

Techniques visuelles de recherche d'information

Fewzi Mokaddem*, Fabien Picarougne*, Hanene Azzag*,
Christiane Guinot**, Gilles Venturini*

*Laboratoire d'Informatique de l'Université de Tours,
École Polytechnique de l'Université de Tours - Département Informatique,
64, Avenue Jean Portalis, F-37200 Tours.
fewzi.mokaddem@etu.univ-tours.fr,
{fabien.picarougne, hanene.azzag, venturini}@univ-tours.fr,
**CE.R.I.E.S.,
20 rue Victor Noir, F-92521 Neuilly-sur-Seine Cedex
christiane.guinot@ceries-lab.com.

Résumé. Nous exposons dans cet article un état de l'art sur les techniques visuelles pouvant être utilisées dans la recherche d'information documentaire sur Internet ou dans un système d'information. Nous détaillons dans un premier temps les techniques qui permettent aux utilisateurs de formuler et d'affiner leurs requêtes de façon interactive et visuelle. Nous présentons ensuite les techniques permettant de représenter un document isolément puis les techniques et systèmes permettant de représenter un ensemble de documents. Nous analysons les atouts et faiblesses de ces méthodes et nous dégageons des perspectives pour ce domaine prometteur.

1. Introduction

La quantité d'information offerte au public sur le Web est très importante et augmente sans cesse. Il est certain que les moteurs de recherche à interface texte dite « classique » vont devenir de moins en moins efficaces face à cette inflation de l'information. En général, avec ce genre de méthodes classiques, les interfaces sont toutes composées sur le même modèle : un champ texte associé généralement à une ligne pour saisir la requête, un bouton pour lancer la recherche, un affichage des résultats sous forme de listes composées de centaines et parfois de milliers de documents relatifs aux critères. L'utilisateur vérifie les résultats et reformule la requête plusieurs fois pour trouver ce qu'il cherche. Une étude [Jansen et al., 1998]¹ a montré qu'en moyenne 20 à 30 documents sont explorés par les utilisateurs, car ils se basent généralement sur le classement. Or ce n'est pas toujours un moyen efficace pour trouver le résultat souhaité rapidement.

Durant cette dernière décennie plusieurs travaux ont été menés en perception visuelle comme par exemple [Myers, 2000] qui a montré que l'être humain a une perception d'abord globale (gestalt-perception) d'une scène, avant de porter son attention aux détails. Ce type de travaux a guidé d'autres recherches sur l'analyse concurrentielle, le data mining, la cartographie de fonds documentaires, vers la conception de nouvelles techniques pour parer les faiblesses des anciennes méthodes, et particulièrement la recherche d'information sur le Web. Des systèmes ont émergé ces dernières années dans le but de permettre aux utilisateurs

¹ 86% de 18 113 utilisateurs n'ont consulté pas plus que les trois premières pages résultant d'Excite (avec 10 liens par page), 77% pas plus de deux pages et 58% pas plus d'une page.

de trouver en un temps raisonnable une liste de résultats optimale et fidèle au but recherché en visualisant les requêtes et les résultats et en permettant des opérations interactives.

Des travaux de synthèse et de classification de ces méthodes ont été accomplis chacun avec des approches différentes. [Shneiderman, 1996] a proposé la TTT (Data Type by Task Taxonomy), une classification composée de sept types de données et les sept tâches qui leurs sont appliquées. Suite à cette étude l'auteur extrait les principes des moteurs de recherche à interface visuelle "Premièrement vue d'ensemble, zoom et filtre, ensuite détails à la demande". Dans [North, 1997] et [Olive, 1997], où a été utilisée la TTT, les types de données ont été augmentés par un huitième appelé "espace de travail". [Zamir, 1998] a proposé une classification en deux groupes : 1) Visualisation des attributs de requêtes (distribution des termes de la requête, attributs prédéfinis et attributs définis par l'utilisateur), 2) Visualisation des attributs de documents (Réseau, intégration de ressorts, groupement et cartes auto-organisatrices). [Mann, 1999] s'est inspiré de la classification de [Zamir, 1998] et des quatre phases de la recherche d'information de [Shneiderman et al. 1997] pour élaborer son schéma qui résume les étapes suivies par un utilisateur devant ce genre de moteur de recherche. D'autres taxonomies peuvent être consultées par exemple dans [Card et al. 1999] et [Chi, 2000].

Dans cet article, nous exposerons les techniques de visualisation utilisées (ou utilisables) à ce jour pour parcourir les résultats d'un moteur de recherche ou d'un système d'information. Nous commencerons par détailler les attributs visuels de requêtes (section 2), les attributs visuels de documents (section 3) et la visualisation d'ensembles de documents (section 4). Nous citerons dans la section 5 quelques applications commerciales opérationnelles et des techniques d'évaluation, et nous proposerons en conclusion (section 6) une classification de toutes ces méthodes et des perspectives.

2. Techniques de formulation visuelle des requêtes

Formuler une requête concise et fidèle au but recherché n'est pas toujours facile pour l'utilisateur et surtout pour ceux qui ne sont pas familiers avec les expressions logiques. L'amélioration à ce niveau a consisté à concevoir des interfaces de formulation visuelle et interactive sur un espace 2D ou 3D. Le plus souvent, après la confirmation de la requête généralement dérivée du langage naturel et l'extraction d'un ensemble de résultats, le système exécute automatiquement certaines opérations comme l'identification des mots parasites et des phrases significatives. Ces éléments seront proposés à l'utilisateur par la suite tout en lui donnant une vue sur la répartition de l'ensemble ou d'une partie des résultats par rapport aux mots-clés retenus. Nous mentionnons parmi ces systèmes AI-STARS [Anick et al., 1990], dans l'interface duquel la première ligne contient la phrase de la requête, chaque mot est dans un pavé. Les mots jugés parasites par le système sont affichés dans des pavés blancs, les autres dans des pavés noirs. Un chiffre en bas à gauche de chaque pavé indique le nombre d'occurrences de chaque terme dans le résultat. Les pavés alignés sur la même colonne sont des synonymes proposés par le système. Nous mentionnons aussi le système de [Stenmark, 1999] qui propose une interface ayant la particularité de permettre de formuler une requête avec peu de mots en langage naturel. Au fur et à mesure que l'on saisit la requête, le système regroupe les résultats en domaines et donne une première vue sur le résultat en pourcentage pour chaque domaine. Ainsi, il est possible de restreindre sa recherche à des domaines précis. Ensuite en cliquant un bouton « Continue » une nouvelle fenêtre s'affiche qui suggère de nouveaux synonymes. Si l'utilisateur le désire, il peut

visualiser les résultats sous forme de listes accompagnées de « tilebars » (voir section suivante). L'utilisateur a toujours la possibilité de faire marche arrière entre les différentes étapes afin d'affiner sa requête.

Nous citons maintenant d'autres systèmes basés sur des principes différents du précédent. Nous les jugeons moins efficaces car ils n'acceptent pas de phrases en langage naturel et n'analysent pas le sens même de la requête saisie. Ils se contentent uniquement de représenter visuellement la requête, améliorée par quelques indications sur le résultat attendu. Parmi ces principes on peut distinguer les diagrammes de Venn. Ils ont été utilisés dans plusieurs cas pour représenter les requêtes logiques. [Jones, 1998] et [Jones, 1998a] exposa une bonne synthèse des diagrammes de Venn et les a intégré dans l'interface du système Vquery (voir figure 1(c)). Dans ce type de système, l'information est codée avec plusieurs éléments : la forme, la proximité, le rang, l'orientation, et la couleur ou la texture. Dans certains cas spéciaux, la taille, la luminance et la saturation sont utilisées. La représentation en diagramme de Venn est limitée par le fait qu'elle ne peut pas montrer d'intersections entre un nombre important d'attributs. Il est impossible de représenter des expressions plus complexes. [Berenci et al., 1998] ont créé le système VIEWER (VIEWS of WEb Results) qui contient un composant qui montre une représentation des sous requêtes (combinaison complète entre tous les mots-clés). [Bürdek et al., 1999] proposa de saisir les mots-clés en cliquant sur les « T » [Eibl, 1999], la longueur des barres des sous-requêtes indique la taille par rapport à l'ensemble des résultats. En cliquant sur une barre est affiché dans une autre fenêtre la liste des documents qui y sont contenus.

D'autres principes existent comme ceux proposés par [Cugini et al., 1998], [Cugini et al, 2000] tels que « la matrice des concepts/mots-clés » ou encore la technique employée pour interroger un système d'information proposée par [Fishkin et Stone, 1995] et nommée filtres mobiles ou filtrage à l'aide de lentilles magiques. L'idée est de représenter l'ensemble des résultats par un « scatter plot » et ensuite de créer autant de fenêtres que de mots-clés. Il est possible de superposer plusieurs filtres afin d'exprimer le but d'une requête en indiquant s'il s'agit d'un ET logique ou d'un OU logique qui est appliqué. Les documents satisfaisant la requête seront facilement distinguables à travers les filtres, puisqu'ils seront colorés différemment des autres (figure 1(a)).

Nous avons regroupé dans le tableau 1 les techniques visuelles employées pour la formulation de requêtes.

3. Visualisation des attributs d'un document

Nous rassemblons dans cette section les techniques visuelles qui permettent de représenter graphiquement un document, soit de manière isolée ou bien parmi un ensemble de documents. Par attribut de documents, nous considérons des informations de nature très diverses comme la proportion des mots-clés formulés dans la requête, la taille du document, sa date de parution ou de dernière lecture, etc.

Les vues en miniature sont l'une des représentations basiques d'un document (figure 1(b)). Ainsi, les systèmes « Graphic History View of the MosaicG » de [Ayers, Stasko 1995] ou « Data mountain » de [Robertson et al. 1998] permettent la navigation graphique à travers l'historique des pages Web consultées. Le système Webview appliqué au navigateur Netscape [Cockburn et al., 1999] permet de naviguer à travers la liste des sites favoris. Il offre en plus de la vue miniature d'autres indices visuels représentant des informations sur le nombre de visites rendues à cette page, sur la fréquence de consultation et sur le temps qui

<i>Composant</i>	<i>Références</i>	<i>Système</i>
Espace de reformulation des requêtes	[Anick et al., 1990]	<i>AI-STARs</i>
	[Stenmark, 1999]	
Filtres mobiles / Filtrage avec des lentilles magiques	[Fishkin et Stone, 1995]	
Diagrammes de Venn	[Hertzum et Frokjaer, 1996]	<i>TeSS</i>
	[Jones, 1998], [Jones, 1998a], [Jones, et al. 1999]	<i>VQuery</i>
	[Spoerri, 1993], [Spoerri, 1993a]	<i>InfoCrystal</i>
Crochets (« Bracket »)	[Bürdek et al., 1999], [Eibl, 1999], [Eibl, 1999a]	<i>GESINE</i>
Graphe de barres (« Bargraph »)	[Berenci et al., 1998]	<i>VIEWER</i>
	[Stenmark 1999]	
Matrice des Concepts/Mots-clés	[Cugini et al., 1998], [Cugini et al., 2000]	<i>NIRVE</i>
Globe des concepts/ vue globale	[Cugini et al., 2000]	<i>NIRVE</i>
Feedback Positif / Négatif	[Veerasamy et Navathe, 1995], [Veerasamy et al., 1995]	<i>Tking</i>

Tab. 1- Composants visuels de représentation des requêtes

sépare la première et la dernière consultation de cette page. Une version équivalente a été implémentée sur Internet Explorer de Microsoft par [Kaasten et Greenberg, 2000], [Kaasten et Greenberg, 2001].

La technique de déformation de la vue (« fisheye ») qui consiste à se concentrer sur une partie de l'image tout en gardant une vue sur le contexte peut être appliquée sur des images réduites en offrant des vues « fermée », « semi-ouvertes » et « ouvertes ». On peut citer Le système DigOut4U² et aussi [Kaugars, 1998], [Ogden et al., 1998], et le système SeeSoft [Eick, 1994], [Wills, 1995]. Dans SeeSoft, tous les paragraphes d'un document sont affichés en barres colorées ordonnées en colonnes et les lignes de chaque paragraphe en rangées. La technique « fish-eye » offre un très bon compromis Espace de présentation/Détails, mais à l'inconvénient de ne pas être facilement comprise par les utilisateurs. Afin d'éviter ce genre de situation, certains auteurs ont repris des représentations inspirées des graphiques utilisés classiquement dans les statistiques tels les histogrammes ou encore l'approche basée sur les feux tricolores du système SQWID [McCrickard et Kehoe, 1997] (figure 3(b)). L'avantage vient du fait que les utilisateurs sont déjà familiers avec ce type de représentations comme pour les propositions de [Veerasamy, 1997] et [Grewal et al., 2000]. [Grewal et al., 1999] ont présenté par exemple le composant R-Wheel (« Roue de pertinence »). Le principe est le même que celui des barres d'histogrammes mais cette fois chaque document est représenté par un disque. Chaque terme a son propre secteur dans ce disque et sa propre couleur. Le secteur est rempli proportionnellement à la pertinence du terme. [Cugini et al., 1998] pour leur système NIRVE 3D ont présenté des icônes sous forme de barre d'histogrammes (voir section 4.1.2).

² Développé par Arisem S.A., Paris, France.

<i>Composant</i>	<i>Références</i>	<i>Système</i>
Vues des images réduites (Thumbnail views)	[Ayers et Stasko, 1995]	<i>MosaicG</i>
	[Robertson et al.,1998],[Czerwinski et al.,1999]	<i>Data Mountain</i>
	[Cockburn et al., 1999]	<i>Webview</i>
	[Kaasten et Greenberg, 2000], [Kaasten et Greenberg, 2001].	
	[Mann, 1999] [Kaugars, 1998]	
Courbes de pertinences	http://www.arisem.com [2001-02-11]	<i>DigOut4U</i>
« Vue semi-ouverte »	[Kaugars, 1998]	
« Thumbnail » et « fisheye »	[Ogden, et al.,1998]	<i>J24</i>
« Bar view »	[Eick et al.,1992], [Eick, 1994], [Wills, 1995]	<i>SeeSoft</i>
« R-Wheels »	[Grewal et al. 1999], [Grewal et al., 2000]	
Représentation iconique	[Cugini et al.,1997], [Cugini et al.,1998], [Cugini et al., 2000]	<i>NIRVE</i>
Histogrammes des historiques des extractions	[Golovchinsky, 1997]	<i>VOIR</i>
Feux tricolores	[McCrickard et Kehoe, 1997]	<i>SQWID</i>
« TileBars »	[Hearst, 1995]	
	[Byrd, 1999]	<i>FancyV</i>
	[Spring et al.,1996], [Heo et al.,1996]	<i>CASCADE</i>
	[Dieberger et Russell, 2001]	
Information murale	[Jerding et Stasko,1995], [Jerding et Stasko,1997]	
« Mural »	[Spring et al.,1996], [Heo et al., 1996]	<i>CASCADE</i>
<i>VQRb</i>	[Byrd, 1999]	<i>FancyV</i>

Tab. 2 - Techniques de visualisation d'un document

[Golovchinsky, 1997] pour le système VOIR a utilisé les barres d'histogrammes pour exprimer la pertinence d'un document par rapport au résultat de recherches issues d'un ensemble de requêtes sur une session de navigation et non par rapport à une seule requête. Ainsi le nombre de barres correspond au nombre de requêtes où le document a été extrait.

Probablement un des plus célèbres composants de la visualisation de document sont les « Tilebars » [Hearst, 1995]. La longueur de la barre représentant un document est proportionnelle à la longueur du document. Des carrés en niveaux de gris indiquent la présence de mots-clés à l'intérieur du document. Plusieurs lignes de carrés peuvent être présentes pour plusieurs mots-clés. Une fois le principe des TileBars compris par les utilisateurs, ces derniers peuvent parcourir visuellement les représentations graphiques et identifier rapidement quels documents sont et ne sont pas intéressants.

On peut remarquer que plusieurs idées ont été influencées par les « Tilebars », comme par exemple le système CASCADE [Spring et al., 1996], [Heo et al., 1996]. Ces auteurs ont combiné les « TileBars » avec une barre de défilement. En cliquant sur un des pavés (situé à gauche), l'écran se positionne à l'endroit correspondant dans le document. Un autre composant appelé le mur (« Mural ») introduit par [Jerding et Stasko, 1995] et [Jerding et Stasko, 1997] utilise en plus des traits qui délimitent les liens hypertextes au sein du document. Un autre exemple sont les courbes de pertinences.

Le composant VQRb du prototype FancyV [Byrd, 1999] a combiné les barres de pavés et une barre de défilement. [Dieberger et Russell, 2001] ont utilisé pour leur système les deux orientations des barres de pavés : horizontale de [Hearst, 1995] et verticale de [Heo et al., 1996].

Nous mentionnons ici un outil un peu différent des précédents : « Document Lens » est un outil développé pour le système « Information Visualizer » de [Robertson et Mackinlay, 1993]. Il permet l'exploration des pages d'un document (ou d'un ensemble de documents) classées sur une grille dans l'ordre de la lecture en images miniatures (« Thumbnails »). La lentille peut être déplacée pour zoomer sur un document. Ceci permet d'avoir une vue plus lisible tout en gardant une vue sur le reste des pages ou des autres documents.

Nous avons regroupé dans le tableau 2 les techniques de visualisation d'un document.

4. Techniques de visualisation d'un ensemble de documents

Montrer une vue d'ensemble du résultat permet à l'utilisateur d'avoir une idée sur la répartition des documents et de discerner les rapports qui existent entre eux. Les documents sont généralement évalués selon trois modes : nominal (égal ou différent d'un autre), ordinal (obéir à une relation d'ordre) ou quantitatif (supportant les opérations arithmétiques). Pour calculer la position de la représentation d'un document dans un espace virtuel 2D ou 3D, il faut définir des méthodes qui permettent de préserver au maximum l'information originaire d'un espace initial de représentation qui comporte souvent n dimensions. Ces méthodes peuvent être classées en six groupes : la sélection de dimensions, les cartes auto-organisatrices, l'utilisation des points d'intérêts, les regroupements implicites, les structures hiérarchiques et la compression des dimensions. Nous parlerons aussi des possibilités de coordination entre plusieurs visualisations et citerons quelques exemples de métaphores.

4.1. Sélection de dimensions

Dans ce cas le système donne à l'utilisateur la possibilité de sélectionner quelques dimensions et les affecte ensuite à des axes X , Y et parfois Z , ou encore dans certains cas à une échelle de couleurs ou à des volumes d'icônes. Les dimensions d'origine ne sont donc pas transformées mais utilisées directement.

4.1.1 Systèmes à une et deux dimensions

Il y a peu de systèmes qui utilisent aujourd'hui une seule dimension pour la visualisation. Citons néanmoins la visualisation introduite dans [Mann, 1999] où par exemple l'axe X de la représentation est la pertinence des documents et l'axe Y le nombre de documents correspondant à cette pertinence. Notons cependant que dans [Mann, 1999] cette visualisation 1D vient en complément à d'autres types de visualisation des mêmes documents (voir section 4.1.3).

Un bon exemple de visualisation en 2D est celui des diagrammes de type « scatterplot ». Le système FilmFinder [Ahlberg et Shneiderman, 1994] [Jog et Shneiderman, 1995], le successeur du système Dynamic Homefinder, sélectionne par exemple pour les deux axes X et Y des variables comme la date et la popularité d'un film. Le genre du film est représenté par des couleurs. L'idée a été implémentée pour les systèmes IVEE [Ahlberg et Wistrand, 1995], [Wistrand, 1995] et Spotfire Pro [Spotfire, 2001]. La figure 2(a) montre un exemple sur une base de données de films employant l'outil de démonstration « Spotfire ». Cette représentation possède cependant le défaut de permettre à des icônes d'en occulter d'autres.

Le système Envision, avec sa matrice des icônes [Nowell et al., 1996], emploie une autre façon de choisir des attributs. Il s'agit de présenter les documents à l'intérieur d'une matrice. Les lignes représentent par exemple l'année de publication des documents et les colonnes représentent les noms des références. Les documents sont classés en employant le mode nominal et le mode ordinal. La taille d'une icône représente sa pertinence par rapport à la requête. Un autre exemple plus récent est celui du moteur de recherche spécialisé AmazonDNA³ fondé sur le paradigme de visualisation WISDNA (Web Information Services DNA) proposé par [Baeza-Yates et al., 2003].

Il est possible aussi d'attribuer aux deux axes des valeurs différentes de celles relatives aux mots-clés comme par exemple les coordonnées géographiques utilisées dans le système VisageWeb [Higgins et al., 1999] qui donne des vues miniatures de pages Web (« thumbnails ») positionnées sur la carte du monde en fonction de leur adresse IP.

On peut utiliser également d'autres métaphores graphiques à deux dimensions comme la technique du « Perspective Wall » de [Mackinlay et al., 1991] qui succède au système « Bifocal » décrit par [Spence et Apperley, 1982]. Le système emploie une technique qui montre une vue d'un ensemble de documents tout en se concentrant sur les détails relatifs à un seul d'entre eux. Les informations sont représentées par des étiquettes placées sur des murs et classées pour l'axe horizontal selon l'ordre dans la liste et l'axe vertical selon un autre attribut comme par exemple la date de parution du document. Bien que leur représentation soit plus claire que celles des « scatter plots », ses systèmes n'affichent qu'un nombre faible de documents au même instant.

4.1.2 Systèmes à trois dimensions

Un exemple en 3D est celui des trois axes de mots-clés décrit dans [Cugini et al., 2000]. L'utilisateur choisit 3 mots-clés pour chaque axe (figure 4(a)). On peut attribuer à un axe un ensemble de mots-clés au lieu d'un seul. Les icônes se mettent dans une position relative aux mots-clés des trois axes. Les défauts de ce système, comme pour les scatterplot, viennent du fait que des problèmes d'occlusion peuvent apparaître : des icônes peuvent en cacher d'autres. Généralement, l'utilisateur résout ce problème en déplaçant quelques éléments pour explorer certaines parties rendues alors visibles. Cependant, il demeure difficile pour l'utilisateur de se rappeler si un document a déjà été consulté ou non.

4.1.3 Systèmes à n dimensions

Dans les histogrammes interactifs de [Tweedie et al., 1994] et [Tweedie et al., 1996], il s'agit d'utiliser une sélection de plusieurs dimensions et de les grouper horizontalement ou verticalement pour représenter l'ensemble des résultats sur un espace multidimensionnel à

³ Lien vers le site d'AmazonDNA: in <http://soap.lemus.cl/>

l'instar du système de visualisation « Parallel coordinates » [Inselberg , 1985]. Avec cette visualisation on arrive à discerner les icônes les uns des autres sans devoir les déplacer.

Citons également une technique de visualisation exploitant les pixels de l'écran. Elle est utilisée par [Keim et Kriegel, 1994] pour le système VisDB ainsi que par [Ankerst et al., 1996] dans une autre alternative combinant l'utilisation des pixels et les segments de cercle. Le but de cette technique était de permettre de trouver des données intéressantes dans les domaines scientifiques, de l'ingénierie et plus généralement les environnements de base de données. Le principe est de cartographier des items de données en pixels colorés en les ordonnant selon leur pertinence par rapport à la requête dont ils sont issus. Les pixels sont arrangés selon le modèle en spirale, l'ordre de positionnement des pixels commence par le centre et va vers l'extérieur. En se concentrant dans ce genre de vues sur l'étendue d'une couleur qui est relative à un terme de la requête on peut déduire lequel des termes est dominant. Cette technique hérite des défauts de la représentation en spirale (section 4.4) et ne permet pas de représenter un nombre important d'attributs sur le même écran (le nombre de couleurs dissociables par l'œil humain étant limité). Par contre on obtient de bons résultats en se concentrant sur la dominance d'un ou de deux attributs parmi un ensemble de documents. Cette méthode permet par ailleurs d'analyser plus de documents simultanément que les autres méthodes grâce à la représentation par un seul pixel (limite d'affichage d'un élément sur un écran).

Les techniques de sélection de dimensions présentent toutes la même faiblesse en ne fournissant pas une mesure précise de la similarité entre les documents deux à deux. Elles se contentent uniquement de quelques attributs comme l'année ou la taille, ce qui peut quelquefois ne pas être satisfaisant pour l'utilisateur.

4.2. Cartes auto-organisatrices

Les cartes auto organisatrices (SOM, « Self-Organizing Map ») ont été introduites pour la première fois dans [Kohonen, 1981]. Les premiers modèles cherchaient tout particulièrement à représenter des données multidimensionnelles. Les applications visées devaient pouvoir concerner de très grands ensembles de données textuelles notamment. Afin de répondre à ces critères, la visualisation par cartes topologiques envisagée par [Kohonen, 1981] cherche par apprentissage à partir des données, à partitionner l'ensemble des observations disponibles en groupements similaires. Les groupements proposés possèdent la particularité d'avoir une structure de voisinage qui peut être matérialisée à l'aide d'un espace discret que l'on appelle « carte topologique ». D'autres applications basées sur les cartes auto-organisatrices se trouvent sur le site de l'université d'Arizona comme par exemple « Adaptive SOM » [Roussinov, 1999] et ET-map de [Chen et al., 1996].

La figure 2(b) montre une vue sur l'applet libViewer du système SOMLib [Raubert et Bina, 1999] qui affiche une partie de la carte auto-organisatrice d'une collection de 420 articles du magazine TIME depuis 1960. La carte utilise une grille de dimensions 10x15 et est représentée par une métaphore de type bibliothèque.

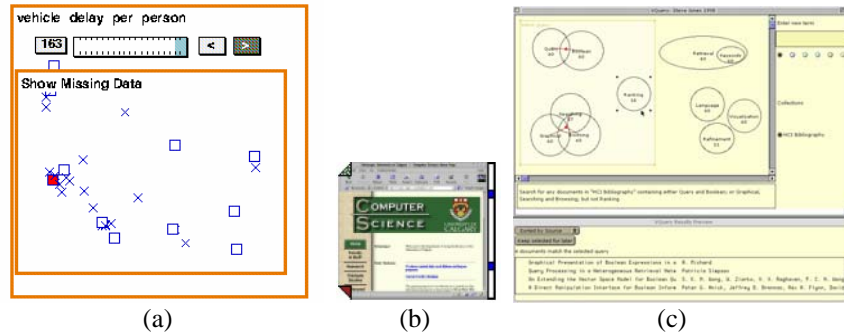


Fig.1 En (a) : « Magic lens » [Fishkin et Stone, 1995]. En (b) : image miniature de documents [Cockburn et al., 1999]. En (c) le système Vquery. (Images fournies gracieusement par K. Fishkin, A. Cockburn et S. Jones respectivement).

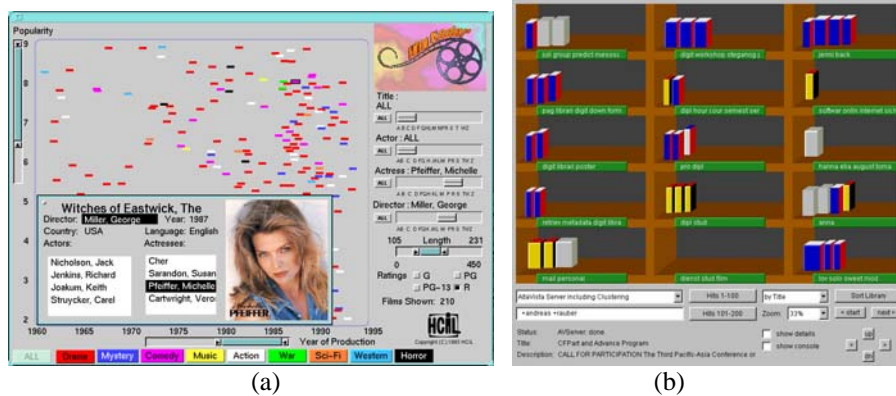


Fig.2 En (a) : le système FilmFinder (image fournie gracieusement par le « University of Maryland Human-Computer Interaction Lab »). En (b) : LibViewer du système SOMLib [Raubert et Bina, 1999] (image fournie gracieusement par A. Raubert)

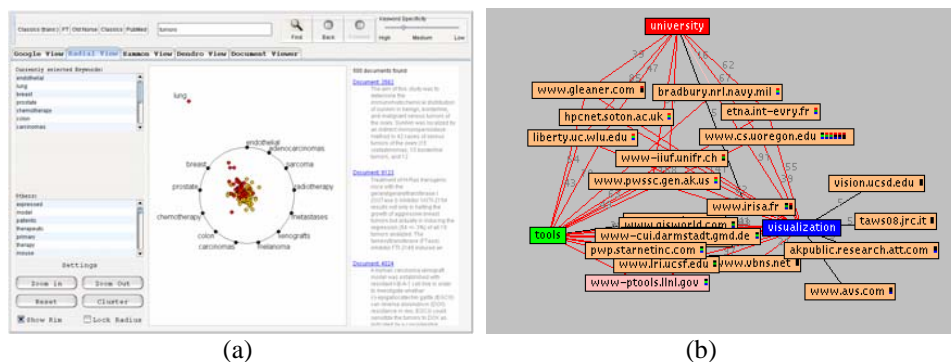


Fig. 3 Visualisation interactive radiale [Au et al., 2000] en (a) et le système SQWID [McCrickard et Kehoe, 1997] en (b) (images fournies gracieusement par S. Rueger et S. McCrickard respectivement)

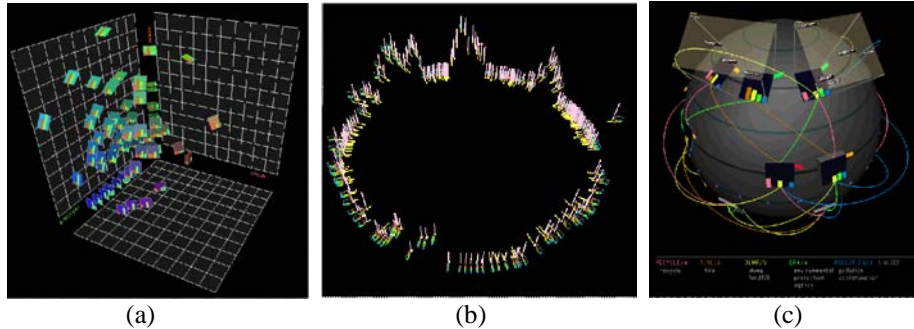


Fig. 4 D'après [Cugini et al., 2000], visualisation de documents selon 3 mots-clés en (a), modèle du plus proche voisin aligné autour du cercle en (b), et modèle du globe 3D en (c) (images gracieusement fournies par le NIST)

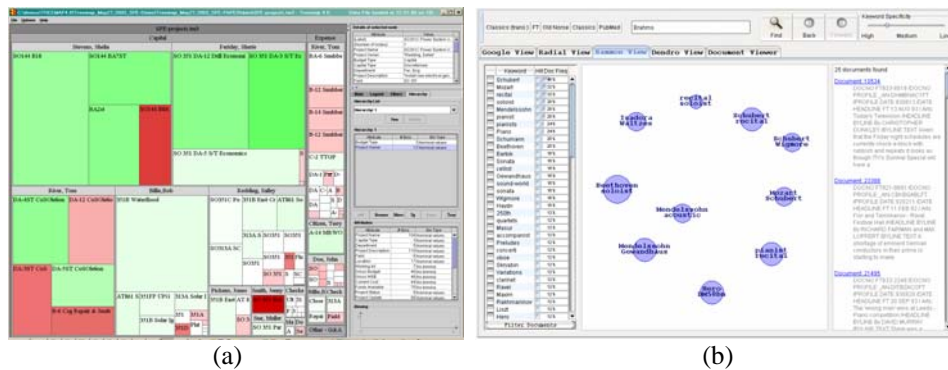


Fig. 5 En (a) : exemple d'une visualisation hiérarchique d'un ensemble de documents avec Treemap [Shneiderman et al 2000] (image fournie gracieusement par le « University of Maryland Human-Computer Interaction Lab »). En (b) : illustration des grappes de Sammon (image fournie gracieusement par S. Rueger)

4.3. Utilisation des points d'intérêts

4.3.1 Principes

Cette méthode est désigné par les termes « points d'intérêts » ou « points de références » (en anglais POIs pour « Points Of Interest »). Il s'agit dans cette technique de positionner sur une surface quelques icônes relatifs aux attributs d'un document et ensuite d'afficher les icônes des documents à des positions déterminées par l'adéquation entre les attributs et les documents. Par exemple, les attributs sélectionnés sont généralement des mots-clés utilisés dans la requête et les documents se positionnent par rapport à ces mots-clés. On utilise pour cela des techniques à base de ressorts et de forces, une méthode générale de visualisation de données (voir par exemple [Wise, 1999]). Les systèmes VIBE [Korfhage, 1991], SQWID [McCrickard et Kehoe, 1997] ou encore la visualisation radiale du système Information Navigator [Au et al., 2000] utilisent ces principes. Parfois il est difficile de voir exactement

vers quel point d'intérêt est attiré un document. Dans ce cas ces systèmes permettent alors de supprimer et d'ajouter des points d'intérêt.

4.3.2 Visualisation Interactive Radiale

Radial [Au et al., 2000] (figure 3(b)) est semblable à VIBE [Korfhage, 1991], à Radviz [Hoffman et al., 2000] et à Lyberworld [Hemmje et al., 1994] (ces derniers auteurs utilisant une sphère). 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 en bas à gauche de l'écran.

Un nuage de points est affiché à l'intérieur du cercle. Un point représente un document. Il n'est affiché que les documents 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 d'information affiche des informations sur ce document. Il est possible de déplacer les termes à l'extérieur du cercle et ainsi de déplacer tous les nœuds de document en rapport avec ces termes. Ceci permet de faire un classement manuel des résultats en catégories.

4.3.3 Système SQWID

L'outil visuel SQWID [McCrickard et Kehoe, 1997] émet une requête à un moteur de recherche (figure 3(b)). Après extraction du résultat, l'outil identifie une série de termes relatifs à ce résultat. Il crée ensuite un graphe composé de nœuds et de liens. Deux types de nœuds existent : les nœuds de termes et les nœuds de pages ou de sites (figure 3(b)).

Les trois premiers termes de la liste sont affichés. Ils sont colorés chacun par une des trois couleurs principales rouge, vert et bleu, ils sont fixés en triangles espacés de sorte à contenir les nœuds de liens qui leur correspondent. Les nœuds de liens sont colorés en orange. La distance entre les nœuds de termes et les nœuds de pages est fonction du classement de celles-ci par rapport à ce terme. Les pages relatives à plusieurs termes sont affichées à une position moyenne, au milieu du triangle. Les nœuds de pages positionnés le long d'une extrémité du triangle dépendent de deux termes (les deux extrémités). Les nœuds de pages positionnés autour d'un sommet dépendent d'un seul terme. Les pages qui ne dépendent d'aucun terme flottent à l'extrémité haute de l'écran loin de tous les termes.

Chaque nœud de page contient une représentation composée de trois barres tricolores relatives aux trois couleurs des trois termes. Leurs intensités montrent la fréquence de chaque terme dans cette page. L'utilisateur peut changer les termes dans une liste proposée dans un menu. Il peut ajuster la vue en sélectionnant le nombre de liens visibles et aussi la date relative aux liens. L'utilisateur a la possibilité de voir les liens entre les sites et de visiter les pages et sites. Ce système est cependant limité à trois mots-clés.

4.4. Cartes à regroupements implicites

Ce sont des cartes sur lesquelles sont affichées des icônes de documents selon une certaine disposition ou une topologie et qui fait que l'on remarque en analysant ces

représentations des regroupements implicites. Ces regroupements sont bien souvent établis à partir d'une classification existante des documents, ou bien plus simplement à partir d'une mesure de similarité entre les documents. Le principal avantage de ces techniques est d'éliminer les excès d'occlusion et d'employer des mesures métriques entre les documents.

Le principe des interfaces visuelles développées dans [Cugini et al. 2000] [Cugini et al. 1997] est de 1) soumettre une requête texte au moteur de recherche PRISE [PRISE], 2) présenter le résultat graphiquement et 3) permettre à l'utilisateur d'interagir jusqu'à trouver les résultats qu'il cherche. Implémentés en plusieurs versions 1D, 2D et 3D, les modèles développés permettent la visualisation de regroupements implicites. Par exemple, le modèle du plus proche voisin aligné autour du cercle : les documents sont arrangés autour d'un cercle. Les voisins d'un document sur le cercle sont ses plus proches voisins. La hauteur du document symbolise la pertinence de celui-ci (figure 4(b)). Un autre exemple en 3D est celui du globe de concept 3D (figure 4(c)). A partir d'un ensemble de concepts (chaque concept est un ensemble de mots-clés), les résultats sont groupés en clusters et présentés sous forme de cubes, rectangles et barres d'histogrammes rangés autour d'une sphère. La latitude d'un cluster dépend du nombre de concepts qu'il contient. Par convention, il y a une position unique au pôle nord pour les clusters qui contiennent tous les concepts et au pôle sud de la sphère pour ceux qui n'en ont aucun. Les clusters sont disposés autour de la sphère de manière à rapprocher ceux qui sont similaires. Les clusters qui diffèrent d'un seul concept sont reliés par un arc de la couleur de ce concept. Ainsi, si un cluster A a les concepts « bateau », « bassin » et « océan », et le cluster B a « bateau », « bassin », « océan », et « tempête », alors ils seront reliés par un arc qui porte la couleur du concept « tempête ». Une matrice interactive sert à former les concepts. Notons également que les documents peuvent être annotés par un label (bon, mauvais, non décidé).

La vue 3D étant une forme sphérique, on ne montre qu'une partie des résultats. Ceci demande à l'utilisateur un effort supplémentaire de mémorisation s'il veut comparer l'ensemble des résultats. On peut également ajouter que les techniques 2D donnent plus d'aisance aux utilisateurs et sont plus simples à appréhender que les représentations en 3D. L'évaluation des trois versions (1D, 2D et 3D) montre qu'au-delà de quatre ou cinq concepts, l'interprétation des résultats devient complexe. Cependant, cette remarque peut être relativisée en considérant, comme les auteurs l'affirment, que n'importe quel sujet peut être caractérisé par au plus cinq à six concepts [Cugini et al., 2000].

Une technique combinant vue d'ensemble, regroupement en clusters et alignement en spirale est utilisée par [Hascoët, 1998] et porte le nom de « Spring display ». En premier l'ensemble des documents est divisé en clusters et ensuite pour chaque cluster est identifié un document prototype (le plus similaire à tous par exemple). Enfin les documents restant sont alignés en spirale autour de chaque prototype. L'inconvénient de l'alignement en forme de spirale est que les utilisateurs ont tendance à juger les icônes appartenant à des rayons différents et rapprochés, comme étant aussi similaires que leurs voisins.

4.5. Structures hiérarchiques

Les techniques de visualisation de structures hiérarchiques de documents sont utilisées par de nombreux systèmes. Citons par exemple Treemap, la visualisation hyperbolique, Dendro-Map, ou encore la navigation à travers des listes et des matrices de [Chimera et Shneiderman, 1994] [Kunz et Botsch, 2002]. Nous allons supposer ici que les documents sont organisés dans un arbre. Les nœuds internes de l'arbre peuvent contenir ou non des

documents. Ces structures hiérarchiques sont particulièrement adaptées pour représenter une arborescence de fichiers, le plan d'un site Web ou encore le résultat d'une classification hiérarchique de documents. Même si les moteurs de recherche actuels pour le grand public ne présentent pas systématiquement les documents sous la forme d'un arbre, on peut observer que de nombreux systèmes de gestion de connaissances utilisent des représentations arborescentes. C'est ainsi le cas des sites portail qui sont une alternative aux moteurs de recherche.

Tree-Map est basé sur une idée de [Shneiderman et al., 2000]. Les résultats sont classés en clusters qui sont représentés par des rectangles. La dimension du rectangle est proportionnelle à la taille du cluster. Les clusters semblables sont regroupés en super-clusters. L'algorithme utilisé pour arranger et dimensionner les clusters est décrit dans [Shneiderman et al., 2000] et consiste à découper récursivement un rectangle initial en autant de sous-rectangles qu'il existe de branches à ce niveau là de l'arbre. Des possibilités d'interaction existent dans ce système (figure 5(a)) : en cliquant avec le bouton droit de la souris sur un rectangle, il s'affiche à droite de l'écran une liste complète des termes classés par ordre décroissant. Il est donc possible de sélectionner ou désélectionner des mots-clés. À côté de chaque terme est affiché une barre pour indiquer en pourcentage sa fréquence dans ce cluster. Le bouton « Filter » filtre les résultats selon les termes sélectionnés. Le bouton « Show all documents » affiche en bas de l'écran la liste complète des résultats (liens et résumé).

Le recours à la géométrie hyperbolique [Lamping et al., 1995] pour représenter des arbres peut être vue comme un cas particulier de l'approche hiérarchique. Selon cette approche les nœuds de l'arbre sont placés sur des cercles concentriques dont le rayon est calculé proportionnellement à la profondeur du nœud dans l'arbre. Le périmètre des cercles hyperboliques croît exponentiellement avec leur rayon ce qui permet de visualiser facilement des arbres contenant un grand nombre de feuilles. Les transformations de la géométrie hyperbolique permettent d'interagir avec la vue et en particulier de déplacer des nœuds d'intérêt au centre. Cependant la perception du sens de l'animation peut se révéler difficile à intégrer par un utilisateur non averti.

La visualisation de Dendro-Map [Carey et al., 2003] imite le principe des transformations non-Euclidiennes des cartes hyperboliques du fait qu'elle représente le résultat d'un moteur de recherche sous forme de nœuds et de feuilles d'un arbre binaire. En raison des contraintes spatiales, la profondeur de visualisation est limitée à cinq niveaux de la hiérarchie, avec à l'extrémité des branches le plus bas niveau représentant des documents ou des sous-clusters.

D'autres systèmes existent. Notons ainsi par exemple que dans [Kunz et Botsch, 2002] les auteurs proposent une interface graphique qui reçoit le résultat d'un moteur de recherche et l'affiche sous forme hiérarchique. L'utilisateur a la possibilité de réduire ou d'augmenter interactivement le nombre de documents résultant et d'affiner la requête jusqu'à obtenir les documents recherchés. L'interface présente deux possibilités : navigation à base de listes ou à base de matrices.

Les structures hiérarchiques limitent souvent le nombre de nœuds affichés simultanément, au détriment des détails de la représentation. Elles permettent en revanche de visualiser l'ensemble des catégories à laquelle appartient un document avec leur contexte hiérarchique.

4.6. Réduction par compression

La réduction d'un nombre important de dimensions en quelques dimensions par compression se fait en prenant en considération tout ou seulement une partie des dimensions. On peut citer la compression par recherche itérative du gradient proposé par [Sammon, 1969]. Ces travaux permettent de projeter une structure multidimensionnelle composée de plusieurs vecteurs sur un espace 2D tout en préservant les distances entre les objets deux à deux (figure 5(b)). Des clusters sont ainsi arrangés en 2D tels que leurs distances mutuelles soient indicatrices de leurs rapports dans l'espace n-dimensionnel.

Les grappes (clusters) sont représentées sous forme de cercles étiquetés par le mot-clé le plus fréquent. La taille des cercles est proportionnelle à la taille des clusters. La distance entre les cercles est en rapport avec la similitude de leurs contenus. Les plus proches sont les plus similaires.

Dans le système Galaxies développé pour le projet MVAB en application sur le système SPIRE [Wise et al., 1995], [Wise, 1999], les documents sont représentés par des « étoiles de documents » (« docustars ») et les clusters par des centroïdes⁴ colorés en orange, le tout sur un fond noir, ce qui donne comme une vue du ciel pendant la nuit.

L'application ThemeScapes qui a été renommée par la suite ThemeView du système SPIRE est basée sur la projection sur un espace 2D de la même manière que celle de Galaxies. Ensuite une surface 3D est construite en faisant élever les endroits positionnés par les documents selon leurs pertinences. Ainsi il est possible de distinguer les plus importants d'entre eux.

Une étude comparative d'un échantillon d'utilisateurs non expérimentés en systèmes visuels basés sur des cartes et du système plus familier Yahoo a montré qu'il était difficile pour eux de comparer les documents représentés dans des cartes sans lire réellement du texte [Kleiboemer 96]. Or la complexité dans toutes les représentations que l'on vient de voir, réside dans l'optimisation de l'espace d'affichage en ayant comme contrainte le compromis « nombre de documents pertinents à afficher simultanément »/ « informations précises pour chacun ». Nous en déduisons que ses systèmes peuvent être plus profitables en permettant d'afficher simultanément l'intitulé de la requête et un résumé court lors du survol, à l'aide d'un outil de pointage, d'une icône ou d'une surface contenant un document.

4.7. Coordination entre plusieurs vues

Des systèmes donnent la possibilité d'avoir plusieurs vues simultanées du même résultat à l'aide de différentes techniques. Ils emploient une combinaison de deux ou trois composants que nous venons d'évoquer précédemment. Citons par exemple le système « Information Navigator », ou encore les travaux de [Mann, 1999]. De nombreux autres articles décrivent des coordinations entre vues (voir par exemple la taxonomie présentée dans [North et Shneiderman, 1997] ou encore [Baldonado et al., 2000]).

[Carey et al., 2000] proposent une implémentation d'un moteur de recherche avec trois interfaces visuelles : les visualisations avec les grappes de Sammon, avec Tree-Map, et avec la visualisation interactive radiale. L'idée est que suite à l'introduction par l'utilisateur d'une requête le système extrait statistiquement en fonction des mots-clés et des résultats une série de termes clés. Ils seront exploités par le système pour faire des regroupements de documents

⁴ Considéré comme l'origine et la destination des déplacements émis et attirés par cette zone.

<i>Composants</i>	<i>Références</i>	<i>Système</i>
Sélection de dimensions		
Vecteur de documents	[Mann, 1999]	
« Scatter-plot »	[Ahlberg et Shneiderman, 1994], [Jog et Shneiderman, 1995]	<i>FilmFinder</i>
	[Spotfire, 2001]	<i>Spotfire Pro</i>
	[Ahlberg et Wistrand, 1995], [Wistrand, 1995]	<i>IVEE</i>
	[Fishkin et Stone, 1995]	<i>Magic Lens Filters</i>
Matrice d'icônes	[Baeza-Yates et al., 2003]	<i>AmazonDNA</i>
	[Nowell et al., 1996]	<i>Envision</i>
Géographique	[Higgins et al., 1999]	<i>VisageWeb</i>
« Perspective Wall »	[Mackinlay et al., 1991]	<i>Information Visualizer</i>
	[Spence et Apperley, 1982]	<i>Bifocal</i>
Axes 3D	[Cugini et al., 1997], [Cugini, et al., 2000].	<i>NIRVE</i>
Histogrammes Interactifs	[Tweedie et al., 1994]	<i>Explorateur d'attributs</i>
	[Tweedie et al., 1996]	<i>Influence Explorer</i>
Pixels	[Keim et Kriegel, 1994].	<i>VisDB</i>
Circle Segments	[Ankerst et al., 1996]	
Cartes auto-organisatrices		
	http://ai.bpa.arizona.edu/resume/dmitri/query.html	<i>Adaptive SOM</i>
	[Chen et al., 1996]	<i>ET-map</i>
	[Rauber et Bina, 2000]	<i>SOMLib</i>
	[Kohonen, 1998]	<i>WEBSOM</i>
Utilisation des points d'intérêts		
Espace de documents 2D avec points de références	[Korfhage, 1991]	<i>VIBE</i>
	[McCrickard et Kehoe, 1997]	<i>SQWID</i>
	[Au et al., 2000], [Carey et al., 2000]	<i>Radial (Information Navigator)</i>
	[Hoffman et al. 2000]	<i>Radviz</i>
<i>Sphère de pertinence</i>	[Hemmje et al., 1994]	<i>Lyberworld</i>

Tab. 3 - Techniques de visualisation d'un ensemble de documents

<i>Composants</i>	<i>Références</i>	<i>Système</i>
Cartes à regroupements implicites		
Document Spiral	[Cugini et al., 1997], [Cugini et al., 2000]	<i>NIRVE</i>
Nearest Neighbor Circle		
Spoke and Wheel / Document Space		
Spring display	[Hascoët 1998]	
Structures hiérarchiques		
Treemap	[Shneiderman et al., 2000]	
	[Au et al., 2000], [Carey, 2000], [Carey et al., 2000]	<i>Information Navigator</i>
Dendro-Map	[Carey et al., 2003]	<i>Info Navigator</i>
Listes et matrices	[Kunz et Botsch, 2002]	
Réduction par compression		
Scatterplot	[Wise et al., 1995], [Wise, 1999]	<i>Galaxies MVAB / SPIRE</i>
Paysages (Landscapes)	[Wise et al., 1995]	<i>ThemeScape/SPIRE</i>
Sammon view	[Au et al., 2000], [Carey, 2000], [Carey et al., 2000]	<i>Information Navigator</i>
Coordination entre plusieurs vues		
	[Mann, 1999]	
	[Au et al., 2000], [Carey et al., 2000], [Carey et al., 2003]	<i>Information Navigator</i>

Tab. 4 - Techniques de visualisation d'un ensemble de documents (suite)

et générer des vues dans le but de permettre à l'utilisateur d'affiner sa recherche dynamiquement. Le système offre à l'utilisateur la possibilité de basculer facilement entre les trois vues pour analyser les résultats.

[Mann, 1999] permet d'explorer les résultats d'un moteur de recherche en proposant à l'utilisateur de basculer entre les différentes visualisations proposées (dans le but notamment d'affiner la requête visuellement). Un document sélectionné dans une vue le restera aussi pour les autres. Les différentes visualisations proposées sont le « vecteur de documents » (une visualisation à une dimension), un « scatter plot » (visualisation 2D), et un affichage d'histogrammes.

La technique de coordination entre plusieurs vues est difficile à mettre en œuvre mais offre à l'utilisateur un grand avantage en comblant les lacunes d'une visualisation par une autre sur une même interface. [Hoffman et al., 2000] ont introduit une primitive qui permet de représenter mathématiquement plusieurs vues rendant ainsi cette technique plus simple à implémenter.

Nous avons regroupé dans les tableaux 3 et 4 les techniques de visualisation d'un ensemble de documents.

4.8. Métaphores

Il existe des tentatives de représentation de l'information dans des espaces inspirés de la vie réelle de l'être humain, appelées métaphores, et certaines utilisant particulièrement des représentations 3D en réalité virtuelle. Cependant toutes ne sont pas appliquées à la recherche d'information sur Internet. La motivation majeure de la conception de ce type de systèmes est de sauter la phase d'appréhension de l'interface à laquelle doit faire face tout utilisateur. Le but d'une métaphore en général est de permettre une perception rapide d'une visualisation en reflétant une vision du monde réel (donc familière). Réciproquement, une métaphore non familière n'en serait pas une à proprement dit [Bederson et Hollan 1994]. Nous citons parmi ces tentatives : la chambre [Henderson et Card 1986], le guide [Hammond et Allison 1987], la ville [Dieberger et Frank 1998], le système « The Information Visualizer » [Robertson et al. 1993], l'aquarium [Bryan et Gershman 2000]. Une métaphore n'est jamais parfaite pour une représentation et a souvent besoin d'être modifiée, ou d'être complétée par d'autres visualisations. Afin d'éviter de troubler la vision de l'utilisateur [Catarci et al. 1995], il faut que le concepteur choisisse les métaphores qui sont fortement liées au monde réel. La difficulté réside notamment dans la partie « mapping » des données à représenter vers la structure de donnée qui soit s'adapter à une visualisation quasiment réaliste d'une métaphore. Nous pensons que ces techniques ont encore du chemin à parcourir pour faire leurs preuves. Elles doivent être étudiées plus profondément du côté psychologique de l'utilisateur pour éviter les sources de confusion.

5. Quelques exemples de logiciels opérationnels et de techniques d'évaluations

En préalable, mentionnons que pour certains des systèmes précédents tels que SOMLib ou MosaicG, le code source est libre. Nous donnons ici à titre indicatif quelques exemples de systèmes commerciaux (ou opérationnels) qui n'ont pas été cités précédemment. Citons par exemple Matrisme qui est un logiciel de cartographie d'ensembles d'information (<http://crrm.u-3mrs.fr/>). Internet Cartographer d'Inventix (<http://www.inventix.com>) est un outil qui aide à rechercher les documents, pages HTML, sites Web aspirés sur votre disque dur et à visualiser leurs relations sous une forme graphique. MapNet est un moteur de recherche en ligne (<http://maps.map.net>) qui offre une navigation visuelle intuitive et pertinente du cyberspace (américain essentiellement). L'un des premiers moteurs à représentation graphique qui permet de faire des associations de concepts est le moteur Webbrain®, de la société TheBrain (<http://www.webbrain.com>). La société WebMap Technologies Inc a développé le plug-in gratuit WebMap Viewer (<http://www.webmap.com>) qui permet à MSIE v5.00 de visualiser une carte de l'internet donnant accès à près de deux millions de sites web d'une manière intuitive. Kartoo (<http://www.kartoo.com>) et Mapstan (<http://Search.mapstan.net>) sont les métamoteurs français à affichage graphique. Ils présentent les sites Web trouvés sous forme de cartes dans lesquelles les sites sont représentés par des icônes de tailles différentes en fonction de leur pertinence par rapport à la requête. Les sites Web sont reliés les uns aux autres par des liens sémantiques générés dynamiquement. Egalement, nous mentionnons le site <http://www.touchgraph.com/> qui

permet de rajouter une surcouche de visualisation à des moteurs de recherche en créant un graphe dynamique expansible.

Comme pour tous les systèmes interactifs, seuls les utilisateurs peuvent juger réellement de l'intérêt d'une méthode donnée par rapport à une autre. Néanmoins quelques recherches se sont axées sur les techniques d'élaboration de ces tests. On cite par exemple l'évaluation faite par [Mann et Reiterer, 2000] pour le système Insyder où ils ont élaboré des tests en prenant en considération cinq facteurs (appelés 5T-Environment) qui influencent l'efficacité d'une technique de visualisation : (1) le groupe d'utilisateurs visés : scientifiques, expert ou « monsieur tout le monde », comme exemples on peut voir les travaux de [Egan, 1988], [Borgman, 1989], [Shneiderman, 1998], ou [Hearst, 1999] ; (2) le type et le nombre de données : examiner 50 documents représentés par des Tilebars peut être très satisfaisant pour trouver les plus appropriés, par contre avec 5000 documents l'utilisateur aimera probablement avoir une étape d'amélioration avec une autre forme de visualisation ; (3) la tâche à accomplir (exemple : classification/recherche d'informations) ; (4) les possibilités techniques : le choix de l'interface par exemple basée sur un browser Web et l'utilisation d'un ordinateur lent peuvent mener à de mauvaises valeurs d'exécution pour une visualisation ; (5) la formation : comme pour toute autre interface utilisateur il est important de trouver le bon équilibre entre la facilité d'apprentissage du principe de la visualisation et l'efficacité [Nielsen, 1998]. Le compromis entre la simplicité et la puissance [Hearst, 1999] n'est pas le seul facteur important ici, où l'expérience semble également être un facteur influençant le succès des visualisations. Par exemple si l'on compare la familiarité des utilisateurs avec les représentations visuelles et avec les représentations traditionnelles des textes, le résultat de l'évaluation peut être biaisée comme c'est le cas pour l'étude [Kleiboemer 96].

6. Conclusion

Nous avons présenté dans cet article un survol des techniques visuelles utilisées dans les systèmes de recherche d'information et notamment ceux liés à Internet. Nous avons commencé par présenter les méthodes visuelles plus spécifiquement dédiées à la formulation de requêtes. Ensuite nous avons décrit les méthodes permettant de visualiser un document pris isolément. Enfin, nous avons décrit les méthodes visualisant un ensemble de documents avec des exemples de systèmes complets. Une vue d'ensemble de ces éléments est présentée sur la figure 6.

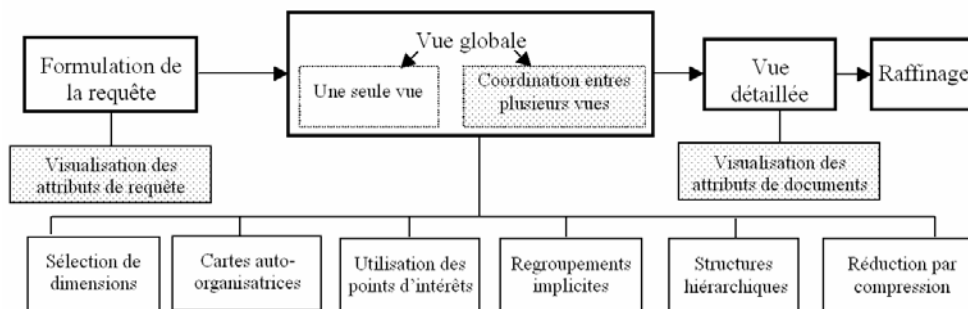


Fig. 6 - Classification des étapes et composants de la visualisation

Ces approches nous apparaissent comme étant très prometteuses pour l'avenir des systèmes de recherche d'information. De nombreuses dimensions peuvent être perçues simultanément ce qui représente une amélioration significative par rapport à une simple liste de résultats. Il est certain également que l'intérêt de cette visualisation augmente par exemple si l'on représente des regroupements entre les documents. Ces groupements permettent à l'œil humain de percevoir des prototypes de classes et d'avoir une vue d'ensemble précise des documents. Le regroupement des documents en classe, leur visualisation sont un des points importants pour le futur de ces techniques.

D'autres perspectives peuvent être envisagées en ce qui concerne le bouclage requête-résultat. De nombreux systèmes visuels ne permettent pas encore de reformuler interactivement la requête à partir de la présentation visuelle des résultats. De même, d'autres attributs pourraient être visualisés, comme des distances entre phrases appartenant à différents documents, en faisant intervenir des techniques de text-mining.

Enfin, une large place est laissée pour l'évaluation et la comparaison des méthodes. Bien souvent, les travaux présentés ne sont pas systématiquement comparés entre eux comme cela est fait par exemple dans [Cugini et al., 2000]. Il faut mentionner également que ces tests nécessitent un public plus mature dans le domaine des interfaces visuelles pour être plus juste. Néanmoins, l'accroissement des besoins en exploration/compréhension des données et l'utilisation pratique de ces systèmes laissent envisager un avenir très prometteur pour toutes ces méthodes.

Références

- [Ahlberg et Shneiderman, 1994] Ahlberg, C.; Shneiderman, B.: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p.313-317.
- [Ahlberg et Wistrand, 1995] Ahlberg, C.; Wistrand, E.: IVEE: An Information Visualization and Exploration Environment. In: Gershon, Nahum; Eick, Stephen G. (Eds.): Proceedings of IEEE Information Visualization 1995. Conference: Atlanta, GA, October 30-31 1995. Los Alamitos, CA (IEEE Computer Soc. Press) 1995. p. 66-73.
- [Anick et al., 1990] Anick, P.; Brennan, J.; Flynn, R.: A Direct Manipulation Interface for Boolean Information Retrieval via Natural Language Query. In: Vidick, Jean-Luc (Eds.): SIGIR 1990: Proceedings of the 13th International Conference on Research and Development in Information Retrieval. Conference: Brussels, Belgium, September 5-7 1990. New York (ACM Press) 1990. S. 135-150.
- [Ankerst et al., 1996] Ankerst, M.; Keim, D. A.; Kriegel, H.: Circle Segments: A Technique for Visually Exploring Large Multidimensional Data Sets. In: Yagel, Roni; Nielson, Gregory M. (Eds.): Proceedings of IEEE Visualization 96 Conference: San Francisco, CA, October 27 - November 1, 1996. Los Alamitos, CA (IEEE Computer Soc. Press) 1996. Hot Topic Session.
- [Arisem, 1999] Arisem S.A., <http://www.arisem.com>
- [Au et al., 2000] Au, P.; Carey, M.; Sewraz, S., Guo Y., Rueger S.: New Paradigms in Information Visualization. In: Belkin, Nicholas J.; Ingwersen, Peter; Leong, Mun-Kew (Eds.): SIGIR 2000: Proceedings of the 23rd Annual International ACM SIGIR

- Conference on Research and Development in Information Retrieval. Conference: Athens, Greece, July 24-28 2000. New York (ACM Press) 2000. S. 307-309.
- [Ayers et Stasko, 1995] Ayers, Eric Z.; Stasko, John T.: Using Graphic History in Browsing the World Wide Web. In: WWW 4: Fourth International World Wide Web Conference. Conference: Boston, MA, December 11-14 1995.
- [Baldonado et al., 2000] Baldonado, M.; Woodruff, A.; Kuchinsky, A.: Guidelines for Using Multiple Views in Information Visualization. In: Di Gesù, Vito; Levialdi, Stefano; Tarantino, Laura (Eds.): AVI 2000: International Workshop on Advanced Visual Interfaces. Conference: Palermo, Italy, May 23-26 2000. New York (ACM Press) 2000. p. 110-119.
- [Baeza-Yates et al., 2003] Baeza-Yates R., Lemus R., Ponceleon R., Srinivasan S., WISDNA: An Information Visualization Paradigm for XML, Proceedings of the First Latin American Web Congress (LA-WEB 2003) , 2003 IEEE, pp205-208.
- [Bederson et Hollan, 1994] Bederson, B.B., and Hollan, J.D., "Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics," Proc. ACM UIST '94, Marina del Rey, CA, Nov. 1994, pp. 17-26.
- [Berenci et al., 1998] Berenci, E.; Carpineto, C.; Giannini, V.: Improving the Effectiveness of WEB Search Engines Using Selectable Views of Retrieval Results. In: Journal of Universal Computer Science, 4 (1998) 9, p. 737-747.
- [Borgman 1989] Borgman, C. L.: All users of information systems are not created equal: An exploration into individual differences. In: Information Processing & Management, 25 (1989) 3, p. 237-252.
- [Bryan et Gershman 2000], Bryan, Dough; Gershman, Anatole: The Aquarium: A Novel User Interface Metaphor for Large, Online Stores. In: Tjoa, A Min; Wagner, Roland R.; Al-Zobaidie, Ala (Eds.): Proceedings 11th International Workshop on Database and Expert Systems Applications. Conference: Greenwich, London, United Kingdom, September 4-8 2000. CA (IEEE Computer Society) 2000. p. 601-607.
- [Bürdek et al., 1999] Bürdek, B., Eibl M.; Kraus J.: Visualisation in Document Retrieval: An Example of the Integration of Software Ergonomics and an Aesthetic Quality in Design. In: Paton, R.; Neilson, I. (Eds.): Proceedings of the VRI'98 - Visual Representations and Interpretations. Conference: Liverpool, United Kingdom, September 22-24 1998. London (Springer Verlag) 1999. p. 387-396.
- [Byrd, 1999] Byrd, Donald: A Scrollbar-based Visualization for Document Navigation. In: Fox, Edward A. et al. (Eds.): Proceedings of the fourth ACM conference on digital libraries. New York (ACM Press) 1999. p. 122-129.
- [Card et al., 1999] Card, S. K.; Mackinlay, J. D.; Shneiderman, Ben (Eds.): Readings in Information Visualization. Using Vision to Think. San Francisco, CA (Morgan Kaufmann) 1999.
- [Carey et al., 2003] Carey M., Heesch D. and Rüger S.. Info Navigator: A Visualization Tool for Document Searching and Browsing, Proc of Intl Conf on Distributed Multimedia Systems (DMS Sep 2003), 2003. Lien vers la version de démonstration : <http://km.doc.ic.ac.uk/pr-m.carey-2000/>
- [Carey et al., 2000] Carey M., Kriwaczek F., Rüger M., A Visualization Interface for Document Searching and Browsing In: Proceedings of NPIVM 2000 Conference: Washington, DC, November 10-11 2000. New York (ACM Press) 2000.
- [Catarci et al, 1995] T. Catarci, M.F. Costabile, M. Matera. "Which Metaphor for Which Database?". In Proc. HCI '95 –Tenth British Conference on Human-Computer Interaction

- , M.A.R. Kirby, A.J. Dix and J.E. Finlay(eds.), Huddersfield, UK, Agosto 1995, pp 151-165.
- [Chen et al., 1996] Chen, H.; Schuffels, C.; Orwig, R.: Internet categorization and search: a self-organizing approach. In: Journal of visual communication and image representation, 7 (1996) 1, p. 88-102. Lien vers la version de démonstration de ET-map : <http://ai3.bpa.arizona.edu/ent/>
- [Chi, 2000] Chi, Ed H.: A Taxonomy of Visualization Techniques Using the Data State Reference Model. In: Roth, Steven F.; Keim, Daniel A. (Eds.): Proceedings of IEEE Information Visualization 2000. Conference: Salt Lake City, UT, October 9-10 2000. Los Alamitos, CA (IEEE Computer Soc. Press) 2000. p. 69 -75.
- [Chimera et Shneiderman, 1994] Chimera R. et Shneiderman B.. An exploratory evaluation of three interfaces for browsing large hierarchical tables of contents. ACM Trans. Inf. Syst., 12(4):383-406, 1994.
- [Cockburn et al., 1999] Cockburn, A.; Greenberg, S.; McKenzie, B., Smith M. Kaasten S.: WebView: A Graphical Aid for Revisiting Web Pages. Proceedings of the OZCHI'99 Australian Conference on Human Computer Interaction.
- [Cugini, 2000] Cugini, J.: Presenting Search Results: Design, Visualization and Evaluation. In: Workshop: Information Doors - Where Information Search and Hypertext Link (held in conjunction with the ACM Hypertext and Digital Libraries conferences) Conference: San Antonio, TX, May 30 2000.
- [Cugini et al., 1997] Cugini J., Laskowski S., Piatko C., "Document Clustering in Concept Space: The NIST Information Retrieval Visualization Engine (NIRVE)" (<http://zing.ncsl.nist.gov/~cugini/uicd/concept-clusters.html>), CODATA Euro-American Workshop on Visualization of Information and Data, Paris, France, June 1997.
- [Cugini et al., 2000] Cugini J., Laskowski S., Sebrechts M., "Design of 3D Visualization of Search Results: Evolution and Evaluation", Proceedings of IST/SPIE's 12th Annual International Symposium: Electronic Imaging 2000: Visual Data Exploration and Analysis (SPIE 2000), San Jose, CA, 23-28 January 2000.
- [Cugini et al., 1998] Cugini, J.; Laskowski, S.; Piatko, C.: Document Clustering in Concept Space: The NIST Information Retrieval Visualization Engine (NIRVE), CODATA Euro-American Workshop Visualization of Information and Data.
- [Dieberger et Frank, 1998] Dieberger, Andreas; Frank, Andrew U.: A city metaphor for supporting navigation in complex information spaces. In: Journal of Visual Languages and Computing, 9 (1998) 6, p. 597-622.
- [Dieberger et Russell, 2001] Dieberger, A.; Russell, D.: Context Lenses - Document Visualization and Navigation Tools for Rapid Access to Detail. In: Hirose, Michitaka (Ed.): Human-Computer Interaction - INTERACT'01. Proceedings. Conference: Tokyo, Japan, 2001 July 9-13 Amsterdam (IOS Press) 2001. S. 545-552.
- [Egan, 1988] Egan, D. E.: Individual differences in human-computer interaction. In: Helander, Martin (Ed.): Handbook of Human-Computer Interaction Amsterdam (North-Holland) 1988. p. 543-568.
- [Eibl, 1999] Eibl, M.: Visualisierung der Recherche in Dokument-Retrievalsystemen: Methoden und Lösungen. In: Nachrichten für Dokumentation, 50 (1999) 2, p. 79-86.
- [Eibl, 1999a] Eibl, M.: Unterstützung der Recherche im Dokumentenretrieval durch den Einsatz von Visualisierung. In: Heuer, Andreas (Ed.): ADI 1999: Agenten, Datenbanken und Information Retrieval. Conference: Rostock-Warnemünde, Germany, September 30 - November 1 2000.

- [Eick, 1994] Eick, S. G.: Graphically displaying text. In: Journal of Computational and Graphical Statistics, 3 (1994) 2, p. 127-142.
- [Eick et al., 1992] Eick, S. G.; Steffen, J. L.; Sumner, Eric E.: SeeSoft: A tool for visualizing line oriented-software statistics In: IEEE Transactions on Software Engineering, 18 (1992) 11, p. 957-968.
- [Fishkin et Stone, 1995] Fishkin, K.; Stone, M. C.: Enhanced Dynamic Queries via Movable Filters. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 23-29.
- [Golovchinsky, 1997] Golovchinsky, G.: Queries? Links? Is there a Difference? In: CHI 1997: Conference Proceedings Human Factors in Computing Systems. Conference: Atlanta, GA, March 22-27 1997. New York (ACM Press) 1997. p. 407-413.
- [Grewal et al., 1999] Grewal, R. S.; Burden, P.; Jackson, M., Wallis, J.: A novel interface for representing search-engine results. In: IEE Informatics Colloquium: Lost in the Web - Navigation on the Internet. Conference: London, United Kingdom, November 2 1999. p. 7/1-7/10.
- [Grewal et al., 2000] Grewal, R. S.; Jackson, M.; Burden, P., Wallis J., A Visual Technique to indicate Overall and Component Relevancy of Search-Engine Results. In: Tjoa, A Min; Wagner, Roland R.; Al-Zobaidie, Ala (Eds.): Proceedings 11th International Workshop on Database and Expert Systems Applications. Conference: Greenwich, London, United Kingdom, September 4-8 2000. Los Alamitos, CA (IEEE Computer Society) 2000. p. 608-612.
- [Hammond et Allison 1987] Hammond, N. V. and L. J. Allison (1987). The Travel Metaphor as Design Principle and Training Aid for Navigating Around Complex Systems. Proceedings of the Third Conference of the British Computer Society, University of Exeter, Cambridge University Press.
- [Hascoët, 1998] Hascoët, M.: Analytical versus empirical evaluation of spatial displays. In: Karat, Clare-Marie; Karat, John; Horrocks, Ian (Eds.): CHI 1998: Conference Proceedings Human Factors in Computing Systems. Conference: Los Angeles, CA, April 21-23 1998. New York (Addison-Wesley) 1998. p. 257-259.
- [Hearst, 1995] Hearst, Marti A.: TileBars: Visualization of Term Distribution Information in Full Text Information Access. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 59-66.
- [Hearst, 1999] Hearst, M.: User Interfaces and Visualization. In: Baeza-Yates, Ricardo A.; Ribeiro-Neto, Berthier (Eds.): Modern Information Retrieval. Reading, MA (Addison Wesley Longman) 1999. p. 257-324.
- [Hemmje et al., 1994] Hemmje, M.; Kunkel, C.; Willett, A.: LyberWorld - A Visualization User Interface Supporting Fulltext Retrieval. In: Croft, W. Bruce; van Rijsbergen, Cornelis J. (Eds.): SIGIR 1994: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval Conference: Dublin, Ireland, July 3-6 1994. New York (ACM Press) 1994. p. 249-259.
- [Henderson et Card, 1986] Henderson, D.; Card, S. Rooms: The Use of Multiple Virtual Workspaces to Reduce Space Contention in a Window-Based Graphical User Interface. ACM Transactions on Graphics, 5(3), 211- 243. July 1986.

- [Heo et al., 1996] Heo, M.; Morse, E. L.; Willms, S., Spring M.: Multi-level Navigation of a Document Space. In: Proceedings of WebNet 1996 - World Conference on the WWW, Internet & Intranet Conference. San Francisco, CA, October 15-19 1996. CD-ROM
- [Hertzum et Frokjaer, 1996] Hertzum, M.; Frokjaer, E.: Browsing and querying in online documentation: A study of user interfaces and the interaction process. In: ACM Transactions on Computer Human Interaction, 3 (1996) 2, p. 136-161.
- [Higgins et al., 1999] Higgins, M.; Lucas, P.; Senn, J.: VisageWeb: Visualizing WWW Data in Visage. In: Will, Graham; Keim, Daniel A. (Eds.): Proceedings of IEEE Information Visualization 1999. Conference: San Francisco, CA, October 24-29 1999. CA (IEEE Computer Soc. Press) 1999. p. 100-107.
- [Hoffman et al., 2000] Hoffman P., Grinstein G., Pinkney D.. Dimensional anchors: A graphic primitive for multidimensional multivariate information visualizations. In NPIVM 99, 2000.
- [Inselberg , 1985] A. Inselberg. The plane with parallel coordinates. The Visual Computer, 1:69--91, 1985.
- [INXIGHT] www.inxight.com
- [Jansen et al., 1998] Jansen, B.. J.; Spink, A.; Bateman, J., Saracevic T.: Real Life Information Retrieval: a Study of User Queries on the Web. In: ACM SIGIR Forum, 32 (1998) 1, p.5-17.
- [Jerding et Stasko, 1995] Jerding, D. F.; Stasko, J. T.: The Information Mural: A technique for displaying and navigating large information spaces. In: Gershon, Nahum; Eick, Stephen G. Eds.): Proceedings of IEEE Information Visualization 1995. Conference: Atlanta, GA, October 30-31 1995. Los Alamitos, CA (IEEE Computer Soc. Press) 1995. p. 43-50.
- [Jerding et Stasko, 1997] Jerding, D. F.; Stasko, J. T.: The Information Mural: A technique for displaying and navigating large information spaces. Graphics, Visualization, and Usability Center, Georgia Institute of Technology, Atlanta, GA, Technical Report GIT-GVU-97-24, December 1997.
- [Jog et Shneiderman, 1995] Jog, N., Shneiderman, B., Starfield Information Visualization with Interactive Smooth Zooming. Aus: Spaccapietra, Stefano; Jain, Ramesh (Eds.): Visual Database Systems 3, Visual Information Management, Proceedings of the third IFIP 2.6 working conference on visual database systems. Conference: Lausanne, Switzerland, March 27-29 1995. London (Chapman & Hall) 1995. p. 3-14.
- [Jones, 1998] Jones, S.: Graphical Query Specification and Dynamic Results Preview for a Digital Library. In: UIST 98: 11th ACM Symposium on User Interface Software and Technology New York (ACM Press) 1998. p. 143-151.
- [Jones, 1998a] Jones, S.: Dynamic Query Result Previews for a Digital Library. In: Witten, Ian H.; Akscyn, Robert M.; Shipman, Frank M. (Eds.): ACM Digital Libraries 1998: Proceedings of the 3rd ACM International Conference on Digital Libraries. Conference: Pittsburgh, PA, June 23-26 1998. New York (ACM Press) 1998. p. 291-292.
- [Jones et al., 1999] Jones, S.; McInnes, S.; Staveley, M.: A Graphical User Interface for Boolean Query Specification. In: International Journal on Digital Libraries, 2 (1999) 2/3, p. 207-223.
- [Kaasten et Greenberg, 2000] Kaasten, S.; Greenberg, Saul: Designing an Integrated Bookmark / History System for Web Browsing. In: Proceedings of the Western Computer Graphics Symposium 2000. Conference: Panorama Mountain Village, BC,

- Canada, March 26-29 2000. <http://www.cpsc.ucalgary.ca/grouplab/papers/2000/00-GrouplabPapers.Skigraph/kaasten/kaasten.pdf>
- [Kaasten et Greenberg, 2001] Kaasten, S., Greenberg, S.: Integrating Back, History and Bookmarks in Web Browsers. <http://www.cpsc.ucalgary.ca/grouplab/papers/2000/00-IntegratedHistory.Report/00-IntegratedHistory-RR.pdf>
- [Kaugars, 1998] Kaugars, K.: Integrated multi scale text retrieval visualization. In: Karat, Clare-Marie; Karat, John; Horrocks, Ian (Eds.): CHI 1998: Conference Proceedings Human Factors in Computing Systems. Conference: Los Angeles, CA, April 21-23 1998. New York (Addison-Wesley) 1998. p. 307-308.
- [Keim et Kriegel, 1994] Keim, D., Kriegel, H.: VisDB: Database exploration using multidimensional visualization In: IEEE Computer Graphics and Applications, 14 (1994) 5, p. 40-49.
- [Kleiboemer 96] Kleiboemer A.J., Lazear M.B., Pedersen J.O. Tailoring a Retrieval System for Naive Users. Proc. 5th Symposium for Document Analysis and Information Retrieval (SDAIR96), Las Vegas, April 1996, pp. 209-216.
- [Kohonen, 1981] Kohonen, T., Construction of similarity diagrams for phonemes by a self-organizing algorithm. Report TKK-F-A463, Helsinki University of Technology, Espoo, Finland.
- [Kohonen, 1998] Kohonen, T: Self-organization of very large document collections: State of the art. In: Niklasson, Lars; Bodén, Mikael; Ziemke, Tom (Eds.): Proceedings of ICANN98, the 8th International Conference on Artificial Neural Networks. Conference: Skövde, Sweden, 1998 September 2-4 Springer (London) 1998. p. 65-74, vol. 1. lien de la démonstration <http://websom.hut.fi/websom/stt/doc/eng/>
- [Korfhage, 1991] Korfhage, R.: To See or Not to See - Is That the Query? In: Bookstein, Abraham; Chiaramella, Yves; Salton, Gerard et al. (Eds.): SIGIR 1991: Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Vol 1. Conference: Chicago, IL, October 13-16 1991. New York (ACM Press) 1991. p. 134-141.
- [Kunz et Botsch, 2002] Kunz, C., Botsch, V. (2002): Visual Representation and Contextualization of Search Results: List and Matrix Browser, Proceedings of the DC-2002, Florence, Italy
- [Lamping et al., 1995] Lamping, J., Rao, R., Pirolli, P., A focus + context technique based on hyperbolic geometry for visualizing large hierarchies. Proc. ACM Human Factors in Computing Systems (CHI'95), pp. 401-408.
- [Mackinlay et al., 1991] Mackinlay, Jock D.; Robertson, George G.; Card, Stuart K.: The Perspective Wall: Detail and Context Smoothly Integrated. In: Robertson, S. P.; Olson, G. M.; Olson, J.S. (Eds.): CHI 1991: Conference Proceedings Human Factors in Computing Systems. Conference: New Orleans, LA, April 27 - May 2 1991. New York (ACM Press) 1991.p. 173-179.
- [Mann, 1999] Mann, T.: Visualization of WWW-Search Results. In: Cammelli, Antonio; Tjoa, A Min; Wagner, Roland R. (Eds.): Proceedings Tenth International Workshop on Database and Expert Systems Applications. Conference: Florence, Italy, September 1-3 1999. Los Alamitos, CA (IEEE Computer Society) 1999. p. 264-268.
- [Mann et Reiterer 2000] Mann, T.; Reiterer, H.: Evaluation of Different Visualization of WWW Search Results. In : Tjoa, A.; Wagner, R.; Al-Zobaidie, A. (Eds.): Proceedings 11th International Workshop on Database and Expert Systems Applications. 2000, pp 586-590.

- [McCrickard et Kehoe, 1997] McCrickard, D. Scott; Kehoe, Colleen M.: Visualizing Search Results using SQWID. In: WWW 6: Sixth International World Wide Web Conference. Conference: Santa Clara, CA, April 7 - 11 1997. Lien vers la version de démonstration de SQWID: <http://www.cc.gatech.edu/grads/m/Scott.McCrickard/sqwid/>
- [Myers, 2000] Myers, D.G., Psychology (6th edition), Worth Publishing.
- [Nielsen, 1998] Nielsen, J.: 2D is Better Than 3D. ... because people are not frogs. <http://www.useit.com/alertbox/981115.html> [1998-10-10]
- [North, 1997] <http://www.cs.umd.edu/~north/infoviz.html>
- [North, Shneiderman 1997] North, C., Shneiderman, B., A Taxonomy Of Multiple Window Coordinations", University of Maryland, College Park, Dept of Computer Science Technical Report #CS-TR-3854, (1997).
- [North, 2000] North, C.: A User Interface for Coordinating Visualizations based on Relational Schemata: Snap-Together Visualization. University of Maryland Computer Science Dept. Doctoral Dissertation, May 2000.
- [Nowell et al., 1996] Nowell, L.; France R.; Hix D., Health L., Fox E., Visualizing Search Results: Some Alternatives to Query-Document Similarity. In: Frei, Hans-Peter; Harman, Donna K.; Schäuble, Peter et al. (Eds.): SIGIR 1996: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Conference: Zürich, Switzerland, August 18 -22 1996. New York (ACM Press) 1996. p. 67-75.
- [Ogden et al., 1998] Ogden, W., Davis, M.; Rice, S.: Document Thumbnail Visualization for Rapid Relevance Judgments: When do They Pay Off? In: Voorhees, Ellen M.; Harman, Donna K. (Eds.): NIST Special Publication 500-242: The Seventh Text Retrieval Conference (TREC-7). Conference: Gaithersburg, MD, November 09-11 1998. Gaithersburg, MD, USA (National Institute of Standards and Technology) 1998. p. 528-534.
- [Olive, 1997] <http://otal.umd.edu/Olive/>
- [PRISE] Lien vers le site du projet NIRVE : <http://zing.ncsl.nist.gov/~cugini/uicd/nirve-home.html>
- [Rauber et Bina, 2000] Rauber A., Bina H.: „'Andreas Rauber'? Conference Pages Are over There, German Documents on the Lower Left, ...": An „Old-Fashioned“ Approach to Web Search Results Visualization. In: Tjoa, A Min; Wagner, Roland R.; Al-Zobaidie, Ala (Eds.): Proceedings 11th International Workshop on Database and Expert Systems Applications. Conference: Greenwich, London, United Kingdom, September 4-8-2000. Los Alamitos, CA (IEEE Computer Society) 2000. p. 615-619.
- [Robertson et al, 1993] Robertson, George G.; Card, Stuart K.; Mackinlay, Jock D.: Information Visualization Using 3-D Interactive Animation. In: Communications of the ACM, 36 (1993) 4, p. 56-71.
- [Robertson et al., 1998] Robertson G.; Czerwinski, M.; Larson K., Robbins D., Thiel D., van Dantzich, Data mountain: using spatial memory for document management. In: UIST 98: 11th ACM Symposium on User Interface Software and Technology New York (ACM Press) 1998. p. 153-162.
- [Robertson et Mackinlay, 1993] Robertson G., Card S., Mackinlay J.: Information Visualization Using 3-D Interactive Animation. In: Communications of the ACM, 36 (1993) 4, p. 56-71.

- [Roussinov, 1999] Roussinov, D., Tolle, K., Ramsey M., McQuaid, M., and Chen, H., "Visualizing Internet Search Results with Adaptive Self-Organizing Maps," Proceedings of ACM SIGIR, August 15- 19, 1999, Berkeley, CA.
- [Sammon, 1969] Sammon, J.: A nonlinear mapping for data structure analysis. In: IEEE Transactions on Computers, 18 (1969) 5, p. 401-409
- [Shneiderman 1996] Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: Proceedings of IEEE Symposium on Visual Languages 96 Conference: Boulder, CO, September 3-6 1996. Los Alamitos, CA (IEEE Computer Soc.Press)1996. p. 336-343.
- [Shneiderman et al., 1997] Shneiderman, B.; Byrd, D.; Croft, W.: Clarifying Search: A User-Interface Framework for Text Searches. In: D-Lib Magazine, (1997).
- [Shneiderman 1998] Shneiderman, Ben: Designing the User Interface. Strategies for Effective Human-Computer Interaction. 3rd edition Reading, MA (Addison-Wesley) 1998.
- [Shneiderman et al., 2000] Shneiderman B., Feldman D., Rose A., et Ferre Grau X., Visualizing digital library search results with categorical and hierarchical axes. In ACM Digital Libraries, 2000.
- [Spence et Apperley, 1982] Spence, R.; Apperley, M.: Data base navigation: An office environment for the professional. In: Behavior and Information Technology, 1 (1982) 1, p. 43-54.
- [Spoerri, 1993] Spoerri, A.: InfoCrystal: A Visual Tool for Information Retrieval In: Nielson, Gregory M.; Bergeron, Daniel (Eds.): Proceedings of IEEE Visualization 1993. Conference: San Jose, CA, October 25-29 1993. Los Alamitos, CA (IEEE Computer Soc. Press) 1993. p. 150-157.
- [Spoerri, 1993a] Spoerri, A.: InfoCrystal: a visual tool for information retrieval management. In: Proceedings of the second international conference on Information and knowledge management. Conference: Washington DC, USA, 1993 November 1 - 5 1993. p. 11-20.
- [Spotfire, 2001] <http://www.spotfire.com> [2004-02-02]
- [Spring et al. 1996] Spring, M.; Morse, E.; Heo M.: Multi-level navigation of a document space. Proc. Leveraging Cyberspace Conference.
- [Stenmark, 1999] Stenmark, D.: To Search is Great, to Find is Greater: a Study of Visualization Tools for the Web, Internal Report.
- [Tweedie et al., 1996] Tweedie, L., Spence, R., Dawkes H., Su H.: Externalising abstract mathematical models. In: Tauber, M. J.; Bellotti, V.; Jeffries, R. et al. (Eds.): CHI 1996: Conference Proceedings Human Factors in Computing Systems. Conference: Vancouver, BC, Canada, April 13-18 1998. New York (ACM Press) 1996. p. 406-412.
- [Tweedie et al., 1994] Tweedie, L. A.; Spence, R.; Williams, D. Bhogal R.: The Attribute Explorer. Video Track. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 435-436.
- [Veerasamy, 1997] Veerasamy, A.: Visualization and User Interface Techniques for Interactive Information Retrieval Systems. PhD Thesis, Georgia Institute of Technology. March 1997
- [Veerasamy et al., 1995] Veerasamy, A.; Hudson, S. E.; Navathe, S. B.: Querying, Navigating and Visualizing an Online Library Catalog . Graphics, Visualization & Usability (GVU) Center at Georgia Tech: Technical Report 95-24. 1995.

- [Veerasamy et Navathe, 1995] Veerasamy, A.; Navathe, S. B.: Querying, Navigating and Visualizing a Digital Library Catalog. In: Digital Libraries 1995: The Second Annual Conference on the Theory and Practice of Digital Libraries. Conference: Austin, TX, June 11-13-1995.
- [Wills, 1995] Wills, G.: Visual Exploration of Large Structured Data Sets. In: Eurostat (Eds.): New Techniques and Trends in Statistics. Amsterdam (IOS Press) 1995. p.237-246.
- [Wise et al., 1995] Wise, J. A.; Thomas, J., Pennock, K., Lantrip D., Pottier, M., Schur, A., Crow, V: Visualizing the Nonvisual: Spatial Analysis and Interaction with Information from Text Documents. In: Gershon, Nahum; Eick, Stephen G. (Eds.): Proceedings of IEEE Information Visualization 1995. Conference: Atlanta, GA, October 30-31 1995. Los Alamitos, CA (IEEE Computer Soc. Press) 1995. p. 51-58.
- [Wise, 1999] Wise, J. A.: The Ecological Approach to Text Visualization. In: Journal of the American Society for Information Science (JASIS), 50 (1999) 13, p. 1224-1233. L'aide de l'application Galaxy ce trouve au http://in-spire.pnl.gov/IN-SPIRE_Help/galaxy.html
- [Wistrand, 1995] Wistrand, E.: Visualization Methods for Dynamic Queries Databases. 1995.
- [Zamir, 1998] Zamir, Oren E.: Visualization of Search Results in Document Retrieval Systems. General Examination. Department of Computer Science and Engineering, University of Washington, September 1998.

Summary

We present in this paper a survey of visual techniques that can be used in information search problems on the Web or in other kind of information systems. We first detail the techniques which allow users to formulate and refine search requests in an interactive and visual way. Then we review the techniques that are intended to represent a single document as well as a set of documents. We analyse the weaknesses and strength of these methods and we highlight some perspectives for this promising domain.