

INSTITUT DE LA FRANCOPHONIE POUR L'INFORMATIQUE



Institut de la Francophonie  
pour l'Informatique

**RAPPORT FINAL**  
**Travail Personnel Encadré**

*Sujet :*

**Visualisation de graphes pour la  
recherche interactive d'images**

**Encadrement : Alain Boucher (IFI)**

**Etudiant : LE Viet Man**

**Promotion 15, IFI**

**Hanoï, June 2010**

# REMERCIEMENTS

Je tiens tout d'abord à remercier mon encadrement Alain Boucher, professeur à l'Institut de la Francophonie pour l'Informatique. Il m'a aidé beaucoup dans la réalisation de mon sujet. Il m'a aussi donné des connaissances de base et des conseils pour réaliser mon sujet.

Je tiens à remercier tous les professeurs de l'IFI de m'équiper des connaissances nécessaires de haut niveau dans le domaine de l'informatique.

Je tiens à remercier mes camarades de promotion 15 de l'IFI pour leurs soutiens et pour l'ambiance familière.

Je tiens enfin à remercier ma famille et ma femme pour l'encouragement tout au long de mes études à l'IFI.

# RESUME

En se basant sur l'approche « mot visuel » - une technique d'indexation et de recherche d'image très connu par le contenu, on peut créer un réseau sémantique des termes (TSN - Term Semantic Network) dans lequel les termes sont les mots visuels et les arrêts sont les relations entre chaque deux termes. A partir de ce réseau, on peut créer deux types de graphes : le graphe des clusters où chaque noeud est un cluster et le graphe d'image où chaque noeud est une image. Ces graphes permettent de découvrir des relations entre des images et celles entre des clusters. Cependant, visualiser le graphe est un défi. Quand la quantité de noeuds est grande, il est difficile d'exploiter le graphe, à reconnaître les noeuds. De plus, les mises en page traditionnelle ne satisfait pas au graphe avec la grande taille. Ce document présentera un modèle de visualisation de graphes d'images où les noeuds du graphes sont des images et les liens représentant les relations existants entre images (ou groupes d'images). En utilisant la combinaison entre la mise en page radiale et la mise en page spirale, ce modèle permet d'exploiter et de traverser facilement ces graphes.

*Mots clés : mot visuel, recherche d'image, réseau sémantique, visualisation de graphes, mise en page radiale, mise en page spirale*

# TABLE DES MATIÈRES

<b>INTRODUCTION</b>	<b>1</b>
1. CONTEXTE	1
2. PROBLEME A RESOUDRE	1
3. RESULTATS ENTENDUS	2
4. CONTENU DU RAPPORT	2
<b>ETAT DE L'ART</b>	<b>3</b>
5. INDEXATION ET PARTITIONNEMENT D'IMAGES PAR LES MOTS VISUELS	3
5.1. Indexation par les mots visuels (Hoang, 2008)	3
5.2. Partitionnement d'images par les mots visuels (Doan, 2009)	4
5.3. Remarques	5
6. VISUALISATION DE GRAPHES POUR LA GRANDE BASE D'IMAGES	5
6.1. Visualisation de larges graphes d'images	5
6.2. Navigation, interaction et animation	7
7. PROJETS DISPONIBLES	8
7.1. MoireGraphs	8
7.2. Google Image Swirl	9
7.3. Vizster	9
7.4. TED Sphere	9
7.5. LJNET: LiveJournal Social Network Browser	10
8. COMPARAISON DES TECHNIQUES	10
9. LIBRAIRIES DISPONIBLES	11

9.1.	Tulip	11
9.2.	Prefuse	11
9.3.	Qt	11
<b>MODELE DE VISUALISATION DE GRAPHE D'IMAGES</b>	<b>13</b>	
10.	MODELE DE VISUALISATION DE GRAPHE D'IMAGES	13
11.	ORGANISATION DE LA BASE DE DONNEES	14
12.	PROCESSUS DE LA CONSTRUCTION DU GRAPHE D'IMAGES	15
13.	ALGORITHME DE LA MISE EN PAGE	16
13.1.	Algorithme de la mise en page spirale	16
13.2.	Algorithme de la mise en page radiale	17
14.	MODULE D'AFFICHAGE	17
15.	NAVIGATION ET INTERACTION	19
16.	OUTILS UTILISES	20
<b>RESULTATS EXPERIMENTAUX</b>	<b>21</b>	
17.	DONNEES EXPERIMENTEES	21
18.	ANALYSE DES RESULTATS	22
18.1.	La classe Dinosaur	22
18.2.	La classe Cheval et Plage	23
18.3.	La classe Montagne	25
18.4.	Les relations entre des images	26
<b>CONCLUSION ET PERSPECTIVES</b>	<b>28</b>	
19.	CONCLUSION	28
20.	PERSPECTIVES	29
<b>REFERENCES</b>	<b>30</b>	

# INTRODUCTION

## 1. CONTEXTE

Dans le développement de l'Internet, la recherche de l'information (en texte, en image, en vidéo, etc.) est maintenant le domaine très attractif. Mais les moteurs de recherche actuels, tels que Google, Yahoo, Bing, sont encore seulement efficaces à chercher des informations sur Internet par texte. Cette méthode est facile à appliquer, mais elle n'est pas optimale de chercher une image, car un texte n'exprime pas toujours exactement le contenu d'une image. La recherche d'images par le contenu (en anglais CBIR : Content-based image retrieval) qui devient maintenant un sujet très actif dans la communauté internationale, est une bonne solution à résoudre ce problème.

Avant-dernière année, HOANG Thanh Lam (promo 13 à l'IFI) a étudié l'approche « sac de mots » (ou bien les mots visuels) dans l'indexation et la recherche d'images par contenu. Et l'an dernier, DOAN Tuan Anh (promo 14) a développé un modèle permettant d'identifier des liens entre images en utilisant les travaux de Lam. Ce modèle peut bien organiser les résultats de recherche en réseaux graphique sémantique. Le travail de Tuan Anh est très intéressant, mais il existe un défaut qui est l'impossibilité de visualiser de grandes bases d'images et d'interagir avec le résultat.

Donc mon TPE a pour but d'étudier et développer une méthode de visualisation de graphes où les noeuds du graphes sont des images et les liens représentant les relations existantes entre images (ou groupes d'images). Cette méthode sera interactive afin de permettre à l'utilisateur de naviguer dans la base d'images, éventuellement d'abstraire ou de zoomer dans les relations existantes. Ce sera un outil permettant d'exploiter et traverser les liens entre des clusters et des images. C'est un TPE plein attrayant qui est partie du domaine traitement d'image, de l'indexation et la recherche d'images, de la visualisation de graphes et de la visualisation interactive.

## 2. PROBLEME A RESOUDRE

Mon travail consiste en deux parties : une partie théorique et une partie pratique.

### Travail théorique

- Etudier l'approche de recherche d'images par mots visuels

- Etudier le domaine de la visualisation interactive de graphes
- Rechercher les outils logiciels permettant de faire la visualisation interactive de graphes
- Développer un modèle de visualisation de graphes pour les grandes bases d'images

## Travail pratique

- Développer un prototype de navigation dans une base d'images par visualisation interactive de graphes.

## 3. RESULTATS ENTENDUS

- Bien comprendre l'indexation et recherche par mots visuels, le réseau graphe sémantique et la visualisation interactive de graphes afin de faire une synthèse de la combinaison entre tous les trois
- Proposer un modèle de visualisation de graphes pour la recherche interactive d'images
- Implémenter un logiciel permettant de visualiser et interagir un graphe des images (à partie du code développé par HOANG Thanh Lam et DOAN Tuan Anh)
- Trouver les relations entre les clusters et les images dans une base d'image en utilisant l'approche « mot visuel »

## 4. CONTENU DU RAPPORT

Ce rapport contient cinq chapitres. Tout d'abord, j'ai présenté le contexte du travail, l'objectif du sujet, les problèmes à résoudre et les résultats entendus dans ce chapitre. Le deuxième chapitre commencera avec le travail de HOANG Thanh Lam et DOAN Tuan Anh dans l'indexation et le partitionnement des images par les mots visuels pour construire des graphes d'images. Par ailleurs, ce chapitre parlera de l'état de l'art de la visualisation de graphes. Ensuite, dans le troisième chapitre, je présenterai mon modèle de visualisation de graphes d'images. Le quatrième chapitre, je vais expliquer quelques expériences pendant le travail sur ce sujet. Enfin, le dernier chapitre, je ferai une conclusion de mon travail avec des avantages et des désavantages. De plus, je proposerai quelques perspectives de mon TPE.

# ETAT DE L'ART

Il y a deux problèmes dans ce sujet. Ce sont :

- Un graphe d'images créé par un réseau sémantique des termes dans lequel les termes sont les mots visuels et les arrêts sont les relations entre chaque deux termes
- Un algorithme de la mise en page qui permet d'exploiter et de traverser facilement ces graphes.

Donc, dans cette partie, je parlerai l'état de l'art de ces problèmes.

## 1. INDEXATION ET PARTITIONNEMENT D'IMAGES PAR LES MOTS VISUELS

### 1.1. Indexation par les mots visuels (Hoang, 2008)

Distingue de la méthode de la recherche traditionnelle sur les mots clés, l'indexation et recherche d'images par les mots visuels utilise les mots visuels. Ce sont des descriptions représentatives d'un ensemble des descriptions invariantes qui sont représentées par des vecteurs.

Dans une image, il y a un ensemble de mots visuels. Ils s'appellent le sac de mots. Pour faire une requête d'image avec les mots visuels, on doit construire un dictionnaire de mots visuels qui contient tous les mots visuels de tous les sacs de mots dans une base d'images.

Pour identifier des mots visuels dans une image, il y a deux étapes : la première étape, on utilise un détecteur pour trouver des régions invariantes qui est une région détectée malgré à la transformation de l'échelle, de la rotation, de la lumière dans cette image et la deuxième, on doit décrire ces régions par une description invariante.

Dans un système de recherche d'images par les mots visuels, le processus a deux phases principales. Ce sont l'indexation d'image (hors-ligne) et la recherche d'images (en-ligne). Le rôle de la première phase est de créer le dictionnaire et le fichier inverse (indexation). La deuxième phase, le système extrait les mots visuels d'image requêtée et compare la similarité entre leur descripteur avec la base des index pour donner le résultat de la requête.

On a des détecteurs comme SA, MS, EBR, IBR, etc. La combinaison SA et MS est utilisé par le TPE de Lam parce qu'ils sont bien complémentaires ensemble dans les types différents d'images. En descripteur, le descripteur SIFT (Scale Invariant Feature Transform) est plus utilisé et bien sûr est utilisé dans le travail de Lam. Pour la comparaison de la similarité entre des descriptions, il y a beaucoup de mesures de la calculer tels que la mesure de similarité Indice de Jaccard, Euclide et Cosinus qui est appliqué par Lam.

## 1.2. Partitionnement d'images par les mots visuels (Doan, 2009)

En se basant sur le dictionnaire de mots visuels, le partitionnement (clustering en anglais) organisera des images dans la base d'images en réseaux graphique pour fournir une représentation cartographique des images et illustrer les liens entre eux.

Tout d'abord, on construit un réseau sémantique des termes (cf. Fig. 1) (TSN - Term Semantic Network) dans lequelle les termes sont les mots visuels et  $Sup(ti, tj) * Conf(ti, tj)$  est le poids des arêtes où  $Sup(ti, tj)$  (Support en anglais) indique les cooccurrences absolus des termes  $ti, tj$  et  $Conf(ti, tj)$  (Confidence en anglais) désigne leurs cooccurrences relatives par rapport aux apparences du terme  $ti$ . Ensuite, on utilise un algorithme pour le regroupement des termes. Selon Doan (2009), l'auteur a utilisé l'algorithme CHAMELEON, car cet algorithme a beaucoup plus d'avantages que les autres.

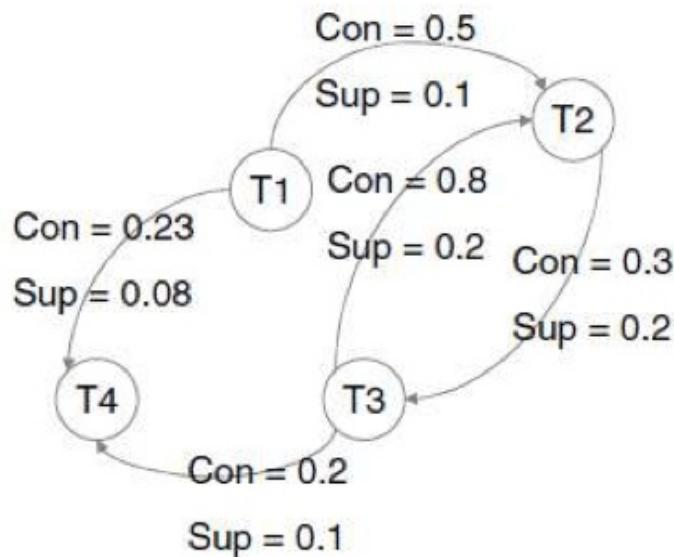


FIG 1 - Réseau sémantique de termes. (Doan, 2009)

Après avoir eu des clusters, on a besoin d'assigner des images de la base d'images dans les clusters qui conviennent. Pour le faire, on doit calculer la similarité entre l'image et le cluster. Une méthode efficace qui est un peu similaire avec la méthode de la computation de tf-idf, a été proposée par Doan (2009).

### 1.3. Remarques

Après le partitionnement, chaque image fera partie d'un ou de plusieurs clusters. Alors, il y a deux types de graphes :

- Le premier, c'est le graphe où chaque noeud est un cluster (le graphe des clusters). Deux clusters ont un lien entre eux quand ils contiennent une même image. Selon l'algorithme CHAMELEON, on ne prévoit pas la quantité de clusters que cet algorithme retournera, mais il est facile à trouver que c'est un graphe général non-orienté.
- Le deuxième, c'est le graphe où chaque noeud est une image (le graphe d'images). Deux noeuds ont un lien entre eux quand ils font partie d'un même cluster. Avec de 1000 à 10000 images, ce graphe est un grand graphe et aussi un graphe général non-orienté.

Donc, le modèle de visualisation de graphes pour les grandes bases d'images doit tout d'abord représenter les liens entre des images et aussi entre des clusters, doit ensuite afficher l'image à chaque noeud. La visualisation doit atteindre les règles esthétiques et l'utilisateur peut facilement exploiter le graphe, zoomer l'image.

## 2. VISUALISATION DE GRAPHES POUR LA GRANDE BASE D'IMAGES

### 2.1. Visualisation de larges graphes d'images

La visualisation de graphes est une partie de la visualisation d'information et a pour but d'exploiter les caractéristiques du système visuel humain pour faciliter la manipulation et l'interprétation de données informatiques variées (Hascoët, 2001). La visualisation de graphes est appliquée dans beaucoup de domaine comme la navigation sur le web, l'état des diagrammes de transition et les structures de données, etc.

Le majeur problème de la visualisation de graphes est la taille de graphe. Quand la quantité de l'éléments est grande, le graphe peut compromettre la performance ou même atteindre des limites du système. Même s'il est possible de mise en page et afficher tous les éléments, le problème de la lisibilité et de l'utilité se pose car il sera devenu impossible de discerner entre les noeuds et les arêtes. C'est aussi un problème qui les mises en page traditionnelles (cf. Fig. 2) ne satisfait pas.

Dans ces mises en page traditionnelles, la mise en page radiale (en anglais, radial layout) a utilisé plusieurs fois pour les graphes d'images dans les recherches récentes de Yee (2001), Jankun-Kelly (2003), Google (2009). C'est une bonne solution associant à la technique Spanning tree pour visualisation de graphes avec les noeuds visuels.

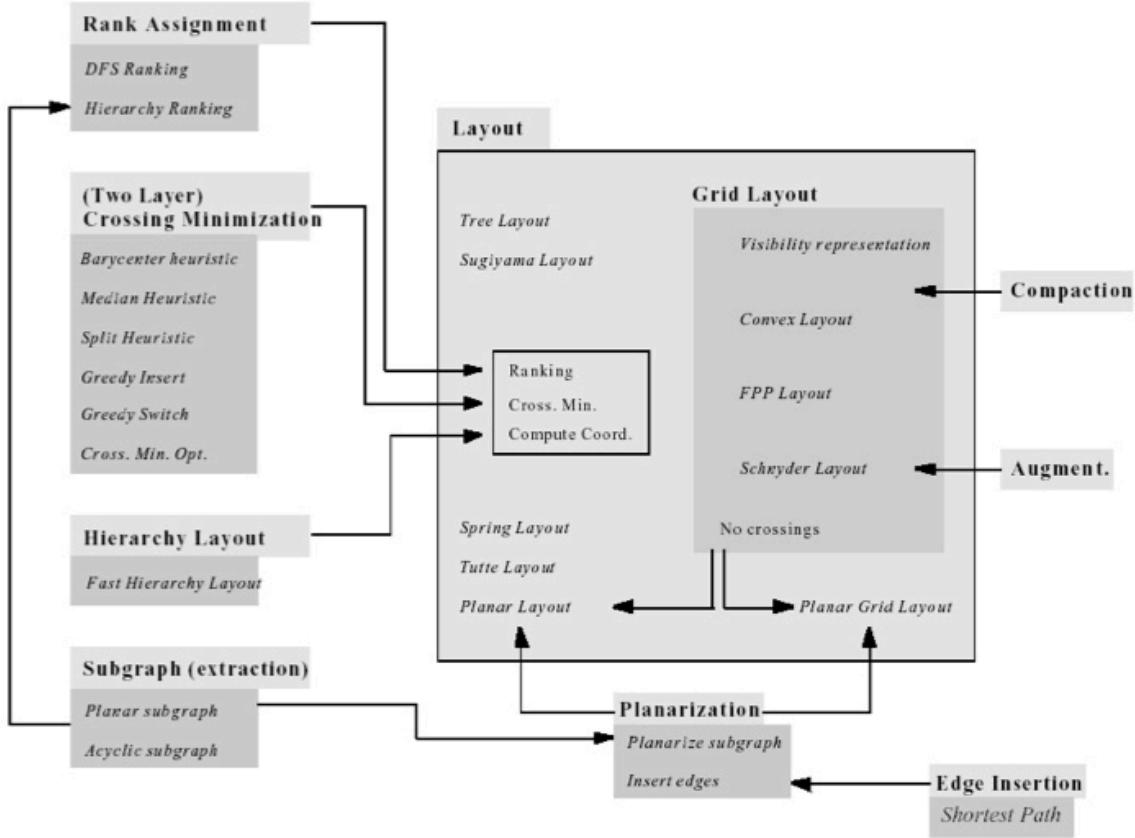


FIG 2 - Aperçu une classification des algorithmes de la mise en page de graphe

La mise en page radiale a été introduit par Eades en 1992. Cette mise en page est prévisible (en anglais, predictable) et a la plus moins complexité. Les graphes utilisant une mise en page radiale ont deux caractères : le noeud «focus» est au centre de la mise en page et les noeuds connecté au noeud «focus» rayonnent vers l'extérieur sur des anneaux séparés uniformément - un anneau pour un niveau de l'arbre. Dans le travail de Yee (2001), la taille variable des noeuds et l'animation en changeant dynamiquement des graphes radiaux ont ajouté. Et Jankun-Kelly (2003) (cf. Fig. 3) a étendu le travail précédent sur le graphe radial en incluant une déformation des niveaux. L'espacement entre les niveaux n'est plus uniforme dans son approche. En

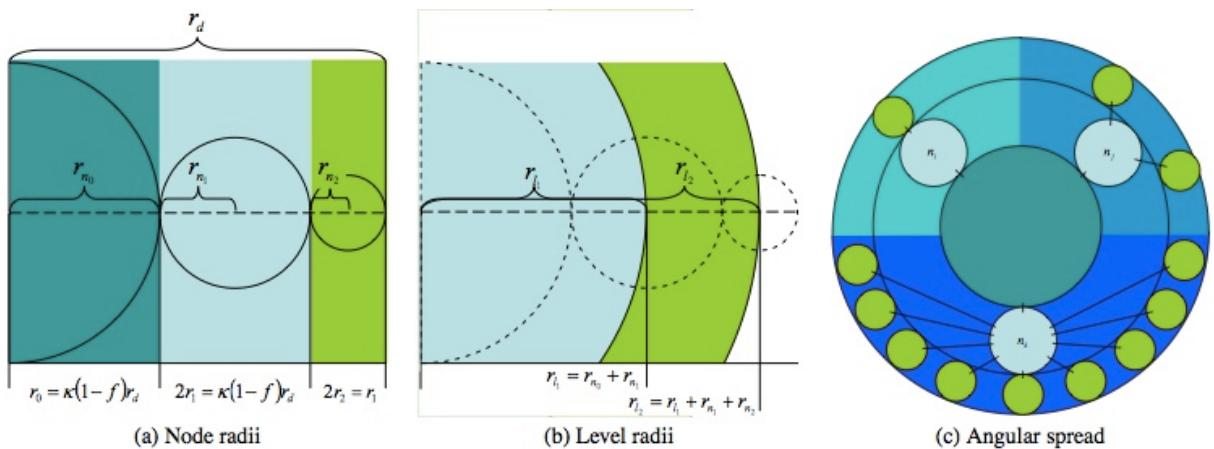


FIG 3 - La calcul de la mise en page Radial Focus+Context de Jankun-Kelly

utilisant cette technique, l'élément ayant le focus est mis en évidence alors que le reste du graphe reste en contexte.

D'autre part, il y a encore trois autres techniques sont proposées pour les grands graphes :

### **Spanning tree**

L'algorithme de la mise en page de l'arbre a la plus moins complexité et est simple à implémenter. Donc, on transforme souvent le graphe en l'arbre. On peut ne pas utiliser une propriété particulière ou utiliser une fonction du poids pour extraire un «spanning tree». Utiliser un «spanning tree» pour mettre en page un graphe peut aussi être une solution pour gagner la prévisibilité de la mise en page.

### **3 dimensions**

Afficher les graphes en 3D au lieu de 2D est une très populaire technique. Il apporte une dimension supplémentaire qui pourrait aider l'affichage de grandes structures. L'approche la plus simple est de généraliser l'algorithme de la mise en page en 2D pour en 3D. Mais l'affichage de graphes en 3D introduit également de nouveaux problèmes. Objets en 3D peuvent voiler les autres et il est souvent difficile de trouver la meilleure position de la caméra dans l'espace.

### **Géométrie hyperbolique**

La mise en page hyperbolique de graphes (principalement les arbres) est une des nouvelles formes de la mise en page qui a été développé avec la visualisation interactive de graphe dans l'esprit. La vue hyperbolique qui peut être implémentée soit en 2D ou 3D, fournit une vision déformée d'un arbre. Il ressemble l'effet de l'utilisation des lentilles fish-eye sur les mises en pages de l'arbre traditionnelles. Cette technique utilise la géométrie hyperbolique.

Les mises en page en 3D et hyperbolique fournissent la plus espace pour afficher les grands graphes, mais il est difficile à exploiter des liens entre les éléments de graphe. Dans la partie 4 de cette chapitre, je ferai une comparaison entre les techniques à choisir de bonnes techniques pour un modèle de visualisation de graphe d'images.

## **2.2. Navigation, interaction et animation**

*La navigation et l'interaction* sont très importantes pour la visualisation de l'information. Aucun seul algorithme de mise en page peut surmonter les problèmes soulevés par la grande taille des graphes.

### **Zoom et expanse**

Zoom et expanse sont des outils traditionnels dans la visualisation. Ils sont tout à fait indispensable dans l'exploration de grandes structures de graphes. Le zoom peut avoir deux formes : géométrique et sémantique. Le zoom géométrique étend simplement le contenu de graphe et dans le zoom sémantique, le contenu change et plus de détails sont montrés. La technique zoom a un problème bien connu. C'est que si l'on zoome sur un foyer, toute l'information contextuelle est perdue.

### **Focus and context Techniques**

Pour résoudre le problème du zoom, on a proposé des techniques qui permettent à l'utilisateur de se focaliser sur un détail sans perdre le contexte. Il y a trois grandes approches dans ce domaine. La première approche déforme le graphe après qu'il a été mis en page (les techniques fish-eye font partie de cette catégorie). La seconde méthode mappe le graphe sur une surface de dimensions supérieures et ensuite reprojecte sur deux dimensions. La dernière approche utilise la géométrie non-euclidienne (comme la géométrie hyperbolique).

Dans le travail de Jankun-Kelly (2003), il a uni la mise en page radiale et la technique focus+context pour créer une nouvelle mise en page pour la visualisation interactive de graphe avec les noeuds visuels.

*L'animation* est aussi une partie importante de la visualisation. Quand on change le foyer (en anglais, focus), la nouvelle vue peut causer un réaménagement très désorienté. L'animation qui effectue une transition lisse, peut réduire cette désorientation. Yee (2001) a proposé une méthode de l'animation dans le graphe radial : les coordonnées polaires de chaque noeud est interpolée entre les positions de départ et d'arrivée lors de l'animation.

## **3. PROJETS DISPONIBLES**

Dans cette partie, je vous introduit quelques projets disponibles qui ont traité des graphes d'images ou des graphes de vidéo.

### **3.1. MoireGraphs**

MoireGraphs (cf. Fig. 4) est une recherche de VisLab à l'Université Mississippi State. Il combine une nouvelle mise en page radiale focus+context avec une suite des techniques d'interaction (changement de la puissance du foyer, la rotation radiale, le marquage du niveau, le foci secondaire, des transitions

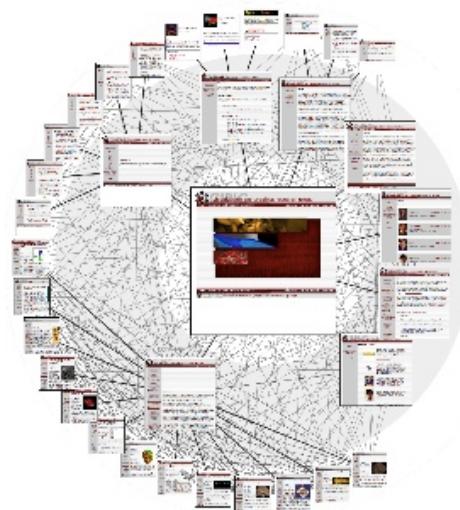


FIG 4 - MoireGraph

animées) pour aider l'exploration des graphes avec des noeuds visuels. La méthode peut être étendue à des centaines de noeuds affichés visuelle. La source code de l'algorithme du MoireGraphs est un module d'extension de la librairie Tulip.

### 3.2. Google Image Swirl

C'est un nouveau outil de Google (cf. Fig. 5) pour la recherche d'images similaires et connexes au travers d'une interface intuitive. Il utilise la mise en page radiale et l'animation spirale. L'utilisateur peut exploiter facilement le graphe d'images, et percevoir les liens entre des images.

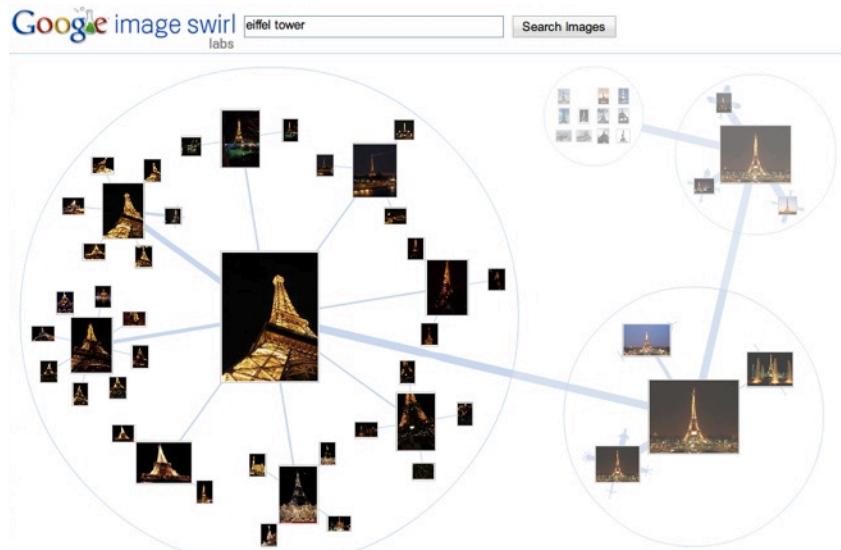


FIG 5 - Google Image Swirl

### 3.3. Vizster

Vizster est un système de la visualisation pour exploiter des réseaux sociaux en ligne. Vizster a été écrit en Java en utilisant la librairie de la visualisation Prefuse. Il utilise les mises en page de réseaux noeud-lien pour l'exploitation de la connectivité dans la structure de grand graphe. La mise en page de réseaux est calculé par l'algorithme spring-embedding. Il soutient la recherche et l'analyse visuelle. Il fournit plusieurs façons pour exploiter le réseaux, telles que le marquage de la connectivité, les vues du couplage, la vue en XRay mode.

### 3.4. TED Sphere

C'est un projet (cf. Fig. 6) du studio Bestiario. Il utilise la mise en page en 3D et la technique de la navigation du réseau basé sur la compatibilité sémantique entre des vidéos de TED.



FIG 6 - TED Sphere

### 3.5. LJNET: LiveJournal Social Network Browser

LJNet (cf. Fig. 7) est un travail de Patrick Barry. Elle a été faite au départ comme un projet en classe du cours de maîtrise de Golan Levin à Carnegie Mellon - Computer Interaction Institute. LJNet est une visualisation interactive des réseaux sociaux des membres de LiveJournal.com. Elle affiche les amis et les amis d'amis d'un membre LJ que satisfait l'esthétique et facile à lire de réseau. Ce système actuel est un logiciel en Java qui utilise la librairie Processing.



FIG 7 - LJNET

## 4. COMPARAISON DES TECHNIQUES

Le tableau suivant vise à comparer les techniques de la mise en page pour la visualisation de graphes. Je l'ai utilisé pour proposer la solution de mon TPE.

Techniques	Avantages	Désavantages	Remarques
<b>Spanning tree</b>	- moins complexité et simple à implémenter - gagner la prévisibilité de la mise en page		visualiser les arbres est plus facile que les graphes
<b>3 dimensions</b>	- fournir la plus espace pour afficher les grands graphes	- voiler les autres noeuds - être difficile de trouver la meilleure position de la caméra dans l'espace	n'est pas bonne pour visualiser les images
<b>Géométrie hyperbolique</b>	- fournir la plus espace pour afficher les grands graphes	- être difficile à exploiter des liens entre les éléments de graphe	n'est pas bonne pour visualiser les images
<b>La mise en page radiale</b>	- prévisible, moins complexité - pouvoir voir le focus et le contexte	- quand la quantité des noeuds « enfant » est grande, la navigation sera devenue difficile	c'est bonne pour visualiser les images

Techniques	Avantages	Désavantages	Remarques
<b>La mise en page « spirale »</b>	- permettre traverser facilement des niveaux de l'arbre	- la technique de l'animation est difficile à implémenter	c'est bonne pour visualiser les images

A partir des avantages et des désavantages ci-dessus, j'utiliserais trois techniques : spanning tree, la mise en page radiale et la mise en page « spirale ».

## 5. LIBRAIRIES DISPONIBLES

J'ai recherché trois librairies suivantes qui résolvent les grandes graphes :

### 5.1. Tulip

Tulip est une plate-forme de visualisation d'information dédié à l'analyse et la visualisation de données relationnelles. Tulip vise à fournir au développeur d'une bibliothèque, appuyant la conception d'applications interactives de visualisation d'information pour les données relationnelles qui peuvent être adaptées aux problèmes qu'il aborde. Tulip est implémenté en basant sur la librairie Qt de Nokia. Cette librairie fournit une librairie de graphique très puissante. De plus, le code source de la mise en page Radial Focus+Context est un module d'extension de Tulip.

Pourtant, Tulip a aussi beaucoup d'erreurs et elle n'a pas d'objets nécessaires pour mon modèle. Donc, je ne l'utilise que pour tester le graphe d'images et j'utiliserais l'algorithme de la mise en page Radial Focus+Context de Tulip.

### 5.2. Prefuse

Prefuse est un ensemble d'outils logiciels pour créer des visualisations de données interactives riches. Il fournit une plate-forme de visualisation pour le langage de programmation Java. Prefuse supporte un ensemble riche de fonctionnalités pour la modélisation des données, de visualisation et d'interaction. Il fournit des structures de données optimisées pour les tableaux, graphiques, et les arbres, les techniques de toute une série de mise en page et le codage visuel, et le soutien pour l'animation, des requêtes dynamiques, de recherche intégré et la connectivité de base de données. Prefuse est écrit en Java, en utilisant la bibliothèque graphique Java 2D, et est facilement intégré dans des applications ou des applets Java.

### 5.3. Qt

Qt est une plate-forme de programmation C++. Qt vous offre une puissante librairie graphique 2D pour la gestion et l'interaction avec un grand nombre d'éléments graphiques. Graphics View utilise un arbre BSP (Binary Space Partitioning) pour fournir la découverte rapide des éléments et il permet de visualiser de grandes scènes

en temps réel. De plus, il supporte le zoom, la rotation, l'animation et de la transformation perspective.

Puisque Qt a des puissantes fonctionnalités, j'utiliserais la plate-forme Qt pour implémenter mon modèle.

# MODELE DE VISUALISATION DE GRAPHE D'IMAGES

## 1. MODELE DE VISUALISATION DE GRAPHE D'IMAGES

Je propose un modèle de visualisation de graphe d'images comme suivant :

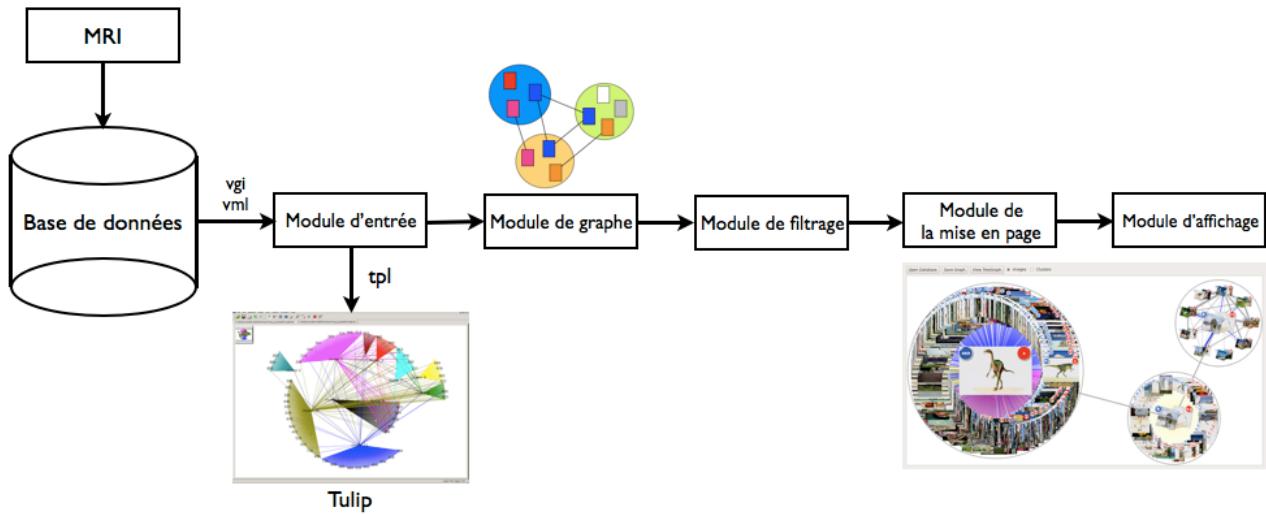


FIG 8 - Modèle de visualisation de graphe d'images

Dans ce modèle, il y a deux phase : la phase hors-ligne et la phase en-ligne.

Tout d'abord, dans la phase hors-ligne, on construit la base de données. J'ai utilisé le logiciel de DOAN Tuan Anh pour construire cette base de données. Ensuite, j'ai créé un fichier de configuration (il s'appelle fichier vgi) qui stocke les chemins des images et les chemins des données.

Dans la phase en-ligne, il y a cinq modules. Chaque module communiquent avec l'autre en envoyant des messages. Le module d'entrée chargera des images et des données à partir de la base de données. Dans ce module, je n'utilise pas de processus légers qui chargent rapidement des images parce que avec ce modèle, chaque moment, on n'affiche que quelques images, pas toutes les images. A partir les images chargées, le module de graphe construira des arbres d'images avec un cluster à la racine. La construction des arbres effectuera en utilisant la technique en largeur d'abord. Après avoir eu des arbres d'images, se baser sur la demande du module d'affichage, le

module de filtrage filtrera des noeuds qui sont nécessaires à afficher. Le module de mise en page mettra des coordonnées à chaque noeuds filtrés en utilisant la combinaison de la mise en page radiale et de la mise en page spirale. Module d'affichage rendra (render en anglais) les noeuds sur l'écran. Il utilise la technique focus+contexte pour aider l'utilisateur dans l'interaction avec le modèle. Quand l'utilisateur change le focus, le module enverra un message au module de filtrage pour refiltrer les noeuds, si nécessaire.

Par ailleurs, le module d'entrée a encore une fonction d'enregistrer le graphe sur le fichier tlp du logiciel Tulip.

## 2. ORGANISATION DE LA BASE DE DONNEES

J'ai utilisé le logiciel de DOAN Tuan Anh pour créer la base de données. Il y a plusieurs données sorties du logiciel de DOAN Tuan Anh, mais j'ai seulement besoin des fichiers suivants :

- des images dans la base d'images
- un fichier de texte qui stocke la quantité des clusters
- un ensemble des fichiers de texte qui sont correspondants avec les images. Chaque fichier stocke des indices de clusters qui cette image appartient.

---

```

vgi1.0
//
// Fichier de configuration pour VGI
//
// chemin du fichier de configuration de la base d'images
image /home/manleviet/TPE/Database/BW/db.conf
// chemin du repertoire de MS
databaseMS /home/manleviet/TPE/Database/database01/MS
// chemin du repertoire de SA
databaseSA /home/manleviet/TPE/Database/database01/SA
// dans les repertoire MS et SA, on doit contenir les fichiers suivants :
// 1. igroups.data
// 2. les fichiers data des images

```

Puisque les entrées sont vraiment beaucoup, j'ai défini un fichier de configuration qui stocke les chemins des données. Le logiciel est simple à lire le contenu du fichier et à prendre les chemins nécessaires. Par exemple, un fichier de configuration du modèle est ci-dessous :

La version du fichier de configuration est 1.0. Cette version soutient seulement les graphes qui n'ont pas de sous-clusters (la version 2.0 la soutient). Dans ce fichier, il y a trois chemins. Le premier chemin est un chemin du fichier de configuration de la base d'images qui stocke des chemins d'images. Le deuxième et le troisième, ce sont

les chemins du répertoire de MS et de SA. Chaque répertoire stocke des données sorties correspondantes avec le détecteur MS ou SA.

### 3. PROCESSUS DE LA CONSTRUCTION DU GRAPHE D'IMAGES

Le processus de la construction du graphe d'images est exécuté comme suivant :

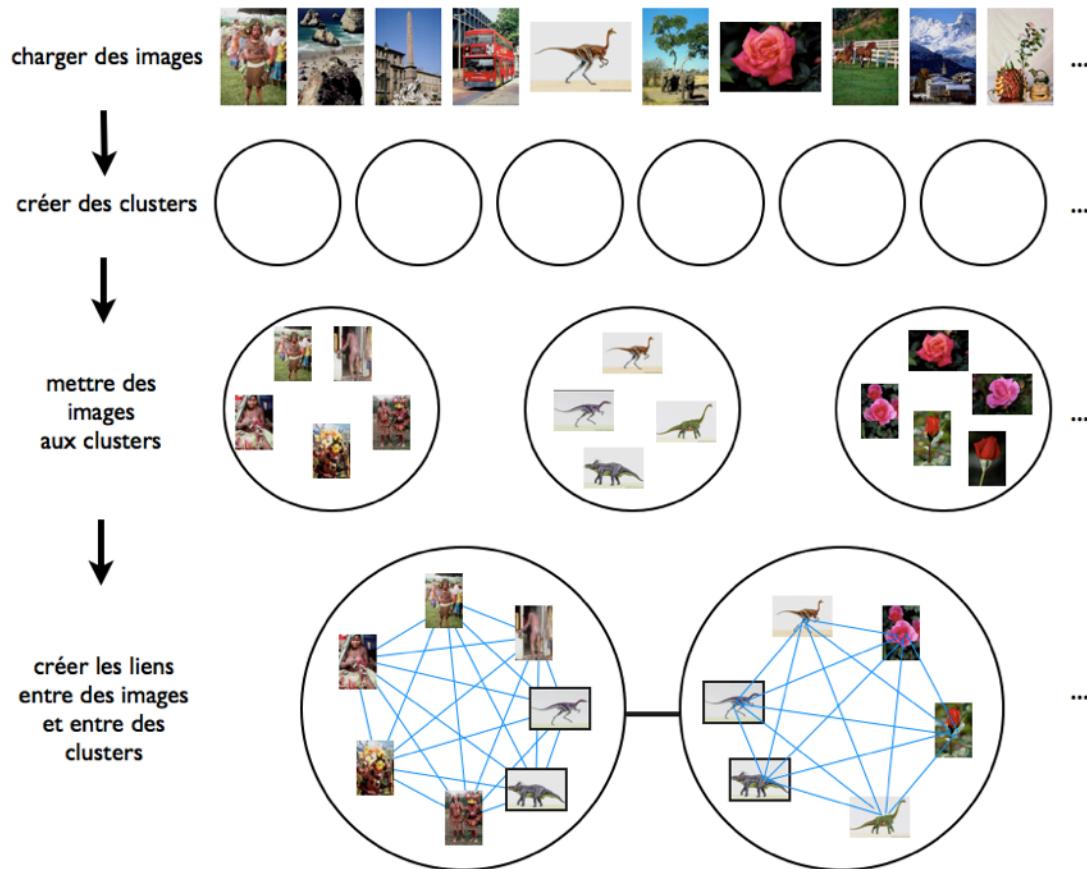


FIG 9 - Le processus de la construction du graphe d'images

Tout d'abord, le modèle prend les chemins des images pour créer des objets qui stockent des images. Dans cette phase, le modèle ne charge pas les images, il ne crée que les objets qui stockent le chemin d'image. Le chargement des images est seulement exécuté quand l'image est affichée. Cela permet de diminuer le temps de l'initialisation du graphe d'images du modèle.

Ensuite, le modèle lit le contenu du fichier qui contient la quantité des clusters pour créer les objets cluster.

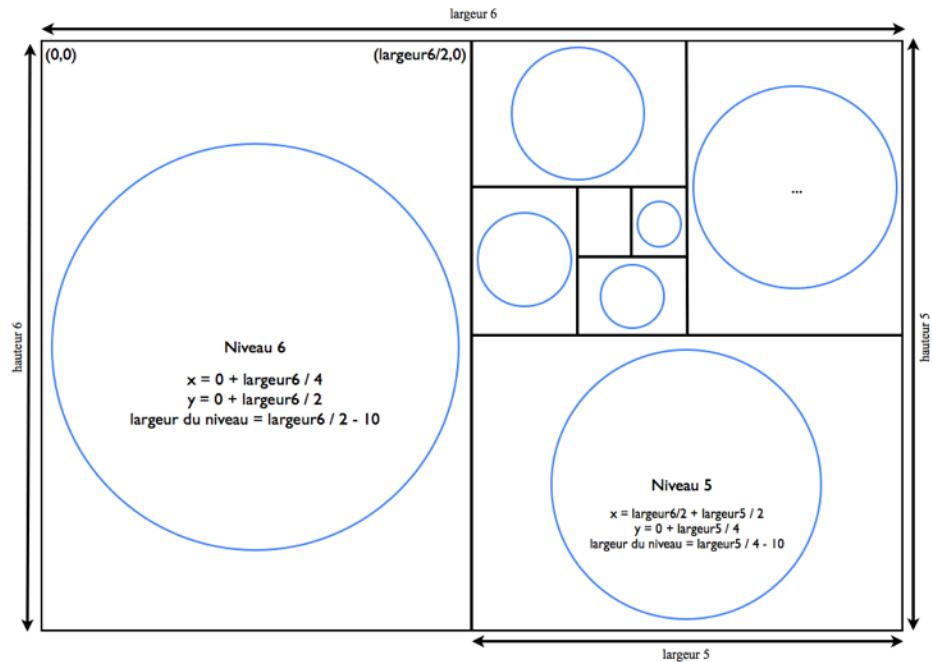
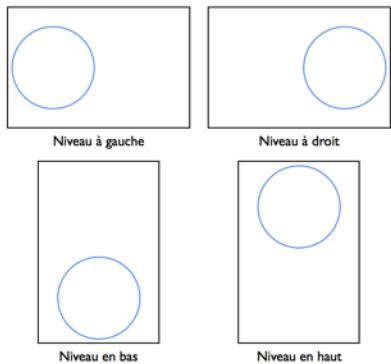
La troisième phase, le modèle lit les fichiers de données qui stockent des indices des clusters pour mettre les objets image aux objets cluster.

Enfin, le modèle créer les liens entre des images et entre des clusters. Les images ont les liens quand elles appartiennent au même cluster. Les clusters ont les liens quand ils ont des mêmes images.

## 4. ALGORITHME DE LA MISE EN PAGE

### 4.1. Algorithme de la mise en page spirale

L'algorithme de la mise en page spirale divise le graphe d'images en plusieurs niveaux (cf. Fig. 11). Ces niveaux se présentent un spirale où le plus haut niveau est à la plus profonde position dans le spirale. Cet algorithme calculera tour à tour la coordonnée (x,y) et la largeur de chaque niveau en commençant au plus bas niveau.



Ici, je trouve qu'il y a quatre types de niveau (cf. Fig. 10) : niveau à gauche, niveau en bas, niveau à droit et niveau en haut. Avec chaque type de niveau, on a une formule propre pour calculer la coordonnée et la largeur du niveau. L'algorithme commencera le calcule au plus bas niveau avec la largeur de la région dessinée est la largeur de l'écran et la hauteur de la région dessinée est la hauteur de l'écran (cf. Fig. 11). Après avoir calculé pour un niveau, on calcule la région dessinée pour l'autre niveau. Cette région est le reste de la région du niveau qui dépend de type du niveau. L'algorithme continue jusqu'au dernier niveau.

En divisant les niveaux en quatre types de niveau, mon algorithme de la mise en page spirale a éliminé la récursive de l'algorithme original. Il est trop simple. Cet algorithme a seulement besoin d'une seule boucle à calculer tous les niveaux. La complexité de l'algorithme est  $O(n)$ .

## 4.2. Algorithme de la mise en page radiale

Pour chaque niveau de l'algorithme de la mise en page spirale, on met les images au niveau en utilisant l'algorithme de la mise en page radiale. En héritant l'algorithme de la mise en page « Radial Focus+Context » de Jankun-Kelly (2003), j'ai modifié cet algorithme pour s'adapter à ce modèle.

Avec mon modèle, il y a seulement deux niveaux radiaux : un focus et un premier niveau. Donc, le calcul de la mise en page radiale est calculé comme suivant :

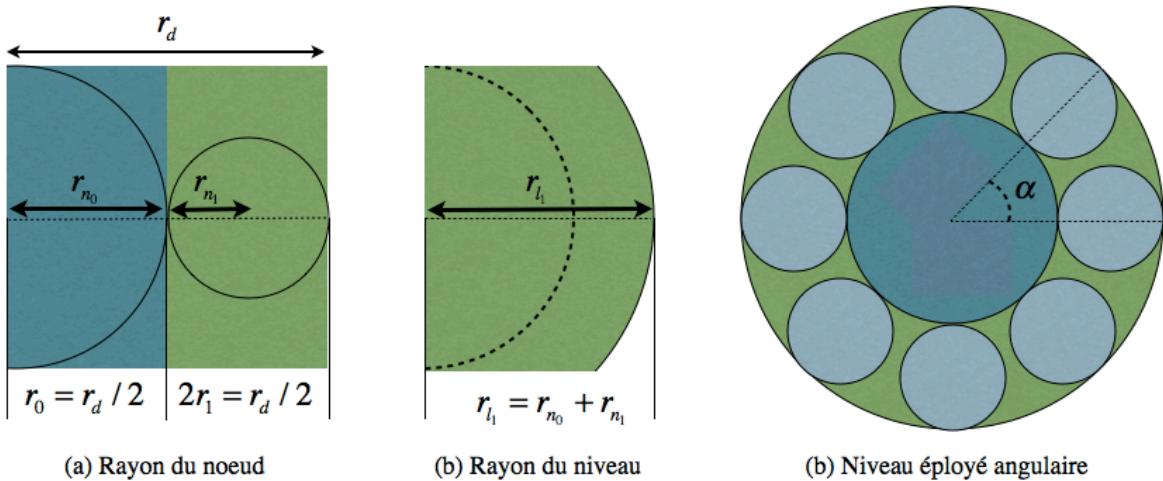


FIG 12 - Le calcul de la mise en page radiale

Tout d'abord, on calcule le rayon du noeud. On calcule le rayon du noeud « focus » est égal à la moitié de la largeur du niveau. Le rayon du noeud postérieur a la moitié de ce rayon. Ensuite, on calcule la distance radiale du focus aux noeuds dans le niveau en sommant les rayons de chaque rayon du noeud précédents. Enfin, les positions angulaires d'un noeud est déterminé par la division en poids du niveau éployé angulaire de son parent. La complexité de l'algorithme est aussi  $O(n)$  avec  $n$  est la quantité des images au premier niveau.

## 5. MODULE D'AFFICHAGE

Après avoir eu le graphe d'images, le premier niveau du graphe (cf. Fig. 13) est mis en page à l'écran. Le premier niveau du graphe est le graphe entre des clusters. Pour chaque cluster, le modèle affiche trois premières images de ce cluster. Il y a deux chiffres sur les clusters. Le premier (avec le fond bleu) est l'indice du cluster et le deuxième (avec le fond rouge) est la quantité des images appartenant au cluster. L'utilisateur peut changer le focus en cliquant le bouton droit (cf. Fig. 14).

Ce modèle supporte aussi les graphes ayant les sous-clusters (cf. Fig. 15). Comme les clusters, les sous-clusters affiche aussi trois premières images de ce sous-cluster et il y a aussi deux chiffres sur les sous-clusters.

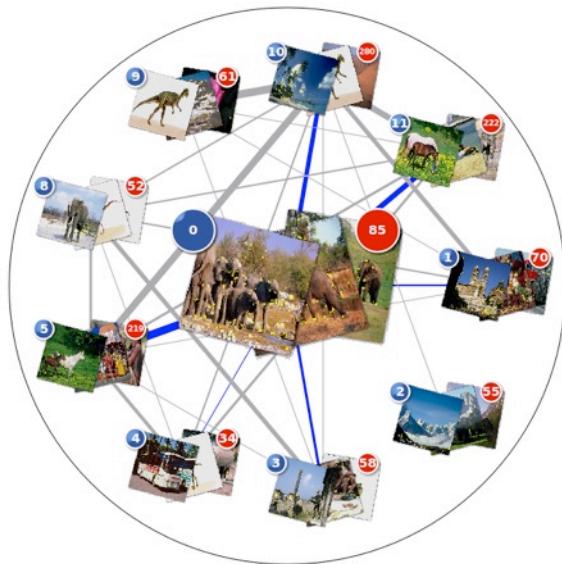


FIG 13 - Le premier niveau du graphe

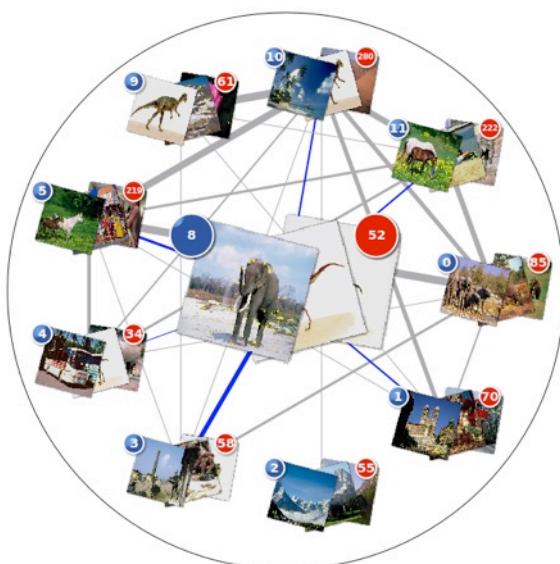


FIG 14 - Changer le focus

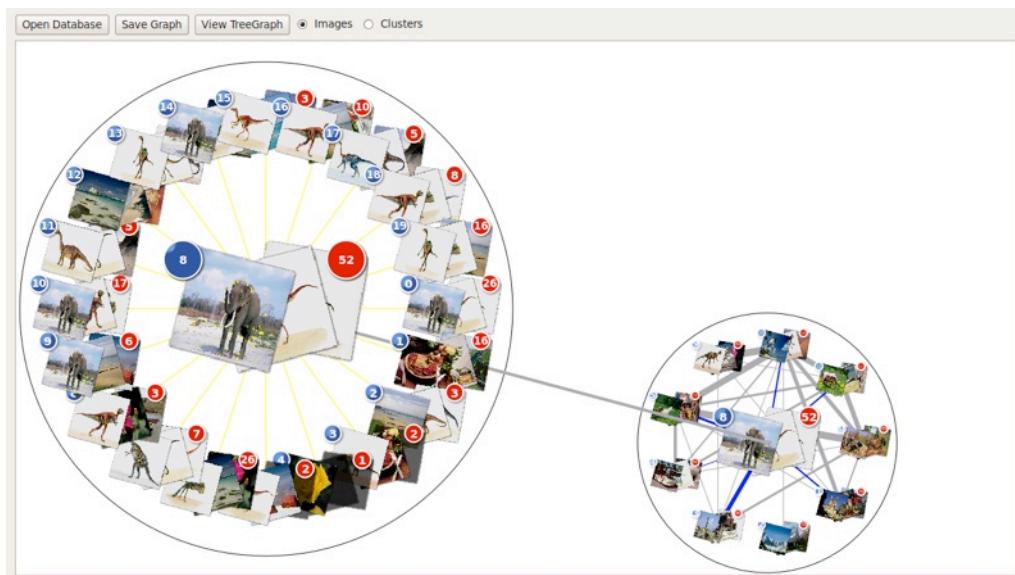


FIG 15 - Visualiser les sous-clusters

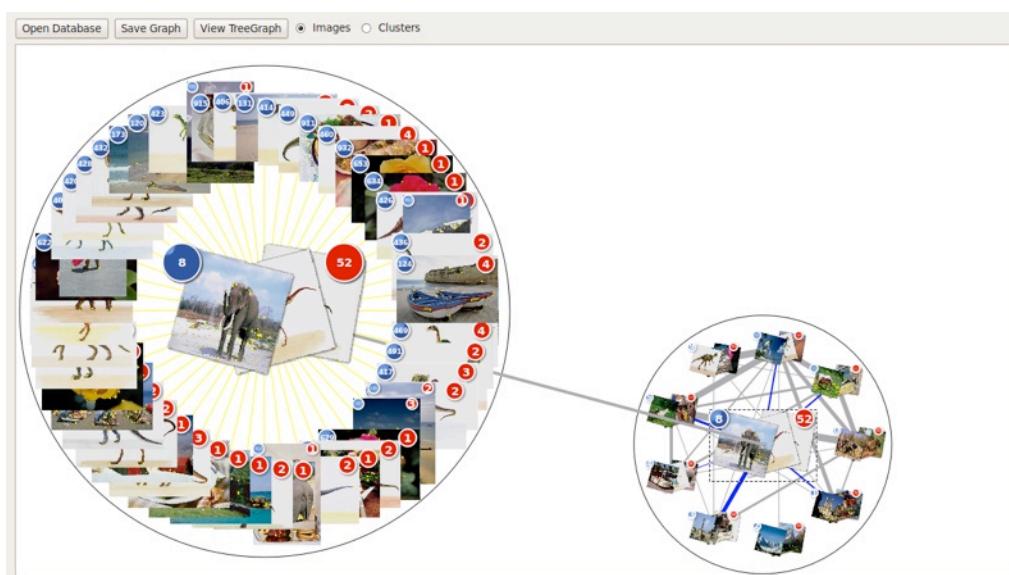


FIG 16 - Visualiser les images du cluster 8

En cliquant sur les clusters ou sur les sous-clusters, on peut voir les images qui appartiennent au cluster (cf. Fig. 16). Sur les images, on a aussi deux chiffres. Le premier chiffre (avec le fond bleu) est le nom du fichier d'image et le deuxième (avec le fond rouge) est la quantité de clusters qui cette image appartient.

Enfin, on peut cliquer sur une image pour trouver les relations entre cette image et les autres images dans la base d'images (cf. Fig. 17). Ici, on peut voir quels clusters appartient-elle cette image en cliquant le bouton radio Clusters (cf. Fig. 18).



FIG 17 - Visualiser les relations entre une image et les autres images

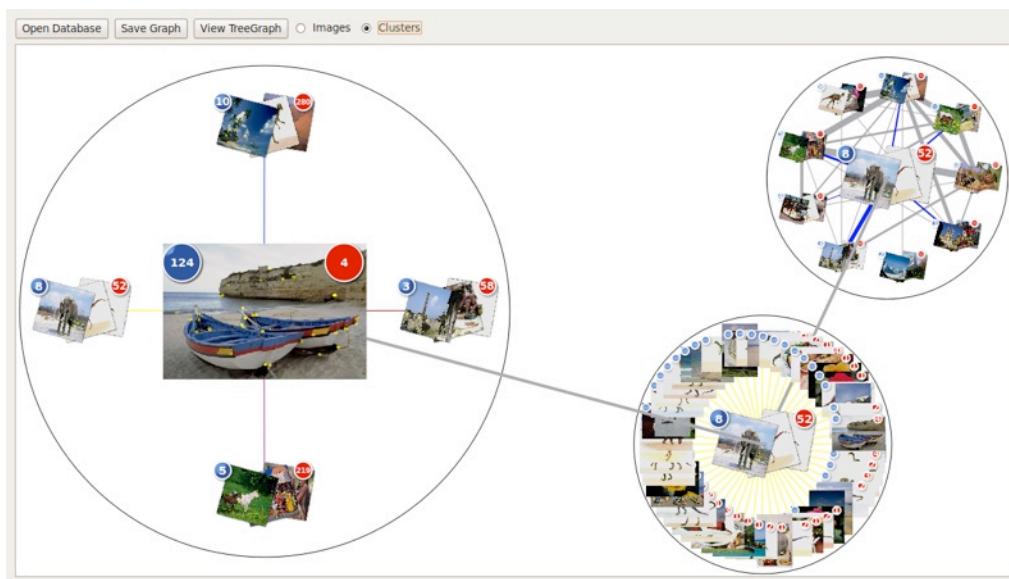


FIG 18 - Visualiser les relations entre une image et les clusters

## 6. NAVIGATION ET INTERACTION

Le modèle fournit quelques techniques de navigation et d'interaction pour permettre d'exploiter facilement le graphe.

### Zoom et expanse

Quand on visualise un niveau avec beaucoup d'images, les images peuvent se voiler. Donc, le modèle fournit la simple technique zoom et expanse. Quand le curseur se passe sur une image, la taille de cette image est élevée. Et le curseur quitte, cette image est restée la taille originale.

### **Exploiter**

L'utilisateur peut seulement exploiter le graphe vers le bas jusqu'au huitième niveau parce que les niveaux après le huitième niveau ne visualise pas bien les images. L'utilisateur peut exploiter vers le haut en cliquant le focus du niveau précédent.

## **7. OUTILS UTILISES**

J'ai utilisé les outils suivants :

- Le logiciel MRI de DOAN Tuan Anh pour construire la base de données
- La librairie Qt de Nokia pour coder ce modèle
- L'algorithme la mise en page radiale de la librairie Tulip

# RESULTATS EXPERIMENTAUX

## 1. DONNEES EXPERIMENTEES

Pour tester mon modèle, j'ai utilisé la base d'images de Wang. Cette base d'images a 1000 images et 10 classes avec 100 images pour chaque classe. C'est la base d'images pour l'indexation et la recherche d'image par le contenu. Mais les classes d'images sont très disparates (cf. Fig. 19). A partir de la comparaison des classes de Wang ci-dessus, on peut trouver que les classes 4 et 7 sont très bon et les classes 8 et 1 sont très mauvais. Donc, je vais tester mon modèle pour trouver ces classes.

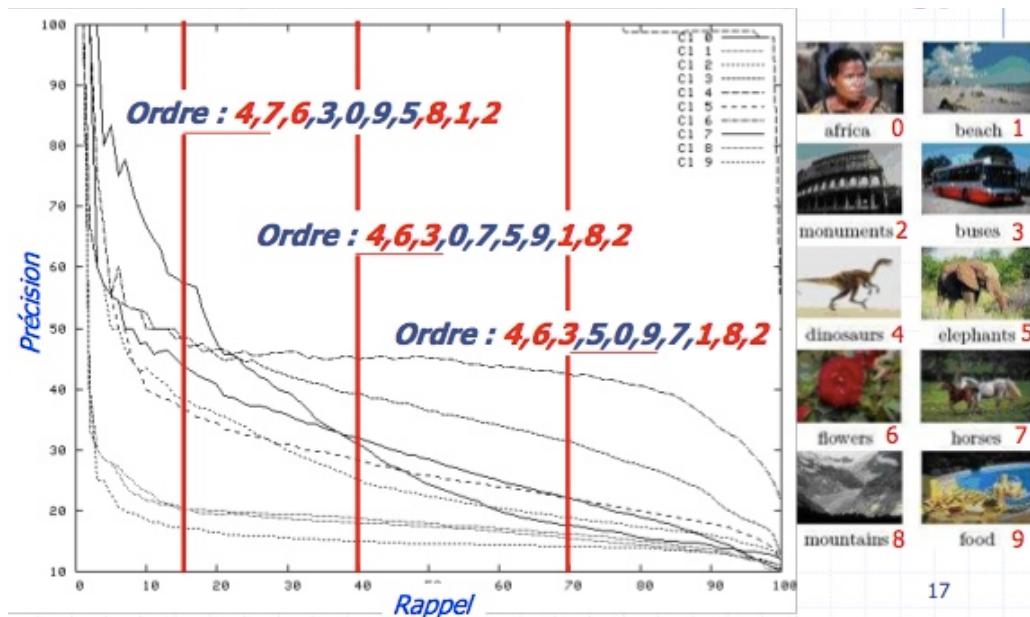


FIG 19 - La comparaison des classes de Wang  
 Source : Alain Boucher, Thanh-Ha Dang, Thi-Lan Le. Classification vs recherche d'information: vers une caractérisation des bases d'images

Les données de l'indexation et du partitionnement a créé avec cette base d'images et la paramètre de la quantité des mots est 500. Mon modèle ne visualise que les clusters détectés par le détecteur SA. J'ai choisi la quantité 500 car selon le résultat de HOANG Thanh Lam, le résultat de la quantité 250 est pareil au résultat de la quantité 500. Par ailleurs, selon les résultats du logiciel de DOAN Tuan Anh, j'ai trouvé que le résultat du détecteur MS n'est pas bon. Il donne seulement un ou deux clusters. Selon moi, son résultat n'est pas bon parce que ce détecteur considère que les mots visuels se sont trop proches et ils ne sont pas séparés.

## 2. ANALYSE DES RESULTATS

### 2.1. La classe Dinosaur

Selon la comparaison des classes de Wang, la classe dinosaure est très bien à rechercher et à classifier parce qu'elle ne contient que des dinosaures et il n'y a pas beaucoup de bruits dans cette image. Pourtant, dans mon résultat, cette classe est divisée en deux clusters : cluster 8 (cf. Fig. 20) et cluster 10 (cf. Fig. 21).



FIG 20 - Cluster 8

FIG 21 - Cluster 10

En examinant les images dans cluster 8 (cf. Fig. 22), on peut dire que ce cluster représente les mots visuels étant au corps du dinosaure et étant à la bordure entre le corps du dinosaure et le fond. Sur les images, les cercles jaunes sont les mots visuels.



FIG 22 - Quelques images dans le cluster 8

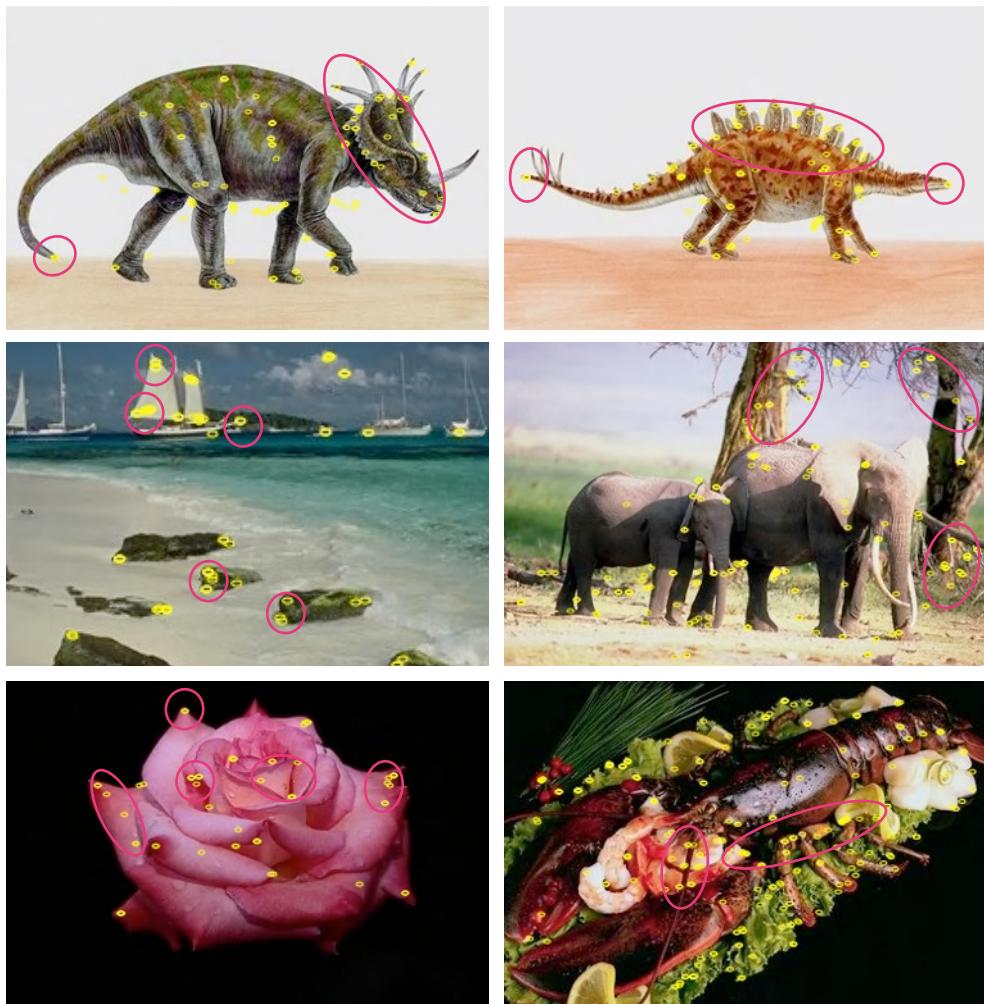


FIG 23 - Quelques images dans le cluster 10

En semblablement, j'ai examiné les images du cluster 10 (cf. Fig. 23), on peut dire que ce cluster représente les mots visuels étant à la ramie, à la bouche et à la queue ou au pied, à la main du dinosaure.

Ces deux clusters peuvent être considérés deux clusters de dinosaure. Chaque cluster représente quelques parties du dinosaure.

## 2.2. La classe Cheval et Plage

La classe cheval est aussi une bonne classe dans la base d'images de Wang. Mais, dans mon test, j'ai trouvé que cette classe se mêle avec la classe plage dans le cluster 11 (cf. Fig. 24). C'est un grand cluster avec 222 images où 52 images de cheval et 40 images de plage. En examinant les images du cluster 11 (cf. Fig. 25), je peux dire que ce cluster représente les mots visuels des chevaux et ceux des arbres. Les images des plages sont considérées comme les chevaux ou comme les arbres.



FIG 24 - Le cluster 11



FIG 25 - Quelques images dans le cluster 11

### 2.3. La classe Montagne

La classe Montagne reçoit souvent les mauvais résultats dans les requêtes de recherche. Les images des montagnes ont plusieurs mots visuels et ont plusieurs types de montagne : montagne avec la neige, montagne avec les arbres verts, montagne avec la pierre. Donc, il est difficile d'obtenir les images de montagne avec bon résultat.



FIG 26 - Le cluster 2



FIG 27 - Quelques images dans le cluster 2

Pourtant, avec mon test, j'ai reçu un cluster qui représente des montagnes. C'est le cluster 2 (cf. Fig. 26).

On peut facilement trouver que toutes les images de montagne du cluster 2 est les montages avec la neige. Il y a quelques images de montagne qui n'ont pas de neige, mais dans ce cas le nuage et le ciel sont considérés comme la neige (cf. Fig. 27).

## 2.4. Les relations entre des images

On n'intéresse pas aux images qui appartiennent à un seul cluster parce que ces images ne sont pas beaucoup significatifs. Donc, j'ai recherché des images qui appartiennent à plusieurs clusters.

Il y a deux images qui appartiennent à quatre clusters. Ce sont l'image 460 et l'image 124. Deux images appartiennent aux mêmes clusters 3, 5, 8, et 10. Donc, pour chaque image, on a quatre chemins pour les trouver (cf. Fig. 28 et 29).



FIG 28 - Quatre chemins pour trouver l'image 460

On peut facilement trouver que l'image 460 et l'image 124 ont moins de mots visuels. Mais, en effet, les images appartiennent à quatre clusters. Donc, ses mots visuels doivent être divisés en quatre parties et ces mots visuels sont très significatifs que les mots visuels des autres images.



FIG 29 - Quatre chemins pour trouver l'image 124

# CONCLUSION ET PERSPECTIVES

## 1. CONCLUSION

L'approche « mot visuel » est très attachant. La découverte des images qui ont des mêmes mots visuels, peut améliorer le résultat des recherches d'images. Avec ce sujet, j'ai proposé et ai implémenté un modèle de visualisation de graphe d'images qui utilise la combinaison de la mise en page spirale et radiale pour permettre d'exploiter et traverser facilement les graphes d'images. Le modèle a des avantages et des inconvénients suivants :

### **Avantages**

- Le modèle a profité des attributs différents de l'affichage du graphe et de l'algorithme de la mise en page à accélérer la phase de l'initialisation du graphe d'images.
- L'algorithme de la mise en page radiale et spirale est très simple avec la petite complexité.
- Le modèle a besoin de moins changements pour s'adapter à la découverte des relations différentes des bases d'images. Les relations peuvent être des relations entre les mots textuels, la couleur, la texture ou le sac de KVR (Keywor Visual Representation).
- Ne pas visualiser en même temps toutes les images, le modèle a diminué le temps d'attente de l'utilisateur.

### **Inconvénients**

- L'algorithme de la mise en page est très bon, mais le temps d'affichage d'un niveau qui a plusieurs images, sont encore trop lent (9-10 secondes).
- Le modèle ne supporte pas les événements du clique et double-clique parce que la librairie Qt a une faute dans le traitement des événements de la souris.
- La technique zoom et expanse n'est pas efficace. Le modèle ne supporte pas de zoomer plus.
- Le modèle dépend plus des données, des types du graphe.

## 2. PERSPECTIVES

Mes expérimentations montrent qu'on peut bien trouver des relations entre des clusters et entre des images, mais elles montrent aussi qu'il y a plusieurs problèmes qu'il faudra améliorer :

- Tout d'abord, avec mon modèle, il est difficile d'expliquer les relations entre des images selon mots visuels parce que l'on ne peut pas voir les mots visuels. Donc, il faudra extraire les mots visuels qui appartient aux clusters et dessiner directement ces mots visuels sur les images. De plus, on peut permettre de trouver les mots visuels correspondants quand on clique sur un mot visuel.
- La combinaison directement le code de DOAN Tuan Anh et ce modèle permettra à l'utilisateur de changer les seuils de l'algorithme partitionnement et ensuite, voir les résultats sur même l'écran. Mais il faudra améliorer le temps du processus de l'indexation et du partitionnement.

# REFERENCES

## RÉFÉRENCES SCIENTIFIQUES

- Doan, Tuan Anh. 2009. «Moteur de recherche graphique d'image». Rapport du travail pratique encadré en ligne, Hanoï, Institut de la Francophonie pour l'Informatique, 39 p. <[http://www2.ifi.auf.org/rapports/tpe-promo14/tpe-doan\\_tuan\\_anh.pdf](http://www2.ifi.auf.org/rapports/tpe-promo14/tpe-doan_tuan_anh.pdf)>. Consulté le 6 juillet 2009.
- Hascoët, M. and M. Beaudouin-Lafon. 2001. «Visualisation interactive d'information». Revue I3 1(1), p. 77-108.
- Herman, I., Melançon, G. et M. Scott Marshall. 2000. «Graph visualization and Navigation in Information Visualisation». *IEEE Transactions on Visualization and Computer Graphics* 6(1), p. 24–43.
- Heer, Jeffrey et Danah Boyd. 2005. «Vizster : Visualizing Online Social Networks». In Proceedings of IEEE Symposium on Information Visualization 2005, p. 32-39.
- Heer, Jeffrey, Stuart K. Card et James A. Landay. 2005. «Prefuse: a toolkit for interactive information visualization». In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 421-430.
- Hoang, Thanh Lam. 2008. «Indexation et recherche d'images par utilisation de mots visuels». Rapport du travail pratique encadré en ligne, Hanoï, Institut de la Francophonie pour l'Informatique, 38 p. <[http://www2.ifi.auf.org/rapports/tpe-promo13/tpe-hoang\\_thanh\\_lam.pdf](http://www2.ifi.auf.org/rapports/tpe-promo13/tpe-hoang_thanh_lam.pdf)>. Consulté le 1 juillet 2009.
- Jankun-Kelly, T.J et Kwan-Liu Ma. 2003. «MoireGraphs : Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes». In Proceedings of IEEE Symposium on Information Visualization 2003, p. 59-66.
- Novak, Ondrej. 2002. «Postgraduate Study Report : Visualization of Large Graphs». Czech, Technichal University in Prague, 35 p.
- Pavlo, Andrew, Christopher Homan et Jonathan Schull. 2006. «A parent-centered radial layout algorithm for interactive graph visualization and animation». In Cornell University Library. <<http://arxiv.org/abs/cs/0606007>>. Consulté le 1 juin 2006.
- Yee, Ka-Ping, Danyel Fisher, Rachna Dhamija et Marti Hearst. 2001. «Animated Exploration of Dynamic Graphs with Radial Layout». In Proceedings of IEEE Symposium on Information Visualization 2001, p. 43-50.

## RÉFÉRENCES AUTRES

Auber, David, «Tulip framework». <<http://www.tulip-software.org/>>, le 16 décembre 2009.

Fry, Ben et Casey Reas, «Processing». <<http://processing.org/>>, le 31 décembre 2009.

Google, «Google Image Swirl», <<http://image-swirl.googlelabs.com/>>, novembre 2009.

Heer, Jeffrey, «The prefuse visualization Toolkit», <<http://prefuse.org>>, janvier 2009.

Nokia, «Qt framework». <<http://qt.nokia.com>>, le 08 juin 2010.