
Vectorisation du modèle d'appariement pour la recherche d'images par le contenu

Hanen Karamti ¹

MIRACL, Université de Sfax
Route de Tunis Km 10 B.P. 242, 3021, Sfax, Tunisie
karamti.hanen@gmail.com

RÉSUMÉ. Le développement rapide des techniques de numérisation et de stockage a engendré une croissance accrue du volume des documents électroniques (textes, images, sons et vidéos). Pour faire face à cette grande masse d'informations, en particulier les images, il est nécessaire de développer des outils qui permettent d'optimiser l'accès à de telles sources de données. Le problème qui se pose est la représentation du contenu d'images, puisque les techniques actuelles ne permettent pas de décrire réellement le contenu. Dans ce papier, nous proposons une approche de recherche d'images par le contenu visuel, qui se base sur l'idée de transformer n'importe quel modèle d'appariement d'images (modèle de calcul de distance entre les vecteurs associés) à un modèle vectoriel fournissant un score. A travers quelques expériences, nous montrons que cette approche entraîne une amélioration importante de la qualité des résultats.

ABSTRACT. The rapid development of digital technologies and storage has led to a large volume of electronic documents (texts, images, sounds, videos). To cope with this information volume, particularly images, it is necessary to develop tools to optimize access to such information resources. The problem is how to represent an image, since current technology can not really describe the content. In this paper, we propose an approach of content based image retrieval (CBIR), which is based on the idea of transforming any matching model (between image vectors) to a vector space model providing a score. Through some experiments, we show that this approach improves the retrieval performance.

MOTS-CLÉS : Recherche d'images, contenu visuel, descripteur, modèle vectoriel, vectorisation.

KEYWORDS: Image retrieval, visual content, descriptor, vector space model, vectorization.

1. Directeur: Faiez Gargouri, Encadrant: Mohamed Tmar

1. Introduction

Malgré les avancées technologiques dans le domaine de la recherche d'images par le contenu, une mauvaise interprétation des contenus, que ce soit de la requête ou bien du document, engendre une estimation approximative du degré de pertinence. Notre travail s'inscrit dans un cadre similaire, soit l'étude et l'intégration d'une nouvelle approche de recherche d'images par le contenu.

Dans cet article, nous proposons une approche dans laquelle le modèle classique d'appariement d'images se transforme en un modèle vectoriel. Cette approche est appelée « la vectorisation ». Le principe de cette technique est de construire une représentation vectorielle des images et de la requête à l'aide du système d'appariement initial. Les composantes de ces vecteurs sont les scores de similarité obtenus, lors d'une recherche initiale, entre un document de la base et un document référence. Une fois que les vecteurs sont construits, la réponse à une requête devient une recherche dans un espace vectoriel classique. Il suffit donc de chercher les documents dont les vecteurs sont proches de celui de la requête. Ainsi, l'appariement dans le nouvel espace vectoriel devient moins complexe que celui dans l'espace original. Les documents de référence sont les centres de gravité des groupes d'images semblables. La suite de cet article est organisée comme suit : dans la section 2, nous présentons les travaux connexes à notre travail ; dans la section 3, nous décrivons notre approche de vectorisation en incluant une fonction de classification ; dans la section 4, nous présentons les résultats que nous avons obtenus sur une collection d'images pour évaluer l'impact de notre approche.

2. Travaux connexes

Ces dernières années, de nombreux systèmes de recherche d'images par le contenu ont vu le jour. La plupart de ces systèmes permettent de naviguer au sein de la base d'images et d'exprimer leurs besoins en information à travers une requête image. Ces systèmes utilisent uniquement des caractéristiques de bas niveau (couleur, texture, forme) appelées aussi des descripteurs (Schettini *et al.*, 2009). Les systèmes appartenant à cette catégorie sont nombreux, comme par exemple le système QBIC (Query By Image Content) du IBM (Flickner, 1997). Le système Frip (Finding Regions in the Pictures) (ByoungChul Ko, 2005) propose de faire des recherches par des régions d'intérêt dessinées par l'utilisateur (Caron *et al.*, 2005). Le système RETIN (REcherche et Traque INteractive) (Fournier *et al.*, 2001) développé à l'université de Cergy-Pontoise du France, sélectionne un ensemble des pixels au hasard dans chaque image pour extraire leurs valeurs de couleur. La texture de ces pixels est obtenue par l'application de la méthode de filtres de Gabor (Rivero-Moreno *et al.*, 2003). Ces valeurs sont regroupées et classées via un réseau de neurones. La comparaison entre les images se fait par un calcul de similarité entre leurs vecteurs de caractéristiques (Tollari, 2006). Certaines études sont proposées pour changer leurs espaces de recherche, par exemple le changement de l'espace de caractéristiques de couleur (Braquelair *et al.*, 1997). Certains travaux ont procédé à la minimisation de l'espace de recherche en

utilisant la technique du calcul des plus proches voisins pour regrouper les données similaires dans des classes (Berrani *et al.*, 2002). Ainsi la recherche d'une image s'effectue en cherchant une classe. L'inconvénient de ces systèmes est que l'utilisateur ne dispose pas toujours d'une image qui exprime son besoin réel, ce qui rend l'utilisation de tel système difficile. Une des solutions à ce problème est la technique de vectorisation, qui permet de retrouver les images pertinentes à une requête et qui ne sont pas retournées par le système initial. Son déroulement nécessite le choix d'un ensemble de documents dits de référence. Ces références sont choisies au hasard par (Claveau *et al.*, 2010), ou bien sont les premiers résultats d'une recherche initiale (Karamti *et al.*, 2012). L'idée de cet article est de proposer une nouvelle méthode permettant de choisir ces références.

3. Recherche d'images avec un modèle d'appariement vectoriel

Notre objectif est de transformer le modèle d'appariement d'images, qui repose sur le calcul de distances entre les vecteurs de caractéristiques, à un modèle vectoriel standard. Dans cette section, nous présentons le principe de la vectorisation. Nous détaillons en particulier la façon dont sont constitués les documents de référence servant de base à la construction de l'espace vectoriel.

3.1. Principe de vectorisation

Cette technique est proposée dans le cadre de la recherche textuelle (Claveau *et al.*, 2010) et elle a montré son intérêt en termes de complexité et de qualité des résultats. Son principe est le suivant : pour chaque document de la collection, on calcule un score de similarité avec m documents de référence. Ainsi, les m scores obtenus forment un vecteur de dimension m et par la suite les n images de la collection seront représentées par les n vecteurs de dimension m . Le même principe est appliqué à la requête, il s'agit de calculer sa similarité avec les m documents de référence pour générer un vecteur de dimension m . La correspondance de la requête avec un document peut donc s'effectuer de manière standard dans ce nouvel espace vectoriel.

3.2. Choix des documents de référence

Le choix de références reste problématique pour la construction de l'espace vectoriel. Dans cette optique, il est nécessaire de sélectionner un ensemble hétérogène de documents de référence. Le nombre de documents de référence est également problématique. En effet, le nombre de documents de référence est la dimension du nouvel espace vectoriel. Si on choisit un petit ensemble de documents de référence, on risque de ne pas pouvoir représenter correctement tous les documents. Par exemple, si on choisit 1 document de référence, 2 documents différents qui en sont équidistants seront confondus (car la distance de chacun au document de référence est la seule co-

Karamti Hanen

ordonnée de celui-ci). Cependant, si on sélectionne trop de documents, on risque de se retrouver avec le problème de l'interdépendance linéaire entre dimensions. Prenant l'algorithme k -means (algorithme 1 où Th est un seuil fixé à 0.5) comme exemple pour regrouper les images homogènes d'une collection de n images.

Algorithme 1 Classification

$\{I_1, I_2, \dots, I_n\}$: la collection d'images

Chaque $C \leftarrow \emptyset$

pour $i \in \{1, 2, \dots, n\}$ **faire**

$C_i \leftarrow \{I_i\}$

$C \leftarrow C \cup \{I_i\}$

fin pour

tant que $\frac{Inertie(C)}{n-|C|+1} < Th$ **faire**

$(C_k, C_l) = \arg \min d(C_x, C_y)$

$(C_x, C_y) \in C^2$

$C_x \neq C_y$

$C \leftarrow C - \{C_k, C_l\}$

$C_k \leftarrow C_k \cup C_l$

$C \leftarrow C \cup \{C_k\}$

fin tant que

Cette méthode représente chaque classe C_i par son centroïde qui est la moyenne des vecteurs caractéristiques des images appartenant à cette classe. L'objectif de cette méthode est de minimiser la somme de l'inertie intra-classe sur l'ensemble des classes ($Inertie(C)$). L'inertie intra-classe d'une classe C_i notée $Inertie(C_i)$ est la moyenne des distances des vecteurs de la classe au centre de gravité g_i de celle-ci :

$$Inertie(C) = \frac{1}{|C|} \sum_{C_i \in C} Inertie(C_i) \quad [1]$$

$$Inertie(C_i) = \frac{1}{|C_i|} \sum_{I_j \in C_i} d(g_i, I_j) \quad [2]$$

$$g_i = \frac{1}{|C_i|} \sum_{I_j \in C_i} I_j \quad [3]$$

Pour mesurer la proximité entre un centroïde et un vecteur, on calculera une distance entre ces deux vecteurs. On pourra utiliser, la distance euclidienne, calculée comme suit, avec p est la dimension du vecteur de caractéristiques :

$$d(g_i, I_j) = \sqrt{\sum_{k=1}^p (g_{i_k}, I_{j_k})^2} \quad [4]$$

Les références sont les centres de gravité g_i des classes obtenues. Tout calcul de similarité entre deux vecteurs de scores v_1 et v_2 est effectué par la fonction Cosinus :

$$\cos(v_1, v_2) = \frac{\sum_{i=1}^n v_{1_i} \times v_{2_i}}{\sqrt{\sum_{i=1}^n v_{1_i}^2} \times \sqrt{\sum_{i=1}^n v_{2_i}^2}} \quad [5]$$

4. Expérimentations

4.1. Données expérimentales

Pour apprécier les différents effets de la vectorisation, nous utilisons une collection composée de 1000 images issue de la bibliothèque de Wang ¹ et indexée par deux descripteurs. Un descripteur couleur appelé Color Layout Descriptor et un descripteur forme nommé Edge Histogram Descriptor.

4.2. Résultats

Les performances sont mesurées en terme précision moyenne (MAP). Le test est effectué à travers 7 requêtes. Le tableau 1 présente les résultats obtenus d'une recherche effectuée par 4 systèmes de recherche d'images différents.

Requêtes/MAP	Système 1	Système 2	Système 3	Système 4
Req1	0.33	0.30	0.43	0.60
Req2	0.20	0.22	0.34	0.36
Req3	0.10	0.08	0.17	0.19
Req4	0.91	0.30	0.94	0.96
Req5	0.48	0.35	0.52	0.56
Req6	0.09	0.04	0.09	0.11
Req7	0.19	0.11	0.17	0.23

Tableau 1. Performances obtenues par différents systèmes de recherche d'images

Le système 1 correspond à une recherche euclidienne entre les vecteurs de caractéristiques des requêtes et des images. Le Système 2 représente une recherche basée sur la technique classique de vectorisation (Claveau *et al.*, 2010) qui sélectionne au hasard les documents de référence. Le Système 3 correspond à une recherche intégrant la technique de vectorisation proposée par (Karamti *et al.*, 2012) qui sélectionne à partir de la recherche initiale les 40 premiers documents jugés pertinents. Enfin, le Système 4 intègre notre nouvelle approche de vectorisation.

1. <http://wang.ist.psu.edu/docs/related/>

On constate un gain important de notre système par rapport aux systèmes existants. La comparaison avec notre approche est particulièrement intéressante puisqu'elle met bien en valeur la qualité des références choisies. Le regroupement des images homogènes dans la collection a donné naissance aux nouveaux vecteurs de référence qui assurent une bonne adéquation du contenu représenté entre les images et les requêtes à traiter. Par contre un choix au hasard des images de référence peut conduire à une baisse des résultats par rapport à celles de la recherche initiale.

5. Conclusion

La technique de vectorisation présentée dans cet article offre un grand intérêt dans la recherche d'images par le contenu. Cette technique permet de changer n'importe quel modèle d'appariement d'images à un modèle vectoriel. Nous avons montré que la méthode k -means de classification peut intervenir dans le choix de documents de référence pour la construction du nouvel espace vectoriel. Les perspectives de notre travail consisteront à comparer différentes méthodes de classification non supervisées pour le choix des références sur de très grosses collections.

6. Bibliographie

- Berrani S.-A., Amsaleg L., Gros P., « Recherche par similarité dans les bases de données multidimensionnelles : panorama des techniques d'indexation », *RSTI*, p. 9-44, 2002.
- Braquelair J. P., Brun L., « Comparison and Optimization of Methods of Color Image Quantization », *IEEE Transactions on Image Processing*, vol. 6, n° 7, p. 1048-1052, july, 1997.
- ByoungChul Ko H. B., « FRIP : a region-based image retrieval tool using automatic image segmentation and stepwise Boolean AND matching », *Trans. Multi*, p. 105-113, 2005.
- Caron Y., Makris P., Vincent N., « Caractérisation d'une région d'intérêt dans les images », *Extraction et gestion des connaissances (EGC 2005)*, p. 451-462, 2005.
- Claveau V., Tavenard R., Amsaleg L., « Vectorisation des processus d'appariement document-requête. », *CORIA*, Centre de Publication Universitaire, p. 313-324, 2010.
- Flickner M., « Query by Image and Video Content : The QBIC System », *Intelligent Multimedia Information Retrieval*, American Association for Artificial Intelligence, p. 7-22, 1997.
- Fournier J., Cord M., Philipp-Foliguet S., « RETIN : A content-based image indexing and retrieval system », 2001.
- Karamti H., Tmar M., Anis B., « A New Relevance Feedback Approach for Multimedia Retrieval. », *IKE, July 16-19, Las Vegas Nevada, USA*, p. 129, 2012.
- Rivero-Moreno C., Bres S., « Les filtres de Hermite et de Gabor donnent-ils des modèles équivalents du système visuel humain ? », *ORASIS*, p. 423-432, 2003.
- Schettini R., Ciocca G., Gagliardi I., « Feature Extraction for Content-Based Image Retrieval », *Encyclopedia of Database Systems*, p. 1115-1119, 2009.
- Tollari S., Indexation et recherche d'images par fusion d'informations textuelles et visuelles, 2006.