

Méthode globale d'Indexation d'images par statistiques de motifs

P. MAKRIS, E. RIOUT, N. VINCENT

Laboratoire d'Informatique, Université de Tours, 64, avenue Jean Portalis, 37200 Tours

{makris, vincent}@univ-tours.fr

Résumé – Dans cet article nous abordons le problème de l'indexation et celui de la recherche d'images par le contenu. Nous ne proposons pas un système complet d'indexation mais présentons quelques indices assez discriminants qui peuvent être utilisés pour la représentation des images. Ces paramètres sont issus de l'application de la loi de Zipf à l'analyse des images. La méthode permet de mettre en évidence un type de structuration globale de l'image par la modélisation d'un phénomène reposant sur une loi puissance. Les paramètres utilisés sont les paramètres d'exposant des lois puissances que l'on peut extraire de l'image. Nous montrons les résultats très satisfaisants obtenus à partir d'une représentation des images dans un espace de dimension 6 et sur une base conséquente d'images.

Abstract – In this paper we tackle the problems of image indexing and content based image retrieval. Here, we do not propose a complete indexing system but we present several indexes that can discriminate in a powerful way between some images. They are used to represent images in a 6D representation. These parameters are defined from the application of Zipf law to image analysis. The method allows highlighting some global structures of the image. A model of a phenomenon is achieved by use of a power law. The parameters used are the power laws exponents that have been computed from the image. We show the very satisfactory results obtained from the image representation in a 6D representation and with large enough image bases.

1. Introduction

La recherche d'images dans des ensembles à grande échelle est devenue parfois une nécessité. Elle est souvent précédée par une phase d'indexation. Pour répondre à cette nécessité de nombreuses méthodes d'indexation et de recherche d'images par le contenu ont été développées et peuvent être classées. On peut les regrouper selon qu'elles reposent sur des histogrammes [1], sur des représentations à l'aide de paramètres de texture ou des moments [2], sur des primitives locales [3] ou sur des méthodes plus complexes en considérant les objets que l'on peut extraire de l'image et les relations spatiales qui existent entre eux [4]. Nous proposons ici une méthode qui peut être classée parmi celles utilisant un histogramme mais aussi une analyse de texture ou plus précisément une mesure de la structure interne des images ou de zones de l'image. Ici, il ne s'agit pas pour nous de présenter une méthode parachevée, mais de montrer qu'une nouvelle famille de paramètres peut apporter une information très efficace pour atteindre les objectifs d'indexation et de recherche d'images par le contenu que nous nous sommes fixés.

Après avoir rappelé l'énoncé de la loi de Zipf définie dans le domaine de signaux monodimensionnels, nous présenterons comment cette loi peut être utilisée dans le cas d'un signal bidimensionnel. Nous verrons comment elle permet dans le cas des images de définir des paramètres globaux qui traduisent la structure interne des images. Dans la dernière partie nous présenterons des résultats obtenus sur une base d'images.

2. La loi de Zipf

2.1 Son énoncé

La loi de Zipf est une loi empirique [5] qui a été observée et vérifiée dans l'étude de nombreux signaux monodimensionnels codés en utilisant un nombre limité de symboles. Elle peut s'énoncer au travers une relation qui repose sur une loi puissance. Elle fait intervenir la fréquence des motifs présents dans le signal et leur rang quand les fréquences d'apparition des motifs sont classées dans l'ordre décroissant. En notant $N_{\sigma(i)}$ la fréquence absolue du motif classé en position i selon l'ordre des fréquences, on peut exprimer cette relation par :

$$N_{\sigma(i)} = k \cdot i^a \quad (1)$$

On voit qu'interviennent deux constantes k et a qui caractérisent la loi. La relation devient linéaire quand on considère les logarithmes des deux membres. La constante k est liée à la taille du signal qui est étudié. Par contre, il est possible d'extraire une information sur la structure du signal. Cette information peut être résumée par l'exposant a de la loi puissance. D'un point de vue pratique, cet exposant peut être estimé par le calcul de la pente de la droite de régression du nuage de points expérimentaux associés aux motifs, les abscisses indiquent le rang du motif et les ordonnées, la fréquence absolue. Un exemple de tel graphe peut être observé sur la figure 1.

2.2 Son adaptation aux images 2D

Dans une application du principe de l'étude de la fréquence des motifs en fonction du rang, au traitement des images, le principe de la loi de Zipf doit être adapté car la notion de motif n'est plus aussi naturelle. Les motifs ne sont plus simplement des suites de symboles successifs. Il est alors nécessaire de respecter les propriétés topologiques de l'image. Nous avons choisi la forme des motifs non pas comme une suite de pixels adjacents mais comme un ensemble de pixels qui forment un voisinage d'un pixel central. La taille du voisinage peut être variée comme on peut faire varier la taille des mots dans le cas des signaux 1D, nous avons choisi des motifs de dimension 3x3.

Il n'est pas possible d'utiliser comme symboles décrivant le signal les données brutes que sont les niveaux de gris. En effet, le cardinal des motifs 3x3 différents que permettent de coder les niveaux de gris serait beaucoup trop important en comparaison du nombre des motifs que l'on peut effectivement observer dans une image de taille raisonnable. Les fréquences relatives seraient toutes pratiquement nulles. Il faut donc réaliser, quand la taille du motif a été choisie, un compromis entre un nombre limité de symboles et une information assez précise sur le signal. Une fréquence non nulle, apporte une information conséquente si le nombre des motifs disponibles est important.

Nous allons, dans le paragraphe suivant, présenter deux codages d'image que nous avons expérimentés. Nous appliquerons sur cette base de codage le principe de la loi de Zipf.

3. Le codage des images

Les deux codages que nous allons décrire reposent d'une part sur des notions différentes du voisinage d'un pixel et d'autre part sur l'utilisation des niveaux de gris dont le nombre est réduit de manières différentes.

3.1 Codage des niveaux de gris

Dans le premier codage nous considérons un voisinage classique carré de taille 3x3 qui est lié à la 8 connexité dans le plan. Il comporte neuf pixels. Il y aurait, en utilisant les 256 niveaux de gris $2^{8 \times 9}$ motifs différents possibles, ce nombre est très supérieur au nombre de pixels d'une image de taille quelconque. Nous diminuons donc le nombre de motifs en réduisant le nombre de niveaux de gris utilisés pour décrire l'image. Nous avons choisi 9 niveaux de gris différents. Nous voulons néanmoins conserver des propriétés qui permettent de caractériser le contenu des images et obtenir un codage qui soit robuste en particulier aux changements d'illumination de l'image. Pour cela nous avons cherché à nous écarter d'une description absolue du contenu de l'image et à rendre la définition des symboles dépendant de l'image elle-même. Dans ce but, un algorithme de classification est utilisé pour définir les 9 niveaux de gris que nous conservons dans la définition. Dans l'application décrite, c'est l'algorithme des k-means qui est appliqué à l'ensemble des niveaux de gris des pixels de l'image. Il pourrait aussi être appliqué de manière similaire à des images

couleur. Nous considérons alors 9 classes. Ainsi, on compte 9^9 motifs différents possibles. C'est un alphabet de grande taille.

3.2 Voisinage 4-connexe

Dans le deuxième codage nous considérons la notion de voisinage d'un pixel au sens de la 4-connexité. Le motif considéré ne contient plus alors que 5 pixels. Dans ce deuxième cas, nous diminuons encore davantage le nombre de motifs différents possibles en réduisant aussi le nombre des niveaux de gris utilisés dans le codage. On ne conserve qu'une description globale de l'image grâce à 3 niveaux de gris. Comme dans le cas précédent, ces niveaux sont obtenus par un algorithme des k-means mais avec 3 classes. Nous avons ici seulement 243 motifs différents possibles dans toute image.

3.3 Définition des paramètres d'indexation

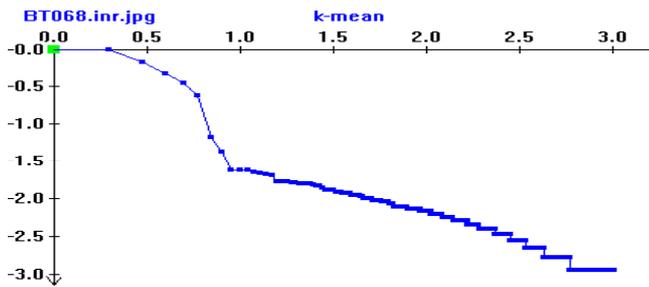
Sur toute l'image, et en utilisant chacun des deux codages, le nombre d'occurrences de chaque motif est calculé et nous pouvons tracer les graphes

$$(\log(N_{\sigma(i)}), \log(i)) \quad (2)$$

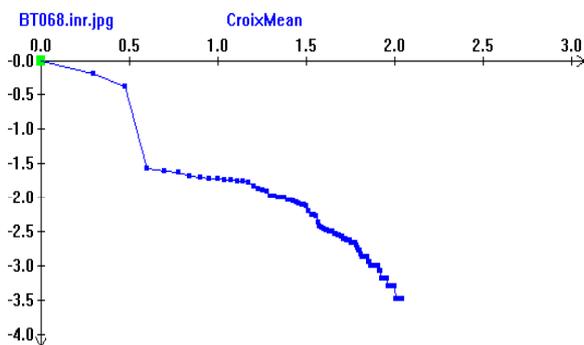
où i varie. Pour une image exemple présentée figure 1 en (a), ces deux graphes ont été tracés figure 1 (b) et 1 (c). Contrairement à ce que l'on aurait pu attendre de l'application de la loi de Zipf, on peut constater que les courbes qui apparaissent ne peuvent être raisonnablement approximées par une droite. C'est-à-dire que la loi sous-jacente à l'image ne peut être représentée par un exposant unique. Par contre, on peut constater que localement la courbe, en plusieurs zones, peut être considérée comme linéaire. De manière à représenter tous les types d'images nous avons opté pour une segmentation de la courbe de Zipf en 3 zones linéaires. Une méthode de segmentation récursive à partir des points de la courbe est utilisée pour l'estimation des trois zones linéaires ainsi que pour celle de leurs pentes respectives. Ces zones correspondent à des structurations de l'image à différents niveaux. Par exemple on peut voir apparaître la structuration des zones contenues dans l'image ou bien celle des contours contenus dans l'image. Dans les deux cas ce ne sont pas les mêmes types de motifs qui interviennent. Ainsi, l'image peut être représentée dans un espace de dimension 6. C'est dans cet espace que la recherche pourra être effectuée.



(a)



(b)



(c)

FIG. 1 : En (a) Image Initiale et en (b) et (c) les graphes de Zipf associés relatifs, en (b) à un codage sur des voisinages 8-connexes 3x3 et en (c) à un codage sur des voisinages 4-connexes

4. La recherche par le contenu

La valeur d'une méthode d'indexation ne peut se mesurer que par la facilité qu'elle fournit de retrouver une image dans une base d'images ou de trouver une image voisine d'une image que l'on peut considérer comme image requête et pour laquelle on cherche à trouver les images ayant même contenu dans la base. Les méthodes les plus efficaces semblent nécessiter un début de description ou d'interprétation des images. Les indices que nous proposons pour représenter l'image sont extraits d'une courbe qui pourrait être comparée à un histogramme mais ils intègrent des connaissances locales pour chaque pixel et une estimation de la structuration globale de l'image.

Nous considérons ici que la base d'images dans laquelle la recherche doit être effectuée est basée sur une image requête. La base est connue a priori et a pu être traitée, c'est-à-dire que chaque image de la base est représentée par un vecteur à 6 composantes, les pentes des 6 zones extraites sur les courbes de Zipf décrites précédemment. Il s'agit alors de comparer une image requête et l'ensemble des images de la base pour en mesurer une proximité.

4.1 Le choix d'une mesure

Tout d'abord, l'image requête doit être traitée de la même manière que les images de la base. C'est-à-dire que l'on calcule les 6 composantes du vecteur des pentes. Les comparaisons sont alors effectuées entre les vecteurs de représentation. Nous avons testé deux distances, la distance infinie et la distance associée à la norme 1 pour construire une mesure. De manière à normaliser chaque composante des vecteurs la différence entre l'image requête et les images de la base est normalisée par la composante correspondante de l'image requête. On note Q l'image requête et P une image quelconque de la base. Les mesures utilisées sont alors données par les formules (2) et (3). La formule (2) fournit un critère de proximité beaucoup plus exigeant que la formule (3). Dans le premier cas on sélectionne les images dont les structures extraites sont toutes semblables à celles de l'image requête alors que dans le second cas, c'est la moyenne des différences relatives qui est considérée.

$$m_{\infty}(Q, P) = \max_{i=1 \text{ à } 9} \left| \frac{Q_i - P_i}{Q_i} \right| \quad (3)$$

$$m_1(Q, P) = \frac{\sum_{i=1}^9 |Q_i - P_i|}{\sum_{i=1}^9 |Q_i|} \quad (4)$$

4.2 Les résultats

Nous présentons ici quelques résultats. Après avoir choisi une image requête et un seuil correspondant à la ressemblance cherchée, nous retenons les images de la base dont la distance à l'image requête est inférieure à ce seuil. Nous présentons figure 2 tout d'abord une image requête, puis les résultats de la requête en utilisant la mesure m_{∞} figurant en (3). L'image requête est présentée en haut de la figure 2, vient ensuite la série des réponses. La base utilisée ici est une base de plus de 3000 images contenant en particulier les images de la base columbia, mais aussi des images très variées de paysages, de visages comme celui de la figure 1, des images de textures, de fleurs. Les résultats montrent que les images jugées les plus semblables à l'image requête sont des images du canard majoritairement vu de dos. On peut comparer les résultats obtenus en utilisant les deux distances. Sur la figure 3 on présente les 24 images les plus semblables. Elles sont ici encore toutes des images du canard. La première intruse apparaît au rang 26, c'est une image de texture. L'exemple que nous avons choisi de faire figurer ici est représentatif des résultats obtenus sur de nombreux essais. On retrouve évidemment que la deuxième mesure renvoie davantage d'images que la première ce qui contribue parmi elles, à introduire plusieurs mauvaises réponses en ce sens que nous ne les attendrions pas.

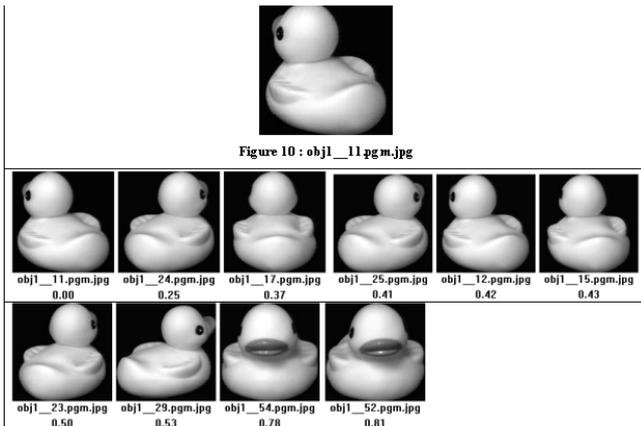


FIG. 2 : en haut l'image requête et en bas les images obtenues par la distance du max

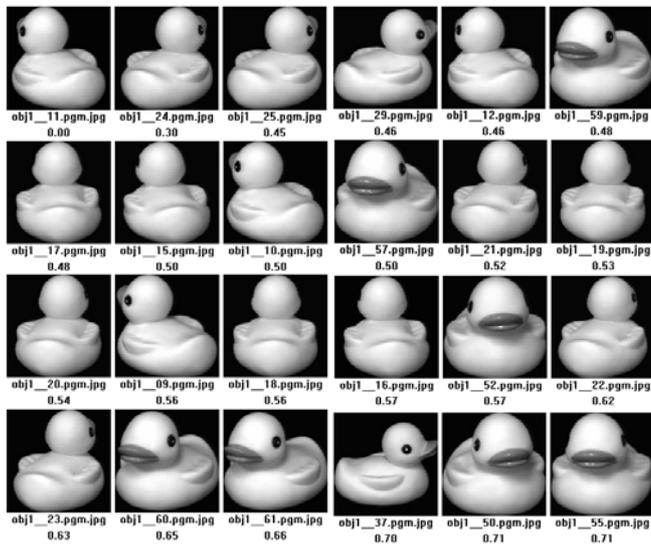


FIG. 3 : les images obtenues par la distance 1 pour la même requête (fig. 2)

5. Conclusion

Nous avons montré ici que la loi de Zipf, même si elle n'est pas suivie par toutes les images, peut néanmoins être utilisée en traitement d'image. Elle permet de mettre en évidence différentes structures internes qui caractérisent bien les images en fonction du codage choisi. Les paramètres extraits des courbes de Zipf peuvent être efficaces dans un problème de recherche d'images par le contenu. Il serait très intéressant d'utiliser ces paramètres en les associant à des paramètres plus classiques pour bénéficier de leurs apports complémentaires.

Références

[1] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, W. Niblack. Efficient color histogram indexing for quadratic form distance functions. *IEEE Transactions*

on Pattern Analysis and Machine Intelligent, n° 17, pp. 729-736.

- [2] J. Z. Wang, G. Wiederhold, O. Firschein, S. X. Wei. Content-based image indexing and searching using daubechies'wavelets. *International Journal on Digital Libraries*, 1(4), 1997.
- [3] A. Guérin-Dugué, A. Oliva, Classification of scene photographs from local orientation features. *Pattern Recognition Letters*, Vol 21, pp. 1135-1140, 2000.
- [4] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, J. Malik, J. Blobworld. A system for region-based image indexing and retrieval. 3rd International Conference on Visual Information Systems, Amsterdam, Netherland, pp. 509-516, 1999.
- [5] G. K. Zipf. Human Behavior and the principle of "Least Effort". Addison-Wesley, New York, USA, 1949.
- [6] A. Cohen, R. N. Mantegna, S. Halvin. Numerical analysis of word frequencies in artificial and natural language texts. *Fractals*, World Scientific Publishing Company, vol. 5 n°1, pp. 95-104, 1997.
- [7] P. Makris, N. Vincent. Zipf Law: A tool for Image Characterization. IAPR Vision Interface (VI'2000), Montréal (Canada), pp.262-268, 2000.
- [8] Y. Caron, P. Makris, N. Vincent. A method for detecting artificial objects in natural environments. In *Proceedings of International Conference on Pattern recognition (ICPR)*, Québec (Canada), 2002.