



Construction d'un graphe de proximité pour l'exploration de larges collections d'images

Frédéric Rayar, Sabine Barrat, Fatma Bouali, Gilles Venturini

► To cite this version:

Frédéric Rayar, Sabine Barrat, Fatma Bouali, Gilles Venturini. Construction d'un graphe de proximité pour l'exploration de larges collections d'images. In the Fourth edition of "Big Data Mining and Visualization" Days. 2015. <hal-01168725>

HAL Id: hal-01168725

<https://hal.archives-ouvertes.fr/hal-01168725>

Submitted on 26 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Construction d'un graphe de proximité pour l'exploration de larges collections d'images

Frédéric Rayar*, Sabine Barrat*
Fatma Bouali*,** Gilles Venturini *

*Université François-Rabelais de Tours, Laboratoire d'Informatique
64 avenue Jean Portalis, 37200 Tours, France,
frederic.rayar@univ-tours.fr, sabine.barrat@univ-tours.fr, gilles.venturini@univ-tours.fr

**Université de Lille2, IUT, Dpt STID
25-27 Rue du Maréchal Foch, 59100 Roubaix, France,
fatma.bouali@univ-lille2.fr

Résumé

Dans ces travaux, nous nous intéressons à la structuration et la visualisation de collections d'images. En effet, durant la dernière décennie, la quantité d'images capturées par tout un chacun a explosé, qu'elles soient générées dans le cadre privé, commercial ou dans celui de projets de numérisation (humanités numériques). De plus, l'avènement d'Internet a accentué le fait que le nombre d'images mises en ligne croît de manière exponentielle, notamment avec les sites d'instituts ou encore les réseaux sociaux.

Afin d'organiser les images au sein d'une structure pertinente, nous avons choisi le *graphe des voisins relatifs* (RNG) défini par Toussaint (1980). Celui-ci fait partie d'une famille de graphe, les *graphes de proximité*, qui permettent d'extraire une structure inhérente aux données étudiées : deux nœuds sont reliés par un arc s'ils sont suffisamment proches pour être considérés comme des voisins. Le RNG ayant la particularité d'être le plus petit graphe de proximité connexe (*i.e.* il existe un chemin entre tout couple de nœud), c'est celui-ci qui a été utilisé dans ces travaux. En effet, cette structure connexe permettra une exploration de la collection par le contenu, en se déplaçant de proches en proches. Par définition, deux nœuds p et q sont des voisins relatifs s'il n'existe aucun autre nœud r plus proche de p et de q qu'ils ne le sont entre eux. Ainsi, on a :

$$E(\text{RNG}) = \{\overline{pq}, p, q \in D, p \neq q, \delta(p, q) \leq \max(\delta(p, r), \delta(q, r)), \forall r \in D \setminus \{p, q\}\},$$

avec D , l'ensemble des nœuds considérés, $E(\text{RNG})$, l'ensemble des arcs du RNG et $\delta : D \times D \rightarrow \mathbf{R}$ une distance. La Figure 1 illustre la notion de voisinage relatif.

L'inconvénient majeur du RNG est la complexité de sa construction, en $O(n^3)$. De fait, une construction classique semble difficilement applicable à de grands jeux de données. Afin de pallier cette limite, nous avons étudié et amélioré un paradigme de construction *incrémentale* du RNG, initialement proposé par Hacid et Zighed (2005). Ce travail fait l'objet d'une

Construction d'un graphe de proximité pour l'exploration de larges collections d'images

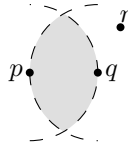


FIG. 1 – *Voisinage relatif de deux points $p, q \in \mathbb{R}^2$ (région grisée).*

soumission à une conférence internationale. Ce paradigme propose, lors de l'insertion d'un nouveau nœud dans un RNG existant, de venir mettre à jour uniquement une région pertinente du graphe, permettant ainsi de ne pas reconsidérer les parties du graphe « éloignées » du nouveau nœud inséré.

Lors de travaux préliminaires (Rayar et al. (2015)), nous avons visualisé les graphes obtenus à l'aide de logiciels existants, nommément Gephi et Tulip. Faisant face à un certain nombre de limitations (mémoire, affichage des images, etc.), nous avons commencé la conception d'une interface de visualisation et d'exploration de graphe d'images. Cette interface, utilisant des technologies web, est en cours de réalisation. La Figure 2 montre un aperçu du prototype actuel.

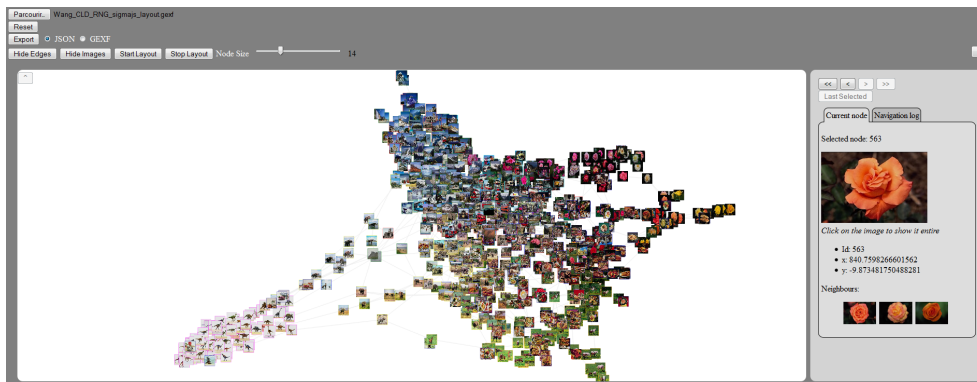


FIG. 2 – *Aperçu de l'interface de navigation de graphes d'images.*

Références

- Hacid, H. et D. A. Zighed (2005). An effective method for locally neighborhood graphs updating. In *DEXA*, pp. 930–939.
- Rayar, F., S. Barrat, F. Bouali, et G. Venturini (2015). Exploration visuelle et interactive d'une large collection d'images en libre accès. In *EGC 2015 - Atelier VIF*.
- Toussaint, G. T. (1980). The relative neighbourhood graph of a finite planar set. *Pattern Recognition* 12, 261–268.