

# Visualisation interactive de données avec des méthodes à base de points d'intérêt

David Da Costa<sup>\*,\*\*</sup>, Gilles Venturini<sup>\*</sup>

<sup>\*</sup>Laboratoire d'Informatique  
Ecole Polytechnique de l'Université de Tours  
64, Avenue Jean Portalis, 37200 Tours, France.  
david.dacosta@etu.univ-tours.fr,  
venturini@univ-tours.fr  
<http://www.antsearch.univ-tours.fr/webtrtic>

<sup>\*\*</sup>Agicom  
Institut d'Etudes  
3, degrés Saint Laumer, 41000 Blois, France.  
ddacosta@agicom.fr  
<http://www.agicom.fr/>

**Résumé.** Nous présentons dans cet article une méthode de visualisation interactive de données numériques ou symboliques permettant à un utilisateur expert du domaine d'obtenir des informations et des connaissances pertinentes. Nous proposons une approche nouvelle en adaptant l'utilisation des points d'intérêts dans un contexte de fouille visuelle de données. A partir d'un ensemble de points d'intérêt disposés sur un cercle, les données sont visualisées à l'intérieur de ce cercle en fonction de leur similarité à ces points d'intérêt. Des opérations interactives sont alors définies : sélectionner, zoomer, changer dynamiquement les points d'intérêts. Nous évaluons les propriétés d'une telle visualisation sur des données aux caractéristiques connues. Nous décrivons une application réelle en cours dans le domaine de l'exploration de données issues d'enquêtes de satisfaction.

## 1 Introduction

Les méthodes de fouille visuelle de données ("Visual data mining") tentent de résoudre les problèmes d'interprétation et d'interaction dans les processus de découverte de connaissances en faisant appel à des visualisations dynamiques et à des requêtes graphiques sur les données et connaissances représentées (Cleveland, 1993), (Shneiderman, 1996), (Wong et Bergeron, 1997). A titre d'exemples classiques, nous pouvons citer les visages de Chernoff (Chernoff, 1973) qui représentent des données sous la forme d'icônes en s'appuyant sur le fait que l'esprit humain analyse facilement les ressemblances et différences entre visages. Nous pouvons citer également les "scatter plots" (Becker et Cleveland, 1987) qui permettent d'obtenir des vues multiples sur les données et d'observer les données à l'aide de techniques graphiques comme le "brushing" qui donne la possibilité de sélectionner des données dans une vue tout en soulignant ces mêmes données dans les autres vues.

Ces méthodes apportent des nouveautés et poursuivent des objectifs qui sont prometteurs pour le domaine de la fouille de données : utilisation de la perception visuelle et souvent de la perception pré-attentive (Healey et al., 1993), interaction dynamique avec les données, simplicité d'utilisation, exploitation directe des résultats. Cependant, ces méthodes ont également des limites en ce qui concerne la fouille de données : les données visualisées sont le plus souvent numériques, les visualisations et leur manipulation nécessitent un apprentissage (comme c'est le cas par exemple pour interpréter des graphiques de types "parallel coordinates" (Inselberg, 1985)), l'interaction dynamique demande beaucoup de ressources de calcul (modifications en temps réel) et doit donc faire appel à des algorithmes les plus rapides possible (mais qui doivent par ailleurs fournir le plus d'information possible).

Dans ce travail, nous proposons une nouvelle méthode de fouille visuelle de données, adaptée elle-même des méthodes à base de points d'intérêt utilisées pour la visualisation de données textuelles. Nos objectifs, outre ceux poursuivis par la fouille visuelle de données, sont les suivants : pouvoir représenter tous les types de données en se basant uniquement sur l'existence d'une fonction de similarité (ou de distance) entre les données, avoir des affichages très rapides lors des interactions dynamiques et traiter si possible de grands volumes de données (algorithmes de complexités temporelle et spatiale linéaires en fonction du nombre de données), utiliser une visualisation nécessitant un temps d'apprentissage le plus court possible (donc compréhensible par la majorité des utilisateurs potentiels qui ne sont pas considérés comme des experts en fouille de données).

L'article est organisé comme suit : la section 2 décrit les techniques initiales utilisant les points d'intérêts dans le contexte de la visualisation de données textuelles. Dans la section 3 nous décrivons notre approche en commençant par spécifier l'utilisation des points d'intérêt pour la fouille de données puis en évaluant cette méthode sur des données aux caractéristiques connues. Dans la section 4 nous décrivons l'application de notre méthode sur des données classiques, puis le contexte d'application réelle de cette étude, à savoir extraire visuellement des connaissances à partir de données issues d'enquêtes de satisfaction. Nous concluons et dégageons des perspectives dans la section 5.

## 2 Survol des méthodes visuelles à base de points d'intérêts

Cette méthode est désignée par les termes "points d'intérêts" ou "points de références" (en anglais POIs pour "Points Of Interest"). Il s'agit dans cette méthode de positionner sur une surface, représentée par un disque, quelques icônes (POIs) relatifs aux attributs d'une donnée et ensuite d'afficher les icônes des données à des positions déterminées par la similarité entre les POIs et les données. Par exemple, cette visualisation a été utilisée comme méthode d'affichage de documents issus d'un moteur de recherche, facilitant la navigation dans l'ensemble des documents retournés par une requête. Les POIs sélectionnés sont généralement des mots-clés utilisés dans la requête et les données sont les documents se positionnant par rapport à ces mots-clés. Le choix des mots-clés dépend de la fréquence de leur occurrence dans les documents. Pour visualiser ces données, on utilise en général pour cela des techniques à base de ressorts et de forces, la force s'exerçant entre un POI et une donnée étant proportionnelle à la similitude entre ce POI et cette donnée. Les systèmes VIBE (Korfhage, 1991), ou un système rigoureusement simplifié de VIBE (Morse et al., 2002), SQWID (McCrickard et Kehoe, 1997), Radviz (Hoffman et al., 1999) ou encore la visualisation radiale du système Information Na-

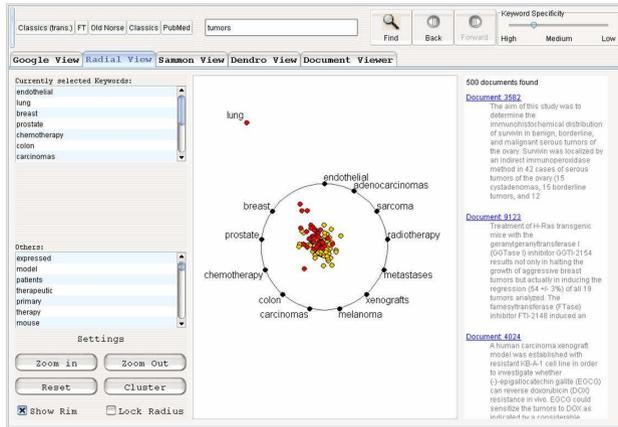


FIG. 1 – Exemple de visualisation radiale (Au et al., 2000) (image gracieusement fournie par Stefan Rürger).

vigator (Au et al., 2000) utilisent ces principes. Parfois il est difficile de voir exactement vers quel point d'intérêt est attiré une donnée. Dans ce cas ces systèmes permettent alors de supprimer et d'ajouter des points d'intérêt sur le disque pour permettre une meilleure représentation des données. Ce sont les principales opérations interactives proposées par ces méthodes.

Radial (Au et al., 2000) est une visualisation qui est très semblable au système VIBE (Korfhage, 1991), à Radviz (Hoffman et al., 1999) et à Lyberworld (Hemmje et al., 1994).

Initialement, après extraction du résultat de la requête, Radial identifie une série de termes clés relatifs à ce résultat. Ensuite les 12 premiers termes les plus recensés dans l'ensemble du résultat sont arrangés tout autour d'un cercle. Il est possible de modifier la liste des termes affichés, le choix se faisant sur deux listes placées à gauche de l'écran. Un nuage de points est alors affiché à l'intérieur du cercle, un point représente une donnée. Il n'est affiché que les données en rapport avec les mots-clés alignés autour du cercle. Un point est comme suspendu par des ressorts reliés aux mots-clés en rapport avec celui-ci. Il est donc impossible de déplacer un point en cliquant dessus, du fait des forces exercées par les ressorts. Par contre, en cliquant sur un point, les mots-clés en rapport avec ce point sont éclairés et une bulle affiche des informations sur cette donnée. Il est possible de déplacer les termes à l'extérieur du cercle et ainsi de déplacer tous les nœuds des données en rapport avec ces termes. Ceci permet de faire un classement manuel des résultats en catégories (voir la figure 1).

Tous ces systèmes ont montré que ce type de visualisation dynamique apporte un grand intérêt pour l'utilisateur qui peut extraire lui-même de l'information en toute simplicité. La rapidité d'affichage couplée à la possibilité d'interaction apportent un plus à ces méthodes. Par ailleurs, elles peuvent a priori visualiser des données de différentes natures comme des données symboliques ou numériques. A notre connaissance, ces méthodes à base de points d'intérêt n'ont pas encore été utilisées pour la fouille visuelle de données comme nous allons le présenter dans les sections qui suivent.

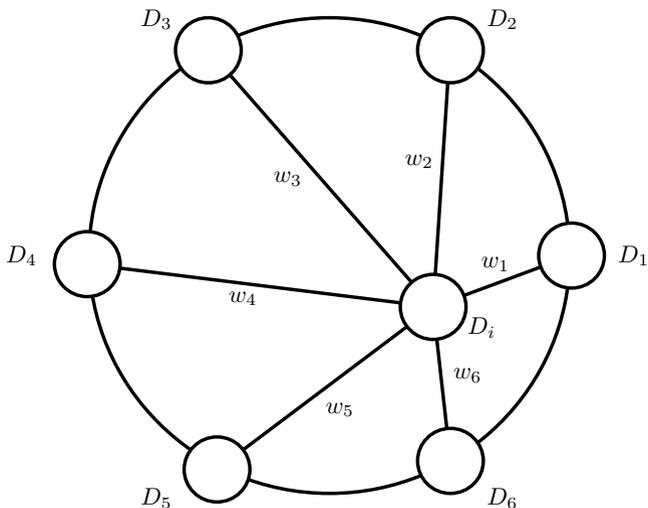


FIG. 2 – Principes de base de la visualisation avec illustration du positionnement d'une donnée  $D_i$  en fonction des  $k$  points d'intérêt.

### 3 Utilisation des points d'intérêt pour la fouille de données

#### 3.1 Principes de base de la visualisation

On considère  $n$  données  $D_1, \dots, D_n$  et une matrice de similarité  $Sim$  entre ces données.  $Sim(i, j)$  est la similarité entre les données  $D_i$  et  $D_j$ , cette matrice étant symétrique et avec une diagonale à 1. On note également que si  $Sim(i, j) = 1$  alors les données  $D_i$  et  $D_j$  sont identiques, et que si  $Sim(i, j) = 0$  alors elles sont totalement différentes.

Dans un premier temps, nous allons considérer que les POIs sont un sous-ensemble de ces données notés  $D_1, \dots, D_k$ . Nous affichons ces  $k$  données sur un cercle avec un arc de longueur constante entre chaque POI (voir la figure 2). On note par  $[(XD_1, YD_1), \dots, (XD_k, YD_k)]$  les coordonnées d'affichage des POIs.

On veut ensuite positionner les  $n - k$  données restantes en fonction de leur similarité aux POIs  $D_1, \dots, D_k$ . On utilise la formule suivante pour calculer les coordonnées d'affichage  $(XD_i, YD_i)$  de la donnée  $D_i$  :

$$\begin{aligned} XD_i &= w_1 \times XD_1 + \dots + w_k \times XD_k \\ YD_i &= w_1 \times YD_1 + \dots + w_k \times YD_k \end{aligned}$$

avec  $w_1 + \dots + w_k = 1$ . Pour cette donnée  $D_i$  et un POI  $D_j$ , le poids  $w_j$  est calculé de la manière suivante :

$$w_j = \frac{Sim(D_i, D_j)}{\sum_{p=1}^k Sim(D_i, D_p)}$$

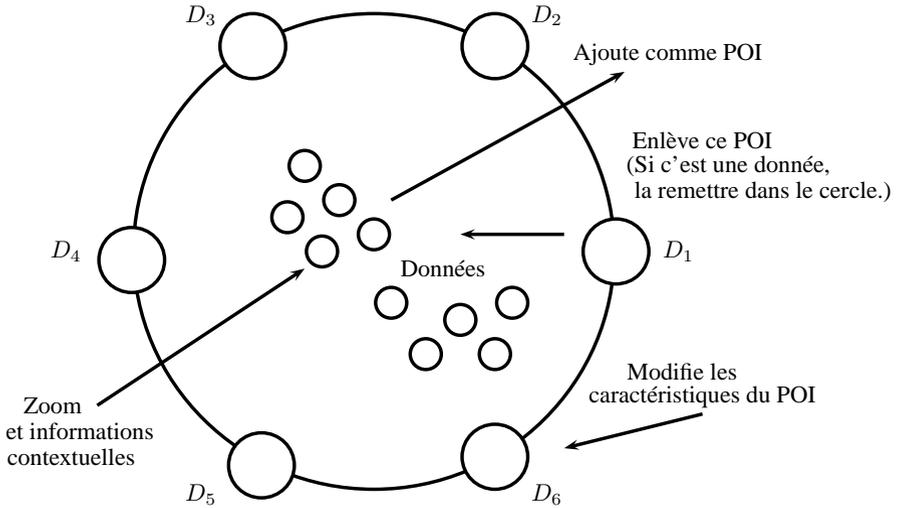


FIG. 3 – Principales interactions avec les données et les POIs.

Si  $D_i$  est identiquement similaire à l'ensemble des POIs, elle sera affichée au milieu du cercle. Inversement, si elle est totalement similaire à un POI et totalement différente des autres, sa position sera confondue avec celle de ce POI. Si sa similarité est biaisée vers certains POIs, celle-ci aura tendance à se rapprocher de ces POIs.

Plus généralement, notre méthode est telle que deux données proches l'une de l'autre dans l'espace d'origine le seront donc également vis à vis des POIs, et elles se retrouveront donc proches dans l'espace 2D. L'espace visualisé devient donc un espace de distance entre des points choisis (les POIs) et les données. C'est de cette manière que cette méthode peut traiter tout type de données. Par contre, la réciproque de cette propriété n'est pas vraie : deux données proches dans l'espace 2D ne le sont pas forcément dans l'espace d'origine (tous les points à distance égale de deux POIs dans l'espace d'origine forment une droite médiatrice, et ne sont pas systématiquement proches les uns des autres). Il faudra utiliser d'autres méthodes pour lever ces ambiguïtés (voir la dernière section).

Enfin, l'affichage, comme nous le recherchons, nécessite très peu de calcul et ne demande que de calculer qu'une partie seulement des similarités ( $k \times (n - k)$ ).

Plusieurs interrogations sont soulevées par cette méthode. Tout d'abord, le choix initial des points d'intérêt doit être effectué. Dans un premier temps, nous considérons que si les données sont supervisées (on dispose d'un label de classe), alors nous prenons le premier représentant de chaque classe comme POI initial. Il y aura donc autant de POIs que de classes dans la première visualisation proposée à l'utilisateur. Si les données sont non supervisées, nous choisissons les  $k$  premières données. D'autres choix automatiques sont possible (et certainement plus judicieux) comme nous le décrivons dans la dernière section, et nous rappelons qu'il

s'agit ici de proposer des choix initiaux que l'utilisateur va pouvoir modifier interactivement et dynamiquement en fonction de ce qui est affiché (voir section suivante). Une deuxième interrogation vient de l'ordre des POIs : si un grand nombre de données sont attirées par deux POIs, alors il y a tout intérêt à ce que ces POIs soient proches les uns des autres sur le cercle. Une situation critique consisterait à placer ces POIs de manière diamétralement opposée, ce qui engendrerait une visualisation peu lisible (beaucoup de données au centre). Nous proposons une solution interactive à ce problème dans la section suivante, mais il est évident que des solutions automatiques peuvent être trouvées comme ordonner les POIs en fonction de leur similarité (voir dernière section). Il serait également possible de ne pas conserver un arc de longueur fixe entre les POIs, afin de représenter les similarités qui existent entre POIs.

### 3.2 Interaction avec la visualisation

Pour être réellement efficace, la visualisation d'information doit être interactive et permettre d'affiner dynamiquement l'affichage et de répondre aux requêtes graphiques de l'utilisateur. Dans la visualisation avec des POIs, l'utilisateur peut se poser au moins les questions suivantes : quelle est cette donnée (ou ce POIs), comment agrandir cette partie de la visualisation (zoom sans perte de contexte), comment changer de POIs (en enlever, en rajouter, changer leur ordre, et comme on va le voir dans cette section, définir des POIs qui ne soient pas nécessairement des données de l'ensemble de départ).

Lors du passage de la souris sur un point, nous indiquons donc quel est ce point. Ensuite, il est possible de "zoomer" sur une donnée par clic de la souris. Le zoom qui se déclenche alors effectue les opérations suivantes : il centre la donnée sur le centre du cercle, il agrandit la zone centrée sur cette donnée et repousse les autres données vers les bords de la visualisation. La déformation est calculée à l'aide d'une fonction hyperbolique. Ce zoom permet de grossir la vue tout en conservant l'ensemble des données. Actuellement cette fonction utilise des coordonnées cartésiennes en considérant le carré dans lequel est contenu le cercle de la représentation, mais nous allons définir un zoom utilisant des coordonnées polaires, afin de laisser toutes les données à l'intérieur du cercle. La zone d'agrandissement sera plus petite mais les données resteront dans la zone habituelle.

En ce qui concerne les POIs, nous avons représenté sur la figure 3 les principales interactions possibles : tout d'abord, il est possible d'enlever un POI. Cela s'effectue très simplement en faisant glisser un POI à l'intérieur du cercle. Ce POI reprend alors sa place au sein des données. La vue est recalculée dynamiquement. Une transition dynamique et progressive est mise en place pour que l'utilisateur puisse suivre le changement de représentation. Ce dernier a la possibilité d'annuler son action, ce qui a pour effet de remettre le POI sur le cercle. Il est possible également de choisir une donnée et de la définir comme POI. Pour cela, on fait glisser la donnée sur le cercle. Si la donnée est placée sur un POI, elle remplace celui-ci, et si elle est placée entre deux POIs, elle s'insère en décalant les autres POIs de manière à maintenir constant les longueurs des arcs entre les POIs. Ces fonctionnalités sont très importantes puisqu'elles vont permettre à l'utilisateur de redéfinir à volonté la représentation.

Enfin, il est possible de généraliser les POIs de manière à ce qu'ils ne soient plus nécessairement des données, mais plus généralement tout point de l'espace de représentation et même tout objet pour lequel il est possible de calculer une similarité avec les données. Ainsi, on peut représenter des données "idéales", n'existant pas réellement, et par rapport auxquelles l'utilisateur voudrait positionner les données réelles. Nous présentons dans la section 4 une application

typique de cette fonctionnalité. Également, il serait possible de représenter par exemple une règle de décision comme un POI, et de placer les cas qui se rapprochent le plus de cette règle. Cette fonctionnalité offre de nombreuses perspectives en visualisant non plus seulement des données mais également des connaissances.

### 3.3 Autres propriétés

Pour illustrer le fonctionnement de notre méthode, nous avons représenté sur la figure 4 un exemple "jouet" de données dont on connaît parfaitement les caractéristiques. Dans cette figure, nous avons utilisé comme POIs les centres de chacune des classes. Après calcul de la matrice de similarité sur la base de la distance Euclidienne (les attributs sont préalablement normalisés), on constate que la représentation respecte l'organisation d'origine des classes et les relations de voisinage qui existent entre les données. À titre de comparaison, nous avons représenté les mêmes données avec les coordonnées parallèles, dont la représentation semble moins intuitive que notre approche. Nous montrons également dans cette figure qu'il peut être important de faire apparaître des rayons particuliers sur notre cercle. Ainsi, les rayons médians (en traits pleins) marquent la séparation entre les données attirées par l'un ou l'autre des deux POIs considérés. Ils permettent de dire qu'une donnée est plus proche de tel POI que de tel autre. Les rayons issus de chaque POI (traits en pointillés) marquent au contraire le fait qu'une donnée est à similarité égale entre les deux POIs opposés.

Sur la figure 5, nous avons utilisé une configuration différente des données "jouet". Une des classes est allongée, et l'on constate que l'on retrouve cette information dans notre visualisation. Cependant, d'autres dispositions des données (voir figure 6) donnent des informations locales plus difficiles à interpréter. Les relations de voisinage sont respectées.

## 4 Résultats

### 4.1 Bases artificielles et classiques

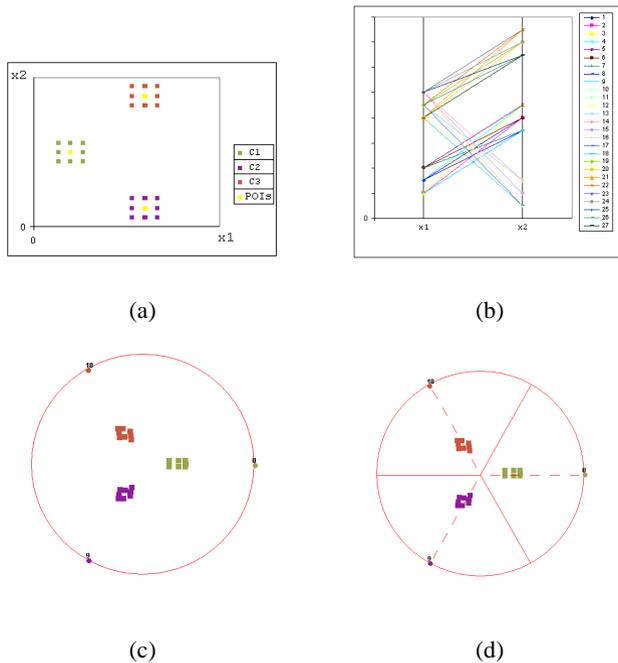
Nous avons évalué cette méthode sur un ensemble de bases composées de différentes données artificielles et classiques. La figure 7 représente la base de données artificielles Art1 composée de 400 données et de 4 classes. Nous illustrons en particulier les effets du zoom. Lorsque l'on dispose seulement de deux classes, les données des deux classes sont positionnées sur le segment de droite allant du POI1 au POI2 (cf. figure 8 où nous utilisons une base de 1000 données avec 2 classes). Pour aider l'utilisateur à mieux visualiser les données, on a permis l'ajout d'un POI supplémentaire tel que dans la figure 8(b) ou de plusieurs POIs (figure 8(c)).

Enfin, nous avons testé notre approche sur des bases classiques (issues du "Machine Learning Repository", (Blake et Merz, 1998)). Nous avons ainsi représenté sur la figure 9 les données Iris (150 données, 3 classes), Wine (178 données, 3 classes) et Segment (2310 données, 7 classes). On retrouve les formes de classes usuelles (comme pour Iris et Wine par exemple).

### 4.2 Application réelle

Dans le cadre de ses activités, Agicom collecte des données issues d'enquêtes de satisfaction à l'aide de questionnaires. Ces données se présentent sous la forme d'un tableau

## Visualisation à base de points d'intérêt



**FIG. 4** – Exemple "jouet" de visualisation de données représentées par deux attributs numériques  $x_1$  et  $x_2$  (voir en a)), avec visualisation sous la forme de coordonnées parallèles (en b)), puis représentation avec des POIs (en c) et d)).

individus  $\times$  variables où ces variables sont qualitatives, i.e. des variables dont les modalités sont symboliques et naturellement ordonnées ("ravi", "satisfait", "insatisfait", "déçu" et "NSP" (Ne Sais Pas)).

Pour qu'un consultant, ou qu'un expert, puisse exploiter ces données, il est important de pouvoir visualiser graphiquement la satisfaction des clients afin de détecter des correspondances possibles entre individus, de connaître l'évolution des clients d'un segment à l'autre, mais aussi de visualiser la relation existante entre une variable définie et les autres variables. Notre but est donc d'élaborer un outil de représentation graphique des résultats d'enquêtes de satisfaction contribuant au but de savoir comment améliorer la satisfaction des clients.

Nous avons évalué et testé notre méthode sur une première base Agicom1 composée de 31 données non supervisées. La figure 10(a) illustre cette première application dans laquelle les POIs ne sont pas des données mais des profils type de variables. Un profil correspond donc à une répartition de modalités (réponses) de cette variable. Ainsi, les POIs représentent différentes typologies connues de variables ou a priori (réponses très positives, combinaisons de modalités de variables).

Nous présentons un deuxième exemple sur la base Agicom2 (cf. la figure 10(b)). Dans cette

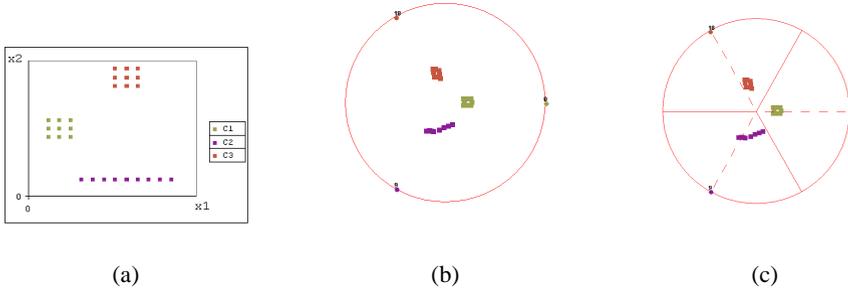


FIG. 5 – Autres données "jouet" initiales (a), puis visualisation avec notre méthode, (b) et (c).

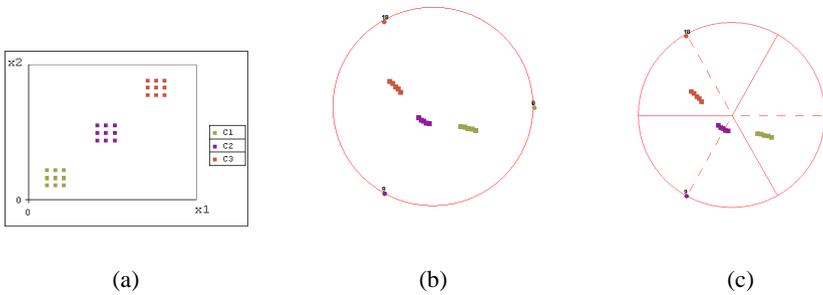


FIG. 6 – Autre disposition de données initiales (a) et visualisation, (b) et (c).

application, nous avons permis aux utilisateurs d'Agicom d'interagir sur les caractéristiques des POIs et sur le zoom (cf. la figure 10(b)). De plus, sur cette base nous avons permis la visualisation des différentes classes. Une validation avec des utilisateurs réels est en cours.

## 5 Conclusions et perspectives

Nous avons décrit dans cet article une nouvelle méthode de visualisation inspirée des travaux réalisés dans le contexte de la recherche documentaire sur les points d'intérêt. Elle consiste à transformer un espace d'origine représenté par une matrice de similarités en une représentation visuelle 2D de ces similarités. Cette méthode possède des atouts comme la rapidité d'affichage, une présentation intuitive des données et d'apprentissage plutôt rapide, des capacités interactives. Nous avons détaillé son comportement sur des données jouet, sur des données classiques et enfin dans le cadre d'une application réelle en cours de déploiement.

Plusieurs perspectives peuvent se dégager. Nous avons mentionné l'importance du choix des POIs ainsi que de leur disposition sur le cercle. Une première extension consiste à étudier

## Visualisation à base de points d'intérêt

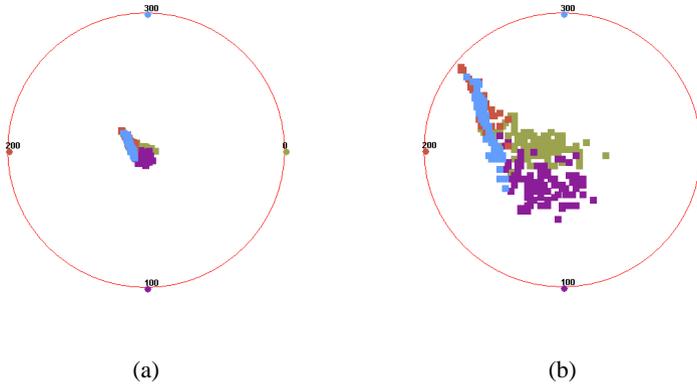


FIG. 7 – Visualisation de la Base Art1 sans zoom (a) et avec zoom (b).

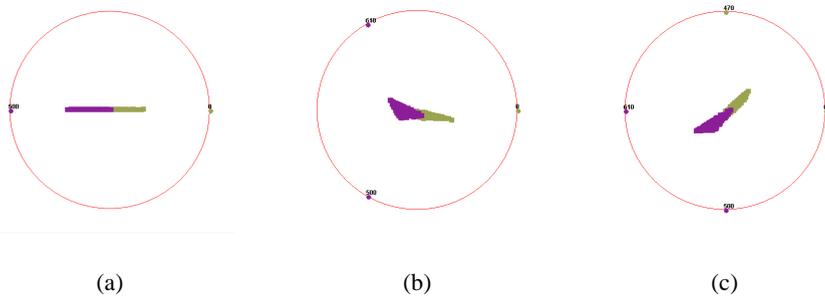


FIG. 8 – Base Art2 sans zoom (a), avec ajout d'un POI (b), puis d'un autre POI (c).

l'utilisation d'un algorithme d'optimisation afin de trouver la disposition de POIs la plus efficace visuellement. Il s'agit donc de trouver, dans l'espace des permutations des  $k$  POIs choisis, l'ordre qui permettra de maximiser certaines propriétés comme la ressemblance entre deux POIs successifs sur le cercle. Une autre perspective importante consiste à étendre la visualisation de manière à enlever les ambiguïtés liées au chevauchement des points. Nous comptons utiliser une méthode d'affichage de graphe à base de forces et de ressorts afin d'éloigner les points qui se trouvent trop proches sur le graphe. Il s'agit aussi de distinguer les points qui sont placés au même endroit mais qui ont une similarité moyenne différente avec les POIs. Une approche 3D sera testée. Enfin, comme nous l'avons mentionné, nous allons ajouter une méthode classique de zoom hyperbolique travaillant sur des coordonnées polaires afin que les points ne sortent pas du cercle.

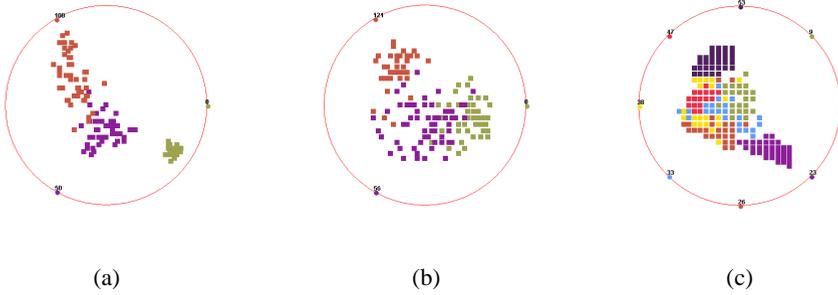


FIG. 9 – Visualisations des bases Iris (a), Wine (b), et Segment (c).

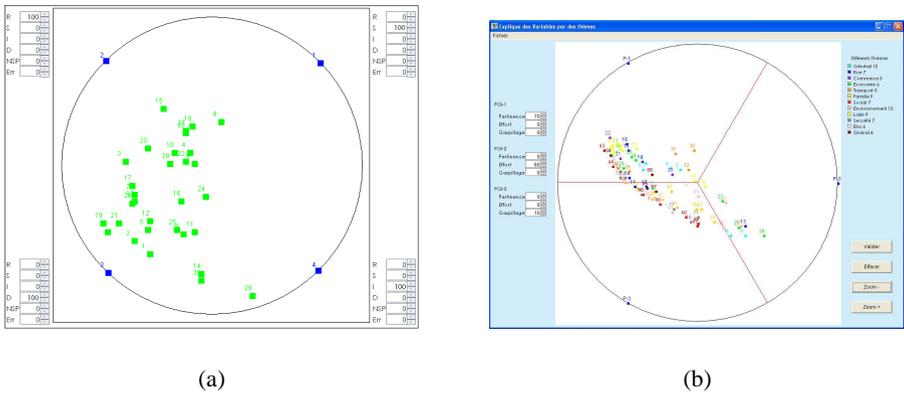


FIG. 10 – Visualisation des bases Agicom1 (a) et Agicom2 (b).

## Références

Au, P., M. Carey, S. Sewraz, Y. Guo, et S. M. Rüger (2000). New paradigms in information visualization. In *Research and Development in Information Retrieval*, pp. 307–309.

Becker, R. A. et W. S. Cleveland (1987). Brushing Scatterplots. *Technometrics* 29, 127–142. Reprinted in *Dynamic Graphics for Data Analysis*, edited by W. S. Cleveland and M. E. McGill, Chapman and Hall, New York, 1988.

Blake, C. et C. Merz (1998). UCI repository of machine learning databases.

Chernoff, H. (1973). Using faces to represent points in  $k$ -dimensional space graphically. *Journal of the American Statistical Association* 68, 361–368.

Cleveland, W. S. (1993). *Visualizing Data*. Summit, New Jersey, U.S.A. : Hobart Press.

Healey, C. G., K. S. Booth, et J. T. Enns (1993). Harnessing preattentive processes for multi-

- variate data visualization. In *Proceedings of Graphics Interface '93*, Toronto, ON, Canada, pp. 107–117.
- Hemmje, M., C. Kunkel, et A. Willett (1994). Lyberworld visualization user interface supporting fulltext retrieval. In *SIGIR '94 : Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, New York, NY, USA, pp. 249–259. Springer-Verlag New York, Inc.
- Hoffman, P., G. Grinstein, et D. Pinkney (1999). Dimensional anchors : a graphic primitive for multidimensional multivariate information visualizations. In *NPIVM '99 : Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth ACM international conference on Information and knowledge management*, New York, NY, USA, pp. 9–16. ACM Press.
- Inselberg, A. (1985). The plane with parallel coordinates. *The Visual Computer 1*, 69–91.
- Korfhage, R. (1991). To see, or not to see : Is that the query ? In A. Bookstein, Y. Chiaramella, G. Salton, et V. V. Raghavan (Eds.), *Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Chicago, Illinois, USA, October 13-16, 1991 (Special Issue of the SIGIR Forum)*, pp. 134–141. ACM.
- McCrickard, S. et C. Kehoe (1997). Visualizing search results using sqwid. In *Proceedings of the Sixth International World Wide Web Conference*.
- Morse, E., M. Lewis, et K. A. Olsen (2002). Testing visual information retrieval methodologies case study : comparative analysis of textual, icon, graphical, and "spring" displays. *J. Am. Soc. Inf. Sci. Technol.* 53(1), 28–40.
- Shneiderman, B. (1996). The eyes have it : A task by data type taxonomy for information visualizations. In *IEEE Visual Languages*, Number UMCP-CSD CS-TR-3665, College Park, Maryland 20742, U.S.A., pp. 336–343.
- Wong, P. C. et R. D. Bergeron (1997). 30 years of multidimensional multivariate visualization. In *Scientific Visualization — Overviews, Methodologies and Techniques*, pp. 3–33. Los Alamitos, CA : IEEE Computer Society Press.

## Summary

We present in this paper an interactive method for numeric or symbolic data visualization that allows a domain expert to extract useful knowledge and information. We propose a new approach based on points of interest but in the context of visual data mining. Points of interest are located on a circle, and data are displayed within this circle according to their similarities to these points of interest. Interactive actions are possible: selection, zoom, dynamical change of points of interest. We evaluate the properties of such a visualization with standard data with known characteristics. We describe a real world application which is under study in the context of data exploration resulting from satisfaction inquiries.