus avec les mains, de ceux qui peuvent être fixés sur la tête. La différence est importante car les dispositifs appartenant à la première catégorie occupent les mains de l'utilisateur, ce qui limite l'utilisation à une activité d'observation uniquement.

Pour tous les modèles répertoriés dans notre liste qui utilisent des modèles de smartphones, nous considérons l'utilisation du smartphone compatible ayant la meilleure définition. Le plus souvent, ces mobiles sont équipés d'un écran d'une définition de 2560x1440 pixels.

Remarque : dans la suite de ce chapitre, les informations sont données à titre indicatif, sans garantie d'exactitude, car elles sont susceptibles d'être modifiées en fonction de l'évolution des versions des produits des constructeurs.

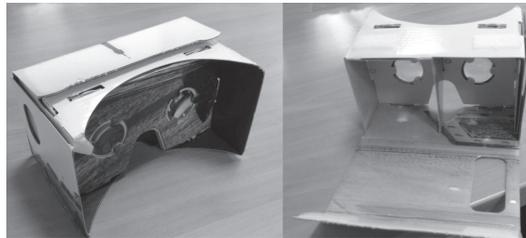
### 3.1.2. Les visiocasques sans fixation sur la tête

Il s'agit des systèmes parmi les plus basiques qu'il est possible de trouver aujourd'hui. Uniquement composé d'une monture, très simple et légère (inférieur à 100g), elle n'est utile que pour maintenir en place un système optique rudimentaire composé de deux lentilles (une pour chaque œil). L'utilisateur doit ainsi maintenir avec ses mains le dispositif sur son visage. Ce genre de matériel ne dispose d'aucun réglage pour s'adapter aux différents utilisateurs et laissent assez facilement passer la lumière extérieure à cause des montures ajourées et d'une mise en place approximative sur le visage.

De nombreux modèles de ce type existent, mais nous avons fait le choix d'en présenter deux. Le plus connu est certainement le modèle de la société *Google*, nommé *Cardboard*. Début 2016, la société annonce avoir vendu 5 millions d'unités en 18 mois. L'application associée a été téléchargée environ 25 millions de fois avec près de 10 millions de téléchargement entre octobre et décembre 2015.

#### *Google Cardboard*

Le produit proposé par la société Google fut présenté initialement au grand public en juin 2014 lors d'une conférence organisée par la société et dédiée aux développeurs. Il est diffusé librement, sous la forme d'un patron en carton, à découper et à plier soi-même, capable ensuite de recevoir un smartphone mobile (différents modèles sont compatibles). D'une conception sommaire, sans réglage pour l'utilisateur, il s'agit d'un modèle excessivement moins coûteux que la majorité des visiocasques. Il est en effet possible de l'obtenir pour quelques dollars sur le site dédié<sup>1</sup> et parfois gratuitement grâce à des opérations commerciales. Il est principalement dédié aux développeurs qui peuvent



l'utiliser pour tester leurs applications. La société fournit une application pour tester le visiocasque disposant de quelques scènes à 360°, ainsi qu'une API permettant aux développeurs de rendre leurs applications compatibles. Disposant uniquement d'un suivi des mouvements de rotation de la tête de l'utilisateur, utilisant les capteurs embarqués dans le smartphone, et immobilisant les mains de l'utilisateur, qui sont utilisées pour maintenir le visiocasque en place sur le visage, il permet uniquement d'observer l'environnement virtuel.

<sup>1</sup> <https://www.google.com/get/cardboard/>

### *Wearality Sky*

Le produit proposé par la société Werality est aussi basé sur l'utilisation d'un smartphone, à l'image du dispositif Google Cardboard. Sa particularité principale cependant est d'offrir un champ de vision horizontal relativement large, d'environ 150°, comparé au champ de vision d'environ 100° du modèle *Cardboard* et de la plupart des modèles de cette catégorie, grâce à l'utilisation de lentilles de Fresnel. Cependant, ce grand champ de vision diminue la densité de pixel puisque le nombre de pixels reste identique. Ainsi, sa densité de pixels se positionne dans la moyenne basse avec une valeur horizontale 17 pppd et verticale de 15 pppd. Ne disposant pas de moyen de fixation au niveau de la tête, et utilisant aussi uniquement les mouvements de rotation, il offre les mêmes fonctionnalités que le produit concurrent de la société Google.



*Image autorisée par Wearality Corporation. Tous droits réservés*

Du point de vue du logiciel, toutes les applications compatibles *Cardboard* peuvent être utilisées avec ce visiocasque puisqu'il est basé sur le même principe qui vise à utiliser les capteurs embarqués dans le smartphone afin de détecter les mouvements de rotation de la tête de l'utilisateur.

### 3.1.3. Les visiocasques avec fixation sur la tête

D'une conception généralement plus robuste que les modèles présentés précédemment, ces dispositifs, toujours basés sur l'utilisation d'un smartphone, offrent cependant à l'utilisateur la possibilité d'ajuster un ou plusieurs paramètres comme la distance intraoculaire ou la netteté. D'un poids relativement faible (inférieur à 350g), mais pourtant plus important que les dispositifs précédents, ils sont le plus souvent maintenus en place sur la tête de l'utilisateur à l'aide d'un bandeau élastique de quelques centimètres de largeur et réglable, offrant ainsi un confort relativement acceptable et un maintien suffisant pour prendre en charge les effets liés à la position assez excentrée du téléphone portable devant les yeux. Le fait de laisser les mains libres à l'utilisateur lui permet ainsi d'utiliser, par exemple, une manette de jeu vidéo sans fil connecté au smartphone et avec une application compatible, offrant ainsi à l'utilisateur la possibilité de se déplacer dans l'environnement virtuel et plus rarement d'interagir sur différents objets virtuels.

Les visiocasques avec fixation sur la tête sont conçus différemment des deux premiers modèles présentés et disposent d'une structure plus importante et fermée. De plus,

au-delà du confort qu'elles apportent, des mousses permettent le plus souvent au visiocasque de s'adapter à la morphologie de l'utilisateur pour réduire encore l'entrée de lumière dans le visiocasque. Ils sont donc moins sensibles aux reflets des rayons de lumière provenant de l'extérieur et permettent de mieux tirer parti de la luminosité et du contraste de l'écran du smartphone. Puisque ce type de dispositif est dédié à une utilisation plus prolongée que les précédents et qu'ils sont relativement bien fermés, il devient nécessaire de prévoir une circulation de l'air à l'intérieur du masque, car dans le cas contraire, de la buée peut se créer et gêner la vision de l'utilisateur, d'autant plus que la chaleur dégagée par le smartphone (dont le processeur et la carte graphique sont extrêmement sollicités) peut être ressentie sur le visage et devenir une source d'inconfort. La difficulté technique est de permettre une circulation de l'air convenable tout en bloquant la lumière qui provient de l'extérieur.

Ces visiocasques proposent en moyenne un champ de vision horizontal très légèrement supérieur à 100° pour une densité de pixel proche de 21 ppd. Nous présentons dans les fiches suivantes les principaux modèles de visiocasque qui rentrent dans cette catégorie mais il n'est évidemment pas possible de faire une liste exhaustive, le nombre de ces dispositifs ayant explosé ces derniers mois. De plus, un engouement certain pour la conception et le développement de ce type de dispositif a eu lieu au cours des années 2014 et 2015. De très nombreuses sociétés se sont positionnées sur ce marché et de nombreux projets ont vu le jour sur certaines plateformes de financement participatives. Il faut noter que nombre de ces projets n'ont pas été finalisés, et certains ont été abandonnés, soit par manque de moyen à cause d'un financement insuffisant, soit à cause de problèmes techniques qui n'ont pas été traités convenablement.

### *BlueSky Pro*

Le produit proposé par la société Immersion Vrelia est aussi basé sur l'utilisation d'un smartphone et grâce à son cadre relativement grand, il est capable d'accueillir deux écrans de 1080x1920. Il dispose d'un casque audio intégré et la composition du système de maintien sur la tête de l'utilisateur est hybride, à la fois souple et rigide. Le champ de vision qu'offre ce visiocasque le positionne au-dessus de la moyenne de la concurrence avec 123°, et malgré un nombre de pixels conséquent horizontalement (jusqu'à 3840), il dispose d'une densité de pixel relativement élevée par rapport à la moyenne de sa catégorie.

À l'image de nombreux casque de ce type, il propose différents réglages comme la possibilité d'adapter l'écartement entre les yeux et le focus, ce qui permet d'adapter le dispositif au système visuel de l'utilisateur.

Uniquement capable de mesurer les rotations de la tête de l'utilisateur, les applications, compatibles cardboard, de ce visiocasque sont capables de se connecter sans fil avec des manettes de jeux et permettent ainsi à l'utilisateur de se déplacer dans l'environnement virtuel.

### *Freefly VR*

Le produit proposé par la société Proteus VR Labs annonce un champ de vision horizontal relativement élevé, environ 120°, par rapport à la moyenne de ses concurrents qui se situe aux alentours de 100°. Il offre également des réglages horizontaux et verticaux pour positionner le smartphone convenablement en face des yeux de l'utilisateur, ce qui implique quand même une phase de réglage. Il dispose d'un système d'aération pour la buée et de deux lentilles d'environ 42mm de diamètres. Si l'on considère la meilleure définition du smartphone compatible avec ce modèle, on obtient une densité de pixel dans la moyenne, d'environ 21 ppd.

### *Gear VR*

Le produit proposé par la société Samsung est peut-être l'un des visiocasques les plus utilisés dans le monde même si le nombre de vente est tenu secret. Il ne dispose pourtant que d'un champ de vision relativement étroit, parmi les moins bons de sa catégorie, avec seulement 96° horizontalement. Mais grâce à ce champ de vision restreint et l'utilisation d'un modèle compatible disposant d'une définition maximum de 2560 pixels, il permet d'obtenir une densité de pixel d'environ 26 ou 28 ppd, suivant le modèle de smartphone utilisé, ce qui est au-dessus de la moyenne.



*Le Gear VR de Samsung, copyright Middle VR*

D'un poids de 379g, il est reconnu pour son maintien sur la tête de bonne qualité, mais ne prend en compte uniquement les mouvements de rotations de l'utilisateur, en utilisant les capteurs embarqués du smartphone via une application dédiée ou compatible cardboard.

## 3.2. Les visiocasques pour la réalité virtuelle

### 3.2.1. Généralités

Ces visiocasques sont différents des modèles précédemment présentés par la nature de leurs composants matériels. En effet, puisqu'ils disposent d'une électronique dédiée et embarquée plutôt que d'utiliser un smartphone, conçue spécifiquement pour répondre aux besoins d'une application de réalité virtuelle, ils sont capables d'offrir de meilleures performances. En ce sens, la centrale inertielle embarquée et(ou) des caméras extérieures permettent de détecter les mouvements de rotation

de la tête beaucoup plus rapidement qu'avec les smartphones et avec une dérive moindre, ce qui permet de limiter les problèmes de latence entre les mouvements de la tête et les effets qu'ils doivent produire sur le contenu à afficher sur l'écran (motion-to-photon ou latency en anglais). De plus, certains de ces modèles sont équipés d'un système capable de mesurer les déplacements en translation, dans les trois directions, de la tête de l'utilisateur. Cependant, alors que les mouvements de rotation sont pris en compte à 360° sur les trois axes, il s'avère que les mouvements de translation ne sont pris en compte que sur quelques centimètres pour certains modèles, à quelques mètres pour d'autres suivant les techniques utilisés (voir le chapitre 5). Les puissances de calculs nécessaires pour atteindre les performances annoncées par les constructeurs nécessitent une électronique embarquée parfois lourde, ce qui a un impact non négligeable sur le poids du dispositif, et encore gourmande en énergie, ce qui rend l'utilisation de câbles obligatoire pour l'instant, qui sont amenés à entraver les mouvements de l'utilisateur. Il est cependant possible d'espérer que les avancées technologiques permettent la suppression des câbles au profit de solution sans fils d'ici quelques années.

Si la plupart de ces dispositifs, de bien meilleure qualité que ceux qui ont été présentés précédemment, offrent en partie les mêmes fonctionnalités, il est cependant possible de les séparer en plusieurs classes selon leur champ de vision horizontal. En moyenne, parmi les 13 modèles répertoriés dans cette section, le champ de vision horizontal est d'environ 120° et vertical d'environ 104°. La densité de pixels moyenne est d'environ 23 ppp, ce qui reste très en dessous des capacités du système visuel humain.

Dans la suite de ce chapitre, nous proposons donc quelques fiches des modèles principaux de visiocasques dédiés au domaine de la réalité virtuelle selon qu'ils détiennent un champ de vision moyen, grand, ou qu'ils disposent d'un système intégré de captation de la direction du regard.

### 3.2.2. À champ de vision moyen

C'est dans cette catégorie que l'on trouve le plus grand nombre de dispositifs. Nous présentons trois modèles proposés par les sociétés les plus actives.

#### *Le Rift*

Le modèle Rift est la version finalement commercialisée début 2016 de la société Oculus, issue du développement des deux versions précédentes (DK1 et DK2). Disposant d'un champ de vision horizontal d'environ 106°, champ *a priori* totalement stéréoscopique comme le DK2, et d'une densité de pixel encore inconnue, il ne fait pas partie des systèmes optiques à grand champ de vision. En revanche, il dispose d'un taux de rafraîchissement élevé pour les données de suivi du mouvement (1000Hz) et d'un taux de rafraîchissement de l'écran de 90Hz. Se faisant, il est possible d'espérer une latence globale relativement faible, mais aucune valeur n'est communiquée par le constructeur malgré les efforts qui ont été fait par la société sur cet aspect.



*Le DK2 d'Oculus*

Ce modèle dispose d'un capteur en translation sur quelques dizaines de centimètres, qui doit être installé face à l'utilisateur et nécessite encore l'usage d'un câble pour l'alimentation du visiocasque et le transfert de certaines données. Il dispose néanmoins d'un système audio spatialisé intégré qui a fait l'objet de plusieurs brevets, d'un réglage de la luminosité et du contraste de l'écran, ainsi

qu'un réglage pour éloigner ou rapprocher l'écran des yeux. Il est fourni en option avec deux contrôleurs permettant à l'utilisateur de manipuler des objets dans la scène, ou encore de se déplacer dans l'environnement. La société propose aussi un kit de développement logiciel (SDK) permettant aux développeurs de rendre compatible leurs applications avec ce dispositif. Ainsi, les éditeurs de jeux ou d'applications industrielles peuvent utiliser cette suite de fonctions afin de créer leur propre produit utilisant ce dispositif.

### *HTC Vive*

Le visiocasque Vive «Pre», qui doit être considéré comme la deuxième version de développement de la société HTC (en attendant le modèle final prévu courant 2016) a été présenté au tout début de l'année 2016. Il dispose d'un écran de même définition que son concurrent Rift, à savoir 2160x1200, et d'un champ de vision affiché identique. La différence se fait surtout au niveau de la technologie utilisée pour suivre les mouvements de rotation et de translation de la tête de l'utilisateur. Dans ce modèle, deux boîtiers, à la fois émetteurs laser et caméra, doivent être positionnés en hauteur des murs de la pièce dans laquelle sera utilisé le produit. Ainsi, les rayons lasers émis viennent se réfléchir sur plusieurs petites diodes (un peu moins de 40) positionnées sur le visiocasque. Ce réfléchissement



*Le HTC Vive, copyright Middle VR*

est capté par les caméras et analysé afin de déterminer l'orientation et la position du casque dans un repère de référence associé à la pièce initialement défini lors du montage des deux boîtiers. Ce principe permet de capter les translations sur trois dimensions et dans un espace d'environ 3 m x 4 m. Ce modèle est muni d'une caméra qui observe

l'environnement de l'utilisateur pour l'intégrer partiellement à l'environnement virtuel, pour éviter les chutes ou les blessures sur des objets physiques. Disposant lui aussi de deux contrôleurs, il offre les mêmes fonctionnalités que le système d'Oculus.

### *PlayStation VR*

S'il partage la particularité de disposer d'un taux de rafraîchissement de l'écran élevé de 120Hz avec le modèle Glyph de la société Avegant, le modèle PlayStation VR, de la société Sony, affiche une définition inférieure de 1920x1080 pixels, fournissant ainsi une densité de pixels inférieure à la moyenne de 20 ppp et d'un champ de vision horizontal de 95°, deux caractéristiques peu en sa faveur. Le constructeur communique sur une valeur de latence de 18 ms, ce qui est très satisfaisant. Elle est difficile à comparer puisque les constructeurs des modèles concurrents ne communiquent pas cette valeur. Le produit dispose d'un système intégré capable de détecter les mouvements de rotation de l'utilisateur et fonctionnant à 1000Hz.

Évidemment dédié pour le jeu vidéo et uniquement compatible avec la console de jeu de la marque, la société ne propose pas de logiciel permettant aux développeurs de rendre compatible leurs produits avec ce visiocasque en dehors du contexte de la console de jeu.

### 3.2.3. À grand champ de vision

#### *Claire FullHD*



Le produit Claire FullHD, proposé par la société VR Union offre l'un des plus grands champs de vision parmi les visiocasques disponibles aujourd'hui, avec près de 170°. Cependant, et malgré les deux écrans intégrés et capables d'afficher 1920 pixels horizontalement, le grand champ de vision implique une densité de pixel de 22 ppp. Pour résoudre ce problème, la même

société propose un nouveau modèle, baptisé Claire 22M et disposant de deux écrans ayant pour définition totale 5210x1440 pixels. Ce nombre de pixels important horizontalement, même s'il n'est pas encore suffisant pour atteindre les capacités du système visuel humain, permet à ce nouveau modèle d'atteindre une densité d'environ 30 ppp, ce qui est très légèrement supérieur à la moyenne des visiocasques de ce type qui fournissent cependant des champs de vision très inférieurs (environ 110° en moyenne).

### *StarVR*

Le produit StarVR, proposé par la société Starbreeze, offre le plus grand champ de vision parmi les visiocasques disponibles aujourd'hui, avec 210° pour la valeur horizontale, dont 90° de recouvrement (overlap), et 130° pour la verticale grâce à l'utilisation de lentilles de Fresnel. Encore en version de développement en début 2016, il est le seul dispositif connu aujourd'hui capable de fournir un champ



de vision cohérent avec les capacités du système visuel humain. De plus, il dispose de deux écrans d'une définition assez importante par rapport à la concurrence de 5120x1440 pixels. Grâce à cette caractéristique, et malgré le champ de vision assez important, la densité de pixels offerte par ce modèle reste au niveau de la moyenne, avec un peu plus de 17 ppp.

#### 3.2.4. Avec suivi des yeux

Pour le suivi des yeux de l'utilisateur, il existe deux possibilités. Soit le suivi s'effectue par l'ajout d'un module de suivi des yeux au visiocasque, soit nativement inclus dans le visiocasque car la fonctionnalité a été prévue dès la conception du produit.

#### *Visiocasques adaptés au suivi des yeux*

Des sociétés dédiées à la conception de dispositifs capables de suivre la direction des yeux peuvent adapter leurs produits ou concevoir de nouveaux modèles pour être compatibles avec certains visiocasques. La société SMI<sup>2</sup>, par exemple, fournit un service de modification d'un modèle Rift de Oculus (la version DK2) pour lui offrir la fonctionnalité de suivi des yeux. Il est donc nécessaire d'acheter le visiocasque directement auprès de la société Oculus, puis le transmettre à la société SMI qui va y apporter les modifications nécessaires. Ensuite, avec l'aide d'un SDK fourni par la société SMI, le développeur pourra obtenir les données de suivi des yeux à une fréquence de 60 Hz, pour un champ de vision horizontal de 80° et un champ de vision vertical de 60°. La société annonce une précision du capteur comprise entre 0,5° et 1° avec une phase de calibration de quelques secondes et moins de trois points. Ce capteur permet, entre autre, d'obtenir la position des pupilles, la

---

<sup>2</sup> <http://www.smivision.com>

distance interpupillaire, le point de regard sur l'écran pour chaque œil et le diamètre de chaque pupille.

La société Starbreeze qui diffuse le modèle StarVR dont nous avons présenté les caractéristiques principales, s'est associée avec la société Tobii<sup>3</sup>, spécialiste de la conception de système de suivi des yeux, pour concevoir un dispositif capable d'être intégré directement au visiocasque de la marque.

### *Visiocasque avec suivi des yeux natif*

Parmi les visiocasques qui ont été commercialisés ou annoncés, seul le modèle FOVE, de la société du même nom, propose un suivi natif du regard. Si le modèle n'est pas encore disponible à la commercialisation au moment de l'écriture de cet ouvrage, il a fait l'objet d'une campagne de financement participative réussie et quelques informations techniques sont déjà diffusées par la société. C'est pourquoi nous le présentons ci-dessous.

### *FOVE*

Le produit FOVE, proposé par la société du même nom, est le seul visiocasque capable de détecter la direction des yeux de l'utilisateur de façon native. Avec deux systèmes infrarouges pour chacun des yeux, le système semble capable à terme de suivre la direction du regard avec une précision de 0,2 degré et une fréquence d'environ 120Hz. Il propose aussi un suivi des rotations de la tête de l'utilisateur et la marque promet de mettre en place un suivi en position à l'avenir sans pour autant donner de date.

## 3.3. Les visiocasques pour la réalité augmentée

### 3.3.1. Généralités

Bien qu'ils offrent des fonctionnalités légèrement différentes, nous pouvons aussi intégrer dans cette revue des visiocasques commercialisés ou en passe de l'être, les modèles dédiés aux fonctionnalités de réalité augmentée. Il existe plusieurs catégories pour ce type de matériel. Qu'ils soient monoculaires ou binoculaires, on distingue les dispositifs basés sur un ou plusieurs écrans classiques (*video see-through*), à vision indirecte, des dispositifs basés sur un ou plusieurs écrans semi-transparents (*optical see-through*), à vision directe. Cette distinction n'est pas uniquement matérielle puisqu'elle conditionne les fonctionnalités atteignables par ces modèles.

---

<sup>3</sup> <http://www.tobii.com>

### 3.3.2. À écran classique

Dans cette catégorie, un écran du même type que ceux utilisés pour les smartphones (LCD ou OLED par exemple), est positionné devant le ou les yeux de l'utilisateur et cache tout ou partie de son champ de vision. Cet écran est maintenu par des montures ou un casque.

#### *À occultation partielle du champ de vision*

Dans ce cas, l'écran occulte qu'une partie du champ de vision de l'utilisateur et il est ainsi capable d'observer son environnement directement lorsqu'il ne dirige pas son regard vers l'écran mais aussi à l'aide de sa vision périphérique. Il s'agit en réalité d'un écran déporté, main-libre, qui fournit une fonctionnalité de «réalité documentée». Par exemple, le modèle Jet, de la société Recon Instruments<sup>4</sup>, se compose d'une paire de lunettes disposant d'un écran très légèrement décalé en dessous de l'œil droit de l'utilisateur. Sur cet écran, s'affichent des informations que l'utilisateur peut consulter sans l'aide de ces mains. Bien que ce modèle dispose d'une caméra, elle ne sert qu'à prendre des photographies de l'environnement et ne rentre pas en compte dans une fonctionnalité de réalité augmentée proprement dite. Puisque l'écran s'éteint lorsque l'utilisateur n'a plus besoin de le consulter (pour des raisons d'autonomie principalement), il peut gêner l'utilisateur en obstruant une partie de son champ de vision. Cependant, ces modèles, dotés d'écran classiques ont l'avantage de fournir une luminosité suffisante pour un usage en plein jour.

#### *À occultation complète du champ de vision*

Dans le cas de ces modèles, les écrans occultant la totalité du champ de vision de l'utilisateur, ils se rapprochent des visiocasques que nous avons présentés dans les premières sections de ce chapitre. La différence majeure provient de la présence d'une ou plusieurs caméras disposées aussi sur la tête de l'utilisateur (souvent au dos des écrans), avec lesquelles l'environnement est filmé et retranscrit sur les écrans, permettant à l'utilisateur de voir ce qu'il aurait dû voir s'il n'avait pas d'écran devant les yeux. Le temps de latence est un critère important pour ce type de modèle qui doit être capable d'afficher les images issues des caméras très rapidement afin de ne pas gêner l'utilisateur lors de ces mouvements de tête. Globalement, les contraintes sont proches de celles rencontrées avec les visiocasques dédiés à la réalité virtuelle. De plus, l'autonomie est aussi un défi de taille, puisque la réalité augmentée est destinée à un usage en mobilité à la différence de la réalité virtuelle, et qu'il paraît donc inapproprié d'être branché à un câble pour obtenir l'énergie nécessaire. Dans cette catégorie, on peut citer le modèle ImpressionPi<sup>5</sup>. Il s'agit d'un modèle de visiocasque, très proche de ceux que nous avons présentés dans la section dédiée

---

<sup>4</sup> <http://www.reconinstruments.com/products/jet/>

<sup>5</sup> <http://www.impressionpi.com>

aux modèles basés sur les smartphones. Mais il dispose de deux capteurs vidéo (une caméra RGB et une caméra de profondeur) capable d'observer l'environnement pour fournir ainsi une fonctionnalité de réalité augmentée en affichant sur l'écran ce qui est filmé par l'une des caméras (RGB) et en utilisant les informations de profondeur pour reconstruire l'environnement proche de l'utilisateur. À l'aide de cette reconstruction, les entités créées numériquement (texte, objets 3D, etc.) peuvent être intégrées convenablement dans l'image en prenant en compte les occultations.

### 3.3.3. À écran semi-transparent

Plusieurs techniques ont été proposées au cours des dernières années et toutes ces solutions peuvent être regroupées en deux classes. Il existe la technologie basée sur des miroirs courbes (*curved mirror*) ou celle basée sur la canalisation des rayons lumineux (*waveguide* ou *light-guide*). La première technologie utilise des miroirs courbés disposés devant les yeux de l'utilisateur mais décalés afin de ne pas gêner la vision. Cette solution implique de grandes distorsions de l'image qu'il est nécessaire de corriger par une électronique dédiée ou à l'aide d'une optique complexe. Ces corrections ajoutent de la complexité au système, du poids, un encombrement non négligeable et ont un impact négatif sur l'autonomie et sur la résolution. Le modèle de la société Vuzix<sup>6</sup>, M100, ou les premiers modèles de la société Laster Technologies<sup>7</sup> utilisent cette technologie. La deuxième technologie est basée sur la canalisation des rayons lumineux et il s'agit de la catégorie dans laquelle on peut trouver le plus de modèles aujourd'hui. Avec cette technologie, l'utilisateur perçoit visuellement l'environnement réel à travers un miroir semi-transparent maintenu devant ces yeux. Cet miroir a la particularité de laisser passer la majeure partie des rayons lumineux provenant de l'extérieur, tout en bloquant une partie des rayons, permettant à l'utilisateur de percevoir des éléments projetés sur cette surface. Ces rayons lumineux projetés sont issus d'une source déportée, guidés au travers d'un canal optique. Ainsi, les positions de la source lumineuse et de l'électronique peuvent être éloignées de la surface de projection perçue par l'utilisateur pour faciliter leur intégration. Plusieurs principes optiques ont été utilisés dans ce cadre comme la diffraction, l'holographie, la polarisation ou la réflexion.

Le modèle de la société Google, baptisé Google Glasses<sup>8</sup>, est un exemple de produit qui utilise le principe optique de la réflexion. Il s'agit d'une paire de lunettes disposant d'un écran semi-transparent monoculaire. Sur cet écran, positionné dans la partie haute du champ de vision de l'utilisateur, est projetée une image issue d'un pico-projecteur positionné à l'arrière du dispositif et conduite depuis la source jusqu'à la surface de projection par réflexion. L'inconvénient principal de ce type de système est la faible luminosité offerte par l'écran, rendant son utilisation très difficile, voire impossible en plein jour. De plus, l'énergie nécessaire pour faire fonctionner un tel projecteur reste encore importante et un compromis est souvent nécessaire entre

6 <https://www.vuzix.com/Products/M100-Smart-Glasses>

7 <http://www.laster.fr>

8 <https://www.google.com/glass/start/>

l'autonomie du dispositif et le poids de la batterie embarquée. La société Française Laster Technologies est aussi sur ce marché depuis 2005. Disposant de plusieurs modèles, elle propose aujourd'hui un produit monoculaire à écran semi-transparent nommé Laster SeeThru. Avec un champ de vision diagonal d'environ 25° et d'une luminosité de 5200 cd/cm<sup>2</sup>, le système est dépendant d'une application dédiée qui doit être installée sur un smartphone mais dispose cependant d'un système de suivi de la tête en orientation (gyroscope, accéléromètres et boussole) et d'un capteur GPS.

C'est au cours de l'année 2016 que la société Microsoft doit commercialiser son dispositif, nommé HoloLens<sup>9</sup>, sous la forme d'une paire de lunettes binoculaire et équipé d'une technologie dont les détails techniques restent aujourd'hui encore peu diffusés. Équipé de très nombreux capteurs (environ une dizaine) dédiés à la captation de l'environnement extérieur et d'un système de son spatialisé, ce modèle se démarque des principaux concurrents par l'utilisation d'un dispositif de retour visuel binoculaire de type *light-field* permettant d'afficher chaque pixel avec une focale variable (voir chapitre 4). Des capteurs sont capables de reconstruire un modèle numérique 3D de l'environnement proche à partir du point de vue de l'utilisateur. Ensuite, un logiciel est en mesure d'utiliser les informations du modèle numérique 3D de la scène réelle afin d'afficher des objets virtuels à incruster. Chaque pixel de l'image d'un objet virtuel est affiché avec la bonne focale en fonction de sa distance de profondeur. Un tel affichage a pour intérêt de supprimer la fatigue visuelle associée au phénomène de conflit d'accommodation-vergence. Bien que le champ de vision ne soit pas clairement communiqué par le constructeur, il semble faible comme tous les modèles des laboratoires de recherche qui ne disposent que d'un champ de vision relativement limité de quelques dizaines de degré comme le Tensor Display<sup>10</sup> du MIT avec environ 20° de champ de vision.

## 4. CONCLUSION

Dans ce chapitre, nous avons proposé un tour d'horizon des principaux modèles de visiocasques destiné à la consultation de vidéo à 360°, à la réalité virtuelle ou à la réalité augmentée. On constate d'abord que le nombre de dispositifs disponibles ou en passe de l'être est assez conséquent et varié avec des solutions techniques hétérogènes. Tous ces modèles ont des avantages et des inconvénients qu'il est nécessaire de bien appréhender afin de choisir convenablement le dispositif en fonction de l'application visée. Il faut noter qu'à l'heure de l'écriture de cet ouvrage, aucun dispositif n'est en mesure d'approcher, dans toutes ces dimensions (résolution, champs de vision, rapidité, etc.) les capacités du système visuel de l'être humain et que de très nombreuses avancées technologiques sont encore nécessaires pour atteindre cet objectif sans faire de compromis sur l'autonomie en situation de mobilité.

9 <https://www.microsoft.com/microsoft-hololens/en-us>

10 Wetzstein, G., Lanman, D., Hirsch, M., Raskar, R. (2012). "Tensor Displays: Compressive Light Field Display using Multilayer Displays with Directional Backlighting", ACM Transactions on Graphics (SIGGRAPH)

## 5. RÉCAPITULATIF DES CARACTÉRISTIQUES DES MODÈLES DE VISIOCASQUE

Nom	Fabricant	Type	Ecran			Res. (ppi)	Définition par oeil (px)		Densité de pixel (ppd)		Rafraîchissement (Hz)	FOC(°)				Tracking	Poids	DIO (mm)		Site web			
			Type	Taille (pouces)			H	V	H	V		2 yeux						Tête	yeux		Réglages	min	max
				min	max							D	H	V	D								
Evomade	Viewbox	Avec Fixation	-	-	-	2560	1440	-	1280	1440	28,4	-	-	-	-	-	-	-	-	http://evomade.com			
Clare FulHD	VR Union	Pour la VR	-	-	-	3840	1080	-	1920	1080	22,6	-	-	-	-	-	-	-	-	http://vrunion.com			
Oculus Rift DK1	Oculus VR	Pour la VR	LCD	7	-	1280	800	216	640	800	12,1	10,78	60	-	-	-	-	-	-	https://www.oculus.com/en=AS/us/			
Oculus Rift DK2	Oculus VR	Pour la VR	OLED	5,7	-	1920	1080	386	960	1080	20,0	15,21	75	-	-	-	-	-	-	https://www.oculus.com/en=AS/us/			
Gameface	Gameface Labs	Pour la VR	OLED	-	-	2560	1440	-	1280	1440	18,3	-	75	-	-	-	-	-	-	http://gamefaceclabs.com			
StarVR	Starbesze	Pour la VR	LCD	5,5x2	-	5120	1440	-	2560	1440	17,1	22,59	60-90	-	-	-	-	-	-	http://starvr.com			
Playstation VR	Sony	Pour la VR	OLED	5,7	-	1920	1080	386	960	1080	20,2	-	120	-	-	-	-	-	-	https://www.playstation.com/en=AS/us/ explaire/playstation=AS/ver/			
OSVR HDK 1.3	Razer/ Senseis	Pour la VR	OLED	5,5	-	1920	1080	401	1080	960	21,6	-	60	-	-	-	-	-	-	http://www.razerzone.com/osvr			
The Rift	Oculus VR	Pour la VR	OLED	-	-	2160	1200	-	1200	1080	-	16,99	90	-	-	-	-	-	-	https://www.oculus.com/en=AS/us/			
HTC Vive	HTC/ Valve	Pour la VR	OLED	-	-	2160	1200	-	1200	1080	20,4	-	90	-	-	-	-	-	-	http://www.htcive.com/us/			
Totem	VRVANA	Pour la VR	OLED	-	-	1920	1080	-	960	1080	18,3	-	120	-	-	-	-	-	-	http://www.vrvana.com/fr			
FOVE	FOVE	Pour la VR	OLED	5,8	-	2560	1440	506	1280	1440	25,6	-	60-90	-	-	-	-	-	-	http://www.getfove.com			
Glyph	Avegant	Pour la VR	Micro- projection	-	-	2560	1440	-	1280	720	56,9	-	120	-	-	-	-	-	-	https://avegant.com			
Clare 22M	VR Union	Pour la VR	OLED	-	-	5120	1440	-	2560	1440	30,1	-	-	-	-	-	-	-	-	-			
Wearality Sky	Wearality	Sans Fixation	-	-	-	2560	1440	-	1280	1440	17,1	15,66	-	-	-	-	-	-	-	http://www.wearality.com			
Cardboard	Google	Sans Fixation	-	-	-	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	-	https://www.google.com/get/cardboard/			

Nom	Fabricant	Type	Ecran			Res. (ppi)	Définition (px)		Rafraîchissement (Hz)	FOC(°)						Poids	DIO (mm)		Site web	
			Type	Taille (pouces)			H	V		1 oeil		2 yeux		Tête	yeux		Réglages	min		max
				D	H					V	D	H	V							
Bluesky Pro	Immersion Vrelia	Avec Fixation	-	-	3840	1080	-	1920	1080	31,2	-	-	-	-	-	-	-	<a href="http://immersionvrelia.com/bluesky/">http://immersionvrelia.com/bluesky/</a>		
AirVR	AirVR 1	Avec Fixation	-	-	1920	1080	-	960	1080	19,2	-	-	-	-	-	-	-	<a href="http://getairvr.com">http://getairvr.com</a>		
FreelyVR	ProteusVR Labs	Avec Fixation	-	4,7	6,1	2560	1440	-	1280	1440	21,3	-	-	-	-	-	-	<a href="https://www.freelyvr.com/">https://www.freelyvr.com/</a>		
Diodlo	H One	Avec Fixation	-	-	2560	1440	-	1280	1440	21,3	-	-	-	-	-	-	-	<a href="http://www.diodlo.com/en/">http://www.diodlo.com/en/</a>		
Cmoar	CmoarVR	Avec Fixation	-	-	2560	1440	-	1280	1440	24,4	-	-	-	-	-	-	-	<a href="http://cmoar.com">http://cmoar.com</a>		
Homido	Homido	Avec Fixation	-	4	5,7	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	<a href="http://www.homido.com">http://www.homido.com</a>		
VR One	Zeiss	Avec Fixation	-	4,7	5,2	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	<a href="http://www.zeissvrone.com">http://www.zeissvrone.com</a>		
Durovis Dive 5	Shoogee	Avec Fixation	-	2	5	2560	1440	-	1280	1440	25,6	-	-	-	-	-	53	77	<a href="https://www.durovis.com/product.html?id=1">https://www.durovis.com/product.html?id=1</a>	
Durovis Dive 7	Shoogee	Avec Fixation	-	-	7	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	<a href="https://www.durovis.com/product.html?id=1">https://www.durovis.com/product.html?id=1</a>	
Visus VR	Visus	Avec Fixation	-	-	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	-	<a href="http://www.visusvr.com">http://www.visusvr.com</a>	
Xingear XH VR	-	Avec Fixation	-	-	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	-	<a href="http://www.linearboard.com/xingear.html">http://www.linearboard.com/xingear.html</a>	
Pinch	Pinch VR	Avec Fixation	-	-	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	-	<a href="http://hellopinch.com">http://hellopinch.com</a>	
Dior Eyes	Dior	Avec Fixation	-	-	2560	1440	-	1280	1440	25,6	-	-	-	-	-	-	-	-	<a href="http://digitallabparis.com/post/119609782829/dior=AS%20eyes">http://digitallabparis.com/post/119609782829/dior=AS%20eyes</a>	
Gear VR (Note 4)	Samsung VR	Avec Fixation	Super AMOLED	5,7	-	2560	1440	515	1280	1440	26,7	-	-	-	-	-	-	55	71	<a href="http://www.samsung.com/global/interoseite/gearvr_specs.html">http://www.samsung.com/global/interoseite/gearvr_specs.html</a>
Gear VR (S6 Edge)	Samsung VR	Avec Fixation	OLED	5,1	-	2560	1440	576	1280	1440	28,4	-	-	-	-	-	-	55	71	<a href="http://www.samsung.com/global/interoseite/gearvr_specs.html">http://www.samsung.com/global/interoseite/gearvr_specs.html</a>
Impression Pi	-	Avec Fixation	-	-	2560	1440	-	1280	1440	28,4	-	-	-	-	-	-	-	-	-	<a href="http://www.impressionpi.com">http://www.impressionpi.com</a>

### 1. INTRODUCTION

L'usage des visiocasques peut poser divers problèmes de confort et de santé. Ces derniers peuvent être induits par certaines catégories d'applications RV, créant des malaises dus principalement :

- à des incohérences sensorimotrices, dont celles dues aux temps de latence trop élevés ;
- à l'activité psychologique du sujet perturbé dans l'environnement virtuel ;
- à un mauvais interfaçage entre le système visuel du sujet et le visiocasque ;
- à des dispositifs technologiques non sécurisés.

Nous développerons les trois premiers cas dans les trois parties suivantes, tandis que la question de sécurité sera traitée dans le chapitre suivant.

Comme nous l'avons déjà indiqué, les techniques de réalité virtuelle induisent presque toujours des incohérences, soit à l'intérieur d'un sens (par exemple, incohérence entre accommodation et convergence des yeux en vision stéréoscopique), soit entre plusieurs sens (par exemple, la locomotion sur un tapis roulant entraînant une incohérence entre la vision et la proprioception), soit entre sens et réponses motrices (par exemple, la manipulation d'objets virtuels sans retour d'effort). Dans le monde réel, l'homme construit une représentation cohérente à partir de tous les stimuli sensoriels reçus. Dans le monde virtuel, l'utilisateur cherchera en conséquence à interpréter avec cohérence ce qu'il perçoit par rapport à son vécu dans le monde réel, malgré les incohérences sensorimotrices. Mais il est difficile actuellement de déterminer *a priori* si le sujet, immergé dans tel ou tel environnement virtuel, arrive à percevoir et rendre cohérents ces environnements sans difficulté, en surmontant les incohérences sensorimotrices.

Une autre difficulté peut rendre inconfortable l'activité de l'utilisateur dans un environnement virtuel : si la réaction de l'environnement virtuel tarde par rapport à l'action de l'utilisateur. Le temps de latence, qui dépend des performances technologiques de l'installation RV, aura donc une incidence sur le comportement de l'utilisateur. On peut dire que c'est aussi une incohérence sensorimotrice, mais avec un critère temporel.

Mais indépendamment des solutions techniques exploitées, même avec un visiocasque parfait, c'est-à-dire qui proposerait une même qualité de vision que celle du monde réel, il y aura toujours des incohérences pour certains types d'applications. Il ne s'agit pas seulement de résoudre des problèmes technologiques dans un proche avenir pour que l'on supprime toute incohérence et que l'on puisse faire n'importe quelle activité sensorimotrice en environnement virtuel. Il y aura toujours des limites. Par principe, on crée un nouveau monde artificiel avec ses règles, ses limites et ses potentialités dont certaines dépassent ou sont différentes de celles du monde réel.

Dans certains cas classiques et bien analysés, nous savons par expérience que la personne peut surmonter certaines incohérences en s'y adaptant. Certaines adaptations se font presque naturellement : l'exemple le plus explicite est celui du déplacement virtuel face à un petit écran d'ordinateur ou de console de jeux. Dans cette incohérence visuo-vestibulaire avec vision périphérique sur l'ER<sup>1</sup>, l'utilisateur ressent plus ou moins son propre déplacement si on affiche des images d'un monde virtuel en mouvement. Ceci est dû au phénomène devection qui impose la sensation de mouvement, bien que la personne immobile dans le monde réel ne détecte point de mouvement par ses systèmes vestibulaires. Il est extrêmement rare de trouver des personnes qui ne s'adaptent pas à ce cas de déplacement virtuel devant un écran ne recouvrant pas leur vision périphérique. Dans le cas de l'observation via un visiocasque, le problème est bien plus complexe, ce que nous détaillerons dans les paragraphes suivants.

Dans les applications RV, tous les sens n'interviennent pas pour créer la simulation souhaitée. Certains sont à prendre en compte dans les incohérences sensorimotrices induites par les techniques RV :

- la vision ;
- l'ouïe ;
- la sensibilité cutanée ;
- le sens kinesthésique (systèmes vestibulaires, récepteurs sensoriels aux articulations, etc.) ;
- la proprioception musculaire.

## 2. INCOHÉRENCES SENSORIMOTRICES

### 2.1. Généralités

Il existe un grand nombre d'incohérences sensorimotrices dans les applications RV. Il n'y a pas que l'incohérence visuo-vestibulaire, même si celle-ci est une des plus compliquées à gérer lorsque l'utilisateur exploite un visiocasque. Il est donc

---

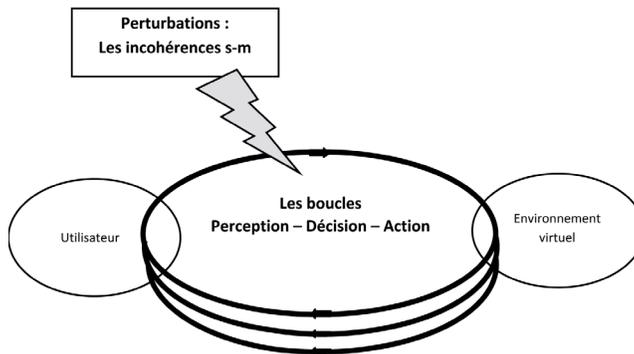
<sup>1</sup> La vision périphérique étant sur l'ER, l'utilisateur est en posture stable dans l'ER grâce à la cohérence entre sa vision périphérique et ses systèmes vestibulaires et ne ressent pas de malaise ou d'inconfort.

souhaitable de classer les incohérences sensorimotrices et de structurer une analyse... pour ne pas avoir un « inventaire à la Prévert » de toutes les recommandations que tout concepteur RV doit suivre pour améliorer le confort de l'utilisateur. Nous allons analyser cette problématique à partir de notre démarche d'interfaçage en RV, le modèle 3I<sup>2</sup>, dont nous rappelons les notions essentielles :

- Le premier niveau, l'immersion et l'interaction sensorimotrices, schématise l'interfaçage entre l'homme et l'environnement virtuel (EV) au niveau physique. Le niveau où l'ordinateur est connecté au corps de l'utilisateur via ses sens et de ses Réponses Motrices (RM - c'est-à-dire ses actions en EV). Ce niveau doit être **transparent** pour l'utilisateur, gage d'une conception réussie de l'interfaçage de l'application RV ;
- Le deuxième niveau, l'immersion et l'interaction cognitives, concerne les processus cognitifs, schèmes comportementaux importés ou métaphores, que le sujet va mettre en œuvre. L'utilisateur devra être assisté dans son interfaçage avec l'environnement virtuel par des Aides Logicielles Comportementales (ALC), composées des aides logicielles cognitives (ALC<sub>Cog</sub>) et des aides logicielles sensorimotrices (ALSM) ;
- Le troisième niveau, l'immersion et l'interaction fonctionnelles, concerne les activités de l'utilisateur en EV. Celles-ci sont toujours décomposables en quelques comportements de base, les « Primitives Comportementales Virtuelles » (PCV). Ces dernières peuvent être regroupées en quatre catégories :
  - observer le monde virtuel ;
  - se déplacer (naviguer) dans le monde virtuel ;
  - agir sur le monde virtuel, principalement manipuler des objets ;
  - communiquer avec autrui ou avec l'application RV.

Les trois premières sont difficiles à traiter aux niveaux des incohérences sensorimotrices et ce n'est que sur celles-ci que nous faisons une analyse détaillée. La dernière PCV pose aussi quelques problèmes spécifiques. De plus, comme nous l'avons indiqué au chapitre 2, les interfaces doivent avoir des caractéristiques métrologiques performantes, si possible, correspondant aux capacités psychophysiologiques maximales des sens ou(et) des réponses motrices de l'homme. En pratique, les qualités des stimuli sensoriels et des réponses motrices doivent être au moins satisfaisantes par rapport aux PCV exigées par l'application.

Comme nous l'avons indiqué au chapitre 1, il est nécessaire de rappeler que la principale difficulté de l'immersion et de l'interaction du sujet provient des **perturbations** de la boucle « Perception, Décision, Action » (PDA). En pratique, on a plusieurs boucles PDA en considérant tous les sens séparément. Grâce à des choix judicieux des PCV exploitées, des dispositifs d'interfaçage, des processus cognitifs et des ALC, les perturbations peuvent être maîtrisées pour permettre au sujet d'agir efficacement en EV, via des boucles PDA artificielles, comme nous l'expliciterons au chapitre suivant. Mais, en premier, nous allons détailler ces perturbations dues aux incohérences sensorimotrices, Figure 7.1.



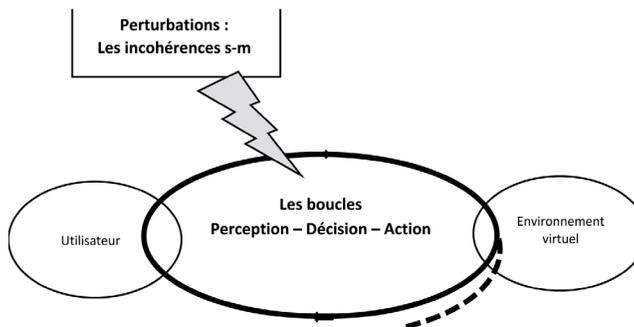
**Figure 7.1 :** les incohérences sensorimotrices perturbent le niveau des  $I^2$  sensorimotrices

Quelles sont les différentes catégories d'incohérences sensorimotrices ? Comme nous l'avons déjà indiqué, il y a des incohérences soit entre deux sens (ou une caractéristique sensorielle)<sup>2</sup>, soit entre un sens et une réponse motrice. Pour ces deux cas, nous pouvons différencier deux sous-catégories avec des incohérences qui sont dues :

- soit à l'**Absence** d'un stimulus ou d'une réponse motrice en EV, par rapport à ce qui se produit en ER ;
- soit à un **Découplage**, c'est-à-dire une différence dans la relation entre deux sens ou entre un sens et une réponse motrice en EV, par rapport à ce qui se produit en ER.

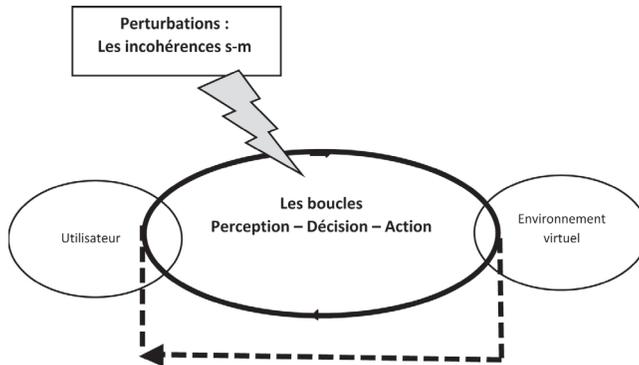
Par ailleurs, une autre catégorie d'incohérences sensorimotrices existe aussi avec le découplage temporel, la **Latence**, qui se produit entre deux sens ou entre un sens et une réponse motrice. Cela fait en tout 6 catégories différentes :

- trois types d'**incohérences sensorielles** :
  - Absence d'un sens (ou d'une caractéristique sensorielle) (noté **As<sup>2</sup>**)

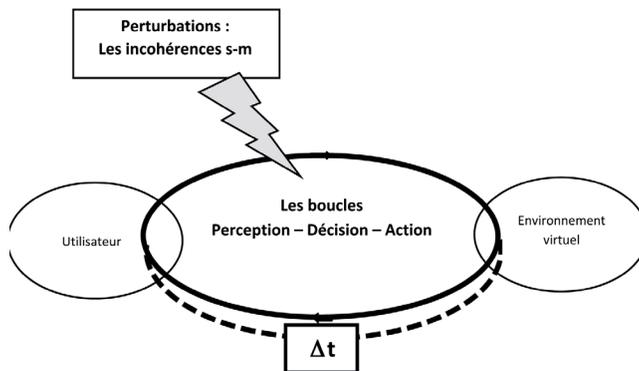


<sup>2</sup> Ce regroupement pour éviter trop de cas différents.

- Découplage entre deux sens (noté **Ds<sup>2</sup>**)

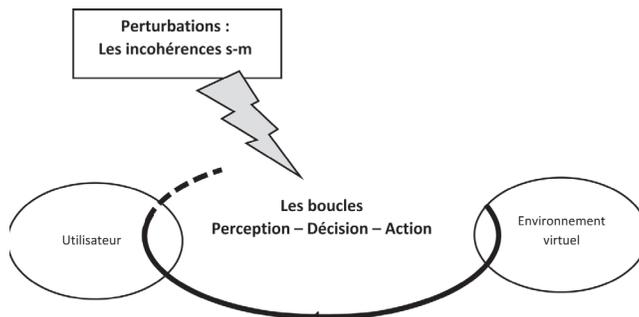


- Latence, découplage temporel entre les stimuli de deux sens (noté **Ls<sup>2</sup>**)

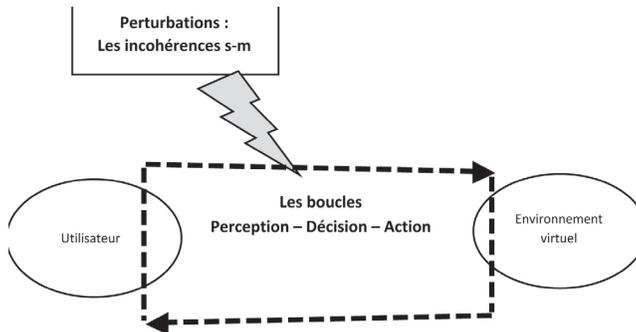


- les trois types d'**incohérences sensorimotrices**

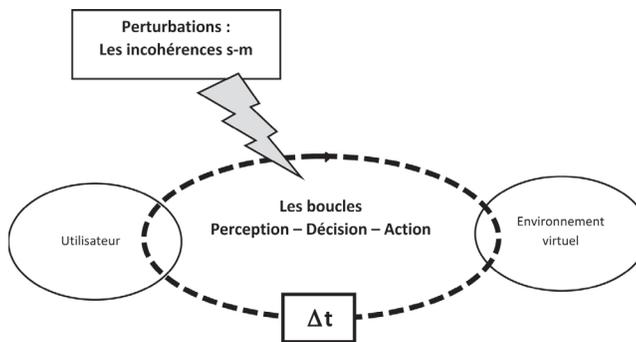
- Absence d'une réponse motrice (noté **As-m**) ;



- Découplage entre un sens et une réponse motrice (noté **Ds-m**)



- Latence, découplage temporel entre une réponse motrice et un sens (noté **Ls-m**).



Les 6 situations sont présentées dans le tableau suivant :

Type d'incohérence Sens ou réponses motrices concernés	Absence	Découplage	Latence
Entre deux sens	As <sup>2</sup>	Ds <sup>2</sup>	Ls <sup>2</sup>
Entre un sens et une réponse motrice	As-m	Ds-m	Ls-m

**Tableau 7.1** : les différents types d'incohérence sensorimotrice

Cette classification est restrictive et discutable car on peut avoir trois sens en incohérence selon le découpage que l'on établit des sens. On regroupe sous le terme « sens de la proprioception » plusieurs récepteurs sensoriels : les systèmes vestibulaires

ainsi que les organes proprioceptifs des muscles, des tendons et des articulations. Nous reparlerons du cas particulier de la proprioception car les interfaces tels que les tapis roulants 1D ou 2D permettent à la personne de marcher donc de stimuler les organes proprioceptifs des muscles, des tendons et des articulations tandis que les interfaces de simulation de mouvement permettent de stimuler tous ces précédents organes proprioceptifs et en plus les systèmes vestibulaires.

Comme nous l'avons déjà indiqué, les perceptions et les actions doivent être perçues cohérentes pour le cerveau de l'utilisateur. C'est à notre avis, la principale hypothèse de travail pour analyser l'influence de tous ces types d'incohérences sensorimotrices. On peut aussi faire d'autres hypothèses, plus ou moins par ordre d'importance :

- H1 : le cerveau cherche à **percevoir un monde cohérent** à partir de tous les stimuli sensoriels reçus et de ses réponses motrices dans la boucle « Perception, Décision, Action » en EV, comme il le fait en ER dès l'enfance, lors du développement de son intelligence sensorimotrice.
- H2 : le cerveau **anticipe ses réponses motrices** en fonction des stimuli sensoriels reçus, comme expliqué au chapitre 2.
- H3 : **la boucle PDA est perturbée** dans toute application de la réalité virtuelle.
- H4 : Les activités du sujet en environnement virtuel **créent presque toujours des incohérences** sensorielles et sensorimotrices, dues aux choix de réalisation des PCV souhaitées et aux difficultés techniques à réaliser les mêmes stimuli et réponses motrices en EV qu'en ER.
- H5 : les techniques RV **perturbent l'instantanéité** de la boucle PDA à cause des temps de latence, toujours présents par principe technique.
- H6 : la **vision est un sens prédominant** dans les incohérences sensorielles et sensorimotrices.
- H7 : si le cerveau n'arrive pas à rendre cohérent ses perceptions et ses actions en EV, il y a au **minimum inconfort, voire malaise et au pire rejet** pour l'utilisateur.
- H8 : si les temps de latences sont trop élevés en EV, l'utilisateur a des **difficultés à anticiper et donc à contrôler** ses réponses motrices, donc son activité.
- H9 : l'utilisateur peut **s'adapter** aux incohérences en EV par apprentissage.
- H10 : l'utilisateur peut s'adapter à contrôler ses actions malgré des temps de latence trop élevés en s'imposant des **actions plus lentes et par apprentissage**.
- H11 : les utilisateurs ont des **capacités d'adaptation très variées**.
- H12 : pour la grande majorité de la population, certaines incohérences sont connues et **exploitables** tandis que d'autres sont **perturbatrices** à des degrés variés. Elles sont plus ou moins caractérisées actuellement, en se basant sur les capacités moyennes de l'ensemble de la population.

## 2.2. Classification des incohérences sensorimotrices

Si la plupart des incohérences peuvent provoquer une gêne, quelquefois rédhitoire, elles peuvent aussi avoir un effet **bénéfique** quand la recherche de cohérence du monde par le sujet fait que celui-ci «croit en la simulation visuelle». On rappelle le cas classique, explicité précédemment, de l'incohérence visuo-vestibulaire avec vision périphérique sur l'ER: l'utilisateur croit se déplacer grâce à la vection sans que ses systèmes vestibulaires soient sollicités. Certaines incohérences sont donc sources de solutions efficaces. C'est avec cette même logique que l'on peut exploiter le **pseudo-haptique**<sup>3</sup> [Lécuyer A., *et al.* 2000]. L'expérience de base est de fournir sur la main de l'utilisateur une force variable par un ressort de *raideur constante*, tout en proposant visuellement une variation de déplacement de l'objet virtuel manipulé par l'utilisateur. Si on change visuellement l'amplitude de déplacement de l'objet virtuel pour une force donnée sur le ressort, le testeur perçoit (ressent) un ressort de raideur différent. Car le cerveau du testeur veut à tout prix rendre cohérent son action en référence au «modèle interne (de mécanique)» acquis par son cerveau lors d'expériences passées. Ce phénomène, exploité dans des applications de réalité virtuelle, génère des sensations haptiques en modifiant le retour visuel au moment où l'utilisateur agit dans l'environnement virtuel, **sans avoir besoin d'une interface à retour haptique** commandée par la simulation. Sur le schéma explicatif du retour pseudo-haptique figure 7.2, nous pouvons observer que le sujet est stimulé par un ensemble incohérent de stimuli visuels et de stimuli haptiques réels, puisque ces derniers ne sont pas variables. Le retour pseudo-haptique correspondrait à une réinterprétation de ces stimuli et à une perception visio-haptique optimale d'un monde qui doit rester ou redevenir cohérent pour le sujet.

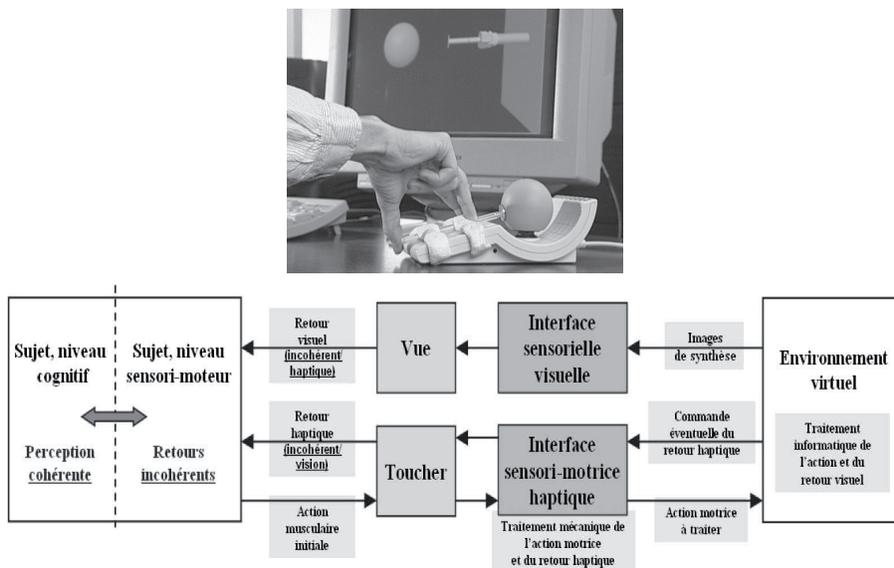


Figure 7.2 : dispositif expérimental et schéma de principe du pseudo-haptique

3 L'haptique concerne le sens du toucher et celui de la proprioception musculaire qui sont stimulés en RV par les interfaces haptiques, composées de deux catégories : les interfaces tactiles et les interfaces à retour d'effort.

Dans les deux cas précédents, la vision est prépondérante sur les autres sens pour la perception de l'utilisateur, « celui-ci croit ce qu'il voit ».

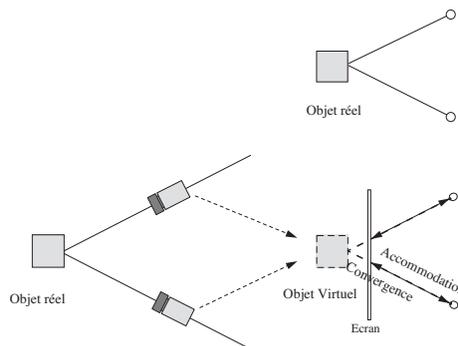
*Ces incohérences sont dites bénéfiques car, tout en ne créant pas d'inconfort et de malaise, elles rendent la simulation plus réaliste au niveau de la perception multisensorielle, plus proche de celle qui se réalise en ER et ceci en omettant de stimuler correctement un sens. À l'opposé, des incohérences sont dites perturbatrices car elles impliquent inconfort ou malaise pour l'utilisateur.*

Certaines sont très perturbatrices, c'est-à-dire qu'elles sont très difficilement supportables par la majorité des utilisateurs. En particulier, il est fortement recommandé au concepteur de l'application RV de ne jamais faire tourner l'EV dans le visiocasque sans que l'utilisateur, immobile dans l'ER, ne tourne la tête, donc sans sa propre action. Car celui-ci ne pourra pas *anticiper* sa perception visuelle dans l'EV, son cerveau n'étant pas renseigné par sa propre action volontaire. Le malaise est presque toujours assuré si vous tentez ce test. Une expérience assez proche dans le monde réel est celle d'être passager dans un véhicule au lieu d'être conducteur. Il est plus difficile d'anticiper les rotations et les accélérations des trajectoires en étant passif par rapport à celles-ci, même si les mouvements sont néanmoins détectés par les systèmes vestibulaires. On est donc plus sensible au mal du transport en étant passager que conducteur. Il est donc fortement recommandé que le concepteur ne fasse jamais de mouvements de l'EV sans commande par la tête (ou les mains ou tout le corps) de l'utilisateur portant un visiocasque.

Nous pouvons établir les listes des incohérences bénéfiques et perturbatrices, en fonction des trois PCV : observer, naviguer et manipuler. Ces listes ne sont pas exhaustives.

#### Liste de 5 incohérences bénéfiques :

- PCV Observation :
  - *Incohérence visio-acoustique ( $Ds^2$ )* : la source sonore d'un objet sera en général localisée par la position affichée de l'objet dans le visiocasque, même si la localisation de la source sonore est différente, ce qui est le cas quand on ne fournit pas un son spatial (3D) dans le visiocasque. Cette illusion du « ventriloque » a été expliquée au chapitre 2. Elle risque ne pas fonctionner s'il y a trop d'écart entre les deux localisations ;
  - *Incohérence oculomotrice ( $Ds^2$ )* : nous avons expliqué que la vision stéréoscopique permet de mieux percevoir les objets sur trois dimensions, même si les images sont affichées sur le plan de l'écran (l'illusion de voir les objets devant ou derrière l'écran). Ceci est réalisé grâce à la convergence des yeux de l'observateur bien qu'il y ait une incohérence « accommodation-vergence » et bien que l'accommodation indique au cerveau que les images des objets sont au niveau de l'écran. Mais cette illusion n'est possible que dans certaines limites (disparité rétinienne limitée). Si on dépasse ces limites, cette incohérence devient une contrainte perturbatrice (voir plus loin).



**Figure 7.3 :** avec des caméras réelles ou virtuelles (images de synthèse stéréoscopiques), l'observateur accommode sur les images au niveau de l'écran et ses yeux convergent à une distance différente, fonction des parallaxes horizontales des images.

- PCV Navigation :

- *Incohérence visuo-vestibulaire ( $Ds^2$ )* : nous avons aussi expliqué que le déplacement virtuel par vection sans déplacement réel de l'observateur permettait à celui-ci de se percevoir en mouvement. Cette illusion est exploitable dans certaines limites cinématiques du mouvement virtuel pour ne pas devenir une contrainte perturbatrice (voir plus loin).

- PCV Manipulation :

- *Incohérence visuo-haptique ( $Ds^2$ )* : le pseudo-haptique permet l'illusion de retour d'effort sans interface haptique. Ce phénomène peut être exploité avec des ressorts dans une manette de commande ou en pressant une balle détectant ses déformations, etc.
- *Incohérence visuo-tactile ( $Ds^2$ )* : il s'agit de réaliser des « images tactiles » donnant l'illusion de percevoir la rugosité d'une surface, représentée sur une image. Cet effet se réalise lorsque l'utilisateur déplace un curseur à la souris sur l'image de la surface plus ou moins rugueuse. Le mouvement du curseur va alors correspondre au mouvement de la souris auquel on rajoute un signal bruité en vitesse de déplacement : si le curseur rencontre une proéminence sur la surface de l'image, il est ralenti et inversement accéléré s'il s'agit d'un creux sur la surface, voir [www.irisa.fr/tactiles](http://www.irisa.fr/tactiles).

Toutes ces incohérences sont de la catégorie  $Ds^2$  et sont bénéfiques grâce à la vision qui influence convenablement un autre sens. La *perception visuelle* l'emporte sur les *sensations* contradictoires reçues par le cerveau. On peut se poser la question pour savoir, en dehors de la vision, si l'ouïe influence le toucher, avec ou sans l'implication de la vision, donc avec ou sans image. En mettant au bout du doigt une bobine vibrante fournissant un stimulus tactile d'intensité et de fréquence constantes, peut-on percevoir des variations tactiles en jouant sur des variations sonores ? L'expérience pourrait consister à déplacer un doigt en l'air sur une surface virtuelle et percevoir la rugosité différente en fonction des vibrations sonores, tout en fournissant au bout du doigt un stimulus tactile figé. Le rajout d'images de la surface touchée influence-t-il beaucoup

la perception tactile? Les résultats de cette expérience psychophysique n'auraient pas d'intérêt pratique car il est toujours aisé et abordable de faire varier en intensité et en fréquence les vibrations tactiles d'une bobine.

À l'opposé des incohérences bénéfiques, il y a donc les incohérences perturbatrices, impliquant inconfort ou malaise pour l'utilisateur dont nous dressons une liste non exhaustive.

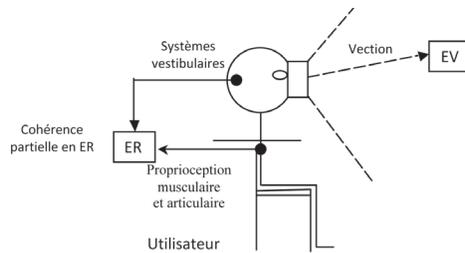
### Liste de 11 incohérences perturbatrices :

- PCV Observation :
  - *Incohérence visuo-motrice temporelle (Ls-m)* : un problème spécifique au visiocasque est la latence entre un mouvement de tête de l'observateur et l'affichage du bon point de vue correspondant au mouvement sur l'écran du visiocasque. Cette incohérence est négligeable, imperceptible si elle est inférieure à une milliseconde, ce qui n'est pas le cas dans les visiocasques actuels, sauf à exploiter un capteur basé sur une structure de liaisons mécaniques entre un support fixe de la pièce et le visiocasque (voir chapitre 4). Rappelons que pour avoir une vision efficace, les yeux doivent être stabilisés sur la zone observée. Le système vestibulaire permet le réflexe vestibulo-oculaire en charge de stabiliser le regard, comme expliqué au chapitre 2. Mais si dans une application RV, la latence entre le mouvement de la tête et l'effet de ce mouvement sur l'affichage à l'écran est importante, les yeux perçoivent un mouvement retardé, qui peut être décalé de quelques millisecondes avec ce qui perçu par le système vestibulaire. Difficile d'anticiper dans ces conditions. Cela cause une instabilité du regard [Stanney, K.M., 2002];
  - *Incohérence visuo-temporelle (Ds<sup>2</sup>)* : il y a incohérence perturbatrice si la fréquence d'affichage des images (FPS : Frames per second) est trop faible par rapport aux besoins du système visuel pour percevoir des images sans scintillement et avec des mouvements continus des objets en mouvement. Cela ne dépend pas de la rémanence des rétines (ancienne fausse théorie) mais de l'effet « phi » et du mouvement « bêta », mécanismes neurophysiologiques : lorsque le cerveau perçoit, par exemple, deux points lumineux alternativement affichés à une distance angulaire proche l'un de l'autre, le cerveau perçoit le premier point se déplacer à la position du second point, ce qui donne une impression de mouvement à partir d'informations ponctuelles. Plus la fréquence est élevée (FPS à 120, 240 Hz), plus les mouvements semblent fluides [Glon E., 2005];
  - *Incohérence oculomotrice (Ds<sup>2</sup>)* : en vision stéréoscopique, l'incohérence « accommodation-vergence » devient une gêne si les disparités rétinienne imposées aux observateurs dépassent une certaine limite qui est fonction du temps d'exposition, des capacités visuelles de l'observateur et d'autres critères. Certaines personnes plus sensibles à cette incohérence ne peuvent même pas fusionner et regarder ce type d'images;
  - *Incohérence visuo-spatiale (Ds<sup>2</sup>)* : une incohérence perturbatrice est créée si les champs de vision du visiocasque sont différents des champs

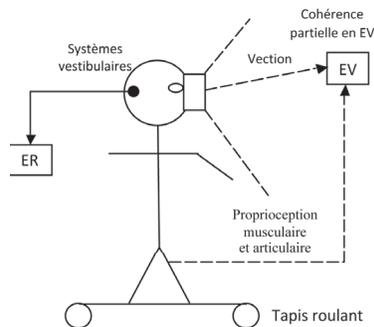
de vision de la caméra filmant l'EV. Certains concepteurs veulent utiliser ce subterfuge pour augmenter les champs de vision pour l'observateur, car la plupart des visiocasques ont des champs de vision trop réduits ;

- *Incohérence visuo-motrice de localisation (Ds-m)* : le mouvement de la tête en ER commande la rotation visuelle en EV qui peut être différente parce que le capteur est imprécis en rotation ou parce que le capteur ne mesure pas les translations. Dans ce dernier cas, il ne pourra pas détecter les petites translations de la tête qui peuvent exister même si l'observateur est assis ou debout en position relativement figée.
  - *Incohérence visuo-motrice spatiale (Ds-m)* : le concepteur de l'application RV peut souhaiter programmer *une observation visuelle non naturelle* :
    - soit une amplification de la rotation virtuelle par rapport à la rotation de la tête, en pratique pour permettre à l'observateur de voir sur un plus grand champ de vision sans trop tourner la tête ;
    - soit une amplification de la translation virtuelle par rapport à la translation de la tête, en pratique pour permettre à l'observateur de « zoomer » artificiellement sa vision ;
    - soit, mais encore plus perturbant, le point de vue affiché ne correspond plus du tout au point de vue de la tête de l'observateur, par exemple la « vue à la troisième personne » de l'EV (ou « vue objective »). L'observateur voit son avatar (sa propre représentation) dans l'EV. L'observateur peut se regarder ou voir ailleurs. Le point de vue peut être aussi celui d'un personnage virtuel, par exemple un personnage lui faisant face. Dans ce cas l'observateur se voit à partir du regard d'une autre personne. Ce type d'incohérence, dû à une vision non naturelle, est difficilement acceptable par la majorité de la population. Mais rien n'empêche en pratique certaines personnes de s'adapter à une telle vision en ayant de la motivation et en prenant le temps nécessaire. Rappelons qu'au milieu du vingtième siècle le professeur T. Erismann testa une vision bien plus complexe à traiter par le cerveau : la vision inversée grâce à une paire de lunettes retournant la vue. En une dizaine de jours, le porteur s'adapta parfaitement à cette nouvelle vision... et dû se réadapter ensuite à la vision normale en retirant la paire de lunettes ;
  - *Incohérence visuo-motrice passive (As-m)* : comme nous venons de l'indiquer, il est fortement recommandé au concepteur de l'application RV de ne jamais faire tourner l'EV dans le visiocasque sans que l'utilisateur, immobile dans l'ER, ne tourne la tête de lui-même. C'est une incohérence très perturbatrice qui ne correspond pas à un comportement réaliste. Mais ce type *d'observation visuelle non naturelle* peut être proposé toutefois avec parcimonie dans certaines applications spécifiques.
- PCV Navigation :
- *Incohérence visuo-vestibulaire (ou visuo-proprioceptive) (Ds<sup>2</sup>)* : c'est le cas classique du déplacement virtuel par vection sans déplacement réel

de l'observateur. Cette incohérence est bien connue et appelé « le mal du simulateur » (motion sickness en anglais). Plusieurs théories expliquent les malaises induits par cette incohérence, présentées en annexe à la fin de ce chapitre. L'incohérence visuo-vestibulaire perturbe l'utilisateur dès le dépassement de certaines limites cinématiques de déplacement (Figure 7.4). La commande de déplacement peut se faire manuellement, par la tête ou par les pieds, ce dernier cas exigeant un tapis 1D ou 2D. Ce cas est plus naturel et plus cohérent avec le déplacement de vection car il y a alors cohérence entre la vision et les organes proprioceptifs des muscles, des tendons et des articulations. Mais les systèmes vestibulaires sont mal stimulés puisque la personne reste immobile dans l'ER (Figure 7.5). En revanche, avec les interfaces de simulation de mouvement, tous les précédents organes proprioceptifs peuvent être correctement stimulés ainsi que les systèmes vestibulaires<sup>4</sup>. Cette incohérence peut être supprimée avec une interface à simulation de mouvement adéquate. Nous en parlerons au chapitre 8 et, dans ce cas, induire une autre incohérence, présentée au point suivant ;



**Figure 7.4 :** la personne étant assise, le déplacement en EV est dû à la vection mais les systèmes vestibulaires sont mal stimulés puisque la personne reste immobile dans l'ER, ainsi que les organes proprioceptifs des muscles, des tendons et des articulations.



**Figure 7.5 :** avec un déplacement par un tapis roulant, il y a cohérence en EV entre la vision et les organes proprioceptifs des muscles, des tendons et des articulations, mais les systèmes vestibulaires sont mal stimulés puisque la personne reste immobile dans l'ER

<sup>4</sup> Nous n'avons pas séparé les deux cas d'incohérences pour ne pas complexifier trop les explications. En théorie, il faudrait parler de l'incohérence « visuo-proprioceptive » et de l'incohérence « visuo-vestibulaire » (le cas où seul les systèmes vestibulaires sont incohérents avec la vision, cas de l'utilisateur se déplaçant via un tapis roulant).

- *Incohérence visuo-vestibulaire temporelle (Ls<sup>2</sup>)* : il est donc possible de stimuler correctement les systèmes vestibulaires par rapport au mouvement en EV en plaçant la personne sur une plate-forme mobile. Cette interface à simulation de mouvement peut être une plate-forme à 6 ddl ou un siège mobile de 2 à 4 ddl, selon les besoins de correspondance entre les mouvements réels et les mouvements virtuels créés par vection. Nous en reparlerons au chapitre suivant. Dans ce cas, le temps de latence entre les mouvements de la plate-forme mobile ou du siège et l’affichage du bon point de vue doit être minimal pour éviter une incohérence visuo-vestibulaire temporelle ;
  - *Incohérence visuo-posturale (Ds<sup>2</sup>)* : quand l’utilisateur est debout en ER, il doit maîtriser sa posture, la position verticale de son corps malgré les incohérences. Celles-ci peuvent survenir s’il y a un déplacement en EV par la vection tandis que le corps est immobile ou presque en ER, le cerveau étant informé de la position fixe du corps par les systèmes vestibulaires *mais aussi* par les autres stimuli proprioceptifs (fuseaux neuromusculaires, organes tendineux de Golgi et récepteurs articulaires) ;
- PCV Manipulation :
- *Incohérence visuo-manuelle (Ds-m)* : en manipulation, si la main réelle de l’utilisateur est mal localisée par rapport à la vision de la main dans le visiocasque (si la main est y représentée !) pour des causes techniques, on obtient une incohérence visuo-manuelle à laquelle l’utilisateur devra s’adapter. Mais la manipulation peut se faire d’une manière non naturelle, irréelle par « téléopération », voir plus loin.

Type d’incohérence Sens ou réponses motrices concernés	Absence	Découplage	Latence
Entre deux sens	<b>As<sup>2</sup></b>	<i>Inc. oculomotrice</i> <i>Inc. visuo-spatiale</i> <i>Inc. visuo-temporelle</i> <i>Inc. visuo-posturale</i> <i>Inc. visuo-vestibulaire</i>	<b>Ls<sup>2</sup></b> <i>I. visuo-vestibulaire</i> <i>temporelle</i>
Entre un sens et une réponse motrice	<b>As-m</b> <u><i>Inc. visuo-</i></u> <u><i>motrice passive</i></u>	<b>Ds-m</b> <i>Inc. visuo-motrice localisation</i> <i>Inc. visuo-motrice spatiale</i> <i>Inc. visuo-manuelle</i>	<b>Ls-m</b> <i>Inc. visuo-motrice</i> <i>temporelle</i>

**Tableau 7.2 :** liste de toutes les incohérences perturbatrices (celle très perturbatrice est soulignée)

**En résumé, la liste de 11 incohérences perturbatrices :**

- PCV Observation :
- *Incohérence visuo-motrice temporelle* (Ls-m)
- *Incohérence visuo-temporelle* (Ds<sup>2</sup>)
- *Incohérence oculomotrice* (Ds<sup>2</sup>)

- *Incohérence visuo-spatiale* ( $Ds^2$ )
- *Incohérence visuo-motrice de localisation* ( $Ds-m$ )
- *Incohérence visuo-motrice spatiale* ( $Ds-m$ )
- *Incohérence visuo-motrice passive* ( $As-m$ )
- PCV Navigation :
  - *Incohérence visuo-vestibulaire* ( $Ds^2$ )
  - *Incohérence visuo-vestibulaire temporelle* ( $Ls^2$ )
  - *Incohérence visuo-posturale* ( $Ds^2$ )
- PCV Manipulation :
  - *Incohérence visuo-manuelle* ( $Ds-m$ )

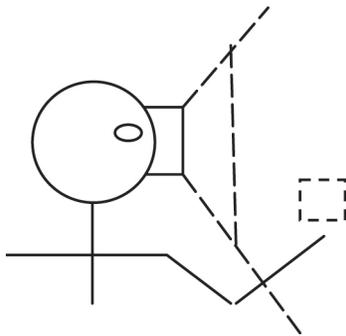
Comme pour les incohérences bénéfiques, les incohérences perturbatrices font toutes intervenir la vision, ce qui correspond à l'hypothèse H5.

Les cinq premières incohérences perturbatrices pour la PCV observation (visuelle), sont toutes dues à des **difficultés techniques** à réaliser le visiocasque « suffisamment parfait » : temps de latence suffisamment court, fréquence des images suffisamment élevée, écrans stéréoscopiques sans incohérence oculomotrice, champs de vision suffisamment larges et suivi de tête suffisamment précis. Dans les cas où **l'observation visuelle est non naturelle, irréaliste** (sixième et septième incohérences), il est évident que les perturbations sont bien plus fortes pour l'utilisateur.

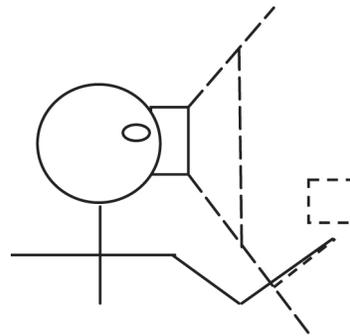
Pour les incohérences perturbatrices de la PCV navigation, il en est de même : les incohérences ne seraient pas perturbatrices si on pouvait aisément exploiter des interfaces de simulation de mouvement à prix modique et suffisamment parfaites techniquement. En revanche, les incohérences sont très perturbatrices que si le concepteur réalise une **navigation non naturelle, irréaliste**, c'est-à-dire que l'utilisateur se voit en train de se déplacer dans l'EV via son visiocasque (déplacement en vue de haut, vue à la troisième personne). Pendant le déplacement, l'utilisateur peut se regarder en train de se déplacer dans l'EV mais peut aussi regarder ailleurs. C'est le cas de l'observation avec Incohérence visuo-motrice spatiale, mais ici avec en plus le déplacement du point de vue dans l'EV.

Pour l'incohérence perturbatrice de la PCV manipulation, là aussi, l'incohérence visuo-manuelle n'est pas perturbatrice si techniquement les mains réelles de l'utilisateur sont parfaitement localisées au même endroit que les mains vues par l'utilisateur via le visiocasque, si bien sûr les mains sont représentées visuellement (autrement, on est dans le cas Figure 7.6a). La correspondance de localisation du corps réel, (ou d'une partie du corps) avec localisation visuelle par l'utilisateur de son corps s'appelle la **colocalisation**, notion importante en réalité virtuelle. La colocalisation est la superposition de l'espace visuel avec l'espace de manipulation pour l'utilisateur de l'application RV (voir Figure 7.6b). Techniquement, cette colocalisation est actuellement réalisable si on prend le soin de choisir de bons

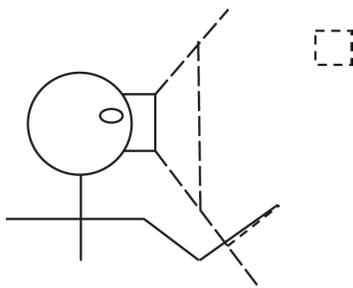
capteurs de localisation de la main et si la vision par visiocasque est bien configurée et calibrée. Mais si on réalise la manipulation d'un objet virtuel par « téléopération », c'est-à-dire en voyant l'objet en EV hors de portée des mains et en le manipulant à distance avec ses mains en ER, on obtient une **manipulation non naturelle, irréaliste** que permettent les techniques RV. Dans cette manipulation par téléopération, la représentation visuelle des mains dans le visiocasque peut se faire au niveau des mains dans l'ER, donc en colocalisation (voir Figure 7.6c), mais la représentation visuelle des mains (réaliste ou symbolique) peut se faire en EV aussi au niveau de l'objet manipulé (voir Figure 7.6d). Il faut tester le cas permettant une plus aisée adaptation à la manipulation par téléopération.



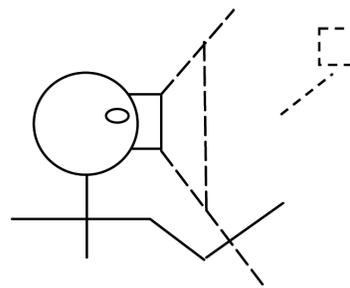
**Figure 7.6a :** manipulation sans voir les mains en EV  
(ER : traits continus et EV : traits pointillés)



**Figure 7.6b :** manipulation en voyant les mains en EV



**Figure 7.6c :** téléopération avec les mains colocalisées



**Figure 7.6d :** téléopération avec les mains en EV vues proches de l'objet

Nous pouvons donc établir la liste des incohérences sensorimotrices perturbatrices en fonction de leur cause :

- Difficultés techniques (Dtec) ;
- PCV non naturelle, Irréaliste, imposées par le concepteur de l'application (PCV Ir).

Les deux causes ne sont pas restrictives l'une par rapport à l'autre.

- *Incohérence visuo-motrice temporelle (Ls-m)* (Dtec)\*
- *Incohérence visuo-temporelle (Ds<sup>2</sup>)* (Dtec)\*
- *Incohérence oculomotrice (Ds<sup>2</sup>)* (Dtec)
- *Incohérence visuo-spatiale (Ds<sup>2</sup>)* (Dtec)
- *Incohérence visuo-motrice de localisation (Ds-m)* (Dtec)\*
- *Incohérence visuo-motrice spatiale (Ds-m)* (PCV Ir)
- *Incohérence visuo-motrice passive (As-m)* (PCV Ir)
- PCV Navigation :
  - *Incohérence visuo-vestibulaire (Ds<sup>2</sup>)* (Dtec) ou (PCV Ir)
  - *Incohérence visuo-vestibulaire temporelle (Ls<sup>2</sup>)* (Dtec)\*
  - *Incohérence visuo-posturale (Ds<sup>2</sup>)* (Dtec)\*
- PCV Manipulation :
  - *Incohérence visuo-manuelle (Ds-m)* (Dtech)\* ou (PCV Ir)

Dans cette liste, les incohérences dépendant des insuffisances techniques des visiocasques qui sont sur le point d'être supprimées grâce à l'évolution actuelle des produits sont notées Dtech\*. Cela concerne la réduction suffisante des temps de latence, l'augmentation suffisante de la fréquence d'affichage et des précisions de localisations de la tête et de la main. En revanche, il faudra attendre plusieurs années avant d'avoir les écrans stéréoscopiques sans incohérence accommodation-vergence et les visiocasques à champs de vision horizontal et vertical identiques à ceux de l'être humain.

#### Listes des incohérences neutres :

*Certaines incohérences sensorimotrices sont neutres, par rapport à notre problématique, dans le sens où elles ne créent pas d'inconfort ou de malaise, mais la simulation de l'activité sensorimotrice en EV est moins réaliste qu'en ER car il manque les stimuli de certains sens, qui ne sont pas compensés par la vision, comme pour les incohérences bénéfiques :*

- *Incohérence visuo-proprioceptive musculaire (As<sup>2</sup>)* : la manipulation d'un objet virtuel sans interface à retour d'effort sur la main et sans gant à retour d'effort fournit une simulation assez éloignée de la manipulation en ER. Cette dernière est plus complexe à gérer, comparable aux difficultés de l'enfant lors de son développement qui met du temps à apprendre la manipulation de pièces avec ses mains et ses doigts en rendant tous les stimuli sensorimoteurs cohérents. Il est nettement plus facile de manipuler en EV des objets sans poids et inertie. En revanche, l'absence de détection proprioceptive d'obstacle peut être un gêne importante pour l'activité de manipulation : par exemple, pour l'assemblage entre objets. Essayer de mettre un tiroir dans son logement en EV sans retour d'effort et sans ALC prévues par le concepteur RV !
- *Incohérence visuo-proprioceptive musculaire temporelle (Ls<sup>2</sup>)* : la manipulation d'un objet virtuel avec une interface à retour d'effort est délicate à réaliser

mécaniquement et informatiquement. Connaissant les capacités psychophysiques de la proprioception musculaire, il faudrait des retours à 1000 Hz avec un temps de latence inférieur à une milliseconde pour simuler un obstacle rigide. En pratique, les contacts avec les obstacles virtuels sont élastiques, même si ce n'est pas ce qui est désiré. Cette incohérence crée une simulation moins satisfaisante mais ne crée pas d'inconfort ou de malaise pour l'utilisateur ;

- *Incohérence visuo-tactile (As<sup>2</sup>)* : de façon identique, la manipulation sans gant à retour tactile fournit une simulation assez éloignée de la manipulation en ER. En revanche, les gants à retour tactile sont aisés à réaliser techniquement avec des temps de réponse très courts, permettant une latence très réduite et donc compatible avec une correcte simulation ;
- *Incohérence visuo-acoustique (As<sup>2</sup>)* : l'absence de son en EV n'est pas une gêne en général, même lors de manipulation d'objets rentrant en contact ;
- Etc.

N'oublions pas que des substitutions sensorielles peuvent aussi pallier à des difficultés techniques pour simuler certains sens dans certaines conditions : empêcher des inconforts et des malaises, éviter des solutions techniques difficiles à réaliser ou trop coûteuses, telles que les interfaces à retour d'effort (dans ce cas, les stimuli sonores peuvent se substituer aux stimuli de la proprioception musculaire pour alerter l'utilisateur d'une collision entre deux objets). Les métaphores avec substitution sensorielle peuvent efficacement servir d'outils pour supprimer certaines incohérences sensorimotrices.

Type d'incohérence Sens ou réponses motrices concernés	Absence	Découplage	Latence
Entre deux sens	<b>As<sup>2</sup></b> <i>I. visuo-proprio- musculaire</i> <i>Inc. visuo-tactile</i> <i>Inc. visuo- acoustique</i>	<u>5 incohérences bénéfiques</u> et <i>Inc. oculomotrice</i> <i>Inc. visuo-spatiale</i> <i>Inc. visuo-temporelle</i> <i>Inc. visuo-vestibulaire</i> <i>Inc. visuo-posturale</i>	<b>Ls<sup>2</sup></b> <i>I. visuo-vestibulaire temporelle</i> <b>I. visuo-proprio- musculaire temporelle</b>
Entre un sens et une réponse motrice	<b>As-m</b> <i>Inc. visuo-motrice passive</i>	<b>Ds-m</b> <i>Inc. visuo-motrice localisation</i> <i>Inc. visuo-motrice spatiale</i> <i>Inc. visuo-manuelle</i>	<b>Ls-m</b> <i>Inc. visuo-motrice temporelle</i>

**Tableau 7.3** : récapitulatif des 20 incohérences listées. En gras les incohérences neutres, soulignées les bénéfiques

Ce tableau n'est pas exhaustif mais représente toutefois les principales incohérences que l'on rencontre sur la plupart des applications RV. Les incohérences As<sup>2</sup>, absence d'un sens, sont toutes neutres, ce qui est normal car l'absence d'un sens rend la simulation, moins réaliste, et n'implique pas de conflit sensoriel pour le cerveau, donc d'inconfort et encore moins de malaise.

Nous expliquerons dans le chapitre suivant comment il est possible de pallier aux insuffisances techniques des visiocasques actuels et comment gérer les incohérences imposées par le concepteur programmant des actions sensorimotrices irréelles (PCV Ir). Celles-ci, dues à une vision non naturelle, sont plus difficilement acceptables. Rappelons que la finalité de la réalité virtuelle n'implique pas nécessairement une simulation réaliste. La simulation peut être irréaliste, imaginaire au niveau de l'application mais aussi aux niveaux des PCV ( $I^2$  fonctionnelles) et au niveau de l'immersion et de l'interaction sensorimotrices ( $I^2$ s-m). Mais dans ce dernier cas, il faut être prudent par rapport aux risques d'inconfort et de malaise.

### 2.3. Contrôlabilité des actions

Au niveau d'Immersion et d'Interaction cognitives, il faut que l'utilisateur puisse exploiter aisément des processus cognitifs par rapport aux  $I^2$  sensorimotrices. En particulier, l'utilisateur doit pouvoir contrôler son (inter)action dans l'EV. Ceci n'est possible que si les temps de latence sont suffisamment faibles pour permettre à l'utilisateur de correctement anticiper ses actions en fonction des stimuli sensoriels reçus, comme pour le maniement classique d'une souris. Quand un temps de latence est élevé, l'utilisateur devra s'imposer de ralentir ses gestes pour maîtriser ses actions. Cela modifie son comportement et il est moins naturellement immergé dans l'environnement virtuel. Si un temps de latence est trop élevé, cela peut générer aussi des instabilités dans le contrôle des actions de l'utilisateur et peut induire un inconfort, surtout si cela est couplé avec d'autres incohérences sensorimotrices. Il est actuellement difficile de déterminer l'influence entre eux de ces deux phénomènes sur l'inconfort et les malaises que peut ressentir l'utilisateur dans ces conditions. Pour une meilleure maîtrise des (inter)actions, il est souhaitable que les temps de latence soient constants, permettant de mieux anticiper les actions. Un temps d'apprentissage pourra aussi s'avérer nécessaire dans certains cas pour permettre à l'utilisateur de s'adapter aux temps de latence qui seraient trop élevés.

Un temps de latence trop élevé n'aura pas le même impact suivant la PCV concernée. Pour la manipulation d'un objet virtuel, celle-ci peut se faire de deux manières principalement :

- manipulation en voyant l'objet au niveau de sa(ses) main(s) avec la colocalisation entre la main réelle et la main virtuelle observée via le visiocasque (et évidemment aussi colocalisation de la tête réelle avec le bon point de vue dans le visiocasque) ;
- manipulation d'un objet virtuel réalisée par « téléopération » en voyant l'objet virtuel hors de portée des mains et en manipulant à distance (manipulation irréaliste que permet la RV) avec, toujours, colocalisation des mains et de la tête.

Dans le cas de la PCV manipulation en colocalisation main et tête (Figure 7.6b), cette colocalisation est fonctionnelle si les capteurs de tracking de la tête et de la main sont assez précis et si le calibrage de colocalisation est bien réalisé. Mais cela ne suffit : il faut que les temps de latence des deux capteurs soient suffisamment bas pour que l'utilisateur maîtrise son action manuelle. En général, les deux tracking sont réalisés avec la même

technologie (voir chapitre 5) et ont donc normalement des temps de latence sont de même ordre de grandeur. On peut faire l'hypothèse (mais qui reste à valider) qu'un temps de latence élevé pour le tracking de la main soit surtout gênant pour le contrôle de l'action manuelle (la coordination œil-main) et qu'un temps de latence élevé pour le tracking de la tête crée de l'inconfort et malaise pour l'utilisateur à cause de l'incohérence visuo-motrice temporelle. Si on est dans le cas de la PCV manipulation en téléopération (Figures 7.6c et 7.6d), la difficulté sera plus grande *a priori* pour le contrôle de l'action si on a des temps de latence trop élevés, que l'observateur voit ou non en colocalisation ses mains (ou un simple repère visuel). En téléopération, on sait que l'orientation et les mouvements de rotation de l'objet sont plus complexes à gérer cognitivement que lorsque l'on «manipule dans ses mains». Nous l'avons vérifié comme d'autres : un petit pourcentage de personnes peut être en difficulté pour orienter un objet manipulé grâce à un capteur à 6 ddl en ER, tout en regardant à distance l'objet dans l'EV.

Dans le cas de la PCV observation, l'analyse est plus simple que dans le cas précédent : il suffit d'avoir une bonne précision et surtout un temps de latence bas du capteur de tracking de la tête pour permettre de contrôler aisément le point de vue.

Pour la PCV navigation, il y a plusieurs latences qui interviennent : celle du tracking de l'interface de commande du déplacement (commande manuelle en général), celle du tracking de tête et éventuellement celle de la plate-forme de simulation de mouvement si on en exploite une. Ces latences doivent être basses et constantes pour permettre de bien anticiper et donc de bien contrôler sa navigation en EV. Nous en reparlerons au chapitre 8.

### 3. PROBLÈMES PSYCHOLOGIQUES INDUITS PAR LES ENVIRONNEMENTS VIRTUELS

Les visiocasques vont être exploités dans des activités professionnelles et de loisirs, en particulier dans les jeux vidéo. L'usage par des professionnels sera en général bien ciblé et ponctuel, ce qui ne devrait pas poser des problèmes psychologiques spécifiques. En revanche, un usage intensif est probable, mais pas encore sûr, pour des activités ludiques et de loisirs. S'il s'agit de s'immerger dans des photographies panoramiques ou des vidéo 360 degrés pour visiter un lieu, le temps d'usage devrait être relativement faible et ponctuel. Il y aura dans ce cas peu de risque d'addiction. En revanche, s'il s'agit d'employer un visiocasque lors d'un jeu vidéo, le joueur peut potentiellement y passer plus d'une heure quotidiennement. Des conséquences psychologiques peuvent survenir suite à une immersion trop prolongée dans un mode virtuel. Mais, comme nous l'avons indiqué, une question de fond mérite d'être posée : est-ce que le risque d'addiction aux jeux vidéo pour certaines personnes peut-être plus élevé si le joueur utilise un visiocasque à la place d'un simple écran ? Cette question, adressée aux psychologues et psychiatres n'a pas été encore débattue car cette pratique n'est pas encore répandue.

Devant un écran d'ordinateur ou de console de jeux, le joueur est-il différemment immergé au niveau mental que lorsqu'il regarde un EV via un visiocasque ? *A priori*,

le joueur sera dans la même situation qui pourra le pousser vers un manque de socialisation, une confusion entre le monde réel et le monde virtuel ou une addiction. Le fait de porter un visiocasque coupe la vue du joueur sur son environnement réel. Nous en verrons les conséquences concernant sa sécurité physique. Mais le joueur face à un écran, en général dans la pénombre, est tout aussi isolé cognitivement. Le degré d'immersion qu'induit les visiocasques de réalité virtuelle ne sont cependant pas les seuls responsables de l'addiction : la conception des scénarios des jeux vidéo créer une certaine forme d'addiction car le joueur cherche à avoir toujours plus d'expérience pour être meilleur.

L'impact des jeux vidéo sur les enfants peut intervenir sur l'imaginaire des enfants en bas âge. Les jeux vidéo et toute autre expérience en réalité virtuelle peut avoir un impact psychologique particulier. Est-ce que l'usage intensif d'un visiocasque peut avoir un impact spécifique sur la psychologie de l'enfant, différente de celle due à l'immersion visuelle via un écran ? (sous réserve que l'on permette cet usage à des enfants, ce qui n'est pas recommandé par les constructeurs de visiocasques). Nous ne débattons pas de cette importante question dans cet ouvrage.

### 3.1. Contraintes optiques et ergonomiques

Des problèmes ophtalmologiques peuvent survenir du fait que l'utilisateur va observer une scène virtuelle via un dispositif optique non spécialement adapté à ses caractéristiques visuelles. À l'opposé de la pratique des lunettes correctrices, les visiocasques vont être utilisés sans que des ophtalmologistes et des opticiens n'interviennent sur l'adaptation du visiocasque à chaque personne, tout en sachant qu'une partie de la population a des anomalies ophtalmologiques. Quelles contraintes optiques vont être imposées aux pratiquants de jeux vidéo avec des visiocasques aux qualités très variables ? Ceux-ci ont peu de réglages pour s'adapter au moins à la morphologie des usagers. Les quelques réglages optiques ainsi que les calibrations seront rarement contrôlés. Ils seront faits selon le bon vouloir des usagers et hors de contrôle de professionnels.

Des problèmes ergonomiques sont connus : le poids du visiocasque peut induire une augmentation des symptômes physiques lorsque le sujet porte un casque lourd trop longtemps. Cependant, l'équipement devenant de plus en plus sophistiqué, ce problème peut être évité par le port d'un casque plus léger. Le visiocasque doit s'adapter à toute morphologie de tête et donc sa fixation est plus ou moins adapté, avec des solutions techniques plus ou moins performantes, présentées au chapitre précédent. Il est amusant de voir sur des vidéos de démonstration, une personne retirant ses mains du visiocasque pour manipuler un objet virtuel et revenir prestement remettre ses mains sur le visiocasque qui n'a pas l'air de tenir de lui-même sur sa tête. Une autre contrainte est la chaleur que peut ressentir l'usager sur son visage. Un manque de ventilation du visiocasque peut aussi entraîner de la buée sur les optiques sur certains modèles de visiocasques, comme expliqué au chapitre précédent.

Les câbles qui relient l'ordinateur au visiocasque sont une gêne pour certains mouvements de la personne. Ils seront présents pendant un certain temps sur les visiocasques performants et même si les visiocasques avec smartphones n'en ont pas. Une solution pour une personne, debout et qui doit pivoter sur elle-même, est d'avoir l'ordinateur à la ceinture, une solution surtout utilisée avec les visiocasques dédiés à la Réalité Augmentée. Les concepteurs doivent donc prendre en compte les câbles et la gêne potentielle pour certains mouvements de la personne. Les câbles peuvent descendre du plafond de la pièce avant d'atteindre le visiocasque pour être moins contraignants.

Il y a plusieurs dispositions de la personne avec une application RV dont les principales sont :

- elle est assise sur un siège fixe ;
- elle est assise sur un siège pivotant sur 360° ;
- elle est assise sur une plate-forme à simulation de mouvement ;
- elle est debout et immobile ;
- elle est debout et se déplace dans la pièce ;
- elle marche sur un tapis roulant 1D ou 2D ;
- elle est allongée sur une balancelle ;
- etc.

Selon la disposition choisie les contraintes ergonomiques sont différentes.

Il y a aussi à étudier la communication avec la configuration de l'application RV, dont celle visiocasque. La communication est en général peu ergonomique car l'utilisateur n'a pas accès à un simple clavier et un écran. Il doit souvent faire des commandes avec des boutons, sur son casque ou sur ses manettes, sans les voir !

## **4. CONCLUSION**

Nous avons développé la problématique des incohérences sensorimotrices dans les applications RV. Dans le chapitre suivant, nous exposerons les solutions pour diminuer ou, mieux, supprimer les incohérences sensorimotrices. Cela se fera techniquement soit via le matériel employé, soit logiquement par des Aides Logicielles SensoriMotrices (ALSM), soit en changeant de Primitives Comportementales Virtuelles (PCV), présentées au chapitre 1. Notre classification a été faite en découplant chaque incohérence sensorimotrice pour plus de clarté mais elles peuvent intervenir entre elles. Dans les faits, la personne peut ressentir une incohérence globalement mais n'est pas en général consciente où se situe chaque incohérence sensorimotrice. Déjà que bien des personnes ignorent la présence et l'utilité de leurs systèmes vestibulaires...

## 5. ANNEXE : THÉORIES DU MAL DU SIMULATEUR ET DU MAL DU TRANSPORT

La première théorie repose sur l'incohérence sensorimotrice visuo-vestibulaire qui serait la source des malaises. C'est actuellement la plus acceptée pour expliquer l'apparition des symptômes causés par le déplacement. Cette théorie suppose que l'orientation de l'humain dans un espace 3D est basée sur un ensemble de stimuli sensoriels provenant :

- des systèmes vestibulaires, donnant des informations sur le mouvement de la tête et son orientation dans l'espace terrestre ;
- du système visuel donnant l'orientation corporelle par rapport à la scène visuelle ;
- de la proprioception articulaire donnant les positions relatives des membres.

Quand l'environnement artificiel (RV ou transport) modifie la cohérence de toutes ces informations, des symptômes peuvent apparaître [Harm, D.L., 2002]. Cette théorie est généralement satisfaisante mais elle n'explique pas certaines situations particulières, par exemple, lorsque l'utilisateur vit un conflit sensoriel sans aucun malaise ou comment il s'adapte à ce conflit.

Stoffregen et Riccio [Stoffregen T., Riccio G., 1991] ont proposé une explication différente qui repose sur le fait que les malaises dus au déplacement dans un environnement virtuel résultent d'une interruption des activités normales du contrôle postural. La personne reçoit des stimuli contradictoires qui perturbent sa manière de maîtriser sa position posturale.

La théorie sur l'intoxication, selon Treisman [Treisman M., 1977], explique que les causes provoquant les malaises dus aux incohérences sensorimotrices découlent d'une réaction ancestrale visant à retirer les toxines de l'estomac, comme lors d'un empoisonnement durant lequel le cerveau perçoit des stimuli sensoriels incohérents.

## 6. BIBLIOGRAPHIE

[Glon, E., 2005] E. Glon. *La persistance de l'erreur ou le mouvement filmique*, Paris IV Sorbonne/Institut Nicod, CNRS, ijn 00000647.

[Harm, D.L., 2002] Harm, D.L., *Motion sickness neurophysiology, physiological correlates, and treatment*. In K.M. Stanney (Eds.) *Handbook of virtual environments: Design, implementation, and applications* (pp. 637-661). Mahwah : IEA.

[Lécuyer, A. *et al.*, 2000] A. Lécuyer, S. Coquillart, A. Kheddar, P. Richard, and P. Coiffet. «*Pseudo-haptic feedback: Can isometric input devices simulate force feedback?*» In *Proceedings of the IEEE International Conference on Virtual Reality* (2000b).

- [Stanney, K.M., 2002] Stanney, K.M., Kennedy, R.S., & Kingdon, K. (2002). *Virtual environment usage protocols*. Dans K.M. Stanney (Éds.) *Handbook of virtual environments: Design, implementation, and applications* (pp. 721-730). Mahwah: IEA.
- [Stoffregen, T., Riccio, G., 1991] Stoffregen T., Riccio G., *An ecological theory of motion sickness and postural instability*. *Ecological Psychology*, 3(3), 195-240.
- [Treisman, M., 1977] Treisman, M., *Motion sickness: An evolutionary hypothesis*. *Science*, 197, 493-495.

---

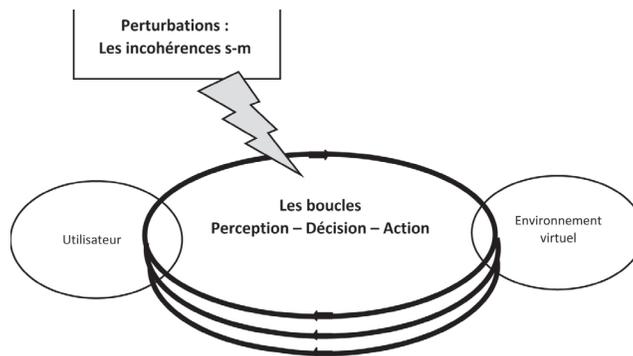
# Conseils pour l'usage d'un visiocasque

---

### 1. INTRODUCTION

Nous avons précisé dans le chapitre précédent qu'il existe des incohérences sensorimotrices perturbatrices dont la cause est due soit aux difficultés techniques à réaliser les bons stimuli sensoriels et moteurs, soit au choix du concepteur de l'application RV qui propose des Primitives Comportementales Virtuelles non naturelles, irréelles (PCV Ir). Nous expliquerons dans ce chapitre comment il est possible de pallier aux insuffisances techniques des visiocasques actuels et comment gérer les incohérences imposées par des PCV Ir. Celles-ci, dues à une vision non naturelle, sont plus difficilement acceptables. Mais rappelons que la finalité de la réalité virtuelle n'implique pas nécessairement une simulation réaliste. La simulation peut être irréaliste, imaginaire au niveau de l'application mais aussi aux niveaux des PCV (I<sup>2</sup> fonctionnelles) et au niveau de l'immersion et de l'interaction sensorimotrices (I<sup>2</sup>s-m). Dans ce dernier cas, il faut être prudent par rapport aux risques d'inconfort et de malaise dus aux PCV Ir.

Il existe encore peu d'études expérimentales sur l'usage de visiocasques par le grand public. Mais il y a eu des recherches sur certaines incohérences sensorimotrices dans les applications professionnelles RV et plus anciennement dans les simulateurs de transport. L'usage d'un visiocasque est peu répandu dans ces applications RV et dans les simulateurs de transport. Certaines incohérences se produisent d'une façon identique pour l'utilisateur, que ce soit devant un écran fixe ou dans un visiocasque. Sous la condition que l'extrapolation soit justifiée, les études expérimentales passées sont sources de solutions et de recommandations. Elles seront analysées dans le cadre de la boucle «Perception, Décision, Action», dont nous rappelons le schéma avec les perturbations créées par les incohérences sensorimotrices (Figure 8.1) :



**Figure 8.1 :** les boucles « Perception, Décision, Action » avec les perturbations par les incohérences sensorimotrices

Dans la suite de ce chapitre, nous allons analyser chaque incohérence sensorimotrice indépendamment pour simplifier l'analyse. Mais le couplage entre les sens peut intervenir dans l'étude des perturbations sur l'utilisateur, même si cela ne sera pas pris en compte. De même, deux PCV sont souvent réalisées en même temps car l'observation visuelle est permanente. L'utilisateur peut faire ainsi deux actions motrices simultanément, par exemple, marcher sur un tapis roulant et tourner la tête pour observer en même temps via son visiocasque. Mais nous étudierons les PCV séparément pour simplifier l'analyse.

Nous présentons les solutions et les recommandations à partir de la liste des incohérences sensorimotrices perturbatrices et pour chacune d'entre elles, il est possible de se poser **quatre questions** :

- Comment peut-on atténuer l'impact de l'incohérence s-m sur l'inconfort ou le malaise de l'utilisateur ?
- Peut-on supprimer l'incohérence s-m en modifiant le fonctionnement de la PCV ?
- Peut-on supprimer l'incohérence s-m en modifiant le fonctionnement de l'interface ou en rajoutant une autre interface ?
- Comment peut-on s'adapter à l'incohérence pour ne plus avoir d'inconfort ou de malaise ?

Les trois premières questions seront posées pour chaque cas d'incohérence. La question d'adaptation sera étudiée ensuite globalement.

Nous présentons de nouveau la liste des 11 incohérences sensorimotrices perturbatrices en rappelant leur catégorie ainsi que leur cause due à des Difficultés techniques (Dtec) ou à une Primitive Comportementale Virtuelle irréaliste, imposée par le concepteur de l'application (PCV Ir) :

- PCV Observation :
  - *Incohérence visuo-motrice temporelle (Ls-m)* (Dtec)
  - *Incohérence visuo-temporelle (Ds<sup>2</sup>)* (Dtec)
  - *Incohérence oculomotrice (Ds<sup>2</sup>)* (Dtec)
  - *Incohérence visuo-spatiale (Ds<sup>2</sup>)* (Dtec)
  - *Incohérence visuo-motrice de localisation (Ds-m)* (Dtec)
  - *Incohérence visuo-motrice spatiale (Ds-m)* (PCV Ir)
  - *Incohérence visuo-motrice passive (As-m)* (PCV Ir)
- PCV Navigation :
  - *Incohérence visuo-vestibulaire (Ds<sup>2</sup>)* (Dtec) ou (PCV Ir)
  - *Incohérence visuo-vestibulaire temporelle (Ls<sup>2</sup>)* (Dtec)
  - *Incohérence visuo-posturale (Ds<sup>2</sup>)* (Dtec)
- PCV Manipulation :
  - *Incohérence visuo-manuelle (Ds-m)* (Dtech) ou (PCV Ir)

*Note de lecture :* 32 solutions vont être présentées, dès plus classiques à d'autres très rarement exploitées ou non encore totalement validées. L'objectif des paragraphes suivants est de faire la liste la plus exhaustive possible. Une grille d'analyse sera ensuite proposée, permettant de faire le tri entre les solutions qu'il faut parfaitement connaître, maîtriser et celles d'usages spécifiques.

## 2. PCV OBSERVATION

### 2.1. Incohérence visuo-motrice temporelle

Il s'agit du problème spécifique de tout visiocasque : la latence entre le mouvement de tête de l'observateur et l'affichage correct du point de vue, correspondant au mouvement du visiocasque.

*On atténue l'incohérence (S1)*

Plusieurs démarches sont possibles pour diminuer l'impact de cette incohérence :

- on diminue *techniquement* le temps de latence, en améliorant le temps d'exécution des programmes, en changeant d'ordinateur, de cartes graphiques, de capteur de localisation ou de visiocasque ;
- à défaut de pouvoir diminuer suffisamment le temps de latence, on peut exploiter des algorithmes de prédiction des mouvements de la tête (filtre de Kalman, etc.). Mais lors d'un changement brusque du mouvement de la tête,

la prédiction est évidemment mise en défaut. À noter que certains fournisseurs de visiocasque, tel qu'Oculus, proposent des algorithmes de prédiction qu'il est recommandé d'exploiter.

Note de lecture : la solution S1 est détaillée après les solutions S2 et S3.

#### *On supprime l'incohérence en modifiant le fonctionnement de la PCV (S2)*

C'est une utilisation rare et détournée du visiocasque qui peut concerner des observateurs très mal à l'aise dans un visiocasque et exploitable pour des applications RV particulières : on n'affiche les images que lorsque la tête est immobile, sollicitant ainsi peu l'observateur. Quand il y a rotation de la tête, les images sont entièrement noires ou de couleur homogène, correspondant à la couleur moyenne de la scène virtuelle. À noter néanmoins qu'avec cette visualisation particulière de séquence d'images, le temps de latence intervient lors du passage de la phase immobile à la phase en mouvement et réciproquement. Mais ces transitions sont moins perturbatrices et peuvent se faire progressivement par un fondu enchaîné des images. L'immobilité de la tête est déterminée par le capteur dont les mesures sont filtrées (ce qui ajoute un retard), pour supprimer les mouvements parasites tant que l'utilisateur veut garder le même point de vue.

Exemple d'usage particulier : on veut observer un lieu, un bâtiment sous un certain nombre de points de vue différents sans avoir besoin d'une immersion dynamique du regard et sans se déplacer dans l'EV, dans l'objectif de peu solliciter l'observateur. Le déplacement en EV peut se faire par un moyen détourné : position donnée sur un plan, etc. Évidemment l'immersion visuelle ressentie est bien plus faible. Cela peut aussi être exploité que temporairement pour reposer le système visuel de l'observateur.

#### *On supprime l'incohérence en rajoutant / modifiant les interfaces (S3)*

Pour ne plus avoir de temps de latence, à la condition que la tête ne déplace pas ou presque pas en translation, on n'utilise plus un visiocasque mais on affiche les images de l'EV sur un ou des écrans fixes par rapport à l'ER. Pour garder la fonctionnalité d'immersion du regard, les écrans doivent entourer suffisamment la tête de l'utilisateur. C'est possible avec un écran parabolique ou avec trois écrans plats en forme de U, entourant la tête de l'observateur. Les images sont calculées indépendamment de la rotation de la tête. Les images seront correctes si la tête reste toujours immobile ou presque en translation, ce qui est souvent le cas quand la personne est assise face au dispositif d'affichage. Dans ce cas, on n'a plus besoin de mesurer la rotation de la tête et l'incohérence est supprimée. De plus, les images peuvent être de meilleure qualité car le nombre de pixels affichés peut être très grand. Ces solutions, comparables à celle d'un visiocasque de bonne qualité en terme de coût, peuvent donc être intéressantes pour certaines applications dans lesquelles la tête est relativement fixe en position, tout en ayant la liberté de tourner.

### 2.1.1. La solution S1 détaillée

En réalité virtuelle, le temps de latence (ou latence, en abrégé) est le temps qui s'écoule entre une action de l'utilisateur et la réponse perçue, conséquence de son action. Cette latence est la somme de plusieurs temps de réponse : celui de l'interface motrice (manettes de jeux, interface à retour d'effort...), celui du traitement informatique du signal de commande reçu, celui du temps de création informatique de la réponse à donner (le rendu des images, ...) et celui du transfert de la réponse à l'interface sensorielle (dans notre cas de base, le temps d'afficher la(les) bonne(s) image(s) dans le visiocasque).

La latence dépend donc au moins de deux matériels qui peuvent ou non être fournis par le même fabricant. Dans le cas du visiocasque, le capteur de tracking (suivi) de la tête est en général inclus. Le constructeur peut donc fournir comme donnée une latence, mais qui sera minimale et atteignable en pratique qu'en prenant la même puissance informatique utile et qu'en exécutant un programme très court. On rappelle que les fréquences de tracking de la tête et celles d'affichage des images dans le visiocasque doivent être très élevées. C'est une condition nécessaire mais pas suffisante. Par exemple, si on a une fréquence de 50 Hz pour des caméras infrarouge détectant la localisation de la tête, une mesure ne pourra être faite que tous les 20 ms. La latence sera automatiquement supérieure à 20 ms. Comme c'est un ordre de grandeur à atteindre, il faut donc que les suivis des mouvements se fassent à fréquences supérieures à 50 Hz.

Le réflexe vestibulo-oculaire<sup>1</sup> se réalisant en 20 ms environ<sup>2</sup>, une latence à 20 ms semble pour beaucoup de spécialistes RV un objectif à atteindre. Moins de 10 ms serait très bien. À plus de 50 ms, la latence est trop grande et le décalage entre le mouvement de la tête et le bon point de vue correspondant est trop grand. L'environnement virtuel ne paraît pas stable et cela peut entraîner une difficulté à contrôler le point de vue. Mais il n'y a pas encore de règles précises sur la valeur de la latence à obtenir. Certains travaux préconisent que la latence soit constante, donc prédictible et dans ce cas contrôlable après un temps d'adaptation<sup>3</sup>.

Comment diminuer la latence ? C'est un problème essentiel pour les fabricants des visiocasques et des capteurs (périphériques) associés. Pour les développeurs d'application, la question doit plutôt être : comment ne pas trop augmenter la latence ? Le programme doit être bien conçu et développé pour que le temps entre la lecture des données venant du système de suivi de la tête, éventuellement des mains, et le rendu des images soit non seulement le plus court mais aussi constant. La fréquence du cycle de rendu des images doit être du même ordre que la fréquence de l'écran du visiocasque, de l'ordre d'au moins 50 Hz. Il faut donc limiter la complexité de l'EV pour ne pas trop augmenter la latence ou paralléliser les actions.

---

1 On rappelle que le réflexe vestibulo-oculaire fait tourner l'œil en sens contraire du mouvement de la tête. Les systèmes vestibulaires, mesurant la rotation de la tête, commandent de produire un mouvement des yeux en sens contraire au mouvement de la tête pour stabiliser l'image rétinienne.

2 Berthoz A., "The Brain's Sense of Movement", Harvard University Press (2002).

3 Draper, M.H., Viire, E.S., Furness, T.A., Gawron, V.J. (2001). Effects of image scale and system time delay on simulator sickness with head-coupled virtual environments. *Human Factors*, 43(1), 129-146.

Il est souhaitable de vérifier la mesure du temps de latence, aussi bien pour les temps dus au matériel que celui dû au programme. L'entreprise Oculus fournit un dispositif, «Latency Tester», qui mesure un certain temps de latence avec un scénario donné. Le dispositif se place sur un oculaire du visiocasque pour détecter un changement d'image. La méthode suggérée se déroule ainsi : l'appui sur un bouton du testeur démarre le test et déclenche l'envoi d'une succession de messages interprétés par le système de rendu comme un ordre de changement de couleur à l'écran. Le temps écoulé entre l'envoi du message et le changement effectif de couleur de l'écran est mesuré pour chaque message, puis moyenné pour obtenir une mesure de latence. La longueur des messages et la vitesse de communication étant connues, la latence peut être calculée. Ce dispositif ne mesure pas le temps de réponse entre le mouvement réel de la tête et sa détection par le visiocasque.

Nous proposons une méthode de mesure indépendant des constructeurs et exploitable pour tout visiocasque, que nous avons déjà utilisé pour des applications RV sur écran ou Cave. Il faut avoir (ou emprunter) une caméra à très haute fréquence, de l'ordre de 1000 Hz, ce qui imposera une précision de 1 milliseconde pour la mesure de la latence globale. Pour des applications RV sur écran, la caméra devra filmer en même temps les images affichées et la main (ou la tête ou les pieds) de la personne, suivant la latence que l'on veut mesurer. Ceci est aisé à faire avec un écran mais demande plus de soin avec un visiocasque qui ne doit pas être porté. La caméra ne peut pas filmer les images dans le visiocasque mais peut détecter au moins un changement de couleur de la sortie d'un des deux oculaires. En incluant dans le programme un simple changement de couleur de la scène de l'EV lors d'une rotation du visiocasque, la caméra filmera le début du mouvement du visiocasque et ensuite l'instant du changement de couleur. La latence sera mesurée grâce au nombre d'images écoulées entre ces deux instants. On peut aussi vouloir mesurer lors d'un mouvement de main, la latence de la commande manuelle. Dans ce cas, la caméra doit filmer la main en plus de la sortie d'un oculaire. Cette méthode externe est un peu complexe à mettre en œuvre mais plus juste et indépendante du matériel et du logiciel.

## 2.2. Incohérence visuo-temporelle

Il y a incohérence perturbatrice si la fréquence d'affichage des images est trop faible par rapport aux besoins du système visuel pour percevoir des images sans scintillement et avec des mouvements continus des objets en mouvement.

### *On atténue l'incohérence (S4)*

Techniquement, il faut que les fabricants exploitent des écrans à fréquence la plus élevée possible, ce qui se réalise progressivement depuis quelques années et cette évolution doit être poursuivie (voir le chapitre 6). En attendant, il est possible de faire du traitement des images avec des filtres spatio-temporels ou d'éviter de créer des contenus produisant des scintillements : par exemple, des textures à hautes fréquences spatiales sur des objets en mouvement, des objets très fins (grille, maillage...), des textures à motifs répétitifs,

des alternances rapides de lumière, etc. S'il y a du texte à lire sur des objets, il ne faut pas hésiter à agrandir leur taille pour qu'ils soient lisibles. Des tests peuvent vérifier cette condition. Ces recommandations sont actuellement importantes car la définition des écrans des visiocasques est relativement faible. On perçoit les pixels dans de nombreux visiocasques actuels, ce qui est déjà désagréable. Si de plus, des scintillements sur des pixels sont observés, la qualité des images sera dégradée.

On peut prévoir un filtrage des mouvements parasites (tremblements) de la tête de l'observateur et du bruit du capteur de localisation de la tête : on supprime partiellement les scintillements des images en figeant leur point de vue tant que l'observateur ne change pas de direction de point de vue, grâce à un filtrage spatio-temporel, du type : si, à partir d'un petit laps de temps (plus ou moins une seconde), l'axe de direction et le centre de la tête ne varient pas d'un petit angle et d'un petit décalage en translation, le point de vue des images est figé. Il ne sera plus figé dès que la tête sortira de la zone neutre. Le côté négatif de ce filtrage est qu'il augmente partiellement le temps de latence et qu'il n'est pas efficace pour les objets en mouvement dans la scène de l'EV. Par ailleurs, la distance des objets texturés par rapport à la caméra est un paramètre qui peut influencer les scintillements des pixels. Elle peut être prise en compte aussi pour améliorer la qualité des images.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV (S5)*

Comme pour la solution S2, on ne fait plus intervenir en continu les images qui ne sont affichées que lorsque la tête est fixe. Le point de vue est figé tant que la tête ne bouge pas. S'il y a des objets en mouvement dans la scène observée, il y aura partiellement du scintillement. L'incohérence n'est donc que partiellement supprimée et, comme on l'a indiqué précédemment, cette solution est rarement exploitée.

*On supprime l'incohérence en rajoutant / modifiant les interfaces*

Requête inappropriée.

### 2.3. Incohérence oculomotrice

En vision stéréoscopique, l'incohérence « accommodation-vergence » devient une gêne si les disparités rétinienne imposées aux observateurs dépassent une certaine limite qui est fonction des capacités visuelles de l'observateur, du temps d'exposition et d'autres critères.

*On atténue l'incohérence (S6)*

Par la diminution des disparités rétinienne et, en plus éventuellement par des traitements d'images, il est possible d'aider l'observateur à fusionner les images

stéréoscopiques avec moins de fatigue oculaire. Si, de plus, les axes optiques des yeux sont suivis en temps réel dans le visiocasque, les traitements des images peuvent être plus performants car on sait à chaque instant où l'observateur regarde précisément dans l'EV. Mais actuellement, peu de visiocasques proposent le tracking des yeux.

Note de lecture : la solution S6 est détaillée après les solutions S7 et S8.

### *On supprime l'incohérence en modifiant le fonctionnement de la PCV (S7)*

Il suffit tout simplement d'afficher en vision monoscopique. Ceci est à proposer systématiquement pour les quelques pourcents de la population qui ne peuvent voir en vision stéréoscopique et ceux qui fusionnent les images stéréoscopiques avec un effort oculomoteur important et donc avec une fatigue oculaire. Cette incohérence est moins importante que face à des écrans stéréoscopiques classiques, car les indices monoculaires, particulièrement la parallaxe de mouvement (une fonctionnalité nouvelle procurée par les visiocasques), permettent de mieux percevoir en trois dimensions que la vision stéréoscopique seule (voir le chapitre 2). Bien sûr, c'est préférable d'associer ces deux indices de perception de la profondeur... si l'observateur voit sans problème en vision stéréoscopique. Le passage à une vision monoscopique n'est donc pas une contrainte très forte. L'utilisateur aura seulement quelques difficultés s'il doit manipuler finement des objets proches de sa tête. Dans ce cas, des ALC judicieusement choisies peuvent aider l'utilisateur à réaliser sa PCV de manipulation. La vision monoscopique pourra aussi être exploitée que temporairement, durant certaines situations sollicitant trop le système visuel de l'observateur (voir solution S6).

### *On supprime l'incohérence en rajoutant / modifiant les interfaces (S8)*

Comme nous l'avons expliqué, la suppression de cette incohérence se réalisera avec l'usage des écrans Light Field ou les écrans de technologie VRD «Virtual Retinal Display» (affichage rétinien virtuel), puisque par principe de ces types d'écrans, l'accommodation et la vergence se font à la même correcte distance pour chaque point de l'EV observé. Les premiers visiocasques utilisant ces technologies sont attendus pour 2016, mais avec des champs de vision relativement limités (voir chapitres 4 et 6).

#### 2.3.1. La solution S6 détaillée

Il y a plusieurs façons de diminuer l'incohérence «accommodation – vergence», la principale étant de limiter les disparités rétiniennes au-dessous d'un seuil. Une première méthode est de fixer le seuil à une valeur précise et définitive durant toute l'application RV (1,5° - 1,2° ou 1° par exemple). Mais comme on sait que la facilité de fusion dépend aussi du type d'images observées, du temps d'observation et des capacités visuelles de chaque observateur, d'autres méthodes plus fines peuvent être employées :

- par le contrôle du temps d'observation des disparités rétinienne et de leurs intensités ;
- par traitements du couple d'images stéréoscopiques ;
- par traitements du couple d'images stéréoscopiques en sachant où regarde l'observateur, exploitable avec les visiocasques ayant le tracking des yeux ;
- par adaptation aux capacités de chaque observateur, en sachant, comme nous l'avons déjà indiqué, que pour certaines personnes, le seuil est de zéro, ce qui implique l'obligation de proposer une vision monoscopique pour ces personnes. La vision monoscopique n'empêche pas de percevoir l'espace en trois dimensions, mais la perception en profondeur est moins performante. Il reste tous les indices monoculaires de perception de la profondeur, explicités au chapitre 2.

La méthode principale consiste à limiter cette incohérence sensorimotrice, donc à ne pas afficher d'images ayant des disparités rétinienne supérieures à un seuil de l'ordre de  $1^\circ$  à  $1,5^\circ$ . Comme nous l'avons expliqué au chapitre 3, les disparités seront créées par des parallaxes horizontales positives ou négatives (croisées ou non) sur l'écran réel ou sur l'écran virtuel dans le cas d'un visiocasque. Il faut donc savoir dans tout visiocasque où se situe par rapport aux yeux, l'écran virtuel, c'est-à-dire le plan d'accommodation ou plan de focalisation. C'est de l'ordre de 1 m ou 2 m selon les visiocasques. Pour le visiocasque DK2 Rift d'Oculus, la distance est de 1,3 mètre environ. Cette entreprise conseille que les objets doivent être positionnés entre 0,75 et 3,5 mètres, ce qui correspond évidemment à limiter les disparités rétinienne. Comme nous l'avons écrit au chapitre 2, la disparité rétinienne horizontale est égale approximativement à la parallaxe angulaire horizontale, ce qui permet d'établir des formules pratiques pour les limites à ne pas dépasser. Comme l'écran virtuel est à 1,3 m, un objet à 0,75 m impliquera une disparité rétinienne  $\delta$  :

En prenant pour distance interpupillaire (DIP) = 0,065 m, soit P la parallaxe horizontale pour les points homologues d'un objet à 0,75m, on a  $P = (1,3 - 0,75) \times 0,065 / 0,75$  et  $\text{tg}(\delta/2) = P / (1,3 \times 2)$ , ce qui donne  $\delta = 1,05^\circ$ . Avec un même calcul, on a pour un objet à 3,5 m une disparité rétinienne  $\delta' = 0,9^\circ$

Cette entreprise conseille donc de limiter les disparités rétinienne à  $1^\circ$  environ ce qui est une limite satisfaisante quand on s'adresse à un large public. Par comparaison, pour le cinéma en relief (3Ds), la disparité rétinienne pour les spectateurs des premiers rangs, les plus exposés à l'incohérence, est de l'ordre de moitié environ, car la durée d'exposition a aussi une grande influence sur l'observateur.

Ce principe d'une limitation de la zone d'observation, bornée à l'avant et à l'arrière de l'écran, doit être bien connu par tout concepteur exploitant la vision stéréoscopique dans des applications RV, avec ou sans visiocasque. Mais cette limitation n'est pas absolue non plus. Il est toujours possible temporairement de positionner des objets plus en avant ou en arrière de la zone à parallaxe nulle. Ils seront alors affichés avec des parallaxes horizontales très grandes et donc observés avec des disparités rétinienne très élevées. C'est ce qui se fait au cinéma 3Ds où, de temps en temps et pour une courte durée, des objets jaillissent de l'écran pour la grande satisfaction

des spectateurs, du moins pour la plupart d'entre eux. Pour le cinéma 3Ds, les stéréographes prévoient un « budget stéréo » pour toute la durée du film afin de ne pas trop solliciter les systèmes visuels des spectateurs. Les durées de chaque séquence avec chaque niveau de disparité sont comptabilisées et ne doivent pas dépasser une certaine limite. Il est souhaitable d'envisager une procédure similaire pour toute application RV dont la durée dépasserait un quart d'heure.

Il faut éviter que l'utilisateur observe trop longtemps un objet très proche avec une disparité supérieure à  $1,5^\circ$ . Ceci peut être le cas par exemple, quand l'utilisateur regarde à travers le viseur d'une arme qui se trouve très proche de sa tête. Il est souhaitable dans ce cas, de passer en vision monoscopique. L'observateur, concentré sur son viseur, ne verra pas trop l'aplatissement relatif de son EV dû à l'absence de vision stéréoscopique. Ce passage doit être progressif. Un autre exemple est le cas où l'on veut transmettre à l'utilisateur des informations qui sont affichées en avant de la scène. Si le plan d'affichage est très proche des yeux et qu'il est vu en relief, l'observateur va fatiguer. Si on recule ce plan d'affichage, il sera plus aisément vu mais il ne faut pas qu'il occulte des objets de la scène virtuelle qui seraient plus proches des yeux car il y aura alors conflit entre l'indice monoculaire d'occultation avec l'indice binoculaire de perception de la profondeur. Il n'y a pas de solution idéale, comme c'est le cas pour le sous-titrage des films stéréoscopiques qui sont en général mis à disparité nulle pour ne pas fatiguer le spectateur sans occulter des objets jaillissant de l'écran. Il peut être opportun d'afficher les informations de l'application RV sur un plan d'un objet si la scène s'y prête.

La vision monoscopique pourra donc être exploitée temporairement, durant certaines situations sollicitant trop le système visuel de l'observateur, dès que l'on sait que l'utilisateur va mettre un objet proche des yeux pendant un certain temps ou que le scénario de l'application impose la même condition. L'entreprise Oculus préconise cela quand les objets sont à moins de 0,5 m, ce qui est cohérent avec les caractéristiques de leur visiocasque. Il peut donc être toujours envisagé de passer en vision monoscopique ou du moins d'atténuer la vision stéréoscopique en diminuant artificiellement les disparités rétinienne.

Une autre considération doit être prise en compte si des objets sont en mouvement. Le système visuel de l'observateur devra plus rapidement fusionner le couple d'images stéréoscopiques. La limite de disparité rétinienne doit être plus petite si les objets sont en mouvement, comme expliqué sur le chapitre sur la vision humaine. Cela doit être pris en compte.

De même, les conditions sont plus restrictives si les fréquences spatiales des images sont élevées. La fatigue visuelle peut donc être diminuée par un contrôle de la parallaxe horizontale entre les images gauche et droite et par un contrôle du contenu en fréquences spatiales des images. La diminution de la fatigue visuelle est basée sur les fréquences spatiales et les disparités rétinienne, en exploitant une fonction de confort visuel dans l'image. Cette fonction permet de déterminer si un point de l'image fatigue l'observateur et si le point image doit faire l'objet d'un traitement. Avec des fréquences spatiales et des disparités rétinienne élevées en un point de

l'image, la fatigue oculaire est plus forte. La fonction de confort a été déterminée expérimentalement par divers recherches<sup>4</sup>. Sans exploiter cette fonction de confort, en première approximation, tout concepteur d'images stéréoscopiques peut au moins rendre flou les zones où les disparités rétinienne sont élevées.

Si on suit le mouvement des yeux de l'observateur, prévu sur certains futurs visiocasques, on peut encore mieux traiter les images stéréoscopiques pour améliorer le confort visuel : les images seront nettes dans la zone observée et floutées en vision périphérique, supprimant des fréquences spatiales élevées qui sont gênantes si les disparités rétinienne sont élevées.

Globalement, il faut noter que la vision stéréoscopique proposée est une vision artificielle qui a ses limites car on impose une distance fixe au plan d'accommodation, le plan de l'écran réel ou virtuel. C'est la grande différence avec la vision naturelle dans laquelle la distance d'accommodation varie constamment. Il serait intéressant de pouvoir changer la distance du plan d'accommodation, le plan de focalisation du visiocasque, en fonction des dimensions de la scène virtuelle dans une application donnée. Mais cela n'est pas prévu dans les visiocasques actuelles, car trop complexe à mettre en œuvre optiquement. L'avenir des visiocasques, comme on l'a souligné, est de supprimer techniquement l'incohérence accommodation – vergence avec les écrans Light Field ou à technologie VRD, (Virtual Retinal Display).

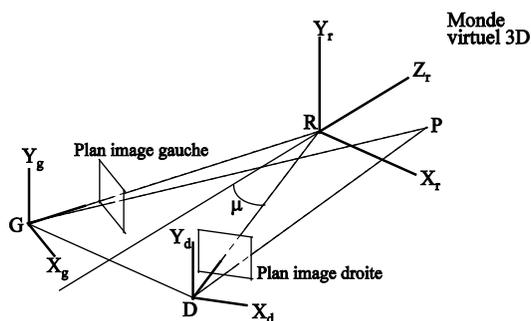
Le domaine de la vision stéréoscopique étant très ancien (la première recommandation date de plus d'un siècle et demi), il y a plusieurs relations bien établies qui sont formulées différemment selon le domaine concerné : la photographie en relief, le cinéma et la télévision en relief, 3Ds (et non « 3D »), la réalité virtuelle et la vision naturelle stéréoscopique, toutes se ramenant à limiter les disparités rétinienne horizontales. J'ai établi les relations entre elles, qui peuvent être consultées dans les pages 89 et suivantes du rapport<sup>5</sup> de la commission sur « les effets sanitaires potentiels des technologies audiovisuelles en 3D stéréoscopique » de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES). Ce rapport complet sur la vision stéréoscopique naturelle et artificielle est très instructif sur les effets sanitaires potentiels.

Sur quels critères doit-on se baser pour positionner les deux points de vue (les caméras virtuelles) générant le couple d'images stéréoscopiques ? *A priori*, on peut penser qu'il est préférable de faire converger les deux axes optiques vers la zone principale à observer pour se rapprocher de la vision humaine. Mais dans ce cas, les projections sur les deux plans donnent des images avec des parallaxes horizontales mais aussi avec des parallaxes verticales (le couple des images correspondent, à un facteur multiplicatif près, aux deux images affichées sur l'écran (Figure 8.2).

---

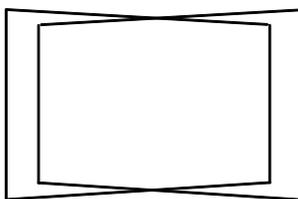
4 Consultable dans la thèse de J. Perrin « Profondeur et binocularité : algorithmie, étude psychophysique et intérêt pour l'ergonomie des interfaces stéréoscopiques ». Thèse de doctorat, Mines ParisTech, 1998.

5 Le rapport, téléchargeable gratuitement, se trouve à <https://www.anses.fr/fr/system/files/AP2011sa0334Ra.pdf>



**Figure 8.2 :** les deux axes optiques des deux caméras virtuelles convergent vers la zone principale à observer

Ce cas est à éviter car il impose des parallaxes verticales qui rendent difficile la fusion des images. Par exemple, les projections droite et gauche de la face avant d'un cube centré sur l'axe  $Z_r$  donnent les images suivantes sur la Figure 8.3 avec des parallaxes verticales au point P qui seront plus grandes quand le point P s'éloigne de l'axe  $Z_r$ .



**Figure 8.3 :** parallaxes verticales créées avec les deux axes optiques des deux caméras virtuelles qui convergent. Les parallaxes verticales sont exagérées sur la figure

Nous avons précisé au chapitre 2 qu'il est difficile de fusionner pour des parallaxes verticales d'angle supérieur à 20' d'arc. Il est donc nécessaire que les axes optiques des deux points de vue soient parallèles pour éviter les parallaxes verticales qui procurent des gênes à l'affichage. Les algorithmes permettant la création d'images stéréoscopiques de synthèse prennent presque toujours la condition de parallélisme<sup>6</sup> des deux axes des deux points de vue (des caméras virtuelles).

Comme nous l'avons déjà indiqué, il est souhaitable que l'écartement des axes optiques des caméras soit égal à la distance interpupillaire de l'observateur. Si cela est possible et souhaitable pour des scènes d'échelle « normale », ceci ne peut être envisagé dans les cas extrêmes :

- pour des vues aériennes stéréoscopiques, l'écartement est de plusieurs centaines de mètres pour rendre un effet de profondeur, qui n'est pas naturel mais utile. Dans cette condition, les objets, tels que des bâtiments sont perçus plus petits que naturellement. C'est l'effet « maquette » bien connu par les stéréographes.

<sup>6</sup> Si les yeux de l'observateur sont suivis en temps réel, il est possible dans ce cas de faire converger les axes des deux caméras vers le point observé, les parallaxes verticales n'étant alors créées qu'en vision périphérique comme dans la vision naturelle. Cette méthode est encore en état de recherche.

Ce cas peut être envisagé pour le pilotage de drones via un visiocasque pour une meilleure perception artificielle des distances, mais peut aussi créer plus de fatigue visuelle ;

- pour une vision d'objets minuscules (insectes par exemple), l'écartement est inférieur à la distance interpupillaire.

En général avec l'usage des visiocasques, les scènes sont observées à l'échelle 1 et donc les perceptions des dimensions avec les indices monoculaires et les indices binoculaires sont en correspondance, c'est-à-dire en vision orthostéréoscopique. La vision stéréoscopique, naturelle comme artificielle dans un visiocasque, est surtout utile en présence d'objets proches sur lesquels on veut interagir, ce qui est souvent le cas dans les applications RV ludiques. La vision hyperstéréoscopique et inversement la vision hypostéréoscopique sont aussi exploitables dans divers domaines, en particulier pour la vidéo 360°. Mais il faut que les disparités imposées soient bien contrôlées avec précision et retenir que le résultat fournit une perception déformée de la réalité.

Un cas particulier est celui de visiocasque en Réalité Augmentée à partir d'un couple de deux caméras stéréoscopiques filmant la scène réelle. Sur un projet de visiocasque RA, les caméras ne sont écartées que de 40 mm et non 65 mm pour des raisons techniques. Pour associer les images de synthèse avec les images réelles, il faut créer les images de synthèse avec une distance entre les deux caméras virtuelles de 40 mm et non de 65 mm.

Dans ces conditions, les yeux distants de 65 mm en moyenne, vont s'adapter mais l'environnement sera perçu plus grand qu'avec une vue réelle dans un rapport théorique de 1,6 (65/40) en moyenne. C'est en théorie seulement car la présence d'objets de tailles connues (les mains de l'utilisateur entre autres), permet au cerveau de compenser grâce aux indices monoculaires.

Concernant tous les types de visiocasques, il faut aussi noter qu'ils ne proposent la vision stéréoscopique que sur une zone centrale du champ de vision horizontal global. Dans certains visiocasques, cette zone centrale de vision binoculaire est relativement peu élevée et l'incohérence sensorimotrice a évidemment moins d'impact... ainsi que la perception de la profondeur et donc l'immersion visuelle. En revanche, pour d'autres visiocasque, le champ de vision total est entièrement en vision stéréoscopique, mais dans ce cas pour des visiocasque dont le champ de vision total n'excède pas 100 à 110° évidemment.

Pour terminer cet ensemble de règles à mettre en œuvre pour la vision stéréoscopique dans un visiocasque, il faut aussi parler de l'adaptation de chaque utilisateur dont leurs capacités sont très variables. Il faudra en tenir compte pour certains en supprimant totalement la vision stéréoscopique car ils ne s'y habitueront pas. Cela est bien compris dans le cinéma 3Ds où un même film est projeté en «3Ds» et en vision normale, monoscopique (ne pas dire «2D»!), bien qu'il y ait quelques années, certains voulaient nous faire croire que les films seraient dorénavant tous en vision stéréoscopique. L'avantage avec les visiocasques, c'est que chaque observateur peut

choisir individuellement ce qu'il préfère... si le concepteur a pensé à lui proposer le choix : vision stéréoscopique normale, visio stéréoscopique modérée ou vision monoscopique.

La distance interpupillaire de l'observateur étant grandement variable comme nous l'avons indiqué, il est indispensable que tout visiocasque permette de régler l'écartement des deux axes optiques pour qu'il s'adapte à la distance interpupillaire de l'observateur. Et évidemment, si on veut que l'observateur perçoive avec haute précision les formes et distances exactes, la valeur de la DIP de chaque observateur doit être incorporée dans le logiciel de création des images stéréoscopiques. Certains constructeurs de visiocasques proposent une méthode pour faire les bons réglages en fonction de la distance interpupillaire (DIP) de la personne, ce qui nous semble indispensable pour toute application dépassant 5 minutes.

Pour conclure sur l'adaptation de l'observateur à l'incohérence « accommodation – vergence », il faut retenir qu'il y a plusieurs paramètres sur lesquels on peut agir : durée et intensité de disparités rétinienne, traitement en fréquences spatiales des images pour les « adoucir », vision monoscopique éventuelle. Il peut être très opportun d'inclure dans l'application RV un test simple d'évaluation des capacités de vision stéréoscopique du porteur du visiocasque pour tenir compte des résultats en vue des réglages de sa vision stéréoscopique.

## 2.4. Incohérence visuo-spatiale

Une incohérence perturbatrice est créée si les champs de vision du visiocasque sont différents des champs de vision de la caméra filmant l'EV. Il est conseillé vivement de garder l'égalité des deux champs de vision. Il faut bien noter que c'est une contrainte que l'on a quand on utilise un visiocasque à la place d'un simple écran. Sur un simple écran monoscopique, il est possible de filmer la scène réelle avec un champ de vision différent du champ de vision d'observation, dépendant des dimensions de l'écran et de la distance d'observation. Le cerveau percevra cognitivement un monde virtuel à dimensions cohérentes. Si l'écran est stéréoscopique, les indices monoculaires et binoculaires doivent être cohérents (vision orthostéréoscopique) pour percevoir un monde cohérent en dimensions. Autrement, en vision hyperstéréoscopique ou hypostéréoscopique, les dimensions sont transformées comme avec l'effet « maquette » précédemment expliqué. En amplifiant le champ de vision artificiellement, le monde observé devient plus incohérent. Le réflexe vestibulo-oculaire (voir chapitre 2) est mis en défaut et l'utilisateur devra s'adapter, ce qui demande un effort important. L'utilisateur aura peut-être bien du mal à manipuler des objets virtuels car la colocalisation de sa main sera mise en défaut. Il faut donc bien réfléchir avant d'augmenter artificiellement le champ de vision trop étroit d'un visiocasque. La solution peut être contre-productive !

*On atténue l'incohérence*

Requête inappropriée.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV (S9)*

On réalise l'égalité des champs de vision, même si le champ de vision du visiocasque est très faible. Il n'y a plus cette incohérence mais la vision est peu immersive, surtout pour les visiocasques à très faible champ de vision où l'on a l'impression de voir la scène à travers un tube. La limitation du champ de vision est gênante pour une fonctionnalité visuelle exploitant la vision périphérique : quand un objet ou toute autre entité est en mouvement en vision périphérique, l'observateur est alerté par ce mouvement, même si ce dernier n'est pas vu nettement (répartition plus grande des bâtonnets en périphérie de la rétine). L'observateur tourne alors la tête pour mieux percevoir l'objet en mouvement. Si le champ de vision du visiocasque ne permet pas cette fonctionnalité d'alerte, il est souhaitable de proposer un subterfuge : en exploitant les écouteurs, on peut créer un son spatial (3D) sur le côté de la tête de l'observateur, à l'endroit de l'objet virtuel que l'observateur doit regarder. Une autre solution est d'attirer le regard sur le côté visuellement, par exemple en mettant un objet attirant le regard (clignotant, etc.) à moitié occulté en limite de champ de vision. L'observateur aura tendance à tourner en partie son regard sur l'objet et ainsi de suite si besoin.

*On supprime l'incohérence en rajoutant / modifiant les interfaces (S10)*

On exploite un *visiocasque intégral* comme le StarVR. Un visiocasque est dit « intégral » lorsque ses champs de vision monoscopiques et stéréoscopiques, horizontaux et verticaux, sont aussi grands que ceux du système visuel de l'être humain.

## 2.5. Incohérence visuo-motrice de localisation

Le mouvement de la tête en ER commande la rotation visuelle en EV qui peut être différente parce que le capteur est imprécis en rotation ou parce que le capteur ne mesure pas les translations.

*On atténue l'incohérence*

Requête inappropriée.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV*

Requête inappropriée.

*On supprime l'incohérence en rajoutant / modifiant les interfaces (S11)*

On utilise un capteur de localisation de la tête, suffisamment précis et mesurant les 6 ddl (voir le chapitre 4): les 3 rotations et les 3 translations.

### **3. PCV « OBSERVATION IRRÉELLE »**

#### **3.1. Incohérence visuo-motrice spatiale**

Le concepteur de l'application RV peut souhaiter programmer une observation visuelle non naturelle : amplification en rotation, en translation ou avec un point de vue non lié à celui de la tête (vision à la troisième personne).

*On atténue l'incohérence (S12)*

On peut limiter le gain d'amplification, pour la rotation comme pour la translation, pour ne pas trop s'éloigner de la vision réelle. Dans le cas de la rotation, si l'amplification est conçue pour permettre à l'observateur de percevoir sur un plus grand champ de vision sans trop tourner la tête (par exemple, si les câbles gênent), une autre solution est de faire une commande avec hystérésis ou une commande non-linéaire :

##### **3.1.1. Commande avec hystérésis**

Quand l'observateur tourne un peu la tête, la relation est linéaire entre l'angle de la tête et l'angle du point de vue (fonctionnement normal du visiocasque). Mais quand l'observateur tourne la tête d'un certain angle  $\alpha$  ( $45^\circ - 60^\circ$ ), l'angle du point de vue tourne avec amplification. Lorsque l'observateur tournera ensuite en sens inverse, il y aura de nouveau la relation linéaire normale, tant que l'angle de rotation de la tête n'a pas atteint un angle  $-\alpha$ . Cela permet d'avoir une vision normale sur une certaine zone angulaire étroite et si on veut observer une autre zone dans une direction éloignée, le changement de zone se fait par cette commande particulière sans tourner exagérément la tête et son corps. Ainsi l'observateur change d'orientation visuelle en EV sans changer l'orientation de son corps en ER (pratique si l'utilisateur est assis et veut par exemple regarder derrière lui – voir aussi la commande suivante). Il faut évidemment un temps d'adaptation à cette commande visuelle non naturelle.

##### **3.1.2. Commande non-linéaire**

Même principe mais dès que l'angle  $\alpha$  est dépassé, la rotation de la tête commande la vitesse de rotation de l'angle de rotation du point de vue. Le changement d'orientation visuelle en EV par rapport l'orientation du corps en ER est plus rapide.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV (S13)*

On peut atténuer les perturbations de la vue à la troisième personne en affichant cette vue sur un écran virtuel, inclus dans le champ de vision du visiocasque, qui est fixe par rapport au visiocasque, donc de la tête. Le changement de point de vue est effectué *manuellement* et non par la tête, comme si l'observateur agissait simplement devant un écran plat, sauf que cet écran virtuel est lié à la tête de l'observateur. Ce type de vision à la troisième personne sur un écran virtuel peut être exploitable si elle est incluse dans une fenêtre de la vision normale du visiocasque, permettant à l'observateur de mieux visualiser la scène, mais avec plus d'effort de perception visuelle car ce n'est pas naturel. Par exemple, vue à la troisième personne d'un lieu par au-dessus, commandée manuellement, tout en voyant en même temps le lieu en vision directe, commandée par la tête. C'est une vision à deux points de vue simultanés, très artificielle et irréaliste, mais que permet un visiocasque, surtout s'il a un grand champ de vision et une bonne résolution. Cette vision est aussi perturbatrice mais différente de la vision commandée par la tête du point de vue à la troisième personne car, dans ce dernier cas, il y a plus de risques de ne pas comprendre la commande et de provoquer une désorientation.

*On supprime l'incohérence en rajoutant / modifiant les interfaces*

Requête inappropriée.

### 3.2. Incohérence visuo-motrice passive

L'EV est en mouvement visuellement (vection) **sans** que l'utilisateur, immobile dans l'ER, ne bouge la tête de lui-même. C'est une incohérence très perturbatrice qui ne correspond pas à un comportement naturel et qui doit être utilisée rarement et avec précaution.

*On atténue l'incohérence (S14)*

Le mouvement de l'EV, non commandé par l'utilisateur, doit être faible en vitesse et en accélération, aussi bien pour des mouvements de translation que de rotation. Les recommandations sur les vitesses, les accélérations et les types de trajectoires seront explicitées en détail plus loin dans le paragraphe sur l'incohérence visuo-vestibulaire quand le déplacement est commandé par l'utilisateur. Comme nous l'avons déjà indiqué, si la personne est passive et non active, comme un passager dans un véhicule, les limitations sur les vitesses, les accélérations doivent être fortement réduites et ce type d'incohérence est strictement réservé à des utilisateurs adaptés à de telles sollicitations. Pour un usage grand public non contrôlé, cette incohérence est à proscrire.

Une autre solution consiste à changer le point de vue de façon discontinue, par palier et en supprimant totalement ou partiellement la vection. Le passage d'une position à une autre du point de vue peut aussi se faire en fondu enchaîné.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV*

Requête inappropriée.

*On supprime l'incohérence en rajoutant / modifiant les interfaces*

Requête inappropriée.

## 4. PCV NAVIGATION

Pour la simplification des explications, on ne sépare pas dans les paragraphes suivants, comme pour la PCV Observation, les incohérences dues aux défauts techniques des visiocasques de celles dues à une PCV navigation irréaliste.

### 4.1. Incohérence visuo-vestibulaire (ou visuo-proprioceptive)

C'est le cas classique du déplacement virtuel par vection sans déplacement réel de l'observateur. Cette incohérence perturbe l'utilisateur dès le dépassement de certaines limites cinématiques de déplacement. La commande de déplacement peut se faire manuellement, par la tête ou par les pieds, ce dernier cas exigeant un tapis 1D ou un tapis 2D<sup>7</sup>. Comme nous l'avons indiqué au chapitre précédent, ce dernier cas est plus naturel et plus cohérent avec le déplacement virtuel par vection car il y a alors cohérence entre la vision et les organes proprioceptifs des muscles, des tendons et des articulations. Mais les systèmes vestibulaires sont mal stimulés puisque la personne reste immobile dans l'ER.

*On atténue l'incohérence (S15, S16, S17, S18, S19 et S20)*

Il y a six solutions différentes, les deux premières étant complémentaires et, de même les quatre dernières entre elles :

- pour ne pas trop solliciter les systèmes vestibulaires, il faut limiter les accélérations en translation et en rotation ainsi que les inclinaisons des mouvements de la caméra virtuelle (le point de vue de la personne en EV) filmant la scène virtuelle (**S15**);

<sup>7</sup> Il existe aussi d'autres dispositifs de commande aux pieds.

- pour ne pas trop solliciter les systèmes vestibulaires, il faut concevoir un EV tel que les trajectoires de la caméra virtuelle ne soient pas trop tortueuses, donc soient à rayons de courbure assez élevés, ce qui limitera les accélérations centripètes, normales à la trajectoire (**S16**);
- la perception du mouvement est plus sensible à la périphérie du champ visuel, détectant les flux optiques dus aux mouvements de vection et aux mouvements des objets dans la scène. Il peut être envisagé de diminuer le champ de vision observé par une occultation des images en vision périphérique, même si la plupart des visiocasques ont actuellement un champ de vision assez restreint. Cette solution est au détriment de l'immersion visuelle (**S17**);
- comme nous l'avons indiqué, si l'utilisateur regarderait l'EV via un simple écran plat et non avec un visiocasque, sa vision périphérique observant l'ER, même inconsciemment, le stabiliserait et lui éviterait de l'inconfort et des malaises. Pour se rapprocher de cette situation, il est envisageable d'atténuer l'incohérence en injectant dans les images en vision périphérique, quelques références spatiales de l'ER pour stabiliser l'utilisateur. Cela diminue les flux optiques en vision périphérique. Il faut bien sûr que le visiocasque ait un champ de vision assez large, au moins 100° pour ne pas perturber la vision centrale. Ces inclusions artificielles et non réalistes peuvent déranger la perception visuelle de l'observateur mais *a priori* pas trop car elles se situent en vision périphérique (**S18**);
- dans la suite de la solution précédente, il est possible d'atténuer l'incohérence en ayant des objets dans l'EV qui soient immobiles par rapport à l'ER, quand l'utilisateur est immobile en ER. L'exemple classique est celui du simulateur statique de conduite : le conducteur est bien stabilisé s'il a en vision périphérique la vue de l'habitacle de la voiture, celui-ci étant immobile en ER (**S19**);
- dans le prolongement de la solution précédente, il peut être intéressant d'utiliser un visiocasque à visière n'occultant pas totalement la vision, les « vidéolunettes », permettant à l'observateur de percevoir directement le monde réel en vision périphérique. Dans ces conditions, les effets perturbateurs de l'incohérence visuo-vestibulaire sont grandement diminués, un peu comme lorsque l'on regarde un simple écran (**S20**).

Les solutions S15, S16, S17 et S18 sont maintenant détaillées.

#### 4.1.1. La solution S15 détaillée

Il est bien évident que pour ne pas solliciter les systèmes vestibulaires, il faut aucune accélération en translation et en rotation ainsi qu'aucune inclinaison de la tête de la personne, vue leur constitution anatomique et leur fonctionnalité. Inversement, si on veut solliciter les systèmes vestibulaires, il faut agir sur les accélérations et les inclinaisons de la tête. Cela impose en simulation de transport comme dans les applications RV d'agir sur le corps de la personne avec des interfaces à simulation de mouvement, solution expliquée plus loin (S24). Cette solution demande un investissement en matériel plus ou moins important. Dans le cas classique sans rajout d'interface, le choix se porte sur la limitation des accélérations et des inclinaisons de la tête pour limiter l'incohérence visuo-vestibulaire. Ces limites ne peuvent pas être déterminées en absolu car elles dépendent de plusieurs paramètres : la durée des

sollicitations incohérentes et les capacités de chaque personne, comme pour les limites des disparités rétinienne pour la vision stéréoscopique.

Les règles pour limiter les perturbations de l'incohérence visuo-vestibulaire ont déjà été bien étudiées dans les simulateurs de transport statique (sans interface à simulation de mouvement). L'hypothèse est que ces règles peuvent aussi s'appliquer au cas des applications RV avec visiocasque, si tous les paramètres de base sont identiques : même champ de vision, même type de trajectoire de déplacement, ce qui limite la correspondance entre les deux situations. La plupart du temps l'affichage est différent, les simulateurs exploitant en général de grands écrans. Mais les spécialistes en simulation de transport et en réalité virtuelle du constructeur automobile Renault ont aussi exploité des visiocasques en simulation de transport<sup>8</sup>, (voir le chapitre 10 présentant leur expérience).

Il est recommandé de régler les vitesses de déplacement aux valeurs habituelles de la marche (3 à 5 km/h) ou, si besoin, de la course (10 à 12 km/h) suivant le contexte. On peut faire l'hypothèse que les vitesses, influençant le réflexe visuo-oculomoteur<sup>9</sup>, détectées surtout en vision périphérique (donc si le visiocasque a des champs de vision larges) interviennent sur le mal du simulateur même si ce sont les accélérations qui sont primordiales. Ce n'est donc pas l'intensité des vitesses qui joue un rôle important dans le mal du simulateur (simulation sickness), mais les intensités d'accélération et même les variations d'accélération : vecteur d'à-coup (ou *jerk* en anglais), qui correspond à des brusques secousses. La durée des accélérations a une influence : il est préférable que la durée d'accélération soit courte, ce qui est compréhensible, l'incohérence durant moins longtemps. Il est donc nécessaire d'adapter autant que possible la programmation afin que les durées d'accélération soient courtes et que leur fréquence d'apparition soit faible aussi. Si c'est l'utilisateur qui contrôle librement et totalement ses déplacements, il n'est pas possible de donner des contraintes.

En revanche, en contrôlant partiellement les mouvements de l'utilisateur, il est possible de limiter les signaux de commandes de l'utilisateur pour son déplacement, de telle sorte qu'il ne puisse pas générer des accélérations trop fortes. Par exemple, la relation entre le signal de commande et la valeur d'accélération, au lieu d'être linéaire, est non linéaire et limitée asymptotiquement. On peut aussi imposer la forme des courbures de la trajectoire commandée par l'utilisateur en imposant des rayons de courbure suffisamment grands pour limiter les grandes accélérations (voir la solution S16) par une commande appropriée, comme précédemment.

Par ailleurs, il est de toute façon préférable que les accélérations soient imposées à l'utilisateur car celui-ci peut anticiper sa réaction et donc être moins sujet au mal du simulateur. Les valeurs acceptables d'accélération sont de l'ordre de 2 m/s<sup>2</sup> pour un

---

8 Simulation sickness comparison between a limited field of view virtual reality head mounted display and a medium range field of view static ecological driving simulator (ECO2), B. Aykent, Z. Yang, F. Merienne, A. Kemeny, Driving Simulation Conference Europe 2014 Proceedings, Sep 2014, Paris, France. Society for Modeling & Simulation International, pp.65-71, 2014.

9 On rappelle que le réflexe vestibulo-oculaire fait tourner l'œil en sens contraire du mouvement de la tête. Les systèmes vestibulaires, mesurant la rotation de la tête, commandent de produire un mouvement des yeux en sens contraire au mouvement de la tête pour stabiliser l'image rétinienne.

mouvement de translation durant quelques secondes et de  $2\text{ }^\circ/\text{s}^2$  pour un mouvement de rotation<sup>10</sup>. Mais une accélération pour un mouvement de rotation de  $13\text{ }^\circ/\text{s}^2$  pour des utilisateurs novices est perturbante et guère acceptable. Il est aussi recommandé de limiter les variations rapides d'accélération, (brusques secousses ou à-coup), difficilement supportable par beaucoup de personnes. Par exemple, il faut éviter de vouloir rendre très réaliste la descente rapide d'un escalier par un mouvement de vection avec des à-coups à chaque marche. Il est préférable que la trajectoire soit purement rectiligne à vitesse constante.

Nous avons précisé que le cerveau anticipe les futurs stimuli sensoriels qu'il reçoit. Si l'utilisateur commande ses mouvements, le cerveau peut anticiper mais s'il est passif vis-à-vis de son déplacement, de sa trajectoire, il est souhaitable de donner des indices sensoriels anticipant tout changement d'accélération. Dans un véhicule réel, pour éviter le mal de transport, il est souhaitable de rajouter un bruit aux changements de vitesse d'un véhicule à boîte automatique pour permettre aux passagers, ne voyant pas d'action gestuelle du conducteur, d'anticiper les accélérations du véhicule. Il peut être intéressant de procéder de même en réalité virtuelle, en aidant l'utilisateur à anticiper son déplacement.

Lors d'un déplacement, l'utilisateur peut en même temps observer l'EV et donc tourner la tête. Le couplage des deux PCV, navigation et observation, va plus solliciter l'utilisateur. Si c'est l'utilisateur qui commande son déplacement, il est recommandé de lui permettre de commander manuellement son déplacement tandis qu'il peut observer dans toutes les directions par commande de la tête. C'est une navigation et observation en « mode tank » qui est intéressant à exploiter<sup>11</sup>. Mais nous sommes habitués à regarder « devant nous » ou légèrement de côté. Il est donc souhaitable de limiter l'angle entre la direction de déplacement et l'angle du point de vue de l'EV. Suivant les aptitudes de l'utilisateur, cet angle peut être plus ou moins limité par programmation ou en incitant l'utilisateur à regarder toujours devant lui, induit par le scénario du jeu vidéo ou par l'aspect visuel ou sonore de l'EV. Si l'utilisateur est assis sur une chaise, il n'aura pas tendance à tourner la tête en dehors de la zone  $-45^\circ$  à  $+45^\circ$ . On peut en tenir compte pour programmer les déplacements (par vection) de telle façon qu'ils restent avec des directions ayant de faibles angles avec l'axe du point de vue de l'utilisateur (et donc pas de « marche en arrière virtuelle / réelle », c'est-à-dire que la direction du déplacement virtuel soit dans la direction arrière du siège réel. Si besoin, il faut prévoir une réinitialisation si l'angle entre la direction de déplacement et l'angle du point de vue de l'EV est trop important. Une autre méthode de déplacement est celle avec un angle toujours nul : la direction du déplacement virtuel est donnée par la direction de la tête de l'utilisateur, ce qui est moins perturbateur car cette navigation se rapproche de la marche réelle. Par ailleurs, le déplacement peut être totalement libre ou inversement être dépendant d'une trajectoire déterminée par le concepteur de l'application RV. Dans ce cas, la trajectoire aura une influence sur les

10 "How to Avoid Simulation Sickness in Virtual Environments during User Displacement" A. Kemeny, F. Colombet, T. Denoual, *Proc. SPIE* 9392, The Engineering Reality of Virtual Reality 2015, 939206 (March 30, 2015)

11 [https://developer.oculus.com/documentation/intro-vr/latest/concepts/bp\\_app\\_ui\\_nav/](https://developer.oculus.com/documentation/intro-vr/latest/concepts/bp_app_ui_nav/)

accélérations subies par la personne, même si c'est elle qui commande ou non la vitesse de déplacement sur la trajectoire prédéfinie. Cette problématique est présentée au paragraphe suivant.

#### 4.1.2. La solution S16 détaillée

On est dans la situation où la trajectoire est déterminée par le concepteur de l'application RV, avec la personne qui commande ou non sa vitesse de déplacement. Toute trajectoire prédéfinie a une influence sur les accélérations subies par la personne dès que la trajectoire n'est plus rectiligne. On rappelle que dans un mouvement circulaire, l'accélération centripète, normale à la trajectoire, est inversement proportionnelle au rayon de courbure et proportionnelle au carré de la vitesse tangentielle. Il faut donc limiter la vitesse dans une trajectoire courbe. Lorsque que l'on prend un virage en voiture, plus le virage est serré, de rayon de courbure petit, plus l'accélération subie est importante. Si la trajectoire n'est pas plane, ni sur une surface (en 2D) mais dans l'espace (en 3D), le problème de limitation des accélérations est plus délicat à gérer. Dès qu'il y a un virage, il y a une accélération en rotation. Pour atténuer les accélérations en rotation, il faut des trajectoires dont les rayons de courbures soient relativement grands et pour éviter les à-coups, les variations brusques d'accélération, il est souhaitable d'avoir un rayon de courbure qui s'établisse progressivement. Cette problématique de la forme des trajectoires sans secousse date de plus d'un siècle, depuis les premières constructions de voies ferrées et ensuite pour les routes et autoroutes : des arcs de clothoïdes<sup>12</sup>. Il faut donc que le concepteur de l'application trace les trajectoires imposées à l'utilisateur de façon identique, qu'il s'agisse de trajectoires 2D ou 3D. Cela a été étudié partiellement pour les routes avec des virages qui ont un certain devers pour le confort des passagers des véhicules. Une relation a été établie entre l'inclinaison du devers, le rayon de courbure et la vitesse pour maximiser le confort. Il faut avoir le même souci en réalité virtuelle si on veut avoir des trajectoires « douces », c'est-à-dire « secouant visuellement » peu la personne.

Si les trajectoires à emprunter dépendent de la géométrie de l'EV, il faut étudier cette géométrie pour éviter le plus possible les rotations et plus généralement les accélérations intempestives. Il faut faire attention à la disposition des obstacles dans l'EV ! Par exemple, dans un magasin virtuel, vu en immersion via un grand écran très proche de l'observateur (ce serait le cas avec un visiocasque), l'implémentation des allées et des étagères a été réalisée de telle sorte que la personne se déplace sans presque aucune rotation. Avec un visiocasque, cela se complique si, en même temps que la personne se déplace, elle tourne la tête, comme on l'a déjà signalé. La géométrie de l'EV qui impose des trajectoires spécifiques et prédéterminées doit donc être bien étudiée pour atténuer l'incohérence visuo-vestibulaire. Les trajectoires

---

12 L'ingénieur Talbot en 1890 a établi les équations de la forme que doivent avoir les rails d'une ligne de chemin de fer au début d'une courbure de la voie pour éliminer les à-coups d'accélération : des arcs de clothoïdes, dont le rayon de courbure varie progressivement. Pour la même raison, on tourne le volant d'une voiture progressivement à l'amorce d'un virage.

peuvent être bornées partiellement mais pas totalement contraintes géométriquement, comme lorsque l'on circule dans un couloir : il y a une marge de manœuvre même si le couloir ne mène qu'à un seul endroit. Dans cette condition, l'utilisateur peut avoir l'impression qu'il est totalement libre de diriger son déplacement. Nous l'avons déjà expérimenté à l'insu des utilisateurs, surpris après essai de savoir que leur trajectoire était contrainte.

Si les déplacements réalistes sont trop contraignants, rien n'empêche de réaliser une navigation PCV irréaliste (PCV Ir). En particulier si on veut que l'utilisateur puisse faire demi-tour, l'opération est très perturbante en RV lorsqu'il est nécessaire de la rendre réaliste. Il peut être préférable d'employer une métaphore et non un schème, comme la métaphore du « tunnel inversé » : quand l'utilisateur se déplace vers l'avant, sur le côté de la trajectoire il y a en permanence une entrée symbolique de tunnel : en se dirigeant vers cette entrée, l'utilisateur y entre et ressort rapidement du tunnel, en une seconde ou deux, en étant sur la trajectoire qu'il a quittée mais en sens inverse ! Il n'a pas donc pas tourner sur lui-même visuellement par vection. Il est souhaitable qu'il y ait des indices d'orientation de l'EV pour éviter la désorientation de l'utilisateur (lumière du soleil, source sonore localisée spatialement, bâtiments différents de chaque côté d'une rue, plantations et nature différente, etc. On peut prévoir d'autres métaphores de déplacements irréels pour tourner d'un quart de tour, de décalage latéral, de montée instantanée, etc.

Dans le monde réel, nous sommes presque toujours en déplacement vers l'avant et très rarement latéralement. Par conséquent, les mouvements vers l'avant sont plus proches de la réalité et moins perturbants. Mais l'espace virtuel créé ne doit pas forcément être de topologie réaliste et géométriquement identique au monde réel ! Cet espace peut être de topologie discontinue, anisotrope, non figé mais évolutif dans le temps, pas obligatoirement insécable, illogique, etc. Cela peut permettre de supprimer ou d'atténuer l'incohérence visuo-vestibulaire comme avec la métaphore du tunnel inversé. Il faut faire appel à son imagination pour trouver des solutions irréelles mais efficaces en terme de confort et d'usage.

Pour finir, il peut être utile d'établir un « budget d'accélération » en translation et en rotation pour le limiter, la borne dépendant des intensités et des durées des accélérations, une méthode similaire au « budget de la vision stéréoscopique ».

#### 4.1.3. La solution S17 détaillée

La perception du mouvement est plus sensible à la périphérie du champ visuel. Il est envisagé de diminuer le champ de vision observé par une occultation en vision périphérique. L'idée de base est de mettre une zone de la vision périphérique en noir. Mais il peut être plus judicieux de mettre cette zone avec la couleur moyenne de la scène (ou la couleur moyenne de la partie du fond de scène) pour moins limiter la perception immersive en EV. Il s'agit de limiter l'incohérence entre les systèmes vestibulaires et la vection créée par les flux optiques dus aux mouvements de la personne par rapport à l'EV et aux mouvements des objets dans l'EV. La réduction

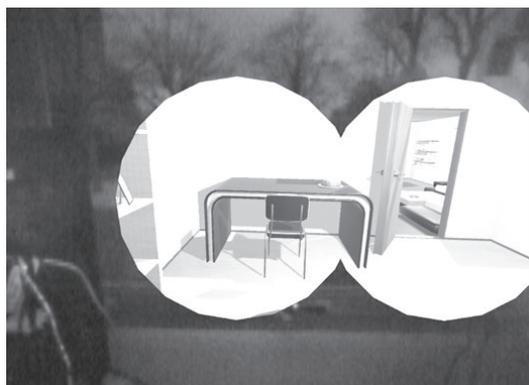
du champ de vision réduit l'incohérence. Mais il peut être judicieux d'occulter seulement les flux optiques dus aux mouvements de la personne (du point de vue) sans occulter en vision périphérique les objets en mouvement par rapport à l'EV, dans le but de mieux comprendre la scène et d'être ainsi mieux immergé. Une expérimentation<sup>13</sup> a été réalisée en supprimant qu'une partie de vision périphérique, celle correspondant approximativement à l'occlusion du nez<sup>14</sup>. Les malaises ont été en partie atténués. Cela revient à faire la solution précédemment décrite, en faisant croire que la pseudo-occlusion du nez interviendrait mieux pour la stabilité de la personne, mais l'occlusion du nez n'intervient pas sur la stabilité en ER d'après les spécialistes. Il y a juste une partie de vision périphérique occultée.

#### 4.1.4. La solution S18 détaillée

Il s'agit d'atténuer l'incohérence en injectant dans les images en vision périphérique, quelques références spatiales de l'ER, des amers, pour stabiliser l'utilisateur. Cela diminue les flux optiques en vision périphérique. Il faut bien sûr que le visiocasque ait un champ de vision assez large, au moins 100° pour ne pas perturber la vision centrale. Il faut aussi que le suivi de la tête soit très précis et bien calibré par rapport à l'ER pour rajouter des amers du monde réel afin qu'ils soient immobiles par rapport à l'ER quand l'observateur bouge sa tête. Cela impose un capteur des mouvements de la tête à 6 ddl et, de préférence, un dispositif de reconstruction 3D de l'environnement réel pour bien recalibrer la personne et son casque dans l'ER. Si le dispositif de reconstruction 3D de l'ER est fixé sur le visiocasque, c'est techniquement intéressant. Ces inclusions artificielles et non réalistes peuvent déranger la perception visuelle de l'observateur mais *a priori* faiblement car elles sont positionnées en vision périphérique et l'utilisateur n'en aura pas forcément conscience. La quantité d'amers est à ajuster en sachant que s'ils occupent une grande partie de la vision périphérique, on se rapproche de la solution S17 précédente. Les amers peuvent être des contours d'objets réels (les meubles d'une pièce, etc.) de l'ER (Figure 8.4). Quand l'utilisateur tournera la tête, il verra bouger ces amers, segments de droites par exemple, dans sa vision périphérique mais dans la vision naturelle en ER. Il peut trouver cela étrange mais cela peut être utile pour atténuer l'incohérence visuo-vestibulaire. Ces amers n'ont pas besoin d'être affichés en permanence *a priori* mais seulement lors de mouvements de rotation ou d'accélération en translation. Des expérimentations sont en cours... Nos premiers résultats, qui seront publiés prochainement, sont positifs et prometteurs. Les amers peuvent prendre différentes formes : segments, ensemble de points, grille partielle, etc.

13 <http://phys.org/news/2015-03-virtual-nose-simulator-sickness-video.html>

14 L'occlusion virtuelle sur le bas, coté intérieur, des images droites et gauches, ne correspond pas véritablement à la faible occlusion d'un nez réel car le champ de vision du visiocasque utilisé est bien plus petit que celui du champ de vision réel.



**Figure 8.4 :** Pour diminuer l'incohérence visuo-vestibulaire, nous avons expérimenté la solution S18 avec des résultats positifs. Lors d'un déplacement en EV, nous incrustons en vision périphérique, la vue de l'ER prise par une Leap Motion fixée sur le visiocasque.

Une utilisation intéressante, avec une reconstruction du lieu réel (l'ER) en nuage de points 3D via une Kinect, a été réalisée par David Nahon et son équipe du iV Lab de Dassault Systèmes. L'expérience<sup>15</sup> «Never Blind in VR» donne un aperçu de l'environnement réel aux utilisateurs d'un visiocasque sous forme de nuage de points 3D correspondant aux surfaces détectées par la Kinect. Cette solution a plusieurs intérêts : percevoir son propre corps, l'environnement réel et les autres personnes en temps réel et bien colocalisés, si la calibration est correctement réalisée. Cette solution permet d'agir en EV tout en étant en sécurité en ER (nous en reparlerons). En n'affichant les points qu'au niveau de la vision périphérique à chaque instant, il est peut-être possible que la personne se sente plus stable malgré l'incohérence visuo-vestibulaire.

L'efficacité de ces solutions reste soumise à validations scientifiques. Elles pourront être améliorées quand les visiocasques auront intégré le suivi des yeux pour intervenir plus efficacement. Nos premières expérimentations donnent des résultats positifs et prometteurs.

#### *On supprime l'incohérence en modifiant le fonctionnement de la PCV (S21, S22 et S23)*

Il y a trois solutions différentes :

- soit le déplacement réel de la personne, debout en ER, est le même géométriquement que le déplacement en EV. Les trajectoires et les vitesses en EV et en ER sont toujours les mêmes. Dans ce cas, les deux environnements ER et EV doivent être de mêmes dimensions, **géométriquement identiques**. En exemple, si l'ER est une pièce de dimensions standards où l'utilisateur est debout dans laquelle il peut se déplacer, le déplacement dans l'EV devra être identique et limité par les dimensions de la pièce. Cela implique une cohérence entre les stimuli de la vision, ceux des systèmes vestibulaires *mais aussi*

<sup>15</sup> <http://www.3ds.com/fr/recits/never-blind-in-vr/>

ceux des autres stimuli proprioceptifs (fuseaux neuromusculaires, organes tendineux de Golgi et récepteurs articulaires), les gestes étant identiques en ER et en EV. La représentation visuelle de l'EV comporte en général l'avatar colocalisé (ou au minimum les mains, si l'utilisateur n'a pas besoin de regarder tout son corps) pour aider l'utilisateur à mieux s'immerger corporellement et aisément manipuler des objets, grâce à la colocalisation de ses mains (S21) (la question de la présence ou l'absence de représentation du corps sera exposée au chapitre 9);

- soit le déplacement en EV se fait par **téléportation** d'un lieu à un autre en EV, tout en restant immobile dans l'ER. Ainsi les mouvements continus sont supprimés et les systèmes vestibulaires ne sont plus sollicités puisqu'il n'y a plus de vitesse et d'accélération. Les deux sens sont donc cohérents grâce à l'immobilité de la personne en ER et en EV. La sélection du lieu d'arrivée est en général faite à l'aide d'un capteur de localisation d'une main, qui indique le point d'arrivée virtuellement (processus métaphorique). Le lieu d'arrivée doit être visible du lieu de départ, sinon il faut pointer sur un plan des lieux. L'utilisateur passe virtuellement du lieu de départ au lieu d'arrivée instantanément. Toutefois, la transition visuelle du point de vue du lieu de départ à celui du lieu d'arrivée peut se faire par un fondu enchaîné pour adoucir la téléportation (S22);
- Une solution radicale est de faire l'application en **Réalité Augmentée**, ce qui impose alors la condition de la solution S21 : les deux environnements ER et EV sont évidemment géométriquement identiques, puisqu'ils sont confondus, et il n'y a donc plus d'incohérence! Cela impose techniquement d'utiliser un visiocasque RA. L'utilisateur voit le monde réel qui le stabilise encore mieux grâce à sa vision périphérique qui est établie sur l'ER, car techniquement il n'y a pas de possibilité d'afficher des images en vision périphérique dans les visiocasques RA... Mais les possibilités de la RA ne sont pas celles de la RV. Chaque domaine, bien distincts dans les usages, a ses avantages et ses limites (S23).

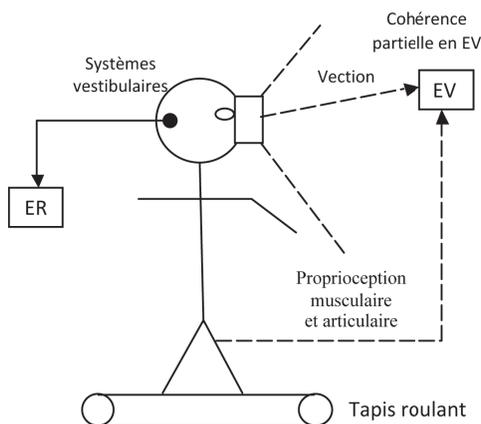
*On supprime l'incohérence en rajoutant / modifiant les interfaces (S24 et S25)*

L'incohérence peut être supprimée en créant matériellement les bons stimuli des systèmes vestibulaires avec une interface à simulation de mouvement ou interface de commande par tapis 1D ou 2D. Les stimuli d'accélération et d'inclinaison sur le corps de la personne doivent être, si possible, constamment en correspondance adéquate avec les déplacements visuels (vection) de la personne :

- Avec une interface à simulation de mouvement, il n'y a pas que les systèmes vestibulaires qui sont sollicités mais aussi les organes proprioceptifs des muscles, des tendons et des articulations qui doivent être aussi cohérents. Suivant les déplacements virtuels désirés, il faut déterminer l'interface à simulation de mouvement qui peut créer les bons stimuli sur les systèmes vestibulaires, sachant que dans certains cas, il n'y a pas de solutions parfaites, même sans s'occuper du prix de l'interface. Par exemple, pour la simulation de conduite d'une voiture, les simulateurs les plus performants des constructeurs automobiles, de plusieurs millions d'euros, peuvent créer correctement la plupart des simulations de conduite, sauf certaines trajectoires, comme celles

autour d'un rond-point routier où les accélérations réelles subies par le corps du conducteur sont trop élevées (S24) ;

- Avec une interface de marche par tapis 1D ou 2D et autres dispositifs similaires, les organes proprioceptifs des muscles, des tendons et des articulations peuvent être correctement stimulés pour diminuer les incohérences, mais les systèmes vestibulaires ne sont pas (correctement) sollicités. L'incohérence visuo-proprioceptive sera diminuée dans ce cas (la proprioception globalement) mais l'incohérence visuo-vestibulaire sera toujours présente (Figure 8.5). Il faudra en tenir compte et exploiter une des solutions S15 à S20 pour atténuer l'incohérence visuo-vestibulaire (S25).



**Figure 8.5 :** avec un tapis roulant, les organes proprioceptifs des muscles, des tendons et des articulations sont correctement stimulés par rapport à la vision, mais l'incohérence visuo-vestibulaire est toujours présente

#### 4.1.5. La solution S24 détaillée

Il y a plusieurs types d'interface à simulation de mouvement (voir le chapitre 5) qui peuvent être exploitées en fonction des déplacements prévus. Il faut donc déterminer l'interface à simulation de mouvement la plus appropriée pour créer les bons stimuli sur les systèmes vestibulaires. Son coût parfois dissuasif, implique quelquefois l'emploi d'une interface moins performante mais plus accessible. Pour des applications tout public, les plates-formes mobiles 6 axes seront réservées à des parcs d'attraction. Des sièges mobiles avec quelques degrés de liberté (2 ou 3 en général) peuvent être aussi utilisés mais sont d'un prix qui ne permet pas l'usage à domicile mais dans des salles de jeux. Des vibrations sur le siège peuvent rendre plus attractives certaines simulations. À noter que l'interaction va se faire en général manuellement et en complément au pied. L'utilisateur a donc la possibilité d'observer indépendamment l'EV tout en se déplaçant. Les mêmes précautions sont à prendre que dans la solution S15 dans le cas où il reste l'incohérence visuo-vestibulaire avec une interface à simulation de mouvement non totalement adéquate avec les mouvements en EV.

#### 4.1.6. La solution S25 détaillée

Avec une interface de marche, telle que les tapis roulants 1D ou 2D ou autres dispositifs similaires (rameur, vélo d'appartement, etc.), il faut exploiter une des solutions S15 à S20. Pour les tapis 1D, il faut prévoir un moyen de tourner qui peut se commander manuellement. Il faut bien sûr faire attention au risque de perte d'orientation mais comme avec les autres façons de se déplacer virtuellement. L'avantage d'un tapis roulant est de mieux percevoir les distances, utile pour certaines applications RV, et d'être plus immersifs sensoriellement. Mais la contrepartie est d'avoir une appréhension pour l'utilisateur qui ne voit pas ses pieds, du moins en vision directe. Il est souhaitable qu'il puisse voir une représentation bien colocalisée de ses pieds ou au minimum qu'il y ait des barres l'entourant pour garder son équilibre.

Avec de tels dispositifs, il faut se poser la question de l'intérêt d'utiliser un visiocasque plutôt qu'un ou plusieurs écrans entourant la tête de la personne. Il est étonnant de voir une application de simulation d'aviron avec un rameur d'appartement et un visiocasque. Plutôt qu'employer un visiocasque, il est plus judicieux d'utiliser deux ou trois écrans plats et fixes, en forme de U, entourant la tête de l'observateur, comme expliqué dans la solution S3. Les images sont de bien meilleure qualité et sont calculées indépendamment de la position de la tête car elle ne bouge relativement que très peu par rapport au paysage vu d'un bateau. Dans ce cas, on n'a plus besoin de mesurer la rotation de la tête et l'incohérence due à la latence dans un visiocasque est de plus supprimée. Il reste toutefois la latence entre la commande de déplacement via les rames et l'affichage des bonnes images avec le bon point de vue. Mais comme celui-ci se déplace lentement et que le décor est relativement éloigné, cette latence n'est pas perceptible pour cette application d'aviron virtuel.

Pour terminer cette longue partie sur l'incohérence visuo-vestibulaire, il peut aussi être opportun d'envisager d'exploiter en complément la stabilisation de la personne par des sources sonores spatialisées et fixes dans l'ER, même si leur localisation est incohérent par rapport aux sons souhaités en EV.

## 4.2. Incohérence visuo-vestibulaire temporelle

Le temps de latence entre les mouvements de la plate-forme mobile et l'affichage du bon point de vue doit être minimal pour éviter une incohérence visuo-vestibulaire temporelle.

*On atténue l'incohérence (S26)*

On diminue le temps de latence techniquement et informatiquement ;

*On supprime l'incohérence en modifiant le fonctionnement de la PCV*

Requête inappropriée.

*On supprime l'incohérence en rajoutant / modifiant les interfaces (S27)*

Pour ne pas avoir de latence entre les mouvements de la plate-forme mobile et l'affichage du bon point de vue, on commande en même temps en synchronisant les mouvements de la plate-forme avec l'affichage des images. Pour la synchronisation, il faut prendre en compte le temps de latence du tracking de la tête, celui de l'interface motrice de commande (par exemple, le volant dans un simulateur de transport) et celui du mouvement de l'interface à simulation de mouvement. Après synchronisation, il restera toutefois le temps de latence entre la commande de navigation et les deux simulations sensorielles synchronisées.

### 4.3. Incohérence visuo-posturale

Cette incohérence peut survenir s'il y a déplacement en EV parvection tandis que l'utilisateur est debout, immobile ou presque en ER. Son système nerveux central est informé de la position fixe du corps par les systèmes vestibulaires *mais aussi* par les autres stimuli proprioceptifs (fuseaux neuromusculaires, organes tendineux de Golgi et récepteurs articulaires).

*On atténue l'incohérence (S28)*

On atténue cette incohérence en intervenant sur le mouvement virtuel (type de trajectoire, cinématique du mouvement) avec les mêmes solutions que celles pour l'incohérence visuo-vestibulaire (visuo-proprioceptive). Il peut être utile de représenter le corps de l'utilisateur dans l'EV, son avatar, pour l'aider à mieux se situer dans l'EV relativement à sa perception proprioceptive dans l'ER. S'il voit virtuellement ses pieds, bien colocalisés en EV, il percevra une cohérence sur sa taille, sur la hauteur de ses yeux et donc sur la position de sa vue en EV par rapport aux stimuli tactiles sous ses pieds en ER. Dans le cas contraire, il peut être perturbé et ne pas se sentir stabilisé.

*On supprime l'incohérence en modifiant le fonctionnement de la PCV (S29)*

On met l'utilisateur assis ou il se tient à un support pour stabiliser sa posture debout. De plus l'utilisateur sera plus rassuré.

*On supprime l'incohérence en rajoutant / modifiant les interfaces*

Requête inappropriée.

## 5. PCV MANIPULATION

Pour la simplification des explications, on ne sépare pas dans le paragraphe suivant, les incohérences dues aux défauts techniques des visiocasques de celles dues à une PCV manipulation irréaliste.

### 5.1. Incohérence visuo-manuelle

Lors d'une manipulation en EV, si la main réelle de l'utilisateur est mal localisée par rapport à la vision de la main dans le visiocasque, on obtient une incohérence à la coordination œil-main à laquelle l'utilisateur devra s'adapter. Elle crée d'elle-même peu de malaise mais un inconfort si la contrôlabilité de l'action est difficile. La représentation de la main peut être réalisée ou non. Avec un dispositif détectant en temps réel les mouvements des doigts, comme le Leap Motion, la simulation est plus réaliste et des contrôles plus fins de manipulation peuvent être effectués. Mais l'avantage en EV est de pouvoir se passer d'une manipulation avec des doigts en employant une PCV de manipulation plus ou moins irréaliste et en utilisant des ALC.

#### *On atténue l'incohérence (S30)*

On atténue le biais entre la proprioception de la main réelle de l'utilisateur et la vision de la main virtuelle dans le visiocasque en améliorant la précision des capteurs et en calibrant correctement entre eux les différents repères : le repère du point de vue dans le visiocasque et les deux repères des positions des deux capteurs de tracking de la tête et celui de la main.

#### *On supprime l'incohérence en modifiant le fonctionnement de la PCV (S31)*

Il est toujours possible, au lieu d'exploiter le schème sensorimoteur de la manipulation manuelle, de manipuler un objet par une action métaphorique. Il y a un grand nombre de possibilités : manipuler avec un joystick, cliquer sur l'objet avec un laser virtuel et le déplacer en indiquant le point d'arrivée, manipuler avec la main non colocalisée mais en téléopération, etc.

#### *On supprime l'incohérence en rajoutant / modifiant les interfaces (S32)*

Il est possible de manipuler manuellement avec une interface à retour tactile et(ou) avec une interface à retour d'effort pour avoir des stimuli cohérents entre le toucher, la proprioception et la vision. En général, avec une interface à retour d'effort, la position de la main est mesurée précisément par l'interface et donc cela supprime l'incohérence visuo-manuelle. Inversement, si on n'exploite pas une interface à retour d'effort, on aura une incohérence sensorimotrice lors de la collision de l'objet avec un autre : pas de force de contact et pas de blocage du mouvement de la main. On peut alors employer

une ALC (Aide Logicielle Comportementale) : on rajoute des contraintes spécifiques lors de la manipulation d'un objet avec un capteur à 6 degrés de liberté (6ddl) : pour aider l'utilisateur à obtenir la motricité désirée, par exemple poser un objet sur une table, les mouvements de l'objet virtuel sont contraints. La face d'appui de l'objet se met parallèlement à la surface de la table quand elles seront proches l'une de l'autre. Ces contraintes irréelles sont appelées guides virtuels (de mouvement) ou effets de magnétisme.

## 6. GRILLE D'ANALYSE DES 32 SOLUTIONS

À partir de la liste des 11 incohérences sensorimotrices perturbatrices, il a été possible de répondre à trois questions :

- Comment peut-on atténuer l'impact de l'incohérence s-m sur l'inconfort ou le malaise de l'utilisateur ?
- Peut-on supprimer l'incohérence s-m en modifiant le fonctionnement de la PCV ?
- Peut-on supprimer l'incohérence s-m en modifiant le fonctionnement de l'interface ou en rajoutant une autre interface ?

Cela a permis de décrire 32 solutions, certaines dépendantes de l'évolution des caractéristiques des visiocasques, d'autres de changement de matériel (interface) et d'autres de la conception fonctionnelle (PCV) de l'application RV. Quelques-unes sont très spécifiques et seront peu souvent exploitées. Il est toujours possible d'envisager au cours d'une application RV de passer d'une solution à une autre, si évidemment l'utilisateur peut s'adapter à un tel changement de fonctionnement de PCV. Il reste donc la dernière question qui sera étudiée dans la partie suivante : comment peut-on s'adapter à ces incohérences pour ne plus avoir d'inconfort ou de malaise ?

En résumé, il y a des solutions qui font :

- changer de domaine : de la réalité virtuelle à la réalité augmentée ;
- changer d'interface visuelle, quand le visiocasque n'est pas la solution optimale ;
- changer de visiocasque pour exploiter un produit plus performant ;
- changer ou rajouter une interface associée au visiocasque ;
- changer d'une PCV irréaliste à une PCV classique ;
- modifier le fonctionnement d'une PCV ;
- supprimer une incohérence sensorimotrice perturbante ;
- atténuer les effets d'une incohérence sensorimotrice perturbante.

Sans oublier qu'il est aussi envisageable que l'utilisateur puisse s'adapter à l'incohérence sensorimotrice perturbante (voir plus loin).

Enfin, il sera parfois nécessaire d'attendre l'évolution des techniques pour exploiter un futur visiocasque aux caractéristiques adéquates, si les solutions précédentes ne sont pas satisfaisantes pour l'application RV souhaitée.

Pour le concepteur d'application RV, en dehors des solutions dépendant des constructeurs des visiocasques, il doit connaître les solutions des trois principales incohérences sensorimotrices perturbatrices qui sont les plus importantes à gérer et les plus connues, (sans ignorer les autres solutions qui interviennent dans certains types d'application RV) :

- l'incohérence visuo-motrice temporelle (latence du tracking de la tête) ;
- l'incohérence oculomotrice ;
- l'incohérence visuo-vestibulaire.

Et pour ces trois incohérences, les solutions sont synthétisées dans le tableau suivant :

	<b>Atténuer l'impact d'une incohérence s-m</b>	<b>Modifier fonctionnement de PCV</b>	<b>Supprimer une incohérence s-m</b>
L'incohérence visuo-motrice temporelle	<b>S1</b> : diminution du temps d'exécution du logiciel	<i>S2</i> : séquence d'images figées	<i>S3</i> : écrans autour de la tête sans translation
L'incohérence oculomotrice	<b>S6</b> : diminution des disparités rétinienne et traitements d'images	<b>S7</b> : vision monoscopique	<i>S8</i> : écrans Light Field ou VRD
L'incohérence visuo-vestibulaire	<b>S15</b> : limiter accélération et inclinaison <b>S16</b> : trajectoires douces <b>S17</b> : diminuer le champ de vision <b>S18</b> : traitements en vision périphérique <b>S19</b> : objets fixes/ER en vision périphérique <i>S20</i> : vision périphérique non occultée	<b>S21</b> : ER et EV géométriquement identiques <b>S22</b> : déplacement par téléportation <i>S23</i> : application en Réalité Augmentée	<b>S24</b> : avec une interface à simulation de mouvement <b>S25</b> : avec un tapis roulant 1D ou 2D

**Tableau 8.1** : synthèse des solutions à envisager pour traiter les trois principales incohérences sensorimotrices

Les 12 solutions importantes à connaître sont pointées avec la lettre **S** en gras. Les solutions secondaires, notées en italique, proposent soit :

- un changement de domaine (RA au lieu de RV), donc de type de visiocasque ;
- un changement d'interface visuelle ;
- la mise en œuvre de solutions très particulières.

### 6.1.1. Considération technico-économique

#### *Certaines solutions sont onéreuses*

Elles ne sont alors guère envisageables pour un **achat individuel, pour les jeux vidéo personnels**, à moins d'être suffisamment à l'aise financièrement, impliquant un marché de niche pour de telles solutions. Elles sont pour les plus onéreuses, réservées à des parcs d'attraction (S24) ou des salles de jeux (S25), même si des entreprises cherchent à vendre des tapis roulants à moins de mille euros.

La solution S21 (ER et EV géométriquement identiques) impose d'avoir une salle dédiée au jeu vidéo de 10 m<sup>2</sup> environ, sauf à limiter l'activité ludique spatialement et à laisser le joueur assis. L'avantage d'une salle dédiée est de l'équiper de caméras de tracking à demeure, pour ne pas avoir à calibrer le tracking constamment. C'est une solution proposée techniquement par les grands constructeurs de visiocasque, car elle permet aussi de résoudre les contraintes des latences très faibles (S1) et de précisions élevées (S11) pour les trackings... si le développeur n'augmente pas les latences ! Cette solution S21 reste assez onéreuse et encombrante, en imposant une salle ou un petit espace dédiée pour une activité à domicile. Et si on veut exploiter un espace plus grand que celui d'une salle, la solution 21 reste réservée à des salles de jeux ou des parcs d'attraction qui peuvent utiliser jusqu'à un hall de grandes dimensions (400 m<sup>2</sup>, par exemple)<sup>16</sup>.

S'il s'agit d'**applications RV professionnelles**, les contraintes financières ne sont pas les mêmes et cela ouvre plus le champ des solutions envisageables... et avec ou sans l'utilisation d'un visiocasque pour l'interfaçage visuelle, solutions réalisées depuis une vingtaine d'années.

Les 9 autres solutions importantes (S1, S6, S7, S15, S16, S17, S18, S19 et S22) sont sous le contrôle du concepteur de l'application RV et n'impliquent pas un coût financier complémentaire, au-delà des coûts engendrés par de nouveaux développements éventuels.

#### *Certaines solutions limitent des PCV*

Les solutions S7 (vision monoscopique) et S6 (vision stéréoscopique bien maîtrisée) ne sont pas de grandes contraintes, qu'il suffit de régler en fonction des capacités individuelles de chaque porteur de visiocasque. La perception tridimensionnelle par stéréopsie de l'EV n'est pas indispensable dans bien des applications. Elle peut être supprimée (S7) ou très atténuée (S6), mais l'immersion visuelle est moins attrayante et les manipulations fines d'objets proches moins aisées. De même, la solution S17, limitant le champ de vision, est aisée à appliquer mais diminue la vision périphérique de l'observateur, qui peut être dans certains cas compensée par du son spatial ou autres subterfuges visuels. Ces trois solutions limitent la PCV observation et la rende moins attrayante. La solution S22 (téléportation) impose à la PCV navigation des déplacements moins attrayants, mais efficaces !

<sup>16</sup> Voir <http://www.journaldugeek.com/2015/08/17/zero-latency-un-centre-de-jeu-utilisant-la-realite-virtuelle/>

### *Solutions sous le contrôle du concepteur*

Il reste les 5 solutions (S1, S15, S16, S18 et S19) :

- la solution S19 impose au moins un objet virtuel fixe par rapport à l'ER et donc impose un certain type d'EV et de scénario du jeu vidéo (cockpit de voiture, etc.).
- les solutions S15, S16 et S18 sont sous le contrôle du concepteur pour limiter les effets négatifs de l'incohérence visuo-vestibulaire. La solution S18 peut être, dans une de ses sous-catégories, avec un coût complémentaire exigeant une reconstruction 3D de l'ER ;
- la solution S1 (latence très faible) impose au développeur de l'application de faire attention à ne pas augmenter la latence. Il doit réaliser son programme avec une phase d'optimisation des fonctions utilisées et avec une architecture logicielle favorisant la parallélisation, pour réduire au maximum la latence globale.

Il reste pour tout concepteur ou développeur à régler des paramètres (seuils acceptables pour telle solution, etc.).

Cela ne peut pas se faire sans connaître le niveau d'adaptation de tout individu à un EV, objet de la partie suivante.

## **7. ADAPTATION EN ENVIRONNEMENT VIRTUEL**

### **7.1. Niveaux de difficultés d'adaptation**

La question de l'adaptation de toute personne à une immersion et une interaction dans un environnement virtuel est plus vaste que seulement l'adaptation aux incohérences sensorimotrices pour éviter inconfort et malaise. En rapport avec notre modèle 3I<sup>2</sup> de l'interfaçage, quatre points sont à traiter, les deux premiers faisant référence au niveau I<sup>2</sup> sensorimotrices :

- adaptation physiologique à l'interface visuelle : le visiocasque ;
- adaptation aux incohérences sensorimotrices ;
- adaptation cognitive à l'interfaçage ;
- adaptation fonctionnelle aux PCV.

L'adaptation physiologique au visiocasque a été traitée au chapitre 4 : comment procéder pour que le visiocasque soit bien ajusté, interfacé avec le système visuel de l'utilisateur, sachant que les capacités optiques sont très variables d'une personne à une autre. Il ne faut donc pas oublier de faire avant chaque application RV ou, au moins, avant chaque première utilisation d'un nouveau visiocasque, les calibrages optiques et ergonomiques et les tests d'évaluation de l'interfaçage visuel. Si ces opérations ne sont pas faites ou trop sommairement, les solutions préconisées pour

s'adapter aux incohérences sensorimotrices risquent de perdre de leur efficacité. En exemple, il est difficile d'atténuer l'incohérence « accommodation – vergence » de la vision stéréoscopique pour des utilisateurs n'ayant pas bien interfacé optiquement le visiocasque, surtout lorsque ceux-ci ont des anomalies visuelles. Et n'oublions pas que, dans un visiocasque, on ne sait pas précisément ce que l'observateur voit exactement...

L'adaptation cognitive à l'interfaçage doit être vérifiée pour s'assurer que la personne a bien compris le processus cognitif d'interfaçage et qu'elle le maîtrise en pratique, que ce soit un Schème Comportemental Importé ou une métaphore employés. Il y a déjà eu des échecs car le processus cognitif utilisé par l'utilisateur n'était pas celui qui avait été prévu par le concepteur de l'application RV.

L'adaptation fonctionnelle aux trois PCV doit être aussi évaluée. Pour les personnes plus sensibles aux incohérences sensorimotrices, il est fortement déconseillé de leur proposer des Primitives Comportementales Virtuelles irréelles (PCV Ir) qui ne sont réservées qu'à des utilisateurs aux capacités d'adaptation élevées.

La principale difficulté en RV est de permettre à toute personne de s'adapter aux incohérences sensorimotrices dans le but d'éviter inconfort au minimum, malaise et, au pire, rejet de l'application RV. Il faut établir un **Niveau Global de Difficultés d'Adaptation (NGDA)** en fonction de la liste des 11 incohérences et des solutions retenues. Plusieurs solutions retenues peuvent être proposées dans une application RV selon les difficultés d'adaptation, ce qui implique plusieurs configurations de solutions à programmer. Certaines des solutions ont des paramètres variables (valeur de seuil, etc.). Ils doivent aussi intervenir dans la détermination du NGDA et dans le paramétrage des configurations. On peut quantifier le NGDA de chaque configuration de l'application. Chaque utilisateur pourra alors choisir sa configuration selon sa sensibilité aux incohérences sensorimotrices.

Des travaux de recherche sont à poursuivre au sujet des niveaux d'impact sur le confort et sur la santé des utilisateurs à partir des onze incohérences et des solutions envisagées. Les niveaux et leurs impacts évolueront en fonction des progrès techniques et des solutions choisies sous la responsabilité du concepteur d'application RV. Celui-ci doit gérer leur intensité et leur temps d'exposition en fonction de chaque personne, ce qui implique la détermination d'un budget. Il peut être prévu que le jeu vidéo s'arrête, ou au moins impose une pause, après un certain temps, indépendamment de la volonté du joueur... comme au cinéma 3Ds où les spectateurs ne choisissent pas la durée du film en relief.

Des tests d'adaptation peuvent être proposés, permettant à chacun de choisir ses niveaux d'impact des différentes incohérences. Des exercices d'apprentissage peuvent aussi être fournis. Certaines solutions et certains paramètres peuvent être réglés par les personnes elles-mêmes, par exemple : passer en vision monoscopique ou seulement diminuer l'impact de la vision stéréoscopique par un réglage personnel de la limite maximale de disparités rétinienne.

## 7.2. Niveaux d'adaptation des utilisateurs

L'adaptation aux incohérences sensorimotrices dépend de chaque utilisateur (hypothèse 11). Dans certains cas classiques et bien analysés, nous savons par expérience que presque toutes les personnes peuvent surmonter certaines incohérences en s'y adaptant. Le concepteur de l'application RV ne doit pas oublier de prendre en compte la morphologie de la personne, entre autres sa distance interpupillaire (DIP), la hauteur de ses yeux par rapport au sol, etc.

L'âge des utilisateurs peut être pris en compte, tout en sachant qu'il n'a pas été établi de relation causale entre l'âge et la capacité d'adaptation car d'autres caractéristiques interviennent : l'expérience des utilisateurs aux applications RV par exemple. Cette caractéristique est un facteur positif qui permet aux utilisateurs d'être moins sensibles aux incohérences, le cerveau ayant un pouvoir d'adaptation reconnu (hypothèse 9). Ceci est vrai aussi pour les concepteurs de jeux vidéo qui s'adaptent progressivement à leurs nouvelles applications lors de leurs propres tests d'évaluation. Ils doivent donc obligatoirement faire tester leurs applications par des utilisateurs novices. Ces derniers doivent faire les tests en commençant par le niveau NGDA le plus bas.

Mais l'adaptation dépend aussi de la motivation du joueur... Certaines séquences dans des applications RV pourront être proposées avec de fortes incohérences sensorimotrices si des joueurs viennent avec la motivation et l'envie de vivre une expérience « pleine d'adrénaline ». Les incohérences seront alors plus aisément acceptées, comme dans les attractions à sensations fortes, de type « montagnes russes ».

## 8. RÈGLES DE SÉCURITÉ

La principale difficulté de sécurité est due à l'isolation visuelle et acoustique du porteur du visiocasque, sauf pour les visiocasques RA évidemment. Comment pallier à cette absence de visibilité direct de l'ER où se trouve la personne, surtout si l'ER change de configuration pendant le déroulement de l'application ? La sécurité physique de la personne est surtout à prendre en compte si elle est debout dans une pièce et non assise sur un siège fixe. Dans cette dernière situation, son isolation visuelle et acoustique l'empêche seulement d'être alertée sur tout événement extérieur non lié à son activité.

En revanche, si la personne est debout dans une pièce, son activité physique liée à l'application RV peut être source de danger, surtout si la personne doit se déplacer, même que sur un ou deux mètres. Dans cette modalité, le minimum de sécurité est d'être alerté dès que l'on s'approche des parois de la pièce ou des meubles. Il est très souhaitable, pour ne pas dire indispensable, que le tracking de tête du visiocasque soit à 6 ddl et que les positions des parois de la pièce, ou d'une zone sans meuble, soient connues par l'application RV. L'information des positions relatives

des parois par rapport au visiocasque peut permettre de sécuriser le déplacement de la personne. Lorsqu'elle se rapproche trop des parois (ou des meubles), une alarme l'avertit du risque de collision. L'avertissement peut se faire par un signal sonore, voire tactile, et surtout visuellement par une grille virtuelle ou tout autre indice visuel indiquant les limites à ne pas franchir. La suppression des images avec arrêt de l'application peut être aussi utilisée avec plus d'efficacité sécuritaire. Toutes ces différentes alarmes peuvent être combinées en même temps ou légèrement décalées dans le temps pour fournir des avertissements progressifs. Indépendamment d'avoir un système de sécurité, il est souhaitable que l'application RV ne se déroule pas dans un lieu avec des potentialités spécifiques de danger : cuisine, salle de bains, pièce donnant sur balcon ou escalier, etc.

La connaissance des positions relatives des parois par rapport au visiocasque peut être établie avant le début de l'application RV lors de sa configuration, mais il est préférable que cela se fasse en temps réel. Cela dépend évidemment de la méthode de tracking de la tête. Par exemple, si celle-ci se fait par des caméras positionnées aux quatre coins du plafond de la pièce, la zone libre de mouvement sans danger peut être définie à la configuration du tracking de tête. Si le visiocasque porte un dispositif incorporé de reconstruction 3D déterminant en temps réel la position des parois, des meubles et des personnes présentes, du type kinect de Microsoft ou la caméra frontale du visiocasque VIVE de HTC (voir chapitre 6), la sécurité est bien assurée. L'application « Never Blind in VR » du iV Lab de Dassault Systèmes permet de bien sécuriser l'application RV car l'ordinateur sait en permanence où se situe l'utilisateur et ce qui se passe autour de son environnement. Dans toutes les situations, il est préférable que les mains et, mieux encore, tout le corps soient représentés virtuellement dans le visiocasque, ce qui ne pose pas de problème technique avec les dispositifs donnant en temps réel une carte de profondeur du lieu.

Si la personne est assise sur une plate-forme mobile, il faut prévoir des sécurités spécifiques, normalement intégrées au dispositif, telles que des ceintures de sécurité, si la plate-forme sollicite beaucoup le corps de l'utilisateur.

Si la personne est debout en restant sur place ou marchant sur un tapis roulant, il faut prévoir une ou plusieurs barres de soutien, surtout si l'EV peut être vu incliné, car en réaction la personne aura tendance dans l'ER à se pencher en sens opposé et pourra perdre l'équilibre avec un risque de chute.

La confusion entre le monde virtuel et le monde réel risque de poser des problèmes de sécurité. Certaines personnes, cognitivement très immergées dans l'EV, oublient qu'elles sont dans un monde simulé. Des comportements inappropriés peuvent survenir. Un comportement, plus amusant que dangereux, s'est souvent produit dans nos applications RV où un utilisateur, « trop » immergé cognitivement, pose son interface de commande (manettes, etc.) sur une table virtuelle qu'il perçoit. Il est alors surpris de voir tomber sur le sol l'interface. En revanche, une personne très effrayée par un événement virtuel peut se mettre à reculer subitement pour échapper au danger virtuel et alors tomber de sa chaise. Il faut donc faire attention aux contenus proposés qui peuvent effrayer trop intensément la personne immergée.

## 9. CONCLUSIONS

L'analyse des incohérences sensorimotrices perturbatrices a permis de présenter dans un cadre précis les solutions qui, au minimum, atténuent les effets négatifs sur le confort et la santé de l'utilisateur. Certaines solutions sont dépendantes des progrès techniques des visiocasques. Certaines ont déjà été validées expérimentalement tandis que d'autres ne sont que des pistes à étudier. Il faut développer de futures expérimentations pour valider des solutions nouvelles dans le cadre de l'usage des visiocasques pour tout public.

Notre analyse a été basée sur notre méthodologie 3I<sup>2</sup> qui est complétée par la prise en compte des perturbations au niveau des I<sup>2</sup> sensorimotrices dans le but d'améliorer le confort et la santé de l'usager, d'où son nom : 3I<sup>2+</sup>. Cette analyse a ses limites car on a considéré chaque incohérence indépendamment des autres et les hypothèses émises sont aussi à justifier et à valider.

Nous n'avons pas exposé les solutions pour la PCV «communication» pour ne pas alourdir la présentation mais des solutions spécifiques sont à concevoir pour la communication entre le système RV et l'utilisateur portant un visiocasque. Par rapport à l'usage d'un écran plat, l'emploi d'un visiocasque est une contrainte non négligeable car l'usager ne peut plus utiliser aisément un clavier et une souris.

Les risques d'impact sur le confort et la santé de l'utilisateur d'une application RV expliquent que les fabricants de visiocasques donnent des recommandations contraignantes pour l'usage de leurs produits. On peut envisager plusieurs raisons à cela :

- ils craignent que le problème de fond des incohérences sensorimotrices en RV puisse perturber le confort et la santé de certains utilisateurs ;
- les techniques de leur visiocasques ne sont pas encore suffisamment au point et craignent des avis négatifs sur leurs produits ;
- ils sont prudents et veulent se protéger juridiquement, en particulier pour l'usage éventuellement néfaste des visiocasques auprès des enfants.

Leurs principales recommandations sont de faire des pauses, de s'arrêter en cas de malaise et, après avoir vécu une application avec un visiocasque, de ne pas faire des tâches physiquement complexes telles que conduire un véhicule. Ils interdisent aux enfants de moins de 13 ans d'utiliser un visiocasque. Il va être important dans un proche avenir de définir des règles précises d'usage des visiocasques et d'étudier les **éventuels effets sur le long terme** pour tous, particulièrement pour les enfants, si l'usage des visiocasques tend à se banaliser. Le contenu des applications RV devra être aussi maîtrisé et contrôlé avec d'éventuels messages d'avertissement, comme cela se fait chez un certain nombre de fabricants de visiocasques.

Si les constructeurs sont prudents, ils sont aussi quelques fois trop prétentieux et exagèrent l'impact de leurs innovations : chacune d'elles supprime au mieux une incohérence perturbatrice, par exemple l'emploi des futurs écrans Light Field annule l'incohérence «accommodation – vergence». Mais cela ne les autorise pas en aucune façon de prétendre qu'ils ont résolu le mal du simulateur, le «mal de la réalité virtuelle», car les autres incohérences n'ont pas disparu par enchantement...

## **Partie II**

## INTRODUCTION AUX APPLICATIONS UTILISANT DES VISIOCASQUES

Avant de présenter en détail quatre applications exploitant les visiocasques, on peut lister les secteurs d'activités, professionnelles ou non, qui sont susceptibles d'exploiter des visiocasques (liste non exhaustive). La liste générale des applications RV, employant ou non des visiocasques, peut être consultée dans le volume « Applications » du « Traité de la réalité virtuelle ».

### Les applications RV pour tout public

- le jeu vidéo et les activités RV ludiques sont concernés en premier. La question fondamentale est de savoir quel sera l'impact des futurs visiocasques commercialisés dans l'industrie du jeu vidéo et des activités ludiques connexes ;
- les arts numériques sont aussi un secteur qui peut profiter du potentiel des visiocasques, présentés au chapitre 11 ;
- les applications en direction des réseaux sociaux sont peut-être susceptibles d'être développées pour des échanges avec plus d'empathie ;
- la téléprésence par visiocasque peut s'avérer intéressant pour quelques cas particuliers. En exemple, un élève qui ne peut être en classe pour des raisons de santé, peut déjà y être téléprésent, via un robot compagnon circulant dans la classe et retransmettant à distance sons et images. L'élève se sentira plus immergé avec plus d'empathie, au milieu du groupe de la classe, en salle comme en cours de récréation, s'il peut visualiser ses camarades via un visiocasque à la place d'un simple écran (si le robot est équipé d'une caméra motorisée, commandée par la tête de l'élève). Mais l'idée qui consiste à utiliser des visiocasques en classe pour la formation avec un groupe d'élèves me semble contre-productive et contraignante. La formation devant un grand écran n'empêche pas le dialogue naturel entre l'enseignant et les élèves, tout en permettant des simulations pédagogiques virtuelles de toutes natures ;
- le pilotage de drones profite des deux fonctionnalités des visiocasques, mais l'impact des incohérences sensorimotrices n'est pas à sous-estimer ;
- les présentations marketing d'un produit peuvent s'avérer intéressantes dans certains cas particuliers : pour des présentations de produit, sans besoin de rendre la qualité esthétique de ce dernier, par exemple pour des produits complexes à utiliser par le client comme l'aménagement ergonomique d'une cuisine à l'échelle 1, ou les informations ergonomiques sur les commandes internes d'une planche de bord de voiture chez un concessionnaire ;
- la vidéo 360°, un nouvel outil, qui n'est pas du cinéma modifié, mais un nouvel art que l'on pourrait appeler « Immerciné », pour exploiter l'immersion dans un monde artificiel, qu'il soit virtuel, issu d'images de synthèse, ou soit un monde réel filmé par des caméras 360. Ce nouveau media peut se développer pour des usages privés, en complément des photos et des vidéos personnelles, et aussi pour des usages publics, même si chaque spectateur visionne isolément la vidéo 360°. Le spectateur peut devenir un « spect-acteur » s'il peut intervenir dans l'EV ou dans le scénario d'une histoire dans un monde réel. Ce n'est donc pas du cinéma, puisque ce n'est pas que le réalisateur du film qui contrôle le point de vue du spectateur mais c'est un nouvel art à créer. Les techniques de prise de vue 360° sont présentées au chapitre 12 ;

- les clips de chanteurs en vidéo 360° («images et musique 360») devraient aussi se développer et se généraliser avec l'emploi des visiocasques équipés de restitutions sonores spatiales. La musique par restitution spatiale s'est peu développée via les oreillettes, mais le pourra avec de nouveaux clips ayant son et images 360.

## Les applications professionnelles

- la réalité virtuelle est un outil expérimental intéressant pour les sciences humaines, permettant d'étudier le comportement humain. Réciproquement ce domaine scientifique doit mener des recherches sur le comportement humain dans les environnements virtuels, entre autres pour mieux exploiter les visiocasques. Le chapitre 9 présente des expérimentations en sciences du comportement ;
- dans le domaine de la santé, les visiocasques sont déjà utilisés pour les thérapies virtuelles, en parallèle aux techniques RV sur grands ou petits écrans : thérapies en psychiatrie, en rééducation fonctionnelle ou sensorimotrice. Pour la formation en chirurgie, par exemple pour les opérations en endoscopie, les visiocasques ne sont pas employés puisque les écrans répondent bien aux besoins. Les visiocasques RA peuvent être, en revanche, exploités pour des opérations chirurgicales en Réalité Augmentée.
- en télécommunication, l'exploitation des visiocasques est envisageable dans certains cas particuliers de visioconférences, tout en sachant que les techniques actuelles suffisent bien souvent pour répondre aux besoins ;
- la formation en RV et les «serious games» sont bien déjà développés depuis 15 ans, sans l'usage des visiocasques et avec grande efficacité pédagogique (voir le chapitre correspondant du volume «Applications» du «Traité de la réalité virtuelle»). Il n'y a pas à attendre de grands changements dans ce domaine avec l'exploitation des visiocasques. Ceux-ci peuvent être adaptés à certains besoins d'immersion visuelle spécifique, tout en rappelant la gêne que procure le visiocasque, séparant visuellement l'apprenant de son formateur (sauf si le formateur est, lui aussi, en situation d'immersion visuelle dans la même scène, ce qui complexifie l'application RV) ;
- Les applications pour les simulateurs de transport aériens, ferroviaires ou terrestre peuvent être exploités de grands écrans mais aussi des visiocasques. Le chapitre 10 présente des usages industriels avec des visiocasques.



# Expérimentation en sciences du comportement

Daniel R. Mestre

### 1. CHANGEMENT DU PAYSAGE DE LA RÉALITÉ VIRTUELLE PAR LES VISIOCASQUES

L'histoire des dispositifs immersifs de réalité virtuelle est relativement courte, à l'échelle de l'aventure scientifique et expérimentale, en particulier dans le domaine de l'étude des comportements humains et de leurs déterminants. Les chercheurs de ce domaine ont mis en œuvre des dispositifs expérimentaux grâce auxquels ils contrôlent et manipulent les informations sensorielles ainsi que les conditions de contrôle de l'activité dont disposent des sujets expérimentaux. De ce point de vue, les techniques de la réalité virtuelle ne sont pour les chercheurs que le développement de techniques déjà existantes permettant de se rapprocher de situations «écologiques» d'analyse de l'activité [Loomis *et al.*, 1999]; [Tarr & Warren, 2002]. L'introduction des technologies de réalité virtuelle s'inscrit donc dans une approche «interactionniste» du bouclage «perception-décision-action» et de l'étude «en temps réel» de la transformation sensorimotrice.

Dans ce cadre, l'invention du dispositif nommé CAVE [Cruz-Neira *et al.*, 1993] a été une avancée très significative dans l'étude expérimentale de situations «écologiques». En effet ce dispositif multi-écrans permet de stimuler la totalité du champ visuel d'un observateur, et surtout de lui proposer, en temps-réel, un retour sensoriel sur ses propres actions (Figure 9.1).



**Figure 9.1 :** un sujet immergé dans le CAVE du CRVM. Le dispositif est constitué de 4 écrans rétro projetés (le sol et 3 écrans verticaux). Dans cette expérimentation, le sujet doit marcher sur une passerelle au-dessus d'un canyon.

En attendant la prochaine évolution technologique, ce dispositif présente cependant des inconvénients sérieux. Il est basé sur un système de projecteurs, couteux et délicats à maintenir (en particulier, la calibration multi-écrans). La plupart des CAVE exploités ont 4 faces et nécessitent au moins 4 vidéoprojecteurs haute-définition et un «cluster» de PC équipé des cartes graphiques professionnelles, pour produire un retour d'information (notamment visuel) cohérent avec les actions du sujet. De plus, les temps de latence de tels systèmes sont, pour un rendu à 60 images par secondes (standard inférieur de nos jours), de l'ordre de 60 à 100 millisecondes. Ce délai est le plus souvent non perturbant, mais peut affecter le contrôle de gestes rapides (on note alors une «instabilité» du monde).

En face du CAVE, les visiocasques sont en fait un concept ancien, si on pense en particulier aux travaux de la NASA, dans les années 1980. Jusqu'à très récemment, à part quelques exceptions prototypiques, les visiocasques étaient (très) lourds, avec une résolution spatiale et temporelle de qualité très moyenne. À ce stade, le CAVE semblait le seul dispositif optimal pour immerger un sujet.

Cependant, les visiocasques qui arrivent sur le marché aujourd'hui possèdent, (outre leur prix bien plus bas que celui d'un CAVE sans nécessité de cluster de PC pour les alimenter) des caractéristiques qui semblent de nature à en faire des concurrents sérieux du CAVE pour des expérimentations sensorimotrices et cognitives. Le champ visuel effectif annoncé est de l'ordre de 100 degrés (inférieur certes au champ fonctionnel d'un CAVE). Le temps de latence de la boucle «perception-décision-action» est annoncé, pour l'Oculus par exemple, de l'ordre de 20 millisecondes, ce qui est largement inférieur à ce qui est atteignable pour la plupart des CAVEs. Ils présentent par ailleurs une caractéristique immersive, sinon décisive, tout au moins intéressante pour le chercheur, qui est d'isoler complètement la vision du sujet de celle du monde extérieur et de son propre corps. Pour obtenir l'isolation visuelle du monde extérieur, il faudrait un CAVE à 6 faces. Or leur nombre dans le monde est confidentiel, étant donné les contraintes techniques et leur coût prohibitif.

Nous avons récemment entrepris de tester ce nouveau type de visiocasque et voulons rapporter ci-après quelques premières observations liées à leur utilisation dans le domaine des sciences du comportement.

## **2. TRAVERSER DES PORTES VIRTUELLES**

La ligne de recherche que nous présentons ici s'intéresse à l'évaluation de la présence comportementale dans les environnements virtuels [Mestre & Fuchs, 2006]. Une étude réalisée préalablement dans le CAVE du CRVM [Lepecq *et al.*, 2009] a recherché une objectivation comportementale de la sensation subjective de présence à l'intérieur d'un Environnement Virtuel (EV). La tâche du sujet consistait à franchir une porte virtuelle de largeur variable, en marchant réellement (Figure 9.3).

Les résultats de cette étude indiquent que les postures locomotrices de sujets devant franchir une porte virtuelle ressemblent fortement à celles de sujets qui doivent franchir une porte réelle [Warren & Chang, 1987]. En effet, chez la plupart des sujets, on observe une transition entre une marche frontale normale et une marche avec rotation du corps au fur et à mesure que la largeur du passage emprunté s'amenuise. L'évaluation de la présence passe donc ici par l'évaluation d'un ajustement locomoteur particulier qui met en relation la métrique corporelle (largeur des épaules) et la métrique de l'environnement (largeur du passage).

Nous avons donc reproduit ce protocole expérimental en utilisant le visiocasque Oculus DK2 au lieu du CAVE (Figures 192 et 9.3). Le sujet était équipé du visiocasque. Il portait une configuration de marqueurs pour mesurer les mouvements de son corps (système ART).



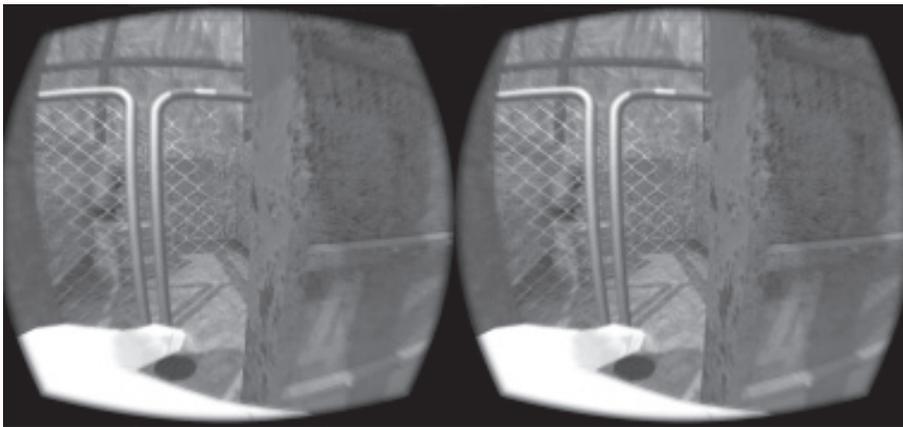
**Figure 9.2 :** un sujet dans le dispositif expérimental. Sa posture correspond à la représentation de la figure 9.3 où il tourne les épaules pour passer à travers une ouverture étroite



**Figure 9.3 :** représentation du sujet traversant la porte virtuelle

Comme on le voit sur la figure 9.2, le sujet était équipé d'une configuration de cibles permettant de mesurer les mouvements de son corps. Cela nous a permis de tester l'importance de la vision de son propre corps dans l'exécution de ce type de tâche naturelle. En effet, par défaut, lorsque le sujet «chausse» le visiocasque, il perd la vision de son propre corps (et du monde réel), ce qui n'est pas le cas dans le CAVE (figure 9.1). On se retrouve alors devant un paradoxe: d'un côté, le visiocasque est décrit comme immersif, car il bloque la totalité de la vision du monde réel (ce qui n'est pas le cas dans un CAVE «4 faces», le plus courant). D'un autre côté, il crée une situation «désincarnée», dans laquelle le sujet ne voit plus son propre corps. Des observations informelles révèlent que les sujets se plaignent souvent d'une sensation de «flottement» lorsqu'ils portent un visiocasque, qui semble liée au fait que le sujet, ne voyant pas ses pieds, a du mal à déterminer sa hauteur par rapport au sol. Cela laisse supposer qu'il pourrait exister alors un problème de perception de l'échelle de l'environnement virtuel.

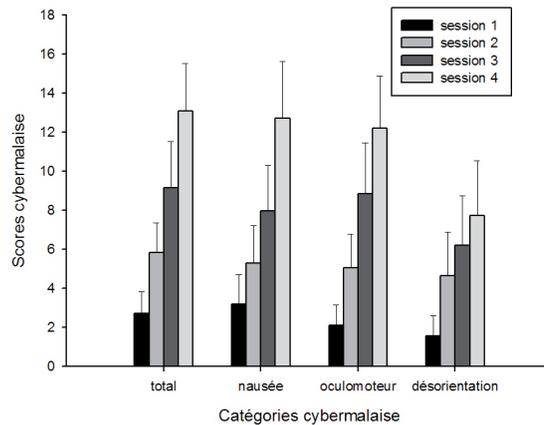
Nous avons donc testé le rôle d'un corps virtuel (image projetée dans le visiocasque du corps du sujet) dans cette tâche. Pour cela, le corps du sujet était traqué, et servait à lui proposer (ou non) une représentation virtuelle colocalisée en temps réel de son propre corps (Figure 9.4).



**Figure 9.4 :** vue subjective projetée dans le visiocasque (avec représentation du corps –en blanc–). Le lecteur peut tenter de fusionner les images gauche et droite pour avoir une idée de la scène en relief.

Les premiers résultats de cette expérimentation sont clairs. Premièrement, l'immersion obtenue avec le visiocasque paraît globalement équivalente à celle obtenue dans le CAVE. Les sujets tournent les épaules pour passer à travers une ouverture étroite [Lepecq *et al.*, 2009]; [Mestre, Louison & Ferlay, 2016]. Deuxièmement, en l'absence de la représentation du corps dans le rendu graphique, les sujets semblent mal calibrer leur perception. Ils «heurten» la porte virtuelle dans près de la moitié des cas (pour de faibles ouvertures; proches de leur largeur d'épaule). La présence d'un corps virtuel réduit significativement des collisions.

La première conclusion est donc que la présence, telle que mesurée dans cette expérience, nécessite la représentation d'un corps virtuel dans le cas du port d'un casque immersif. Un deuxième résultat de cette expérimentation est lié à la notion de cybermalaise. Ce trouble (qui peut être sérieux) est lié aux incohérences sensorimotrices observées dans les environnements virtuels [Kennedy *et al.*, 1993] dues à des difficultés techniques (temps de latence trop élevé, imprécision des tracking). Lors de cette expérimentation, les sujets subissaient 4 sessions successives, entrecoupées de courtes pauses. Nous avons utilisé ce temps de pause pour les soumettre à un questionnaire de cybermalaise. Les résultats sont représentés dans la figure 9.5.



**Figure 9.5:** représentations des scores moyens observés de cybermalaise moyenne sur 18 sujets, avec représentation de l'erreur-type)

Les résultats sont clairs et montrent une augmentation régulière des symptômes de cybermalaise au fil des sessions. On observe, pour toutes les catégories, une augmentation régulière du score au fil des répétitions du questionnaire, passé après chacun des 4 sessions. Chaque session durait environ 10 minutes. Toutes les catégories sont sensibles à l'exposition répétée. Les moments d'évaluation du cybermalaise étaient espacés d'environ 10 à 15 minutes. Il existait une certaine variabilité selon les sujets, mais on peut raisonnablement affirmer que tous les sujets rapportaient des symptômes significants après la deuxième session (environ 20 à 30 minutes d'exposition), ce qui est cohérent avec les données de la littérature [Stanney *et al.*, 2002].

### 3. CONCLUSION

Ces résultats montrent l'efficacité des visiocasques de nouvelle génération pour l'induction de présence chez leurs utilisateurs. Ils montrent aussi que, dans toutes les tâches où le comportement spatial des sujets est impliqué (c'est le plus souvent le cas), une représentation du corps du sujet (colocalisée dans l'espace et dans le temps)

semble fondamentale pour une interaction « efficace » (éviter de rentrer dans les portes dans notre cas).

Dans d'autres protocoles en cours dans notre équipe, en particulier de protocoles de thérapie comportementale par exposition à la réalité virtuelle [le journal du CNRS, 2015], voir figure 9.1, on observe que la vision du corps est fondamentale (ce qui, encore une fois, est naturel lorsqu'on utilise un Cave, figure 9.1). Cet aspect des choses rend l'utilisation d'un visiocasque moins aisée qu'on pourrait le penser initialement, car créer un corps virtuel spatialement et temporellement correct peut s'avérer techniquement complexe.

Deuxièmement, on a pu, lors de présentations publiques, entendre que la réduction de latence (effective) allait résoudre, entre autres, les problèmes de cybermalaise. Dans l'expérience rapportée ici, on voit qu'il n'en est rien, au moins qualitativement, puisqu'on voit que les symptômes de cybermalaise augmentent systématiquement lors d'une exposition répétée aux stimulations présentées dans le visiocasque. On pourrait observer que les valeurs atteintes sont moindres qu'avec des dispositifs type simulateurs de conduite [Mestre, 2016].

Cependant, pour conclure sur ce point, il faudrait tester des conditions expérimentales similaires dans un CAVE et avec un visiocasque. Dans l'expérimentation présentée ici, le sujet marchait réellement. Il n'y avait donc pas d'incohérence visuo-vestibulaire puisque l'environnement réel et l'environnement virtuel étaient « géométriquement identiques » (voir la solution S21 au chapitre 8). Nous pouvons cependant remarquer que le cybermalaise (voir figure 9.5) a « survécu » à la réduction de latence permise par le visiocasque utilisé, mais peut-être pas suffisamment pour supprimer les malaise dus à l'incohérence visuo-motrice temporelle. Il reste donc à évaluer d'autres facteurs de déclenchement du cybermalaise, par exemple la fatigue visuelle provoquée par le port de ce type de dispositif due aux possibles incohérences suivantes : Incohérence visuo-temporelle, Incohérence oculomotrice et Incohérence visuo-motrice de localisation (voir le chapitre 7).

## 4. BIBLIOGRAPHIE

[Cruz-Neira *et al.*, 1993] Cruz-Neira, C., Sandin, D. & DeFanti, T. Surround-screen projection-based virtual reality: The design and implementation of the CAVE, *Proceedings of SIGGRAPH'93*, 135-142.

[le journal du CNRS, 2015] *Journal du CNRS*. Guérir le vertige grâce à la réalité virtuelle. <https://lejournald.cnrs.fr/articles/guerir-le-vertige-grace-a-la-realite-virtuelle>

[Kennedy *et al.*, 1993] Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3, 203-220.

- [Lepecq *et al.*, 2009] Lepecq, J., Bringoux, L., Pergandi, J., Coyle, T., Mestre, D. Afforded actions as a behavioral assessment of physical presence in virtual environments. *Virtual Reality*, 13, 141-151.
- [Loomis *et al.*, 1999] Loomis *et al.*, 1999] Loomis, J.M., Blascovich, J.J. & Beall, A.C. Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments and Computers*, 31, 557-564.
- [Mestre, 2016] Mestre, D.R.. Perceptual calibration in virtual reality applications. Proceedings of Electronic Imaging 2016, *The Engineering Reality of Virtual Reality* (in press).
- [Mestre & Fuchs, 2006] Mestre, D.R. & Fuchs, P. Immersion et Présence. In P. Fuchs, G. Moreau, A. Berthoz, J.L. Vercher (Eds.), *Traité de la Réalité Virtuelle, Volume 1 L'homme et l'environnement virtuel*. Paris: Ecole des Mines de Paris. pp 309-338.
- [Mestre, Louison & Ferlay, 2016] Mestre, D.R., Louison, C., Ferlay, F. The contribution of a virtual self and vibrotactile feedback to walking through virtual apertures. *Proceedings of the 18th International Conference on Human-Computer Interaction* (in press).
- [Stanney *et al.*, 2002] Stanney, K.M., Kingdon, K.S., Graeber, D. & Kennedy, R.S. Human Performance in Immersive Virtual Environments: Effects of Exposure Duration, User Control, and Scene Complexity. *Human performance*, 15, 339-366.
- [Tarr & Warren, 2002] Tarr, M.J. & Warren, W.H. Virtual reality in behavioral neuroscience and beyond. *Nature Neuroscience*, 5, 1089-1092.
- [Warren & Chang, 1987] Warren, W., Whang, S.: Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*. 13, 371-383.



---

## Usages industriels des visiocasques

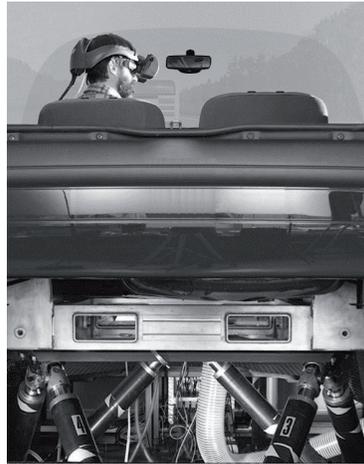
---

Andras Kemeny

### 1. INTRODUCTION

Si les premiers casques de réalité virtuelle apparaissent dès la fin des années 1960, leurs usages industriels larges ne voient le jour que dans les années 1990. En effet, le premier casque, développé par Ivan Sutherland (Sutherland, 1968) n'utilisait que des images file-de-fer. La génération d'images temps réel de qualité suffisante, c'est-à-dire texturées et anti-aliasées devient seulement accessible à la fin des 1980 avec des machines Silicon Graphics (Kemeny, 1987). Les cartes nVidia compatibles PC finissent par imposer la génération d'images de synthèses accessibles pour tous à la fin des années 1990, permettant ainsi le développement des visiocasques industriels, même si leurs usages restent confinés aux grands groupes, leurs prix restant jusqu'à récemment élevés.

L'industrie automobile s'intéresse très tôt à cette nouvelle technologie et Daimler compare son usage avec celui des projecteurs vidéo (dans un dôme de simulateur de conduite) dès 1997 (Schiller 1997) puis Volvo compare l'usage de la réalité virtuelle avec des essais réels avec un visiocasque nVision intégré dans un cockpit Volvo 850 et mesure le temps de saccade (Burns and Saluäär, 1999). Les deux expérimentations concluent une validité relative, mais l'usage réel ne démarre que quelques années plus tard, chez Renault, avec des visiocasques SEOS 120, spécifiquement développés dans le cadre d'un projet européen, CARDS (Coates *et al.* 2002). Une fois de plus, il s'agit d'une installation en simulation de conduite (Figure 10.1), annonçant une convergence progressive entre technologies de simulation de conduite et réalité virtuelle, depuis largement confirmé (Kemeny A., 2014).



**Figure 10.1 :** simulateur de conduite dynamique avec visiocasque SEOS 120

Chez Renault, après plusieurs tentatives d'intégration dans des simulateurs de conduite, y compris avec le visiocasque ProView VL50, ayant une résolution de 2.3 arc min à comparer avec la résolution de l'œil de 1 arc min, la conduite semble difficile, à la fois liée à des performances de génération d'images insuffisantes et de champ visuel instantané limité (Figure 10.2). Enfin, quelques années plus tard, dès 2004, à travers le projet MOER, d'abord avec un le visiocasque SEOS, puis Crescent et enfin nVisor, une application de réalité augmentée voit le jour et introduit la technologie de réalité virtuelle et augmentée dans le cycle de développement de véhicule pour les études et validation d'architecture de véhicule (Kemeny, 2008). Le visiocasque nVisor, représentait un compromis acceptable avec une technologie LCOS et stabilité colorimétrique acceptable et un champ de vision suffisamment large, permettant d'afficher les montants de baie simultanément. Le large champ de vision instantané est une exigence critique pour ce type d'application car il permet de simuler une exploration naturelle du poste de conduite, parfois avec peu de mouvement de tête et plus de mouvement de saccade avec les yeux et donc ne profitant pas du champ visuel dynamiquement 360° que les mouvements de tête autorisent (cependant contraints par des câbles vidéos et des gênes potentielles si le délais de réponse n'est pas suffisant).



**Figure 10.2 :** conduite sur conformateur avec visiocasque ProView XL50

## 2. SIMULATION DE CONDUITE ET RÉALITÉ VIRTUELLE

### 2.1. Convergence des deux domaines

La convergence entre les technologies de simulation de conduite et la réalité virtuelle démarre dès l'apparition de visiocasque, concomitant avec celle des simulateurs de conduite, les deux à la fin des années 1960, avec des simulateurs de conduite, il est vrai, employant encore des images de vidéo. Cependant, les deux domaines utilisent très vite les mêmes technologies, la génération d'images de synthèse, l'interactivité avec retour d'effort et produisant une immersion à la fois physique et mentale pour une utilisation efficace, ces critères étant considérés comme ceux de la Réalité Virtuelle (Shermann W. R. and Craig, 2003).

Le terme de Réalité Virtuelle est introduit par Jaron Lanier en 1987 et les premières installations de salles immersives (CAVE) apparaissent un peu plus tard (Cruz, 1992 et 1993). En même temps sont construits également les premiers simulateurs industriels utilisés déjà pour les études de sécurité automobile (Allen, 2011 ; Drosdol and Panik, 1985 ; Nordmark, 1994). Enfin, les premières installations CAVE arrivent en France au début des années 2000, notamment à l'Institut Image de Arts et Métiers en 2001, suivi rapidement par celui de PSA (Voillequin, 2006), alors que les premiers grands simulateurs commencent à être installés chez BMW, Renault et Volvo (Kemeny, 2000 et 2001).

### 2.2. Conflit visuo-vestibulaire

Si les installations de simulation de conduite et de réalité virtuelle utilisent les mêmes technologies, elles expérimentent les mêmes difficultés également, comme le mal de simulation ou VRISSE (Virtual Reality Induced Sickness Effect) lors des déplacements en situation de conduite ou en navigation utilisant les interfaces de commande de mouvement tel le joystick, wiimote, etc... Ceci est causé en particulier par le conflit visuo-vestibulaire, dû aux incohérences entre le système visuel et les systèmes vestibulaires, voir le chapitre 7, (Kennedy *et al.*, 2001).

### 2.3. Délais de réponse ou latence

Une autre cause peut être les délais de réponse, la latence, entre l'action et le rendu virtuel correspondant. Ce délai dépend de l'acquisition des mouvements de l'observateur (notamment de ses mouvements de tête), de la fréquence d'images et de la plateforme de simulation de mouvement et peut causer des incohérences perceptuelles fortes (Berthoz, 2002). Lorsqu'on observe une scène dont les images sont rendues en fonction des mouvements de tête, les délais de restitution peuvent induire des malaises liés au réflexe RVO (réflexe vestibulo-oculaire) qui est très rapide, approximativement 20 ms. Si les derniers visiocasques, tels que le HTC Vive, gèrent avec une rapidité de réponse suffisante cette latence (notamment grâce à une chaîne de calcul d'image

à 90 Hz), le délai de réponse global, notamment en conduite, peut encore poser des problèmes, en particulier lorsqu'une simulation de mouvement physique est fournie. Cet effet peut être encore plus important pour des usagers subissant ces mouvements, par exemple par des passagers en conduite sur simulateur (Dagdelen *et al.*, 2002), comme expliqué au chapitre 8. Pour éviter ces effets, la latence devra être inférieure à 50 ms en situation de conduite et à 20 ms pour des visiocasques ou salles immersives.

## 2.4. Perception de distance, de vitesse et d'accélération à l'échelle 1

La validité absolue en utilisation des systèmes de simulation n'est pas toujours requise. Par exemple dans l'étude de facteurs humains d'un système d'aide à la conduite, la validité relative peut suffire. Cela veut dire que si le comportement de conducteur et sa charge mentale peuvent se comparer valablement entre deux situations, les caractéristiques absolues de qualité de rendu ne sont pas nécessaires (Kemeny, 2009). Cependant la perception de la vitesse et de l'accélération est cruciale pour ces mêmes études, comme la charge mentale, l'attention du conducteur et son comportement en dépend (Barthou *et al.*, 2010). La perception de la vitesse dépend du champ visuel, à cause du rôle du champ périphérique dans la perception des mouvements mais aussi par le niveau d'immersion visuelle offert (Jamson, 2000). La perception correcte de l'accélération à partir d'un certain niveau (par exemple, en conduite sur route en virage ou lors des accélérations et/ou freinage) requiert l'emploi des plateformes à simulation de mouvement (Siegler *et al.*, 2001), ce qui est rarement possible avec des systèmes de réalité virtuelle de type salle immersive, mais compatible avec l'utilisation des visiocasques.

La perception des distances et de la taille des objets est essentielle dans l'utilisation des systèmes de réalité virtuelle, en particulier lors de l'évaluation des postes de conduite en revue de projet pour le style, l'ergonomie et les éclairages extérieurs ou intérieurs. Des études de perception ont démontré, dès le début d'utilisation des systèmes de réalité virtuelle, les différences de perception de distance et d'échelle entre les situations réelle et virtuelle, confirmées ensuite par des études plus récentes (Loomis and Knapp, 2003, Kemeny *et al.*, 2008).

L'un des plus importants indices de profondeur est la parallaxe de mouvement et son rôle a été démontré amplement en environnements réel comme en virtuel (Rogers and Graham, 1979; Panerai *et al.*, 2002; Kemeny and Panerai, 2003). Cependant le rôle précis de la vergence binoculaire (Paille *et al.*, 2005; Kemeny *et al.*, 2008) et des facteurs cognitifs (Glennerster *et al.*, 2008) restent à évaluer pour la perception de distance et d'échelle en environnement virtuel.

## 3. APPLICATIONS AVEC VISIOCASQUE OU AVEC SALLE IMMERSIVE

L'architecture de véhicule est l'un des domaines pour lesquels les techniques de réalité virtuelle et augmentée sont particulièrement adaptées. En effet, l'objectif est d'évaluer les choix d'architecture de poste de conduite, par exemple l'impact des montants de

baie sur la visibilité. Une vision immersive d'architecture 3D est nécessaire, à la fois avec une vision binoculaire pour estimer correctement les occultations des voitures et des autres usagers vus par le conducteur (Figure 10.3). Mais aussi une perception 3D des distances et des volumes est également critique pour bien apprécier l'espace conducteur ou conductrice en vue de juger l'habitabilité de l'intérieur du véhicule, par exemple, le volume de la boîte à gants. En même temps, la qualité de rendu visuel n'est pas aussi essentielle que pour les deux domaines « style » et « qualité perçue », pour lesquels un affichage de très haute qualité, avec une définition 4K peut s'avérer nécessaire.



**Figure 10.3 :** vision de l'intérieur de véhicule avec un visiocasque à large champ de vision

En l'absence de modèle de l'intérieur de véhicule et de vision immersive 3D, la seule possibilité est de faire appel à des prototypes physiques, alors que leurs diminutions en nombre, voire leur suppression, est l'objectif constant des constructeurs. Les évaluations peuvent alors de faire à partir des données CAD, par exemple Catia, aussi bien pour la visualisation de l'intérieur, y compris l'éclairage, que pour l'extérieur, y compris pour le style, lors d'une phase de début de projet (Combe *et al.*, 2008). Il est donc normal que l'usage des visiocasques a été déployé chez Renault dès 2004, comme un outil systématique dans l'évaluation des prestations d'architecture et d'ergonomie de véhicule. Cependant, la vision simultanée des montants de baies en vision stéréoscopique pour étudier leur impact visuel pour le conducteur, impose un large champ de vision, supérieur à 80°, avec un recouvrement de champ stéréoscopique d'au moins 50%. Pour y arriver, Renault a fait appel à l'utilisation des visiocasques SEOS (champ 120° x 40°), puis au visiocasque nVisor (SW111 – champ 102° x 64°), tout en testant d'autres solutions, comme par exemple les visiocasques Sensics, également utilisés par Airbus, mais plus difficiles à régler, à cause de son architecture d'éléments d'affichage multiple.

Les visiocasques de réalité augmentée (RA) trouvent une application majeure dans le domaine de l'architecture de véhicule. Par exemple, grâce à une composition de l'image d'environnement virtuel calculée avec la vue reproduite par une caméra d'un poste de conduite réel, on peut tester différentes solutions d'architecture de véhicule et comparer la vue extérieure, autorisée par le montants de baie avec différentes propositions de montant de baie (épaisseur, angle, forme, etc.). Ainsi on peut évaluer la prestation de vision et ceci sans utilisation de prototype physique. Cette technique

a été appliquée, grâce à l'utilisation de codes couleurs (permettant de distinguer le poste de conduite à afficher et l'environnement extérieur à remplacer par des scènes virtuelles avec du trafic), (voir Figure 10.5), et du logiciel D'fusion de Total Immersion chez Renault. En outre, la vue obtenue par des caméras attachées au visiocasque du véhicule réel peut être aisément comparée avec celle d'un véhicule numérique et ainsi comparer plusieurs versions de prestations de visions physiques et numériques.

L'avènement des salles immersives avec les technologies d'affichage 4K a relancé cependant l'utilisation des CAVE à la place des visiocasques. Renault vient d'acquérir un CAVE de très haut niveau dès 2012-2013 avec 5 faces 4K à 120 Hz (avec des projecteurs Sony SXRD multiples initialement, remplacés depuis par des projecteurs Christie), ce qui met l'entreprise à la tête des constructeurs automobiles, devant les allemands, Volkswagen et Daimler et les autres, de Ford à Jaguar. PSA Peugeot-Citroën suit en 2015, Renault en construit 2 autres et des CAVE à technologie 4K se répandent progressivement (Jaguar, Volvo, Renault, Ford, PSA...). La qualité d'affichage 4K avec un niveau de contraste élevé et l'introduction de l'ingénierie collaborative intra et inter CAVE rendent ces installations, malgré leur prix, très attractives. L'arrivée des visiocasques à bon niveau de qualité visuelle, comparable à celle des CAVE de génération ancienne, ou des nouveaux CAVE, à bas coût et avec de très courts délais de réponse comme Valve de HTC en 2016, annonce une évolution qui amènera probablement une utilisation partagée et complémentaire de ces technologies, dans des rapports pas encore établis.

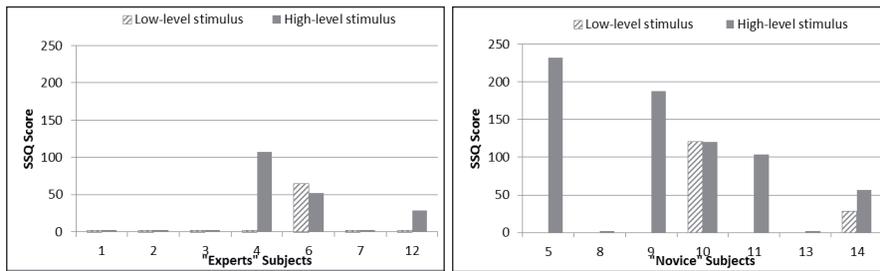
Le mal de simulation (cyber-sickness ou motion sickness) subit par les utilisateurs et en particulier par les utilisatrices (Flanagan *et al.*, 2006), restent un frein majeur de l'usage de réalité virtuelle en milieu industriel. En outre, le côté intrusif des visiocasques joue également un rôle important, à la fois par la restriction temporelle dans son usage à cause de l'inconfort généré et de la limitation dans les mouvements lorsque le casque dispose des câbles vidéo. Malheureusement la plupart des visiocasques de bonne qualité ne fonctionnent pas encore sans câbles. La taille des visiocasques pose un problème de mouvement et d'utilisation, notamment à l'intérieur d'un véhicule en exploitation RA.

#### **4. ÉTUDE SUR LE MAL DE SIMULATION (VRISE)**

Nous avons vu que le conflit visuo-vestibulaire joue un rôle prépondérant dans l'apparition de cyber-sickness, le délai de réponse global du système devant pouvoir se réduire à des valeurs faibles. Pour certains visiocasques, nous avons déjà des délais de réponse inférieurs à 20 ms pour la chaîne de calcul d'image et pour les systèmes plus complets des techniques de prédiction devraient progressivement réduire les effets indésirables correspondants. Cependant d'autres facteurs peuvent également jouer un rôle, tels que les attentes basées sur des expériences précédentes (Reason and Brand, 1975) ou d'autres paramètres physiologiques et psychologiques qui peuvent influencer sur la stabilité posturale. Une expérience passées, avec de maux de tête lors d'une séance de conduite antérieure avec des virages pris à grande vitesse peut induire une crainte

ou un réflexe conditionnel, rendant malade le sujet pour une conduite future dans une configuration à priori ne présentant pas de stimuli incohérents.

Parmi les études menées sur le mal de simulateurs, certaines visent à mesurer les seuils physiologiques au-delà desquels les mouvements pilotés n'induiraient pas d'effets indésirables. Une étude menée dans le CAVE IRIS à haute performance de Renault a comparé des niveaux de vitesse et d'accélération différents lors des déplacements pilotés. Le mal de simulateur a été quantifié avec des questionnaires SSQ (Kennedy, 1993), ainsi que la stabilité posturale avec une plateforme Techno-concept Stabilotest (Riccio, 1991).



**Figure 10.4 :** valeurs SSQ pour des rotations pour des sujets experts (gauche) et novices (droit)

Les niveaux d'accélération forte (jusqu'à  $13^\circ/s^2$ ) ont généré des valeurs SSQ significativement plus importantes que les ceux plus faibles ( $2^\circ/s^2$ ) et ceci pour les sujets experts et novices (Figure 10.4). En outre, nous pouvons constater une différence significative entre sujets experts et novices. Les premiers n'expérimentent presque jamais de malaises et si oui, seulement avec des niveaux d'accélération forte, les derniers sont presque toujours gênés avec des niveaux forts, mais rarement avec des niveaux faibles.

Il apparait ainsi que les sujets inexpérimentés sont souvent sujet au cybersickness si l'on ne sait pas générer des niveaux d'accélération adéquate, alors que les joueurs habitués et les sujets expérimentés sont en majorité exemptes de ces effets, même si pour ceux-ci mieux vaut utiliser des niveaux acceptables pour des durées d'expérimentations plus importantes.

Nous avons déjà noté que les sujets mâles sont moins sujets de ces effets et nous voyons que l'apprentissage joue également un rôle important. Cette capacité d'adaptation (voir le chapitre 8) semble être liée à une expérimentation connue. Chaque nouvelle situation nécessite un nouvel apprentissage ou la pratique de niveaux d'accélération adaptés pour éviter les effets indésirables (Reason, 1975). Ces résultats sont à compléter par des expérimentations avec des visiocasques, mais on peut supposer que des niveaux d'accélération comparables vont résulter dans des effets comparables, que ce soit pour des visiocasques de réalité virtuelle ou de réalité augmentée « see-through » (Nelson *et al.*, 2000). Des différences de perception peuvent être mesurées cependant entre visiocasques « optical see-through » et « video see-through » puisque pour ces derniers, le mixage des sources d'images de qualité différente peuvent également

induire des effets indésirables et même des écarts de perception en terme de distance et d'échelle (Kemeny *et al.*, 2008) (voir chapitre 8, § 2.3.1). Les visiocasques «optical see-through», utilisant des lames semi-transparentes pour l'affichage des images virtuelles, tout en permettant de voir les objets réels bien éclairés, souffrent également des limitations de visibilité, soit pour des objets pas assez éclairés, soit d'une lumière extérieure trop forte, écrasant les images virtuelles. Une calibration dynamique adaptée à un large éventail de condition d'éclairage reste encore à construire pour des usages industriels déployables.

## 5. ÉTUDE SUR LA PERCEPTION DE L'ÉCHELLE 1

L'un des enjeux majeurs dans l'utilisation des systèmes de réalité virtuelle ou de réalité augmentée est la perception correcte des échelles (Paille *et al.*, 2005; Kemeny and Panerai, 2003). Loomis and Knapp, 2003 reportent des écarts de perception de taille en environnement virtuel que l'on retrouve, bien qu'à des niveaux bien plus faibles quelques années plus tard, avec des mesures effectuées sur un écran stéréoscopique avec des champs de vision et des conditions d'observation différents (Paillé *et al.*, 2005). Il en ressort que la perception des distances est sous-évaluée en vision stéréoscopique et moins le champ de vision est large, plus cette sous-estimation est importante. Ces résultats sont cohérents avec ceux observés en environnement réel avec des champs de vision différents (Alfano and Michel, 1990; Hagen *et al.*, 1978; Watt *et al.*, 2000) bien que le débat reste ouvert en environnement virtuel (Knapp and Loomis, 2004).

En effet, la parallaxe de mouvement, indice essentiel de la perception de la profondeur, exploitable si on a la capture des mouvements de tête, est de plus en plus performant avec l'avancement de la technologie (Gogel and Tietz, 1973; Howard and Rogers, 1995; Kemeny and Panerai, 2003; Panerai *et al.*, 2002).

Une étude menée chez Renault a permis de comparer la perception d'échelle dans trois différents systèmes de réalité virtuelle (Figure 10.5):

- un écran cylindrique avec un champ de vision de 210° x 50° (Galaxy Warp Barco projectors), voir Figure 10.7.: SCREEN;
- un visiocasque avec un champ de vision de 120° x 40° (SEOS 120 HMD) équipé avec des caméras (Sony 768x582) alignées avec les yeux de l'observateur, utilisé en mode RA (réalité augmentée – video see-through) voir Figure 10.7: AR;
- Le même visiocasque utilisé en mode RV (sans vision du corps), voir Figure 10.7, et ce dernier avec (VR HT) ou sans (VR no HT) capture de mouvement de tête.

L'observateur devait comparer la taille perçue d'un poste de conduite de Renault Scénic, vu juste avant, avec des différentes tailles affichées, dans un ordre de présentation aléatoire (Figure 10.6).



**Figure 10.5:** observateur devant un écran cylindrique et avec un visiocasque VR ou AR

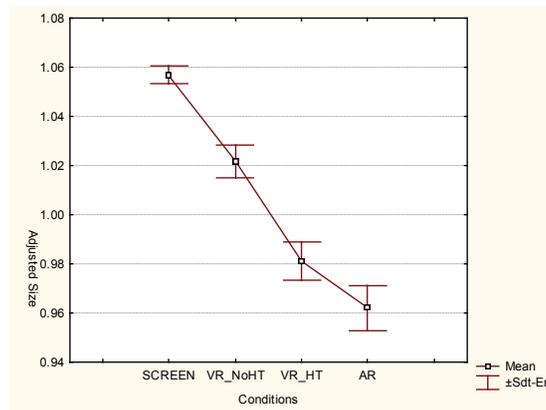
Les résultats montrent que la perception de l'échelle a été comparable avec un visiocasque de réalité virtuelle immersive, c'est-à-dire sans vision de l'extérieur (y compris son corps). Les tailles des éléments de poste de conduite ont été surestimées sur l'écran cylindrique et sous-estimées en condition de réalité augmentée, c'est-à-dire avec la vision de son corps.

Les moins bonnes estimations avec l'affichage sur l'écran cylindrique peuvent être expliquées par l'absence de capture de mouvement de tête (Rogers and Graham, 1975), due à l'absence de correction de distorsion temps réel (et donc un seul point optiquement correct de l'écran Szymanski, 2002).



**Figure 10.6:** cockpit virtuel : échelle 1, échelle inférieure à 1, échelle supérieure à 1

En revanche la faible performance d'estimation en condition de réalité augmentée semble contrintuitive, la vision de corps devant aider la bonne calibration de l'espace par l'observateur. Une explication peut être avancée par les différentes conditions de luminosités des sources d'images différentes. En effet il a déjà été montré en environnement réel l'effet de la luminosité sur la perception de distance (Coello and Greal, 1997). Une autre explication peut venir des traitements de vision stéréoscopique pour les images de caméra et les images calculées, les indices stéréoscopiques ainsi modifiées induisant de l'inconfort et une perception altérée (Hale and Stanney, 2006).



**Figure 10.7 :** perception d'échelle sur écran cylindrique, en réalité virtuelle sans et avec tracking et en réalité augmentée

Ainsi cette expérience met en évidence les problèmes liés à des traitements spécifiques d'un système de réalité augmentée «video see-through». Un système de réalité augmenté «optical see-through» n'aura pas les mêmes origines d'altération d'images. Cependant les écarts de luminosité et de qualité d'images seront toujours observés, induisant des altérations comparables. Les conflits entre vergence et accommodation, dus à la différence de distance binoculaire et d'accommodation, induisent également des effets de fatigue et de mal le simulateur potentiel (Hoffman *et al.*, 2008 ; Watt *et al.*, 2005). Enfin, le poids relatifs des différents indices restent encore à déterminer, étant impactés par les indices cognitifs associés (Bradshaw *et al.*, 2000 ; Glennerster *et al.*, 2006).

En conclusion, les visiocasques, qu'ils soient immersifs ou augmentés avec des caméras ou en vision directe, sont sujet encore à beaucoup de sources d'altération de qualité d'image ou de perception d'échelle. Comme expliqué au chapitre 7, ils sont susceptibles d'induire des effets d'inconfort ou de malaise visuellement induits, rendant leur utilisation industrielle complexe à mettre en œuvre.

## 6. BIBLIOGRAPHIE

- Alfano, P. L. and Michel, G. F. 1990, Restricting the field of view: perceptual and performance effects, *Perceptual and Motor Skills* 70, 35-45.
- Allen R. W., Rosenthal T. J., Cook M. L. 2011, A short history of driving simulators, In *Handbook of driving simulation for Engineering, Medicine and Psychology*, Fisher D.L., Rizzo M., Caird J.K., Lee J.D., Boca Raton, FL:CRC/Press Taylor and Francis, Ch. 2, pp. 2.1-2.16.

- Barthou A., Kemeny A., Reymond G., Merienne F., Berthoz A. 2010, Driver trust and reliance on a navigation system: effect of graphical display, In A. Kemeny, F. Mérienne, S. Espié (Eds.), *Trends in Driving Simulation Design and Experiments*. Les Collections de l'INRETS, pp. 199-208.
- Berthoz A. 2002, *The Brain's Sense of Movement*, Harvard University Press.
- Bradshaw M. F., Parton A. D., Glennerster A. 2000, The task-dependent use of binocular disparity and motion parallax information, *Vision Research* 40, 3725-3734.
- Burns P.C. and Saluäär D.1999, Intersections between driving in reality and virtual reality, *Proceedings of the Driving Simulation Conference 1999*, pp.155-164.
- Coates N., Ehrette M., Blackham G., Heidet A., Kemeny 2002, A Head-Mounted Display in Driving Simulation Applications in CARDS, *Proceedings of the Driving Simulation Conference 2002*, Paris, France, pp 33-43.
- Coello, Y. and Greal, M. A. 1997, Effect of size and frame of visual field on the accuracy of an aiming movement, *Perception* 26, 287-300.
- Cruz-Neira C., Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon and John C. Hart 1992, The CAVE: Audio Visual Experience Automatic Virtual Environment, *Communications of the ACM*, vol. 35(6), 1992, pp. 64-72.
- Cruz-Neira C., Daniel J. Sandin and Thomas A. DeFanti 1993, Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE, *SIGGRAPH'93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 135-142.
- Dagdelen M., Reymond G., Kemeny A. 2002, Analysis of the visual compensation in the Renault driving Simulator, *Proceedings of the Driving Simulation Conference*, Paris, France, 2002, pp 109-119.
- Drosdol J., Panik F. 1985, *The Daimler-Benz driving simulator, a tool for vehicle development*, SAE paper, n° 850334.
- Fang Z., Reymond G., Kemeny A. 2011, Performance identification and compensation of simulator motion cueing delays, *Journal of Computing and Information Science in Engineering*, Special Issue in Driving Simulation, 11(4): pp 1003-1 - 1003-4.
- Flanagan M.B., May J.G., Tobie T.G. 2006, Sex differences in tolerance to visually induced motion sickness, *Aviation Space and Environmental Medecine*, 2005, 76(7), pp. 642-646.
- Glennerster A., Tcheang L., Gilson S. J., Fitzgibbon A. W., Parker A. J. (2006): Humans ignore motion and stereo cues in favour of a fictional stable world. *Current Biology* 16(4), pp. 428-432.

- Gogel, W. C. and Tietz, J. D. 1973, Absolute motion parallax and the specific distance tendency, *Perception and Psychophysics* 13 (2), 284-292.
- Hagen, M. A., Jones, R. K., Reed, E.S. 1978, On a neglected variable in theories of pictorial perception: truncation of the visual field. *Perception and Psychophysics* 23, 326-330.
- Hale K. S., Stanney K. M. (2006): Effects of low stereo acuity on performance, presence and sickness within a virtual environment. *Applied Ergonomics* 37, 329-339.
- Hoffman, D. M., Girschick, A. R., Akeley, K., Banks, M. S. 2008, Vergence accommodation conflicts hinder visual performance and cause visual fatigue, *Journal of Vision* 8(3), 1-30.
- Howard, I. P. and Rogers, B. J. 1995, *Binocular vision and stereopsis*, Oxford University Press, New York.
- Jamson A. H. 2000, Driving Simulation Validity, *Proceedings of the Driving Simulation Conference*, Paris, France, 2000, pp. 57-64.
- Kemeny A. 1987, Synthèse d'images fixes et animées, *Techniques de l'ingénieur*, E 5 530, pp. 1-21.
- Kemeny A. 2000, Simulation and Perception of Movement, *Proceedings of the Driving Simulation Conference*, Paris, 2000, pp.13-22.
- Kemeny A. 2001, Recent developments in visuo-vestibular restitution of self-motion in driving simulation, *Proceedings of the Driving Simulation Conference*, Sophia Antipolis, France, 2001, pp.15-18.
- Kennedy, R. S., Lane, N. E., Grizzard, M. C., Stanney, K. M., Kingdon, K., & Lanham, S. 2001, Use of a motion history questionnaire to predict simulator sickness, *Proceedings of the Driving Simulation Conference*, Sophia Antipolis, France, pp. 79-89.
- Kemeny, A. and Panerai, F. 2003, Evaluating perception in driving simulation experiments, *Trends in Cognitive Sciences*, 7, pp. 31-37.
- Kemeny A., Combe E., Posselt J. 2008, Perception of Size in Vehicle Architecture Studies, *Proceedings of the 5th Intuition International Conference*, Torino, Italy.
- Kemeny A. 2009, Driving simulation for virtual testing and perception studies, *Proceedings of the Driving Simulation Conference Europe 2009*, Monte Carlo, pp.15-23.
- Kemeny, A. 2014, From driving simulation to virtual reality, *Proceedings of the VRIC'14 Virtual Reality International Conference*, Laval, France, 2014, art. No. 32.
- Kemeny A. Colombet F., Denoual T. 2015, How to avoid simulation sickness in virtual environments during user displacement, *Proceedings of conference: IS&T/SPIE*

*Electronic Imaging, The Engineering Reality of Virtual Reality*, volume 9392, pp. 939206.1-9.

- Kennedy R. S., Lane N. E., Berbaum K. S., Lilienthal M. G. 1993, Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness, *The International Journal of Aviation Psychology*, 3(3), pp. 203-220 (1993).
- Knapp, J. M. and Loomis, J. M. 2004, Limited Field of View of Head Mounted Displays is not the cause of distance underestimation in virtual environments, *Presence: Teleoperators & Virtual Environments*, 13, 572-577.
- Loomis, J. M. and Knapp, J. M. 2003, Visual perception of egocentric distance in real and virtual environments, In *Virtual and Adaptive Environments*, eds. L.J. Hettinger and M.W. Haas, Erlbaum, Mahwah NJ., pp. 21-46.
- Nelson W.T., Roe M. M., Bolia R.S., Morley R.M. 2000, Assessing Simulator Sickness in a See-Through HMD: Effects of Time Delay, Time on Task, and Task Complexity, *Proc. IMAGE Conference ASC-00-1047* (2000).
- Nordmark S. 1994, Driving simulators, Trends and Experiences, *Proceedings of the Driving Simulation Conference, Real Time Systems*, Paris, 1994, pp.5-13.
- Paillé D., Kemeny A., Berthoz A. 2005, Stereoscopic Stimuli are not used in Absolute Distance Evaluation to Proximal Objects in Multi-Cue Virtual Environment, *Proceedings of SPIE* Vol. 5664, pp.596-605.
- Panerai F., Droulez J., Kelada J-M., Kemeny A., Balligand E., Favre B. 2001, Speed and safety distance control in truck driving: comparison of simulation and real-world environment, *Proceedings of the Driving Simulation Conference*, Sophia Antipolis, France, 2001, pp.21-32.
- Panerai, F., Cornilleau-Peres, V. and Droulez, J. (2002), Contribution of extraretinal signals to the scaling of object distance during self motion, *Percept Psychophys*. 64, pp.717-731.
- Reason J.T., Brand J. 1975, *Motion Sickness*, London: Academic Press.
- Riccio G.E., Stoffregen T.A. 1991, An ecological theory of motion sickness and postural instability, *Ecological Psychology* 3, pp. 195-240 (1991).
- Rogers B., Graham M. 1979, Motion parallax as an independent cue for depth perception, *Perception* 8, pp. 125-134.
- Schiller V. *et al.* 1997, Car Research using virtual reality at Daimler-Benz, *Proceedings of the Driving Simulation Conference 1997*, Lyon, France, pp .35-44.

- Shermann W. R. and Craig A. B. 2003, *Understanding Virtual Reality*, Morgan Kaufmann, San Francisco.
- Siegler I., Reymond G., Kemeny A., Berthoz A. 2001, Sensorimotor integration in a driving simulator: contributions of motion cueing in elementary driving tasks, *Proceedings of the Driving Simulation Conference*, Sophia Antipolis, France, 2001, pp.21-32.
- Sutherland I. E. 1968, *A Head-mounted Three Dimensional Display*, FJCC Thompson Books, Washington, D.C., pp.757-764.
- Szymanski, M. 2002, *An introduction to Curved Screen Displays*, White paper. vRco, [www.vrco.com](http://www.vrco.com)
- Voillequin T. 2006, *First steps of Haptics at PSA Peugeot Citroën*, Special session of industrial applications of force feed-back, VRIC'06, Laval.
- Watt, S. J., Bradshaw, M. F., Rushton, S. K. 2000, Field of view affects reaching, not grasping, *Experimental Brain Research* 135, 411-416.

---

## L'usage des visiocasques pour les arts numériques

---

Judith Guez et Jean-François Jégo

La réalité virtuelle donne les moyens d'immerger l'homme dans un monde virtuel, tout en lui donnant la possibilité d'y interagir en temps réel. Du fait qu'il est possible d'y représenter un monde virtuel réaliste, symbolique et imaginaire, la réalité virtuelle inspire de nombreux domaines, il n'est alors pas étonnant de le retrouver dans les Arts. Ce médium permet au spectateur d'y plonger pour vivre des expériences interactives particulières et il ouvre le champ aux artistes à des explorations pour la création. Dans ce chapitre, nous décrivons en première partie l'apport des visiocasques dans la création artistique d'un point de vue historique et vis-à-vis des autres dispositifs. Nous proposons dans une deuxième partie une méthodologie de création d'installations artistiques de réalité virtuelle à partir des spécificités et propriétés des visiocasques.

### **1. L'APPORT DES VISIOCASQUES DANS LA CRÉATION ARTISTIQUE**

#### **1.1. La réalité virtuelle comme médium de création artistique**

##### **1.1.1. La relation du spectateur à l'œuvre dans l'histoire de l'art**

Les notions d'immersion et d'interaction inhérentes à la réalité virtuelle sont questionnées et expérimentées par les artistes avant même l'arrivée du numérique, en interrogeant des jeux sur les perceptions et des possibilités d'interaction avec l'œuvre. Nous proposons un survol rapide de l'emploi de ces notions dans les arts modernes et post-modernes.

À la fin des années 1950, un art environnemental sort l'art des musées, vient rencontrer le spectateur dans son lieu de vie, l'entourer et faire partie de son environnement. Il propose d'englober le spectateur dans un espace artistique : par des installations<sup>1</sup>, par notre espace naturel modifié<sup>2</sup>, le spectateur est ainsi au centre de l'œuvre<sup>3</sup>. Parallèlement, il se développe dans cette même période le courant de la participation

---

1 Par exemple l'installation : *Le Cyclop*, 1969-1994 de Jean Tinguely et Niki de Saint Phalle.

2 Citons les artistes du Land art : Mike Heizer, Dennis Oppenheim ou Dan Graham.

3 Isoler, transporter ailleurs, englober le spectateur, l'art environnemental « concerne l'agencement d'un espace intérieur, où l'objet d'art peut entrer en relation avec le milieu ambiant. » (Popper 1985) (Guez 2015).

où les artistes proposent d'impliquer davantage le spectateur en prenant directement en compte sa relation avec l'œuvre. L'œuvre devient plus accessible et proche du public (Popper 1985). La relation devient une forme artistique à part entière (Tramus 2001) qui est conceptualisée sous le nom d'esthétique relationnelle (Bourriaud 1998). Par exemple, les artistes de l'Op art et l'art cinétique explorent plus en profondeur la manière dont le spectateur peut être attiré par l'œuvre. Ils interrogent l'instabilité «des mécanismes perceptifs» (Couchot 1998), afin de se rapprocher du public. Dans l'œuvre de Marcel Duchamp *Rotary Glass Plates* (1920), le spectateur est invité à tourner la manivelle d'une machine pour qu'il puisse ensuite observer les effets moirés que le mouvement tournant provoque, établissant ici un jeu entre perceptions et action sur l'œuvre.

### 1.1.2. Art et réalité virtuelle

L'arrivée de l'analogique, notamment pour l'Art vidéo, va proposer d'autres moyens d'immerger le spectateur dans un environnement artificiel. Il va ensuite apparaître dans les années 1970 l'art interactif dans lequel le spectateur se retrouve face à des œuvres qui réagissent en temps réel par le biais d'interfaces d'abord analogiques puis numériques. Myron Kruger dans son installation *Vidéoplace* (1970) propose notamment au spectateur de dessiner en temps réel avec sa silhouette. Dans les arts, l'interaction prend alors plusieurs aspects et devient propice aux expérimentations artistiques.

Dès les années 1980, apparaît l'interaction dite de commande qui consiste, à permettre au spectateur d'actionner lui-même l'œuvre d'art numérique par déclenchement<sup>4</sup>, rappelant la démarche de Marcel Duchamp dans l'œuvre *Rotary Glass Plates* (1920). Puis dans les années 1990, l'interaction devient plus relationnelle en proposant des œuvres en temps réel dans lesquelles le spectateur peut modifier dans le temps de l'expérience l'œuvre et s'y inscrire. Des installations artistiques exploitant les dispositifs tels les visiocasques et les CAVE émergent alors. En 1995, l'installation *Osmose* de Charlotte Davies propose au spectateur grâce à une ceinture qui détecte le rythme respiratoire et un visiocasque, d'explorer en temps réel des paysages artificiels basés sur la nature et le vivant. Dans l'installation de Maurice Benayoun, *World Skin* (1997) utilisant un système CAVE, le spectateur qui prend la peau d'un photographe de guerre se retrouve à arracher des morceaux du monde virtuel à chaque fois qu'il le photographie à l'aide de l'interface «appareil photo» qui est suivie en temps réel. À la sortie de l'expérience, le spectateur peut récupérer une version imprimée de sa photo. L'artiste questionne ainsi le médium lui-même par son hybridation avec le réel.

Dans les années 2000, certains artistes, en prenant comme exemple le vivant, ont créé de nouvelles relations œuvres/spectateurs. «*Un art qui se vit*» (Guez 2013) apparaît alors, plus libre laissant place à l'improvisation. C'est ainsi que Michel Bret, Marie-Hélène Tramus en collaboration avec Alain Berthoz, définit une «seconde interactivité» avec *La Funambule virtuelle* (2000) (Bret, Tramus, and Berthoz 2005).

4 Notons l'œuvre de J-L Boissier, *Pékin pour mémoire*, 1986.

Cet être virtuel ne se contente plus de répondre aux stimuli du spectateur, mais il va prendre en compte son environnement et y répondre en temps réel selon sa perception et son apprentissage. C'est dans ce rapport dialogique et éactif qu'une relation va se construire entre l'œuvre et son utilisateur : espace d'émergence de nouvelles formes artistiques.

La réalité virtuelle s'avère être un formidable medium qui permet de penser et de façonner ces différentes formes d'immersion et d'interaction. Les artistes peuvent se jouer de son potentiel tout comme de ses limites qu'il convient de connaître pour mieux se les approprier, les contourner et, donc, les dépasser. Tout l'enjeu de la création est là !

## 1.2. Particularités des visiocasques vis-à-vis des autres interfaces pour la création artistique

La réalité virtuelle offre de nombreuses interfaces sensorielles, motrices et sensorimotrices (voir les chapitres 1 et 5) et chaque année, l'innovation numérique apporte des variantes qu'il est intéressant d'explorer. Nous notons l'arrivée sur le marché depuis la dernière décennie de nombreux visiocasques moins chers et donc plus accessibles. Ils ont chacun des spécificités et qualités différentes qu'il convient de connaître. En effet, outre les influences sur les aspects sensorimoteurs et cognitifs, dont la possibilité de PCV irréelles, ils offrent au niveau artistique des possibilités d'expériences bien différentes que celles rendues possibles par un écran ou un CAVE.

### 1.2.1. Isoler le spectateur du monde réel

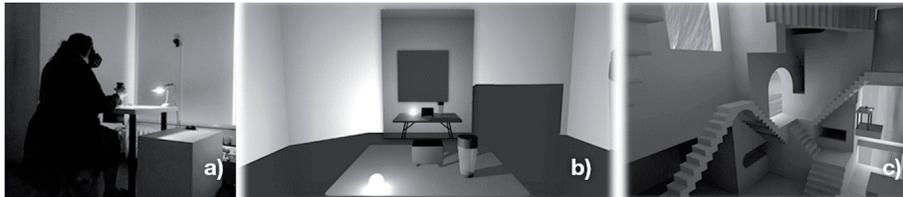
La première particularité du visiocasque est qu'il permet d'isoler complètement le spectateur aux niveaux visuel et sonore vis-à-vis de l'environnement réel extérieur. Le sujet peut être totalement immergé à 360 degrés dans l'environnement virtuel. Les capteurs de position et d'orientation de tout le corps, de la tête, des mains et éventuellement des pieds, permettent à l'utilisateur d'explorer l'environnement virtuel autour de lui de manière pseudo-naturelle. L'ajout de la stéréoscopie permet des effets de profondeur et de jaillissement de l'image qu'il convient de bien doser.

Cet isolement peut permettre au spectateur de se sentir plus « libre » de vivre l'expérience à huis clos, dirigeant lui-même son observation, sans le regard des spectateurs extérieurs. L'installation *InterACTE* (Batra, Jégo, Guez, Tramus, 2015) propose au spectateur d'interagir avec un Pierrot virtuel en deux actes. Dans le premier acte, il s'agit d'improviser par des gestes des bras avec l'ombre du personnage virtuel qui y répond (figure 11.1a). Dans le second acte, le spectateur se retrouve plongé dans le monde virtuel du Pierrot (figures 11.1b et 11.1c). Les auteurs ont noté que les spectateurs semblent lâcher prise davantage lorsqu'ils improvisent gestuellement en immersion dans un visiocasque d'Oculus que lorsqu'ils interagissent avec l'ombre seule, percevant ainsi le public qui les observe (Batra *et al.* 2016).



**Figure 11.1 :** installation *InterACTE* ; a) improvisation avec l'ombre du Pierrot Virtuel, b) improvisation en immersion dans l'univers virtuel du personnage, c) vue de l'intérieur du casque

Dans l'installation *Lab'Surd : le LABORatoire de la SURvirtualité* (Guez, Bertinet, Wagrez, Costes, 2015), le spectateur est immergé par le visiocasque dans une salle virtuelle qui ressemble à la salle réelle dans laquelle il s'est assis pour mettre le casque (figure 11.2). Une fois le casque enfilé, le spectateur est transporté dans un autre monde, qu'il est libre d'explorer (a). Ce monde virtuel qui au début représente de façon réaliste une partie de la pièce réelle dans laquelle il se trouve (b), va petit à petit se transformer, se déconstruire dans le temps de l'expérience, jusqu'à atteindre un environnement plus spacieux à l'architecture imaginaire paradoxale (c). Le spectateur équipé d'un casque audio entend une musique tout au long de l'expérience qui – en cohérence avec les éléments visuels – a pour objectif de l'emporter plus loin dans l'imaginaire. De plus, du son spatialisé est utilisé au cours de l'expérience pour guider le spectateur et l'inciter à observer autour de lui. Le son se révèle être un élément important à exploiter pour diriger le regard et la narration.



**Figure 11.2 :** installation *Lab'Surd : le LABORatoire de la SURvirtualité* ; a) scénographie de l'installation ; b) vue de l'intérieur du casque en début d'expérience ; c) vue de l'intérieur du casque en fin d'expérience

### 1.2.2. Le spectateur dans un corps invisible

À noter qu'il existe une différence fondamentale entre un visiocasque et un système de projection vidéo ou CAVE : l'utilisateur ne voit pas son propre corps. Le sujet immergé peut alors s'aider des autres sens comme le toucher et la proprioception pour percevoir son propre corps ou l'environnement extérieur par palpation ou par tâtonnement (voir chapitre 2 pour une description détaillée). Il peut être intéressant artistiquement de pouvoir créer un jeu entre ce que voit et ce que ressent le spectateur notamment entre le visuel et la proprioception. Il faut cependant faire attention aux conflits sensorimoteurs (notamment entre la vision et les systèmes vestibulaires) qui peuvent être à l'origine d'inconforts ou de malaises (voir chapitre 7).

Dans l'installation *Lab 'Surd* précédemment décrite, le spectateur est assis sur une chaise et place le casque de réalité virtuelle. Dans le monde virtuel, il ne voit plus son corps. Les auteurs ont observé que ce sentiment est ressenti comme étrange au début de l'expérience<sup>5</sup>. Cependant, il a été proposé au spectateur de tenir un verre qui est réel et virtuel et qui suit les mouvements de sa main. Grâce à lui, il peut avoir une action sur l'environnement, car le verre prothétique devient une extension visible de son corps invisible dans ce nouveau monde artificiel. On note que, très vite, dès la prise en main du verre, le spectateur retrouve ses repères et accepte l'absence de son corps dans le virtuel. Ce jeu entre le visible et l'invisible, et entre le réel et le virtuel, renforce l'aspect surréaliste et mystérieux.

### 1.2.3. Corps virtuel ou accepter un autre corps

Afin de pallier l'absence de corps, une solution simple est de placer le spectateur dans un corps virtuel ressemblant ou différent du sien, à la manière d'une nouvelle peau ou d'un scaphandre. Ce corps virtuel sera d'autant plus accepté qu'il suivra les mouvements du spectateur au moyen d'un système de capture de mouvements. À noter que le corps virtuel n'a pas d'obligation de ressemblance littérale avec celui du spectateur qui va très vite «accepter» et adopter cette nouvelle enveloppe ou avatar comme le montrent les travaux de (Slater *et al.* 2009).

Dans l'installation artistique *Liber* (Guez, Jégo, Wagrez *et al.*, 2015), le spectateur se retrouve dans la peau d'un être humanoïde et végétal face à un miroir virtuel. Les bras et la tête du spectateur sont suivis en temps réel lui permettant de résoudre des énigmes en touchant des tatouages disposés sur son corps virtuel. Le spectateur se réfère alors à sa vision, sa proprioception et le toucher de son propre corps pour interagir. Ces jeux de présence-absence de corps et de ressemblance-différence par rapport à son propre corps permettent différentes variantes et se révèlent être des principes intéressants à explorer dans la création.

### 1.2.4. Possibilités d'interaction avec le visiocasque

Les visiocasques sont le plus souvent équipés d'un système de suivi de la tête de l'utilisateur. Ce système de suivi de la tête peut être exploité seul, à des fins d'interaction, ce qui permet de ne pas surcharger l'utilisateur d'interfaces.

#### *Navigation dans un monde virtuel sans déplacement dans le monde réel*

Le suivi de la rotation de la tête peut être un moyen simple de naviguer en environnement virtuel, notamment si la caméra virtuelle suit un parcours prédéfini. Par exemple, dans l'œuvre *I need a Haircut* (Fleuryfontaine, 2014), la caméra est sur un rail qui boucle dans le monde virtuel. Le spectateur peut quand il le souhaite

---

5 Cf. Analyse de l'installation *Lab 'Surd* dans (Guez 2015).

mettre le visiocasque et plonger dans l'univers de l'artiste à n'importe quel moment de la progression de l'expérience. Il décide également de lui-même quand enlever le casque et quitter l'environnement virtuel. Cette installation s'affranchit d'une interface pour la navigation, l'aspect interactif ne résidant ici que dans l'observation. Ceci se révèle être très similaire aux nombreuses installations immersives d'observation de vidéos 360 degrés.

Dans le jeu artistique musical *PlayHead* (Auxietre *et al.*, 2014) utilisant le visiocasque Samsung GearVR, toutes les interactions dans le monde virtuel se font grâce à l'orientation de la tête, à l'aide d'un viseur qui permet plusieurs types d'actions (enclencher des boutons, moduler des pistes musicales, etc.).

De même dans l'installation *5<sup>e</sup> sommeil* (Auxietre 2014), l'artiste propose une exploration artistique de l'intérieur d'un corps humain. Le spectateur est invité à observer autour de lui et à diriger la capsule virtuelle dans laquelle il se déplace, en tournant la tête vers le lieu où il souhaite avancer.

#### *Déplacement dans un monde virtuel identique géométriquement au monde réel*

Certaines installations avec un visiocasque ont un espace de captation qui peut permettre un déplacement du spectateur sur une surface pouvant aller de 2 à 9 mètres carrés. Le spectateur peut ainsi se déplacer librement et marcher dans l'espace virtuel identiquement à sa marche dans l'environnement réel. Il n'y a plus dans ce cas d'incohérence visuo-vestibulaire (solution S21, voir chapitre 8).

#### *Action et commande*

Certains visiocasques comme le Samsung GearVR possèdent des boutons qui peuvent être utilisés pour effectuer des actions dans le monde virtuel (presser des boutons virtuels, actionner ou manipuler des objets virtuels, naviguer, etc.) ou du contrôle d'application (relancer ou quitter l'application sans enlever le visiocasque). À noter que le micro du casque est aussi utilisé pour déclencher des événements. Dans le jeu *Panopticon* (Spiers *et al.* 2015), il faut siffler pour contrôler les prisonniers d'une prison. Le micro peut être employé plus traditionnellement à des fins de communication avec des agents virtuels ou avec autrui.

Au-delà de la navigation, le tracking du visiocasque peut servir pour avoir une estimation de ce que voit le spectateur, et ainsi en jouer. L'expérience *Sightline: the Chair* (Mariancik 2013) pour Oculus Rift exploite cette technique pour faire changer les objets et le décor lorsque l'utilisateur ne les regarde pas, permettant d'avancer par étape dans la narration vers différents environnements qui se modifient. Ce procédé a également été utilisé dans l'installation *Lab'Surd* pour faire déplacer des objets quand le spectateur ne les regarde pas, en jouant sur des effets visuels, pour mieux le surprendre.

À noter qu'il existe quelques visiocasques ayant un tracking des yeux (voir chapitre 6) permettant cette fois d'avoir une visée directement par le suivi du regard plutôt qu'en tournant la tête. Cela donne des perspectives d'interaction intéressantes pour la création.

## 2. UNE MÉTHODE POUR CRÉER AVEC UN VISIOCASQUE

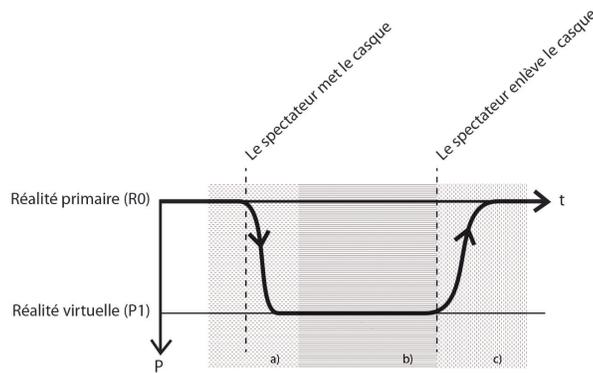
### 2.1. Introduction

Il convient d'utiliser le visiocasque en fonction des besoins de l'expérience et qu'il s'intègre dans la narration désirée par l'artiste. Chaque artiste peut construire une expérience qui correspond à son univers. Les œuvres utilisant les deux fonctionnalités spécifiques du visiocasque (immersion visuelle et suivi du regard) sont ainsi très différentes. La manière de construire le scénario dans le temps et l'espace de l'œuvre restent à imaginer. Cependant, nous notons de l'observation et l'analyse d'œuvres utilisant des visiocasques qu'il existe des schémas récurrents. Nous proposons de les détailler dans les prochaines parties.

Première considération : c'est le spectateur qui vit l'œuvre. Il convient alors de concevoir l'expérience selon son point de vue. La présence ressentie lors du vécu dans un monde virtuel ou mixte peut ainsi être conceptualisée « *comme forme artistique de la relation œuvre-spectateur* », et prendre ainsi en compte plusieurs problématiques de création (Guez 2015), à savoir :

1. Comment capter le spectateur dans son premier contact avec l'œuvre ?
2. Comment entretenir ce contact ?
3. Comment il en sort ? Et quelles influences alors cette expérience pourrait avoir sur lui ?

Pour y répondre, nous proposons trois étapes de conception : le protocole d'entrée (a), l'expérience (b), le protocole de sortie (c) (figure 11.3). Ces étapes peuvent se schématiser à l'aide du « *modèle de plongée dans le monde virtuel* » (Guez 2015), pour servir de guide pour la création artistique. L'illustration suivante représente comment une expérience peut être conçue dans le temps (t), en partant du point de vue du spectateur. Elle permet d'identifier les moments de passage entre le vécu du spectateur dans sa réalité habituelle « primaire », et son vécu dans la réalité virtuelle. Ici, le visiocasque permet d'être ce point de passage entre les deux mondes : de permettre la plongée dans les profondeurs (P) d'un autre monde réaliste, imaginaire ou symbolique.



**Figure 11.3 :** schéma de plongée dans le monde virtuel par l'intermédiaire du visiocasque (Guez 2015); t: le temps; P: profondeur dans les illusions réelles/virtuelles; a) protocole d'entrée; b) expérience : présence dans le monde virtuel; c) protocole de sortie

À noter que ce schéma a vocation d'être une aide empirique à la conceptualisation d'une expérience utilisant le visiocasque. Il permet de visualiser dans le temps comment le spectateur pourrait vivre une expérience. Bien évidemment, chaque spectateur peut vivre l'expérience à sa manière. Il est alors intéressant de réaliser des retours d'expériences afin de comparer les attentes voulues dans l'intention artistique lors de la création, et le réel impact de l'œuvre. Ceci se révèle être une source inépuisable d'inspiration pour continuer à explorer le médium. Nous décrivons en détail dans la partie suivante comment s'approprier ces trois étapes, à savoir : le protocole d'entrée (a), l'expérience (b) et le protocole de sortie (c).

## 2.2. Protocole d'entrée : les aspects scénographiques et le début de la narration

Le premier contact avec l'œuvre, c'est le moment où le spectateur va passer de sa réalité habituelle à la «réalité de l'œuvre». Il est important, car il permet de rendre crédible le contexte de l'expérience et d'inviter la personne à accepter de s'y plonger, selon le principe de la suspension consentie de l'incrédulité (Coleridge 1817).

### 2.2.1. Créer la scénographie réelle de l'œuvre

L'espace dans lequel s'inscrit l'œuvre est un élément essentiel : il convient de penser comment le spectateur va se placer, comment il arrive dans l'espace de l'œuvre, comment il progresse (physiquement), et enfin comment il va en sortir. Par exemple, si l'installation prend en compte seulement le vécu dans le monde virtuel en s'affranchissant du monde réel, une chaise seule dans un espace calme peut suffire. Si en revanche un décor réel est pensé autour de l'expérience virtuelle, le visiocasque doit pouvoir y avoir pertinemment sa place.

Nous proposons trois manières d'imaginer la progression physique du spectateur dans l'œuvre :

- La position allongée va favoriser un positionnement statique et plus enclin à des expériences contemplatives ;
- La position assise va favoriser peu de déplacement : les mouvements pourront être plus concentrés par le haut du corps (bras, tête) : manipulations, explorations, etc. ;
- La position debout va favoriser un déplacement dans l'espace de tout le corps : un cheminement, une recherche, etc.

Et voici quelques exemples de scénographies possibles associées :

- La personne est debout/assise ou allongée et enfle un casque qui est directement visible dans le lieu de l'exposition. Elle est moins isolée du monde réel et les autres spectateurs peuvent la voir ;
- La personne est debout/assise ou allongée et enfle un casque situé dans un espace isolé dans le lieu de l'exposition (rideaux, salle, boîte, etc.). Seul un spectateur après l'autre peut rentrer dans le lieu de « passage » vers le virtuel ;
- La personne est debout/assise ou allongée dans un décor réel scénographié selon la narration voulue, dans lequel est installé le visiocasque.

À noter que le visiocasque a ses contraintes technologiques qui sont à intégrer dans cet espace selon le parti pris artistique. Il convient selon le type de visiocasque de penser ces variantes de scénographies : penser à la longueur des câbles, au placement des caméras de suivi de mouvement ou des autres interfaces, etc.

### 2.2.2. Prendre en compte le lieu de l'exposition

Le lieu d'exposition de l'œuvre est également à prendre en considération. En effet, si l'installation est réalisée pour un lieu en particulier, elle peut s'inscrire dans l'architecture du lieu. Et si elle est ensuite exposée autre part, la scénographie se verra probablement être modifiée.

Aussi, la manière de placer le visiocasque vis-à-vis des autres œuvres et de l'espace d'exposition est à prendre en considération.

La présence ou l'absence de médiateur est importante : l'installation nécessite-t-elle une médiation pour notamment montrer à la personne comment placer le visiocasque ? L'installation est-elle autonome ? Dans tous les cas un mode d'emploi ou une fiche technique récapitulative s'avèrent toujours pratiques lorsque l'installation n'est pas présentée par son auteur.

### 2.2.3. Créer une scénographie propice au passage du réel au virtuel

Revenons sur les différents aspects scénographiques liés à l'espace d'exposition de l'œuvre. Cet espace va permettre le passage du monde réel au monde virtuel, c'est-à-dire de plonger dans l'expérience de réalité virtuelle (*cf.* schéma de plongée, figure 11.3). Il convient de créer un lieu propice à ce passage, un lieu qui peut attirer le spectateur, l'interroger, voire de commencer à lui faire vivre un contexte particulier lié à l'œuvre. Ce lieu peut-être :

- Soit complètement réel ;
- Soit réel et virtuel, avec par exemple de la projection vidéo sur des objets réels ;
- Soit majoritairement virtuel : par exemple le spectateur porte un visiocasque au milieu d'une projection vidéo (systèmes Wall ou CAVE, voir figure 11.1 *InterACTE*).

Le passage du réel au virtuel, qui se fait notamment par l'intermédiaire du visiocasque, peut être pensé en rendant cohérents l'espace réel et l'espace virtuel. Lorsque le spectateur place le visiocasque, tout est à imaginer selon l'intention artistique, il peut :

- Soit créer un passage progressif en s'aidant de la ressemblance entre l'environnement réel/virtuel ;
- Soit créer un passage progressif en s'aidant d'un corps virtuel qui vient remplacer le corps réel du spectateur ;
- Soit créer un passage abrupt en ne créant aucune correspondance entre les deux environnements.

## 2.3. L'expérience : entretenir la relation du spectateur avec l'œuvre

Une fois que le premier contact est établi avec l'œuvre et que le spectateur a mis le casque, il se retrouve au cœur de l'expérience. Il semble nécessaire de garder une cohérence afin qu'il reste dans l'expérience le temps décidé par l'artiste.

### 2.3.1. Les différentes formes de narration de l'expérience

Tout un scénario peut être à construire afin d'emporter plus loin le spectateur dans l'œuvre. Ce scénario peut prendre plusieurs formes :

- Soit suivre un schéma narratif linéaire ou à choix multiples, dont la trame est bien définie dans le temps de l'expérience ;
- Soit n'avoir aucun temps prédéfini : le spectateur dans l'interaction peut choisir librement du temps à vivre dans l'expérience.

À noter que la forme de l'expérience peut être abstraite, laissant ainsi au spectateur le choix d'explorer, d'errer, de contempler ou d'improviser, etc.

### 2.3.2. Garder la présence dans le monde virtuel

Le concept de présence, très étudié dans le domaine de la réalité virtuelle, permet de donner des pistes pour prolonger l'expérience in virtuo. Cette «*illusion d'être dans un monde alternatif*», tel que le définit Jaron Lanier (Lanier 1990) va permettre de se focaliser sur l'aspect subjectif du vécu du spectateur dans l'œuvre. La prise en compte des nombreuses études sur la présence en réalité virtuelle (Slater 2003) (Nannipieri *et al.* 2014) (Wiederhold and Bouchard 2014) (Bouvier 2009) (Guez 2015) va permettre de fournir des outils intéressants dans la création d'expériences. Dans le domaine, il est habituellement conseillé de conserver cette présence à un niveau élevé – bien que cela reste difficile pour des raisons techniques ou cognitives – pour que la personne focalise ainsi son attention<sup>6</sup> sur son vécu dans le monde virtuel.

### 2.3.3. Jouer avec le concept de présence

Il est habituel de ne pas créer de « cassures de présence », c'est-à-dire de faire revenir l'attention du spectateur sur le monde réel. Cependant, il peut être intéressant de jouer sur ce concept de présence. Les travaux de Mel Slater (Kiltner *et al.* 2012) donnent de bons exemples quant aux possibilités de jeux d'illusions dans le monde virtuel. D'un point de vue artistique, il peut être pertinent de créer des cassures volontaires pour ainsi interroger cette oscillation d'attention entre le monde réel et le monde virtuel. Par exemple, dans l'installation *Lab'Surd*, le spectateur passe de mondes virtuels en mondes virtuels. L'enjeu se situe ici dans la création de passages cohérents entre les mises en abîme de vécus possibles, et donc sur le jeu de « cassure de présence » pour passer d'un environnement à un autre. Le schéma (figure 11.3) qui nous a servi de référence peut par ailleurs être utilisé pour aller plus loin par exemple dans la conception de ces jeux de mise en abyme<sup>7</sup>.

## 2.4. Protocole de sortie

La fin de l'expérience n'est pas à négliger. Elle clôturera l'expérience et peut être un élément important de l'œuvre.

---

6 Cf. D.Mestre et P. Fuchs. Chapitre 14, « Immersion et présence » ; et M. Auvray chapitre 7 « Théorie de la perception : le monde comme mémoire externe » dans : P. Fuchs, G. Moreau, A. Berthoz, J.-L. Vercher, *Le traité de la réalité virtuelle*, volume 1, 2006. Les auteurs identifient le passage du monde réel vers le monde virtuel à un déplacement de l'attention du sujet. Ils postulent que c'est l'attention qui module la perception. Malika Auvray, dans le traité étudie cette attention qui peut être par exemple facilement dérivée par notre faculté à être aveugle à certains changements.

7 Cf. différents exemples de créations et de schémas de plongés lié au concept d'oscillation entre le réel et le virtuel, et de profondeur dans les illusions (Guez 2015).

### 2.4.1. La scène de fin : le passage du virtuel au réel

En comparaison au protocole d'entrée qui permet le passage du réel au virtuel, le protocole de fin va permettre le passage du virtuel au réel. Il va créer le lien entre le dernier moment vécu dans le monde virtuel et le retour au monde réel. Selon l'intention artistique, plusieurs types de sorties sont possibles. À la manière d'un film, la scène de fin peut être particulière, vectrice d'une émotion, comme un aboutissement, et finir par exemple avec un écran noir. Mais elle peut aussi ne pas être différente du reste de l'expérience et finir quand le spectateur le souhaite. Le retour peut ainsi être plus ou moins progressif ou abrupt.

À noter par ailleurs que la sortie d'une expérience de réalité virtuelle peut dérouter le spectateur, avec une certaine perte de repères pouvant conduire à des effets secondaires, dus à la réadaptation sensorimotrice au monde réel après que la personne se soit adaptée aux incohérences sensorimotrices de l'environnement virtuel. Le protocole de sortie peut aider à éviter cela en aidant à revenir progressivement à la réalité, en prenant par exemple le temps de se lever tranquillement.

La scénographie peut aussi accompagner la sortie, tout comme elle a invité le spectateur à y entrer. Un son ambiant différent ou une lumière plus feutrée qu'au début de l'expérience peuvent être des éléments de transition au monde réel. Enfin, l'accompagnement par l'auteur, ou par un acteur dans le cas d'une performance, ou encore d'un médiateur, peuvent faciliter dans les meilleures conditions le « réveil » du retour au réel.

### 2.4.2. Retour à la réalité et collecte des impressions

Ça y est ! Le spectateur a enlevé le visiocasque. Il quitte l'espace de l'installation. Mais ce n'est pas fini. Tout juste sorti de l'installation, le souvenir de son vécu est encore bien là. Selon comment il a vécu ce moment, il sera plus ou moins enclin à avoir envie de raconter ce qu'il a vécu. Le partage de son expérience vécue peut-être porteur de précieuses informations pour l'artiste. En effet, elles peuvent permettre par ces impressions relevées de continuer l'exploration du médium et de la démarche artistique. La méthode qualitative de l'entretien d'explicitation (Vermersch 2012) est un moyen récent, mais déjà éprouvé pour laisser émerger de nouveaux chemins d'explorations<sup>8</sup>. Il permet d'accompagner le spectateur dans la mémoire de son vécu et de relever ainsi des traces sensibles de l'expérience.

---

<sup>8</sup> Cf. Méthode d'analyse des retours du public sur la base d'enregistrement d'entretien d'explicitation (Guez 2015).

### 3. CONCLUSION ET PERSPECTIVES

La réalité virtuelle est un médium offrant de riches possibilités pour les artistes. Le visiocasque permet d'emporter le spectateur dans des mondes inventés ou réinventés, dans lesquels l'espace et le temps peuvent être complètement modifiés. Il y est aussi possible d'expérimenter d'autres manières de ressentir notre réalité et de questionner ainsi nos propres perceptions et le monde dans lequel nous vivons.

Nous avons principalement pris en exemple des installations artistiques de réalité virtuelle, mais le visiocasque est également utilisé dans les arts de la scène. Par exemple, dans la pièce de théâtre *Cassandre-Matériaux* (Chabalier *et al.* 2016)<sup>9</sup>, l'actrice utilise le visiocasque pour traduire les dires de la prophétesse grecque Cassandre, assaillie de visions du futur. Le visiocasque représente alors ce point d'entrée vers le monde de ces visions qu'elle seule peut voir. Il se crée alors un jeu de mise en abîme sur scène, où l'actrice réelle se retrouve actrice du monde virtuel dans lequel elle improvise avec un acteur virtuel (figure 11.4). Cela questionne la présence même des acteurs virtuels visibles ou cachés du public, et des jeux scéniques possibles.



**Figure 11.4** : pièce de théâtre *Cassandre-Matériaux*. Ici l'actrice (au centre) est en immersion dans un casque Oculus Rift et interagit en temps-réel par ses gestes avec un acteur virtuel. Cet acteur virtuel est également rétro-projeté sur un tulle pour être visible du public (à droite) qui peut percevoir ainsi les « visions » de l'actrice. Crédit photo : Emmanuel Valette.

Le monde du cinéma s'intéresse également à ce médium, notamment au travers du développement de la captation et de la visualisation de vidéos 360 degrés (voir le

<sup>9</sup> *Cassandre-Matériaux*, février 2016 théâtre de la Commune Aubervilliers; d'après Alexandra de Lycophron de Chalcis, traduit du grec ancien par Pascal Quignard. Conception, adaptation et mise en scène Clara Chabalier. Avec Clara Chabalier, Venia Stamatidi, Jules Turlet (chansigne), Alvisé Sivinia (Musicien). Assistant à la mise en scène Thomas Morisset. Création lumière: Emmanuel Valette. Sound-design Vicky Bisbicki. Recherche et conception sur l'interaction et l'avatar: Jean-François Jégo, Judith Guez, Dimitrios Batras. Médiatrice en langue des signes française Claire Danet.

chapitre suivant). Le réalisateur Chris Milk réalise depuis 2013 des films de ce type, à expérimenter au moyen de visiocasques<sup>10</sup>.

Il y a également de nombreuses possibilités à partir de captures vidéos temps réel : par exemple, l'expérimentation *The Machine to be another* (2015) du BeAnotherLab propose de se glisser littéralement dans le corps d'un autre, en montrant un flux vidéo correspondant à son point de vue. Ces images réelles peuvent aussi être hybridées avec des éléments virtuels (voir le domaine de la réalité augmentée).

Ainsi, la création artistique permet de s'affranchir de nombreuses barrières et d'explorer de nouveaux usages via le vécu d'expériences inédites. L'artiste peut conceptualiser toutes sortes d'approches et de rêves. Il peut remettre en doute l'utilisation même d'un tel dispositif en interrogeant, par exemple, l'objet (esthétique) « casque », les immersions et les interactions possibles, et les hybridations de matériaux (tangibles ou immatériels). Il convient alors à l'artiste de définir librement le cadre et les actions de « son Art ».

#### 4. BIBLIOGRAPHIE

- [Batras *et al.*, 2016] D. Batras, J-F. Jégo, M-H. Tramus and J. Guez, *A Virtual Reality Agent-Based Platform for Improvisation between Real and Virtual Actors Using Gestures*. ACM VRIC Laval France.
- [Bourriaud, 1998] N. Bourriaud, *L'esthétique relationnelle*, Les Presses du réel.
- [Bouvier, 2009] P. Bouvier, *La présence en réalité virtuelle, une approche centrée utilisateur* Thèse, Université Paris-Est, dir. Gilles Bertrand.
- [Bretet *et al.*, 2005] M. Bret, M-H. Tramus and A. Berthoz. « Interacting with an Intelligent Dancing Figure : Artistic Experiments at the Crossroad between Art and Cognitive Science », *Leonardo for the Art Sciences and Technology* (38) : 46-53.
- [Coleridge, 1772-1934] S. T. Coleridge, *Biographie literaria, 1772-1934*, [e-Book en ligne 2004 : consulté le 06-02-2015, [http://www.gutenberg.org/files/6081/6081-h/6081-h.htm#link2H\\_4\\_0002](http://www.gutenberg.org/files/6081/6081-h/6081-h.htm#link2H_4_0002)].
- [Couchot, 1998] E. Couchot, *La technologie dans l'Art : de la photographie à la réalité virtuelle*. Jacqueline Chambon.
- [Fuchs *et al.*, 2006] P. Fuchs, G. Moreau, A. Berthoz, J-L. Vercher, *Le traité de la réalité virtuelle. volume 1 – L'Homme et l'environnement virtuel*. Presses des Mines.

<sup>10</sup> Cf. Chris Milk, « comment la réalité virtuelle peut créer un dispositif empathique exceptionnel », TED2015, [https://www.ted.com/talks/chris\\_milk\\_how\\_virtual\\_reality\\_can\\_create\\_the\\_ultimate\\_empathy\\_machine?language=fr](https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine?language=fr)

- [Guez, 2013] J.Guez, « De l'interaction à la présence – Un Art qui se vit ». Proteus, *Le spectateur face à l'art interactif* 6.
- [Guez, 2015] J.Guez, « Illusions entre le réel et le virtuel (IRV) comme nouvelles formes artistiques : présence et émerveillement ». Thèse, Esthétique, Science et Technologie des Arts, Université Paris 8, dir. Marie-Hélène Tramus.
- [Kilteni *et al.*, 2012] K. Konstantina, J.M. Normand, M.V. Sanchez-Vives, M. Slater, « Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion » *PLoS ONE* 7(7).
- [Lanier, 1990] J. Lanier, « Life in the Data-Cloud », *Mondo* 2000, 2, 44-51.
- [Nannipieri *et al.*, 2014] O. Nannipieri, I. Muratore, D. Mestre and J-C Lepecq, *La présence dans la réalité virtuelle : quand la frontière se fait passage*. Frontières Numériques.
- [Popper, 1985] F. Popper, *Art, action et participation : l'artiste et la créativité aujourd'hui*, Klincksieck.
- [Slater, 2003] M. Slater, « A Note on Presence Terminology. » *Émotion* 3 : 1-5.
- [Slater, 2009] M. Slater, « Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. » *Philosophical transactions of the Royal Society of London. Série B, Biological sciences* 364(1535): 3549–57.
- [Tramus, 2001] M-H. Tramus, « Recherches, Expérimentations et Créations Dans Les Arts Numériques : Interactivité, Acteurs Virtuels. » HDR. Université Paris 8.
- [Vermersch, 2012] P. Vermersch, *Explication et phénoménologie : vers une Psychophénoménologie*. PUF.
- [Wiederhold *et al.*, 2014] B. Wiederhold and S. Bouchard, *Advances in Virtual Reality and Anxiety Disorders*. Springer US.

## Œuvres citées

- [Auxietre *et al.*, 2015], B. Auxietre, *Le 5<sup>e</sup> sommeil*, Installation de réalité virtuelle, Visiocasque, Innerspace.
- [Auxietre, *et al.*, 2015], B. Auxietre, *Playhead*, jeu musical artistique de réalité virtuelle, Innerspace.
- [Batras *et al.*, 2015] D. Batras, J-F. Jégo, J. Guez and M-H. Tramus, InterACTE, *Installation de réalité virtuelle, casque et projection vidéo*, Ars Electronica 2015.

- [BeAnotherLab, 2015] Collectif BeAnotherLab, *The machine to be another*, installation, visiocasque et caméra vidéo.
- [Benayoun, 1998] M. Benayoun, *World Skin*, Installation de réalité virtuelle, CAVE, Ars Electronica Center.
- [Boissier, 1986] J-L. Boissier, *Pékin pour mémoire*, installation interactive.
- [Bret et al., 2000-2006] M. Bret and M-H. Tramus, *La Funambule virtuelle*, Installation de «seconde interactivité», Wall.
- [Chabalier et al., 2016] C. Chabalier et al., *Cassandra-Matériaux*, pièce de théâtre numérique, la Commune Aubervilliers.
- [Davies, 1995] C. Davies, *Osmose*, installation de réalité virtuelle, visiocasque.
- [Duchamp, 1920] M. Duchamp, *Rotary Glass Plates*, installation.
- [Guez et al., 2015] J. Guez, G. Bertinet, K. Wagrez, F. Costes, *Lab 'Surd: le LABORatoire de la SURvirtualité*, installation de réalité virtuelle, Ars electronica, 2015.
- [Guez et al., 2015] J. Guez, J-F. Jégo, K. Wagrez, G. Gilly-Poitou, J. Loosveld, C. Moussier, D. Rivoire, *Liber*, Jeu artistique de réalité virtuelle, Bibliothèque Publique d'Information (BPI), au Centre Pompidou à Paris.
- [Fleuryfontaine 2014] Fleuryfontaine, *I need a haircut*, installation de réalité virtuelle, Galerie YGREC.
- [Kruger, 1970] M. Kruger, *Videoplace*, installation de réalité «artificielle».
- [Mariančík 2013] T. Mariančík, L. Arrigoni, O. Pultar, J. Ceipek, *Sightline*, installation de réalité virtuelle, visiocasque.
- [Spiers et al., 2015] J.Spiers, L. Eveillard and F. Dervieux, *Panopticon*, VRJam.
- [Tinguely et al., 1969-1994] J.Tinguely and N. de Saint Phalle, *Le Cyclop*.

---

## Prise de vues 360 pour visiocasque

---

Stéphane Faudeux

Les contenus pour les casques d'immersion 360 peuvent être réalisés en image de synthèse ou en prise de vues réelles. Produire le contenu en images de synthèse (en CGI<sup>1</sup>) n'est pas complexe en soi, nombre d'applications scientifiques ou films de divertissement sont produits ainsi depuis de nombreuses années. En revanche, la production de films Vidéo 360 en images réelles constitue un challenge d'une tout autre nature.

### 1. PRISE DE VUES 360 EN IMAGES RÉELLES

La prise de vues réelle en 360 est vraiment complexe, notamment pour des contenus stéréoscopiques. Dans ce cas, l'incohérence «accommodation –vergence» est présente et la règle de limitation maximale de disparités rétiennes, donc de parallaxes horizontales sur les images stéréoscopiques, doit être respectée sur l'écran de focalisation du visiocasque, quelle que soit la direction du point de vue (voir chapitre 8, solution S6). Il convient de s'assurer de la qualité afin d'éviter les déconvenues qui ont surgi avec certains films 3D Relief (3Ds) ne respectant pas cette règle. Pour produire des contenus qualitatifs, il faut s'entourer d'une équipe aguerrie, surtout quand on n'a jamais tourné ce genre de film. L'expérience utilisateur est cruciale, une personne qui aura vécu une immersion 360 lui donnant la nausée ne voudra plus tenter l'expérience une seconde fois.

#### 1.1. Les outils de prise de vues, de 250 à plusieurs milliers d'euros

La prise de vues 360 voit le mélange des genres. À la différence de la prise de vues traditionnelle, il règne une certaine hiérarchie souvent dictée par le prix des produits : d'un côté les caméras grand public, de l'autre les modèles professionnels. En 360, les frontières s'avèrent parfois plus floues puisqu'il n'existe pas, à proprement parler, de caméra 360 dites «professionnelles». Les producteurs de contenu utilisent ce qu'ils ont sous la main. Pour le moment, nous en sommes encore aux balbutiements des caméras 360 «moyen ou haut de gamme». Des caméras sont annoncées, beaucoup émanent de projets faisant appel au financement participatif, mais encore rien de tangible à l'heure actuelle.

---

<sup>1</sup> Computer-Generated Imagery

## 1.2. Le choix de la caméra

Les producteurs de contenus plus professionnels, ceux qui disposent d'un « vrai budget », utilisent un rig avec plusieurs caméras, et les plus néophytes des caméras 360 simple d'usage. « Rig » est devenu un terme générique pour décrire un support sur lequel on va fixer de 2 à un nombre X de caméras. Dans le cadre de la VR, ce sont des caméras légères et petites qui sont placées sur un rig (dans la majorité des cas, il s'agit de GoPro). Autrement, il existe des caméras compactes capables de produire un contenu directement 360. Ces caméras peuvent elles-mêmes revêtir plusieurs natures : des caméras mono capteur avec un dispositif optique qui va « sphériser » l'image, et des caméras multicapteurs (multi-optique) avec un assemblage numérique interne à la caméra qui créera une seule image plus ou moins sphérique.

## 1.3. Caméras et Rig

Le rig est un support qui permet d'accueillir un certain nombre de caméras. Certains opérateurs (ou superviseurs VR) fabriquent leur propre dispositif, notamment grâce à des imprimantes 3D. Parmi les studios les plus réputés sur le marché de la VR, figure Felix & Paul Studios à Montréal. Ces derniers fabriquent leur propre caméra qu'ils se gardent bien de montrer, aucune photo de plateau sur laquelle on verrait ce dispositif n'existe.

Hormis 360 Heros<sup>2</sup> il n'existe pas de fabricant de rigs commercialisés sur le marché. Toutefois, il y a fort à parier que les usines chinoises proposeront progressivement des modèles dans les prochains mois.

Comme le souligne Thomas Villepoux, superviseur VR pour Real Insight Gentlemen, « Généralement les rigs ne se vendent pas, car il s'agit d'une prestation globale. Le savoir-faire des stéréographes et des superviseurs VR réside, pour une grande part, dans la connaissance pointue et personnelle de son rig. Il faut bien connaître son matériel pour s'adapter rapidement à toutes les conditions de prises de vues. »

360 Heros propose des rigs de 6 à 14 caméras pour les GoPro Hero. En terme de ratio caméra-image 360, il faut en moyenne 10 à 12 caméras pour réaliser des images stéréoscopiques sur 360°, alors que 8 caméras seront suffisantes pour des images monoscopiques sur 360°. Pour obtenir une qualité optimale, il est préférable d'utiliser des GoPro Hero 4 (qui sont 4K). Ces rigs peuvent être livrés avec des logiciels et accessoires pour piloter, déclencher les caméras, gérer les fichiers. Les rigs 360 Heros affichent une forme carrée pour couvrir un champ le plus hémisphérique possible.

---

2 <http://www.360heros.com>

Jump, un autre dispositif développé par Google, est une technologie de stitch<sup>3</sup> permettant d'assembler 16 caméras GoPro placées sur le rig circulaire GoPro Odyssey<sup>4</sup> stéréoscopique. Le tout pour produire des vidéos immersives 8K30 en 3Ds.

## 1.4. Caméra simple mono capteur

Ces caméras dotées d'un seul capteur donnent une image d'assez basse qualité. Dans le cadre de ce type de caméras, l'optique se trouve au-dessus. Ces caméras reposent sur des compromis techniques, qualitatifs, l'image n'est pas réellement hémisphérique.

La caméra Kodak SP360 4K<sup>5</sup> procure une image hémisphérique, mais Kodak propose un rig pour y placer deux caméras dos à dos. La société commercialise également un Pack Dual Pro avec deux caméras à l'intérieur. Kodak va plus loin avec le développement d'un logiciel de stitch et une application pour commander la caméra (compatible iOS et Android). Il existe un support « spécial drone » pour placer une caméra au-dessus et en dessous d'un drone. Le capteur 4K CMOS comporte 12 millions de pixels et l'enregistrement s'effectue sur une carte micro SD.

La caméra VSN<sup>6</sup> 360 peut enregistrer des vidéos avec un angle de 60° vertical et une résolution maximale de 6480 x 1080 HD. Il n'y a pas besoin de stitching, mais la qualité est tout de même limitée et l'angle de vision peu important de 60°.

## 1.5. Caméra multicapteurs

Ricoh Theta S<sup>7</sup> est une caméra bon marché, capable de filmer en HD 1080p avec une capacité de stockage de 25 minutes (30 i/s). La caméra se tient à la main avec sa forme simple parallélépipédique, peu encombrante très différente des autres modèles et possède deux capteurs de 12 millions de pixels. Sa forme allongée offre une bonne tenue. La caméra est pilotable depuis une App permettant d'être guidée depuis un smartphone ou une tablette. Elle possède une sortie mini HDMI et autorise une diffusion en direct (streaming via la WiFi). Ricoh Theta S est le produit grand public par excellence, pour preuve Theta m 15, le modèle entrée de gamme, se décline en plusieurs couleurs.

La Giroptic 360cam<sup>8</sup>, sans faire de comparaison ou d'effet de style, pourrait être la GoPro pour la vidéo 360. Cette caméra, conçue en France, est capable de réaliser des

---

3 La technologie de Stitch consiste à assembler des images issues de plusieurs caméras filmant une même scène sans que l'on s'aperçoive des raccordements entre images.

4 <http://fr.gopro.com/odyssey>

5 <http://kodakpixpro.com/Europe/fr/cameras/actioncam/sp3604k/>

6 <http://www.vsnmobil.com/products/v360>

7 <https://theta360.com/fr/>

8 <http://eu.360.tv/fr/>

prises de vues 360 en 2K (2048 x 1080) et des photographies 4K. Elle est dotée de trois capteurs et peut produire des fichiers directement utilisables avec des visiocasques. Il n'y a pas besoin de postproduction et d'assemblage des séquences. Il est possible de streamer les contenus 360 via une connectivité WiFi et Bluetooth.

Nikon étonne avec ce modèle bifocal Keymission 360<sup>9</sup> qui, produisant une image sphérique en UHD 4K, fera certainement de l'ombre à des modèles compacts. Taillée pour l'aventure, cette caméra est étanche jusqu'à 30 m, résiste aux chocs et possède un système intégré de stabilisation. Elle est WiFi et Bluetooth, pour être pilotée à distance et faire du streaming.

Vuze VR<sup>10</sup>, développée par HumanEyes Technologies, comprend 8 capteurs capables de filmer en 360, en 4K et en 3D. La caméra est légère (250 g) et compacte (de la taille d'un CD). Elle peut se piloter depuis une App (pilotage, prévisualisation et gestion des contenus). La Vuze VR est livrée avec une suite logicielle, Vuze Studio, qui repose sur la technologie Adaptive Blending, laquelle va produire un « stitching » en quasi temps réel. Vuze Studio permet d'importer les contenus réalisés avec la caméra, d'appliquer des effets 3Ds, de monter les contenus, de créer des titres et gérer l'audio associé.

Samsung Gear 360 demeure un prototype connu sous le nom de Project Beyond. Cette caméra revêt un peu la même forme que la Vuze, mais est dotée de deux fois plus de capteurs (16 capteurs HD sur sa circonférence) et un capteur supplémentaire au-dessus. Les images pourront être exploitables directement avec le visiocasque Samsung Gear VR (streaming en direct via WiFi).

La caméra bicapteur ALLie Go<sup>11</sup> (chaque capteur affiche une résolution de 8 millions de pixels) chasse sur les terres de GoPro. Cette caméra aux formes arrondies est en effet conçue pour les sports extrêmes et dispose d'accessoires pour se fixer sur des casques. Elle est compatible WiFi, mais propose un port pour une carte microSD. Il existe un modèle pour la maison qui permet de faire de la vidéosurveillance. La résolution maximale de cette caméra est de 2048 x 2048 à 22 i/s.

## 1.6. Caméras professionnelles

Nokia Ozo<sup>12</sup> affiche un design de science fiction, 8 caméras avec des capteurs 2K et 8 microphones synchrones. Cette caméra est plutôt destinée aux applications professionnelles. Elle peut capter la vidéo en 3Ds et en hémisphérique. Sur un plan esthétique, le tout est plutôt réussi. La caméra possède un enregistreur vidéo propriétaire qui s'insère sous sa base.

---

9 <http://www.nikonusa.com/en/nikon-products/action-cameras/index.page>

10 <http://vuze.camera>

11 <https://alliecam.com>

12 <https://ozo.nokia.com>

La technologie Jaunt<sup>13</sup> fournit une solution complète (logiciel et matériel). La caméra capture des images stéréoscopiques dans tous les axes, avec des microphones 3D d'ambiance. Jaunt a développé des algorithmes pour réaliser un stitch automatique, une correction géométrique et colorimétrique. L'entreprise propose un module d'export pour être visible sur tous les visiocasques.

## 2. TOURNAGE DE VIDÉO 360

On compte deux types de tournage : les tournages pour des contenus qui seront postproduits (fiction, documentaire, communication) et les tournages dont les contenus sont diffusés en temps réel, en « live » (sport, musique). Pour une diffusion « live », il convient d'admettre un compromis sur la qualité (les films sont encodés en MP4) et sur le nombre d'images.

### 2.1. La préparation et les conseils de mise en scène

Dans le cadre d'un tournage de fiction ou de documentaire, il est important de procéder à un repérage préalable. Il faut placer le rig dans un endroit qui permettra de faciliter l'invisibilité de raccord entre les caméras. Aucun objet ne doit bouger devant les caméras. Comme le souligne Julien Levy, *executive manager* de Digital Immersion, « il faut penser à la chorégraphie des actions et des personnages et choisir son rig selon le lieu dans lequel on filme ».

Les caméras largement utilisées sur les rigs sont les GoPro. Elles sont très peu sensibles en basse lumière, il faut donc bien éclairer la scène. Qui dit éclairage, suppose projecteur, d'où le besoin de trouver des astuces pour cacher les pieds ou accrocher les projecteurs en hauteur.

Dans les conseils de prise de vues, il est important de se poser la question de ce qui va fonctionner en terme de VR et éviter des effets ou des erreurs de prise de vues qui pourraient créer des nausées (voir les chapitres 7 et 8). Si les caméras bougent, il est nécessaire de les stabiliser en postproduction.

Le superviseur VR travaille sur le scénario avec le réalisateur, il pourra donner les règles à respecter sur les limitations techniques. Cela, pour trouver une nouvelle cinématographie, parce que, en vidéo 360, le montage est différent, les gros plans sont impossibles.

---

13 <http://www.jauntvr.com/#about>

## 2.2. Tournage en prise de vue stéréoscopique ou monoscopique

Il existe deux types de projets 360 : des contenus en monoscopie (appelé à tort « vidéo 360 2D », voir le chapitre 2) et en stéréoscopie (« vidéo 360 3D »). Il est relativement facile de produire des contenus monoscopiques, que ce soit avec des caméras grand public 360 ou des rigs professionnels. Produire des contenus en stéréoscopie (3Ds) s'avère beaucoup plus complexe et nécessite une prise en charge par une équipe aguerrie, laquelle doit avoir la maîtrise de la stéréoscopie sur 360 degrés. Il faut éviter que la vidéo 360 commette les mêmes erreurs que le cinéma 3Ds, avec une production de contenus parfois de mauvaise qualité. Déjà trop de contenus vidéo 360 monoscopique ne sont pas acceptables.

## 2.3. Comment visualiser les rushes

Les caméras compactes disposent, pour la plupart, d'une connexion WiFi. Elles autorisent donc un monitoring sur une tablette ou un smartphone. Toutefois, ces caméras restent utilisées pour des applications grand public ou semi-professionnelles. Pour le moment, les professionnels, qui doivent produire des contenus les plus qualitatifs possible, ont recours à des rigs équipés de plusieurs caméras. Ces caméras sont compactes (GoPro ou Blackmagic) pour éviter les problèmes d'entraxes. Ces rigs comprennent généralement 8 caméras, même si certains fabricants ont développé des concepts de prototypes à 42 caméras. La société 360 Designs a développé le concept Eye VR qui permet de capturer réellement 360 degrés dans les 3 axes. Le système comprend 42 caméras Blackmagic Micro Cinema Camera et 24 microphones. Il peut également fonctionner sur 2 axes (30 caméras en 3D).

À la différence des caméras compactes qui délivrent un flux vidéo immédiatement exploitable, les systèmes de rig multicaméra (pour visualiser l'image finale) nécessitent de recomposer l'image à partir des X caméras. Bien évidemment, plus il y a de caméras et plus elles ont des capteurs CMOS avec une grande résolution (4K), et plus il est long et complexe de pouvoir visualiser l'image 360. Il est quasi impossible de voir en temps réel l'image 360.

Pour voir ces contenus, il faut faire l'acquisition de chacun des flux vidéo en provenance de chacune des caméras. Il est nécessaire d'avoir une station de travail puissante pour traiter en direct les images et faire un « stitch » des différentes images vidéo. Une contrainte toutefois, des câbles vidéo doivent être placés derrière chaque caméra, ce qui alourdit le dispositif et complexifie le tournage. Qui plus est, lors d'un tournage 360, les caméras voient tout le champ, il convient donc de masquer lesdits câbles. Et si la caméra se déplace, les contraintes s'accumulent. Des tentatives d'associer des émetteurs HF Vidéo derrière chaque caméra se sont avérées complexes à mettre en œuvre et les résultats peu concluants. Placer X émetteurs HF les uns à côté des autres alourdit en vérité le dispositif et des interférences apparaissent entre les signaux vidéo.

Pour la prise de vues 360 depuis un drone, avec un rig multicaméra, il est fréquent de placer deux petites caméras qui ont une optique sphérique 180°. Les deux caméras sont agencées dos à dos ; dotées d'un émetteur HF, elles vont aider l'opérateur 360 à cadrer.

Sur le plateau vidéo VR 360, une équipe technique prend en charge la gestion des images. Le *data manager*, dont la fonction est de gérer les fichiers, décharge les fichiers des différentes caméras sur un ordinateur, il réalise un stitch en basse résolution permettant de corriger les positions de caméra.

## 2.4. Les outils de stitching temps réel

Il existe des outils de stitching de couture numérique qui permettent de créer des images panoramiques plus ou moins hémisphériques. Le but est de ne pas voir les points de couture. Ce sont deux sociétés françaises qui sont en pointe sur ce sujet, en attendant la solution Jump de Google.

Videostitch et Kolor (Kolor a été rachetée par GoPro) proposent ainsi des applications logicielles pour du stitching en temps réel et également en postproduction. Les deux principales applications sont Vahana VR de Videostitch et Autopano Video de Kolor GoPro.

Vahana VR est compatible avec des contenus 360 2K et 4K (30 images/s). Le logiciel est capable de délivrer un format de sortie vidéo sphérique 360° par 180° (équirectangulaire) qui est compatible avec la majorité des visiocasques (dont l'Oculus Rift) et les plates-formes vidéo 360, comme YouTube ou Littlstar.

Autopano est un logiciel qui permet de créer des panoramiques à partir d'un ensemble de photographies. La manière la plus simple pour créer des contenus est, selon GoPro, d'utiliser un rig 360 avec des caméras GoPro et Kolor Autopano Vidéo pour les assembler. La société a développé un player Kolor Eyes qui est compatible avec les tablettes, smartphones et ordinateurs.

## 2.5. Le son

Si l'image est complexe en 360, le son ne l'est pas moins. Certaines caméras (compactes) ont des microphones intégrés. Mais, en ce qui concerne les tournages professionnels, le son peut être géré de plusieurs façons : soit des microphones ambiophoniques, soit des microphones placés dans différents points de la scène et ensuite mixés pour recréer un espace sonore. Pour la diffusion de l'audio 3D, c'est souvent le moteur de jeu Unity qui est utilisé.

### 3. LA POSTPRODUCTION POUR FINALISER LES IMAGES

Première étape de la postproduction, le superviseur VR et le réalisateur sélectionnent les meilleures prises, qui seront ensuite «stitchées» avec les outils de compositing. Pour cette opération, ne sont plus utilisés les outils de stitching temps réel, mais des outils plus précis.

La postproduction des contenus VR 360, issus de plusieurs caméras, consiste à recréer une image hémisphérique à partir d'un composite de séquences vidéo. Cette étape de postproduction est fastidieuse et manuelle. Pour le moment, les superviseurs VR estiment qu'il n'y a pas de logiciel spécifique suffisamment performant. Ils préconisent, par conséquent, d'utiliser des outils de compositing pour travailler sur le stitching des images. Les logiciels les plus couramment utilisés sont Adobe After Effects et Nuke. Les étapes sont la calibration, la stabilisation, la correction colorimétrique (les caméras n'étant pas appairées, des différences de couleur entre les vidéos apparaissent), la correction d'exposition. L'un des réglages les plus importants porte sur la synchronisation des différentes séquences. Une fois encore, la plupart des caméras utilisées n'étant pas synchronisées, il faut donc les synchroniser en postproduction. Cette synchronisation peut se faire par une image (flash) ou par le son.

Videostitch Studio est l'outil de postproduction de Videostitch. Il comprend de nombreux réglages pour finaliser les contenus VR, avec notamment un preview temps réel.

Pour améliorer la qualité des images, et surtout corriger les défauts liés à la prise de vues, des éditeurs développent des plug-ins, à l'instar de ceux proposés par Dashwood, la suite 360VR Toolbox. La société de Tim Dashwood possède une solide expertise en matière de science et géométrie des images. Ce dernier a conçu et développé des outils de postproduction pour la 3D stéréoscopique, comme Stereo3D CAT. Ses connaissances sont désormais au service de la réalité virtuelle.

Une fois les images parfaitement corrigées, elles peuvent suivre le parcours classique d'un film. Les images définitives sont conformées, puis étalonnées. Enfin, elles sont mastérisées pour être adaptées à un player VR (player intégré à un casque ou une App). Pour le moment, il n'y a pas encore de normalisation ou de standardisation, il faut donc préparer le contenu vidéo pour chacun des visiocasques ou dispositifs VR disponibles sur le marché.

### 4. LES NOUVEAUX MÉTIERS

#### *Superviseur VR ou Stéréographe*

Il doit connaître les règles de la VR360, de la 3Ds, mais aussi de la postproduction. Il est là pour aider l'équipe technique qui n'a pas de connaissances sur la VR360.

### *Opérateur 360*

Il gère le tournage, les caméras 360, il peut avoir la fonction de Data Manager, décharger les caméras 360 et faire le stitch des différentes caméras.

### *«Stitcher»*

Cette personne, qui vient du monde des effets spéciaux et du compositing, travaille principalement sur la fabrication des images 360.

## **5. LES PRINCIPALES APPLICATIONS DES VIDÉOS 360 PRÉVISIBLES**

- Des reportages d'actualité pour mettre le spectateur au cœur de l'événement (par exemple, dans une zone de combat, une épreuve sportive). Le spectateur se sent plus présent dans l'événement.
- Des documentaires de découvertes pour être immergé visuellement dans des lieux ou monuments touristiques. Ceux-ci seront mieux appréciés par l'augmentation du champ de vision et par le contrôle personnel du point de vue observé, les deux fonctionnalités des visiocasques.
- Des contenus en «direct», notamment des retransmissions sportives ou des concerts de musique. Le spectateur se retrouve au milieu du terrain de football, de tennis ou sur la scène avec son chanteur préféré.
- Des programmes de fiction immersifs avec un nouveau type de scénarisation, laissant du temps au spectateur d'observer la scène à sa guise et non sous le seul contrôle du réalisateur. Le spectateur peut vivre une expérience et non simplement regarder un film. C'est un changement de paradigme de placer le spectateur dans le scénario qui est narré et sur lequel il peut interagir.
- Des performances artistiques à partir d'une immersion visuelle du spectateur dans un monde réel (filmé) et dans lequel il peut plus ou moins agir avec l'œuvre artistique (voir le chapitre 11 sur les arts numériques avec un monde virtuel);
- Des vidéos immersives pour des souvenirs personnels comme on a des photos, panoramiques ou non.

Cette nouvelle interface visuelle, le visiocasque, peut permettre la naissance d'un nouveau procédé pour capter la réalité dans lequel on peut plus ou moins s'y activer selon les moyens techniques mis en œuvre. Cela impose une nouvelle réflexion sur l'expérience que l'on pourra proposer au public.



---

# Conclusion et perspectives

---

## 1. CONCLUSION

Au début de cet ouvrage, nous avons formulé la problématique suivante : quel est le comportement de l'homme, équipé d'un visiocasque et confronté à des incohérences sensorimotrices perturbantes ? Et quelles sont les solutions envisageables ? Nous avons essayé de clarifier cette récente problématique avec l'introduction probable des visiocasques auprès du grand public. Nous avons bien conscience de n'avoir pas répondu totalement à ces questions ouvertes. Le lien éventuel entre les différentes incohérences sensorimotrices n'a pas été abordé.

Les chercheurs n'ont pas suffisamment de recul sur les techniques de la réalité virtuelle, juste une vingtaine d'années, et encore moins concernant l'usage des visiocasques. Les futurs développeurs auront à expérimenter des applications RV à destination de tout public, sans avoir toutes les connaissances nécessaires. Ils vont devoir expérimenter mais il ne faut pas trop attendre de réponses de leur part, car ils ne pourront faire que des conclusions circonstanciées sur des cas particuliers qu'il sera difficile de généraliser. Il y a tellement de paramètres à contrôler et d'autres à faire varier pour analyser correctement une solution effective et généralisable. L'expérimentation sur le comportement humain demande du temps et des moyens conséquents à laquelle les développeurs auront peu de temps à consacrer. De plus, la quantité de testeurs doit être élevée pour valider des résultats pertinents et pérennes.

Ce sont donc les chercheurs en réalité virtuelle et surtout en sciences cognitives qui doivent poursuivre leurs travaux de recherche, en prenant en compte les spécificités de cette nouvelle interface visuelle et du public auquel elle s'adresse. Il faut que les chercheurs en sciences cognitives, spécialement en neurosciences, analysent le comportement humain en environnement virtuel : comment les stimuli sensoriels sont-ils traités ? Comment analyser l'importance des fonctionnalités de la vision centrale et de la vision périphérique dans un espace virtuel ? Quels paramètres des stimuli, intervenant dans les traitements neuronaux, sont à prendre en compte : intensité, contraste, fréquence spatiale, fréquence temporelle ? On ne peut pas se contenter de tests ergonomiques, nécessaires mais non suffisants, pour concevoir des solutions fiables et pérennes. De même, la question de la « présence » en environnement virtuel est à analyser mais elle est loin d'être suffisante pour inventer des solutions nouvelles et efficaces. Nous espérons que cet ouvrage donnera aux chercheurs en neurosciences le désir d'analyser méthodologiquement et en profondeur la question de l'activité de l'homme dans un environnement virtuel.

Les développeurs doivent mieux comprendre le fonctionnement de l'être humain immergé dans un environnement virtuel pour offrir à leurs clients des applications RV irréprochables en terme de confort et de santé. Les tests avec des utilisateurs seront cruciaux pour l'évaluation de leur application RV dans le but de s'assurer que leurs expériences novatrices seront acceptées et captivantes. La réalité virtuelle est un domaine encore trop jeune pour pouvoir compter sur des normes établies. Par manque de recul, toute nouvelle expérimentation en environnement virtuel n'est pas simple à concevoir. Les concepteurs ne devront pas réduire l'utilisateur à une «boîte noire» avec ses entrées et ses sorties... Cet ouvrage devrait leur fournir un nouvel éclairage sur la réalité virtuelle et ses enjeux. Et faudra-t-il demain des «virtuographes» comme on a aujourd'hui des stéréographes pour le cinéma en relief?

Nous pouvons prédire un *changement de culture* pour les concepteurs de jeux vidéo. Ils doivent dès maintenant se préoccuper de trouver la meilleure solution de l'immersion et de l'interaction de la personne, avec une approche plus large qu'auparavant : il y a plus de variantes d'interfaçage, même si celles-ci ne vont pas toutes devenir des standards d'Immersion et d'Interaction dans les univers virtuels. Fini le sacro-saint ordinateur avec son joystick et la console de jeux... Une plus grande panoplie de techniques d'interfaçage, matérielles et logicielles, peuvent être exploitées, offrant une multitude variée d'activités sensorimotrices et cognitives. Un début d'évolution a déjà eu lieu avec la Kinect... mais, à l'instar de l'australopithèque à l'homme, peu de joueurs se sont redressés de leur siège ! L'arrivée des visiocasque et des potentialités de la réalité virtuelle fera-elle mettre debout les joueurs ? Et remettre en cause le paradigme du jeu vidéo classique ? Mais tous les Jeux vidéo ne sont pas à développer en employant un visiocasque. Cette interface visuelle ouvre de nouveaux horizons pour un certain type d'activités ludiques et artistiques à l'adresse du public, celles qui exploiteront judicieusement les deux fonctionnalités des visiocasques pour de nouvelles activités.

La réalité virtuelle est un domaine scientifique et technique qui tente de répliquer l'expérience du monde réel. Les usagers souhaiteront *a priori* vouloir interagir avec le monde virtuel de la même manière qu'ils le font dans la réalité. Il va falloir les convaincre que ce n'est pas la seule potentialité de la réalité virtuelle. Les développeurs peuvent utiliser les scénarios familiers du monde réel pour guider les utilisateurs, mais pourquoi en rester là ? Des expériences pourront aller au-delà d'interactions virtuelles naturelles, avec des PCV irréelles, couplées à des mondes symboliques ou imaginaires.

Il y a plus d'un siècle un nouveau média, le cinéma, ouvrait de nouveaux horizons. Une nouvelle innovation technique sans grande perturbation sur la santé des spectateurs. Tout juste furent-ils effrayés d'un train qui faillit sortir de l'écran ! Puis vint le cinéma en relief qui faisait jaillir virtuellement des objets de l'écran mais créait déjà une première gêne visuelle étrange, une première incohérence sensorimotrice, pour le spectateur. Maintenant, il s'agit d'être acteur et, sous la dépendance de la boucle « Perception, Décision, Action », les incohérences perturbatrices affluent en tous sens (au propre comme au figuré) ! Faut-il attendre toutes les solutions en provenance des concepteurs et des fabricants de visiocasques ? S'il est primordiale qu'ils réalisent des produits de meilleures qualités, au regard des capacités humaines, ils ne sont pas les seuls maîtres

du jeu. Nous avons précisé en détail ce qui est du ressort du concepteur de visiocasque et du développeur de l'application RV.

La réalité virtuelle est un «art», avec de nombreuses «recettes», mais il faut de l'originalité, aussi bien sur l'interface, sur l'interfaçage, sur des PCV, réelles ou irréelles pour des activités, réalistes ou non, dans des mondes réalistes ou irréalistes spatialement, temporellement et physiquement...

## 2. PERSPECTIVES

Il est difficile de prévoir l'avenir, surtout pour des techniques innovantes proposant de nouvelles activités humaines, comme celles initiées par la réalité virtuelle mais il est cependant instructif de relire les anciennes prévisions. Il y a vingt ans, les commentaires sur l'essor de la réalité virtuelle étaient enthousiastes pour certains journalistes : «Avez-vous déjà voyagé en réalité virtuelle? Ceux qui ont tenté l'expérience en sont revenus éberlués : ils ont vu l'Avenir!» de Sciences et Vie, N°990, juin 1993<sup>1</sup>. Il est nécessaire de faire la part du rêve et de la réalité, sans trop être conditionné par la science-fiction : bien des fois, j'ai été questionné sur la comparaison entre la réalité virtuelle et les célèbres films de science-fiction, Matrix, Minority Report, Avatar... Il va falloir clarifier auprès du public ce que peuvent faire les techniques de la réalité virtuelle et ce qu'elles ne peuvent réaliser.

Il a fallu attendre vingt ans environ pour que la réalité virtuelle soit bien maîtrisée et utilisée quotidiennement pour des applications professionnelles. Elles se banalisent progressivement, même si certains secteurs professionnels n'ont pas encore exploité ses potentialités. Faudra-t-il attendre encore vingt ans, ou plus, pour avoir des visiocasques de qualité très satisfaisante, après avoir attendu vingt ans pour passer de visiocasques onéreux et bas de gamme aux actuels visiocasques plus accessibles financièrement et de moyenne gamme (pour certains)?

Si, pour le matériel, il est encore aisé de faire des prévisions, il est plus délicat d'en entreprendre sur les futurs vrais usages grand public des applications RV, car l'homme est au centre de l'activité proposée. Combien d'années faudra-t-il attendre, dix ans, vingt ans ou plus, pour que les spécialistes RV et les chercheurs en sciences humaines aient suffisamment de recul et d'expérimentations à leur actif, pour faire le point sur cette problématique?

L'évolution rapide de certaines techniques récentes ne doit pas nous influencer, bien que les enjeux financiers soient importants. D'un côté, l'évolution a été très rapide entre l'usage des premiers téléphones portables, très chers, lourds et à faible autonomie, à l'usage quotidien des smartphones actuels, utilisés par tous. D'un autre côté, la télévision 3Ds en relief, peu onéreuse, n'a pas eu le succès escompté, certains pensant à tort que

---

<sup>1</sup> p118 – extrait du livre «La réalité virtuelle», de Bernard Jolival, édition «Que sais-je», 1995, ISBN 2-13-047290-7

toutes les émissions seraient en relief! Nous avons pronostiqué que l'usage de la télévision en relief resterait et restera marginale car elle répond à un besoin secondaire de la vision humaine, comme nous l'avons expliqué précédemment. Ce besoin, attrayant et exploité judicieusement dans les films en relief (3Ds), est sans commune mesure avec le besoin impérieux que nous avons tous quotidiennement pour communiquer avec nos smartphones. Nous pourrions en dire de même pour l'échec prévisible des lunettes Google Glass : nous n'avons pas besoin quotidiennement des usages de la Réalité Augmentée...

Concernant les visiocasques, ils ne répondent pas à un besoin aussi important que celui des smartphones. Il ne faut donc pas s'attendre à une ruée sur les premiers produits qui vont sortir prochainement, sachant qu'en plus, à la différence des smartphones, ils peuvent créer de l'inconfort et des malaises. La lecture de sites internet rend compte de ce problème et donne une image mitigée de la «réalité virtuelle grand public». Il s'agit donc de savoir si les concepteurs des futures applications des visiocasques sauront proposer des expérimentations d'immersion en EV toujours vécues positivement par les usagers. Les prix des premiers visiocasques commercialisés sont relativement élevés, avec l'équipement complémentaire souhaitable (plus de sept cents euros). Mais rien n'empêche que les prix baissent comme pour les smartphones si la demande de masse sera présente dans le futur.

Les vidéo 360° ouvrent des perspectives intéressantes, au-delà des photos et vidéos familiales que nous possédons tous, même si l'usage est plus contraignant. Est-ce par cet usage que le public va acquérir des visiocasques? La réponse n'est pas certaine, mais, couplé au désir d'expérimenter de nouveaux types de jeux vidéo, la demande d'équipement sera peut-être assez forte pour créer un marché de masse, mais réservé aux milieux aisés financièrement. Des sondages ont été entrepris pour connaître la demande sur les futures activités en RV à domicile, tel que ce sondage sur l'implantation<sup>2</sup> d'application RV à domicile : 69% des personnes interrogées ne sont pas disposées à aménager une pièce spécialement pour exploiter une application avec visiocasque et 42% des sondés ne sont pas prêt à changer de configuration pour du matériel plus puissant nécessaire à une application RV, etc. L'avenir des visiocasques est plus prometteur pour ceux qui seront conçus pour la Réalité Augmentée (voir le chapitre 6), puisqu'il y aura moins d'incohérences sensorimotrices perturbantes, comme nous l'avons expliqué, source de rejet pour une part non négligeable de la population. Mais ces produits ne sont pas encore sur le marché.

Les applications professionnelles continueront d'utiliser des visiocasques quand ils sont plus adéquats que les autres interfaces visuelles et elles profiteront des derniers progrès techniques. Les arts numériques vont profiter aussi de l'évolution des visiocasques pour créer des œuvres symboliques et imaginaires que permettent les environnements virtuels. Mais ce n'est pas ce secteur qui portera le marché des visiocasques.

Les jeux vidéo et les activités ludiques ont de l'avenir si les capacités, sensorimotrices et cognitives, personnelles de chaque joueur seront bien pris en compte.

---

2 Voir <https://www.surveymonkey.com/results/SM-YCHJCG32/>

## 1. PRÉSENTATION DE L'AUTEUR

Philippe Fuchs, professeur à Mines ParisTech, dirige des recherches sur la réalité virtuelle depuis plus de vingt ans. Ses axes de recherche concernent les aspects théoriques et techniques de l'interfaçage en environnement virtuel, en particulier sur l'immersion visuelle avec ou sans visiocasques. Il collabore avec les entreprises à travers de nombreux projets de recherches et de développement. Il a animé la communauté des chercheurs et professionnels en réalité virtuelle en tant que président de l'AFRV (Association Française de Réalité Virtuelle et Réalité Augmentée). Il a rédigé plusieurs livres sur la réalité virtuelle, dont l'ouvrage collectif «Le traité de la Réalité Virtuelle», en cinq volumes et 101 auteurs, dont il a été à l'initiative et le directeur de l'ouvrage.

## 2. ONT COLLABORÉ À L'OUVRAGE

**Stéphan Faudeux**, diplômé de l'ESRA (Ecole Supérieure de Réalisation Audiovisuelle), est le Directeur de Génération Numérique, éditeur de Mediakwest et Sonovision et le Directeur de Think Factory Formation. Il travaille comme réalisateur, producteur sur de nombreux films. Il intervient comme formateur, consultant et journaliste sur les nouvelles technologies. En 2007, il crée Dimension 3, le Forum International de l'Image 3D relief. En Avril 2012, il développe le premier magazine sur l'écosystème de l'audiovisuel et des technologies, Mediakwest. En 2015, le Club HD organise Screen4All Forum, dont une journée est dédiée à la VR 360.

**Olivier Hugues** est ingénieur en informatique et docteur en informatique de l'Université de Bordeaux I. Ses axes de recherche ont porté sur un outil d'aide à la navigation maritime utilisant la réalité augmentée. Membre de l'équipe «RV&RA» du Centre de Robotique de Mines ParisTech, il contribue à caractériser la pertinence de la réalité virtuelle comme un outil d'évaluation de l'acceptabilité de la collaboration homme-robot dans le cadre d'une chaire de recherche et participe à de nombreux projets scientifiques et techniques autour de la réalité virtuelle.

**Jean-François Jégo** est artiste et enseignant-chercheur dans l'équipe INREV de l'Université Paris 8. Après des études en art contemporain, il obtient une thèse en informatique temps-réel de l'école Mines ParisTech. Ses recherches en réalité virtuelle questionnent le rapport du corps à l'environnement virtuel en permettant au spectateur-acteur d'interagir par des gestes incarnés et expressifs dans des dispositifs immersifs.

Il est l'auteur d'une dizaine d'installations en art numérique primées et scénographe de quatre performances scéniques numériques.

**Judith Guez**, Artiste-chercheuse en réalité virtuelle et mixte, et docteure en art numérique (INREV Paris 8). Ses axes de recherche explorent l'utilisation des illusions entre le réel et le virtuel comme nouvelles formes artistiques, à travers les concepts de présence et d'émerveillement. Entre art, science et innovation technologique, elle a exposé plusieurs installations interactives et participé au développement de plateformes de réalité virtuelle et augmentée. <https://judartvr.wordpress.com>

**Andras Kemeny**, Docteur ès Sciences, est Directeur du Laboratoire d'Immersion Virtuel, laboratoire de recherche commun entre Arts et Métiers et Renault. Il est également Responsable du Centre de Réalité Virtuelle et de Simulation Immersive de Renault et Professeur des Universités Associé à Arts et Métiers. Son centre d'intérêt principal porte sur la simulation de conduite et réalité virtuelle, englobant les aspects liés à la perception et au comportement des usagers ainsi qu'aux technologies immersives et aux applications automobiles correspondantes.

**Daniel Mestre** est directeur de recherche, Centre national de la recherche scientifique (CNRS). Il est affilié à l'Institut des Sciences du Mouvement (CNRS et Université Aix-Marseille II). Il est responsable scientifique du Centre de Réalité Virtuelle de la Méditerranée ([www.crvm.eu](http://www.crvm.eu)). La sensation de présence dans les environnements virtuels constitue un centre d'intérêt prioritaire de ses travaux actuels, avec des aspects fondamentaux (neurosciences cognitives et comportementales) et appliqués (monde industriel, médical et artistique).

---

# Table des matières

---

<b>PRÉFACE</b> .....	7
<b>AVANT-PROPOS</b> .....	9
<b>INTRODUCTION ET ENJEUX</b> .....	11
<b>PARTIE I</b> .....	17
<b>CHAPITRE 1 - CONCEPTS DE BASE EN RÉALITÉ VIRTUELLE</b> .....	19
1. Définitions de la réalité virtuelle .....	19
2. Modèle de référence pour l'immersion et l'interaction .....	23
3. Immersion et présence.....	31
4. Bibliographie.....	32
<b>CHAPITRE 2 - QUELQUES NOTIONS SUR LES SENS HUMAINS</b> .....	35
1. Introduction .....	35
2. Le système visuel .....	38
3. Le système auditif.....	53
4. La sensibilité cutanée.....	54
5. La proprioception .....	56
6. Bibliographie.....	58
<b>CHAPITRE 3 - LES INTERFACES VISUELLES</b> .....	61
1. Introduction .....	61
2. Fonctionnalités pour la perception d'un espace tridimensionnel .....	61
3. Les interfaces à écrans fixes .....	64
4. Les interfaces portables.....	68
5. La restitution visuelle stéréoscopique .....	68
<b>CHAPITRE 4 - ÉTUDE FONCTIONNELLE DES VISIOPORTABLES</b> .....	71
1. Introduction .....	71
2. Les différents types de visiocasques .....	72

3. La conception optique .....	75
4. Les écrans .....	76
5. Le tracking de la tête .....	79
6. La conception ergonomique .....	84
<b>CHAPITRE 5 - LES INTERFACES ASSOCIÉES AU VISIOCASQUE .....</b>	<b>89</b>
1. Introduction .....	89
2. Interfaces de commande et d'interaction .....	89
3. Interfaces acoustiques .....	92
4. Tapis roulants 1D ou 2D .....	93
5. Interface de simulation de mouvement .....	96
6. Interfaces de suivi des yeux .....	101
<b>CHAPITRE 6 - CARACTÉRISTIQUES TECHNIQUES DES VISIOCASQUES .....</b>	<b>103</b>
<i>Olivier Hugues</i>	
1. Introduction .....	103
2. Caractéristiques générales .....	103
3. Caractéristiques techniques des visiocasques .....	106
4. Conclusion .....	118
<b>CHAPITRE 7 - CONFORT ET SANTÉ .....</b>	<b>121</b>
1. Introduction .....	121
2. Incohérences sensorimotrices .....	122
3. Problèmes psychologiques induits par les environnements virtuels .....	140
4. Conclusion .....	142
5. Annexe : théories du mal du simulateur et du mal du transport .....	143
6. Bibliographie .....	143
<b>CHAPITRE 8 - CONSEILS POUR L'USAGE D'UN VISIOCASQUE .....</b>	<b>145</b>
1. Introduction .....	145
2. PCV observation .....	147
3. PCV « observation irréaliste » .....	160
4. PCV navigation .....	162
5. PCV manipulation .....	174
6. Grille d'analyse des 32 solutions .....	175
7. Adaptation en environnement virtuel .....	178
8. Règles de sécurité .....	180
9. Conclusions .....	182

<b>PARTIE II</b> .....	183
<b>CHAPITRE 9 - EXPÉRIMENTATION EN SCIENCES DU COMPORTEMENT</b> .....	187
<i>Daniel R. Mestre</i>	
1. Changement du paysage de la réalité virtuelle par les visiocasques.....	187
2. Traverser des portes virtuelles.....	188
3. Conclusion.....	191
4. Bibliographie.....	192
<b>CHAPITRE 10 - USAGES INDUSTRIELS DES VISIOCASQUES</b> .....	195
<i>Andras Kemeny</i>	
1. Introduction.....	195
2. Simulation de conduite et réalité virtuelle.....	197
3. Applications avec visiocasque ou avec salle immersive.....	198
4. Étude sur le mal de simulation (VRISE).....	200
5. Étude sur la perception de l'échelle 1.....	202
6. Bibliographie.....	204
<b>CHAPITRE 11 - L'USAGE DES VISIOCASQUES POUR LES ARTS NUMÉRIQUES</b> .....	209
<i>Judith Guez et Jean-François Jégo</i>	
1. L'apport des visiocasques dans la création artistique.....	209
2. Une méthode pour créer avec un visiocasque.....	215
3. Conclusion et perspectives.....	221
4. Bibliographie.....	222
<b>CHAPITRE 12 - PRISE DE VUES 360 POUR VISIOCASQUE</b> .....	225
<i>Stéphan Faudeux</i>	
1. Prise de vues 360 en images réelles.....	225
2. Tournage de vidéo 360.....	229
3. La postproduction pour finaliser les images.....	232
4. Les nouveaux métiers.....	232
5. Les principales applications des vidéos 360 prévisibles.....	233
<b>CONCLUSION ET PERSPECTIVES</b> .....	235
<b>AUTEURS</b> .....	239

