

Fouille visuelle de données en 3D et réalité virtuelle : état de l'art

Zohra Ben Said*, Fabrice Guillet*, Paul Richard**

* LINA, UMR 6241 CNRS
(zohra.bensaid,fabrice.guillet)@univ-nantes.fr
www.lina.univ-nantes.fr

** LISA, EA 4094
paul.richard@istia.univ-angers.fr
www.istia.univ-angers.fr

Résumé. La fouille visuelle de données (ou Visual Data Mining, VDM) a pour objectif de faciliter l'interprétation des résultats issus d'une fouille de données, grâce à l'usage de représentations graphiques. Au cours de la dernière décennie, un grand nombre de techniques de visualisation d'information ont été mises au point, permettant la visualisation de données multidimensionnelles dans des environnements virtuels. Lors des travaux antérieurs, les chercheurs ont proposé des taxonomies pour classer les techniques de VDM (Chi (2000), Herman et al. (2000)). Toutefois, ces taxonomies ne prennent en compte que partiellement les techniques récentes relatives à l'utilisation de la 3D et de la réalité virtuelle. Le but de cet article est de faire un état de l'art récent et spécifique à ces techniques. Celles-ci sont détaillées, classées et comparées selon différents critères : les applications, l'encodage graphique, les techniques d'interaction, les avantages et les inconvénients de chaque approche. Ces techniques sont présentées dans des tableaux accompagnées d'illustrations graphiques.

1 Introduction

Une des limites sérieuses de la fouille de données provient des difficultés rencontrées par le décideur à interpréter les résultats produits. La fouille visuelle de données (visual Data Mining, VDM) propose de faciliter l'interprétation des résultats grâce à des représentations visuelles pertinentes. En effet, selon Card et al. (1999), l'esprit humain traite facilement les informations visuelles et en extrait rapidement un grand nombre de connaissances. Au début, les techniques de VDM étaient essentiellement basées sur des représentations 2D telles que les coordonnées parallèles (Inselberg et Dimsdale (1990)). Avec l'augmentation de la taille, de la complexité et de l'hétérogénéité des données stockées, de nouveaux paradigmes et métaphores de visualisation ont été développés. En effet, les résultats expérimentaux de Carswell et al. (1991) font état d'une amélioration des temps de réponse grâce à la 3D par rapport à la 2D pour la perception de tendances générales dans les données, nécessaires pour plusieurs applications telles que : (Tee et al. (2004), Schreck et al. (2007), Marroquín et al. (2008)). La 3D permet, grâce