Sujet de thèse en Informatique

Apport de l'apprentissage machine aux prochaines générations d'algorithmes de visualisation haute performance

Employeur : Université de Reims Champagne-Ardenne (URCA)
Lieu : IUT de Reims-Châlons-Charleville, Reims
Durée : 36 mois
Direction de thèse : Laurent LUCAS et Hervé DELEAU (co-encadrant)

Mots clés : Visualisation haute performance, maillage, interaction, out-of-core, intelligence artificielle

Thèse en partenariat avec le CEA DAM Ile de France.

La visualisation scientifique de données est devenue un outil indispensable pour répondre aux principaux enjeux sociétaux en matière de climat, de prédiction des risques, de santé publique, de bioéconomie et/ou d'industrie, pour ne citer que ces exemples. Les progrès réalisés en particulier au cours des 20 dernières années en matière de calcul intensif (HPC), d'informatique graphique (CG) et de traitement générique sur processeur graphique (GPGPU) ont transformé le principe même de visualisation pour passer d'un simple outil de présentation à un outil de découverte et d'analyse permettant de percevoir des phénomènes potentiellement non envisagés. Ceci est d'autant plus vrai que les volumes de données traitées ne cessent d'augmenter pour atteindre des tailles ne permettant plus à un ingénieur d'effectuer une analyse pertinente et complète par lui-même.
En parallèle, l'apprentissage machine a connu dans le même temps un grand succès dans l'exploration de données massives sur base d'outils graphiques ad-hoc. Or, il est très étonnant de constater encore aujourd'hui que l'usage de l'apprentissage machine, au service de la visualisation scientifique, reste aujourd'hui encore - malgré l'engouement actuel pour les technologies autour de l'intelligence artificielle (IA) - un domaine de recherche assez confidentiel.
C'est dans ce cadre que nous proposons de mener les travaux à venir. Il s'agira notamment d'étudier comment enrichir des codes de calcul de sorte à ce qu'ils puissent apprendre à partir de données d'entrée afin d'optimiser leurs propres performances. Dans le contexte de la visualisation scientifique, l'utilisation de l'apprentissage machine peut potentiellement libérer de tout un ensemble de contraintes inhérentes aux algorithmes employés comme aux données sur lesquels ils opèrent. Ce projet doctoral vise à développer de nouveaux outils de visualisation intelligente de données massives (maillages 3D à raffinement adaptatif résolus dans le temps dits Hyper Tree Grid (HTG) dans trois directions différentes mais néanmoins complémentaires : (1) la gestion out-of-core [1,2,3] de maillages HTG, (2) la classification et la visualisation [4,5,6] progressive de maillages 4D, (3) l'extraction et le suivi des caractéristiques spatiotemporelles [7,8,9].

Le candidat aura de préférence un profil informatique avec de bonnes connaissances en informatique graphique, en visualisation et en calcul scientifique. Il devra également avoir des bonnes compétences en programmation C++, GPGPU (CUDA), OpenGL, le goût pour l'algorithmique et une formation de master avec une forte dominante en visualisation à haute performance. Des connaissances en machine learning comme sur OptiX et/ou VTK seraient appréciées.

Les travaux de recherche envisagés se dérouleront à Reims au sein du laboratoire CReSTIC de l'Université de Reims Champagne-Ardenne. Des déplacements réguliers à Bruyères-le-Châtel (91) sont à prévoir sur le site Ile-de-France du CEA DAM.

Le travail sera encadré par Laurent LUCAS (laurent.lucas@univ-reims.fr) et Hervé Deleau (herve.deleau@univ-reims.fr) pour le CReSTIC en relation avec Franck LEDOUX (Franck.LEDOUX@CEA.FR) et Jérôme DUBOIS (Franck.LEDOUX@CEA.FR) pour le CEA DAM

Merci d'attacher à votre réponse un CV incluant le cas échéant des références de personnes pouvant faire état de votre potentiel de chercheur (encadrants de stage, enseignants, etc.).
Références :

[1] J. Sarton, N. Courilleau, Y. Remion and L. Lucas. Interactive Visualization and On-Demand Processing of Large Volume Data: A Fully GPU-Based Out-Of-Core Approach. IEEE Transactions on Visualization and Computer Graphics, 2019. DOI:10.1109/TVCG.2019.2912752
[2] P. Muigg, M. Hadwiger, H. Doleisch, and E. Gröller. Interactive Volume Visualization of General Polyhedral Grids. IEEE Transactions on Visualization and Computer Graphics, 17(12):2115–2124, 2011.
[3] S.-E. Yoon, B. Salomon, R. Gayle, and D. Manocha. Quick-VDR: out-of-core view-dependent rendering of gigantic models. IEEE Transactions on Visualization and Computer Graphics, 11(4):369–382, 2005.
[4] J. Chen,D. Silver, and L. Jiang. The Feature Tree: Visualizing Feature Tracking in Distributed AMR Datasets. In 2003 IEEE Symposium on Parallel and Large-Data Visualization and Graphics (PVG), pp. 103-110, 2013.
[5] G. Harel, J-B. Lekien and P.P. Pébaÿ. Visualization and Analysis of Large-Scale, Tree-Based, Adaptive Mesh Refinement Simulations with Arbitrary Rectilinear Geometry. arXiv:1702.04852, 2017.
[6] J. Dubois, G. Harel and J-B. Lekien. Interactive visualization and analysis of high resolution HPC simulation data on a laptop with Vtk. IEEE VIS 2018 Scientific Visualization Contest(2018), IEEE VIS'18 SciVis Contest. 1.
[7] E. Bittar, A. Benassarou, L. Lucas, E. Elias, P. Tchelidze, D. Ploton and M-F. O'Donohue. How ReV4D helps biologists study the effects of anti-cancerous drugs on living cells. IEEE Visualization (VIS03), Poster Compendium, pp 46-47. 2003.
[8] N. Chentanez, M. Müller, M. Macklin, and T-Y. Kim. Fast grid-free surface tracking. ACM Transactions on Graphics, 34(4): 148:1-148:11, 2015.
[9] D. Kouril, L. Cmolik, B. Kozlikova, H-Y. Wu, G. Johnson, D.S. Goodsell, A. Olson, M.E. Gröller, and I. Viola. Labels on Levels: Labeling of Multi-Scale Multi-Instance and Crowded 3D Biological Environments. IEEE Transactions on Visualization and Computer Graphics, 25(1):977–2124, 2019.