

Etude statistique des motifs présents dans une image

Pascal MAKRIS¹, Jean Pierre BONNEFOY² et Nicole VINCENT¹

¹ Laboratoire d'Informatique
Université de Tours
64, avenue Jean Portalis, 37200 Tours, France

² Laboratoire d'Informatique et d'Imagerie Industrielle
Université de La Rochelle
Pole Sciences et Technologie, Avenue Marillac, 17042 La Rochelle cedex 1 France

(makris@univ-tours.fr, jpbonnef@gi.univ-lr.fr, vincent@univ-tours.fr)

Résumé - L'information contenue dans une image exerce d'abord sur le spectateur une impression générale globale. Cette impression est le résultat d'une accumulation d'information contenue dans des zones locales. C'est la variété et la répartition des motifs apparaissant dans les zones locales qui constitue la base de notre étude. Elle permet d'en déduire la complexité de la structuration d'une image. Nous montrons que le caractère "naturel" d'une image se traduit par une complexité de la répartition des motifs, plus grande que dans une image géométrique. C'est au moyen du théorème fondamental de Zipf que nous abordons ce problème de classification des images.

Abstract - The information contained in an image has, at first, a global and general impact on the observer. The global feeling is the result of an accumulation of information contained in local areas. The variety and the distribution of the patterns that appear in the local areas are used as the basic tools of our study. Then, it is possible to deduce the complexity of the image structure. We show that the natural aspect of an image implies many different patterns to appear in the image and gives way to some complexity in the distribution of pattern histogram, a complexity, greater than in geometrical images. Zipf theorem is the fundamental base used to come to the problem of image classification.

Introduction

A l'heure où, dans les documents, sont associés de plus en plus au texte, ou aux hypertextes, des images fixes ou vidéo, on est souvent amené à choisir entre photographies et images de synthèse dont la qualité, c'est à dire la ressemblance avec la réalité, est de plus en plus frappante. La difficulté du choix est encore accrue lorsque les images apparaissent en sous impression dans différents documents, tels que des chèques postaux ou bancaires. Certaines problématiques nécessitent de pouvoir distinguer les deux styles d'images. C'est en particulier le cas lorsque l'on souhaite distinguer une image contenant un paysage d'une image contenant du texte [1] ou encore lorsqu'il est nécessaire de recourir à la compression d'images et donc d'adopter une méthode efficace en fonction du contenu de l'image.

C'est pour résoudre les problèmes de ce genre que nous avons étudié les relations qui peuvent exister entre les motifs présents dans une image.

Dans une première partie nous exposons le principe de la méthode qui prend appui sur le théorème de Zipf. Dans la seconde partie, nous expliquons comment cette méthode peut être appliquée à l'analyse des images. Nous terminons en donnant quelques uns des résultats significatifs obtenus, notamment en abordant la distinction entre ce que nous appellerons abusivement image naturelle et image artificielle.

1. Le principe de la méthode

Cette étude repose sur l'application d'une loi empirique qui a été énoncée il y a 50 ans : la loi de Zipf [2] et qui a depuis fait l'objet de nombreuses vérifications dans différents domaines [3]. Elle peut s'énoncer ainsi : Dans un ensemble de symboles organisés topologiquement, les n-uplets de symboles ne s'organisent pas de manière aléatoire. On peut constater que les fréquences N_1, \dots, N_n d'apparition des n-uplets présents M_1, \dots, M_n sont en relations avec ces motifs. Plus précisément, si l'on classe les motifs suivant l'ordre décroissant des fréquences, la suite $(N_{\sigma(1)}, \dots, N_{\sigma(n)})$ avec $i = 1$ à n , vérifie la formule fondamentale :

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

Cette loi puissance, si elle peut être mise en évidence, est caractérisée principalement par la valeur "a" de la puissance. Cette valeur est évidemment négative. Le moyen le plus aisé d'estimer cette valeur "a" consiste à étudier le lien qui existe entre les logarithmes respectifs de $N_{\sigma(i)}$ et de i . Dans les cas où la loi est vérifiée, les deux grandeurs sont liées par une relation linéaire et la valeur de "a" peut être estimée par le coefficient directeur de la droite de régression approximant, au sens des

moindres carrés, les couples $[\ln i, \ln (N_{\sigma(i)})]$. Pratiquement, il est facile de constater, sur un graphe où l'abscisse indique $\ln i$ et l'ordonnée $\ln(N_{\sigma(i)})$, les points former un alignement. La pente de la droite correspond alors à la valeur de a .

Une classification des phénomènes étudiés peut alors utiliser, la valeur de l'exposant "a", mais aussi la qualité de l'alignement des points, c'est à dire l'adéquation de la loi de Zipf au phénomène étudié. Pour notre problématique, nous introduirons d'autres paramètres.

C'est dans le domaine des signaux mono-dimensionnels que cette loi a fait le plus souvent l'objet d'investigations. Notre intention est maintenant de l'appliquer à l'analyse des images. Cela nécessite évidemment quelques adaptations et surtout des choix que nous allons exposer.

2. L'adaptation à l'analyse d'image

Il est évident que l'adaptation posera des problèmes plus ou moins délicats selon qu'elle s'appliquera à une image plus ou moins structurée. L'interprétation sera plus aisée lorsque des rapports invariants existeront entre différents éléments de l'image tandis que dans le cas contraire des difficultés plus nombreuses se présenteront.

Une image est un ensemble de pixels organisés dans le plan sous la forme d'une matrice. Les symboles considérés pour la vérification de la loi de Zipf sont donc ici les niveaux de gris utilisés pour coder les pixels (2 pour une image binaire et 256 pour la plupart des images en niveaux de gris). Les n-uplets dont il a été question dans la partie précédente sont choisis comme la suite des niveaux des pixels appartenant à des masques susceptibles de prendre des formes variées, par exemple des masques carrés 3x3 ou des masques linéaires verticaux 3x1, 7x1 ou horizontaux, 1x3, 1x7. Ce choix est déterminé par le type de motif que l'on souhaite rechercher : motif linéaire ou motif surfacique, la forme du masque traduit le type de vision que l'on veut modéliser. Dans le cas par exemple d'un matériau composite à fibres linéaires, on choisira un masque linéaire.

On peut aussi considérer ce masque comme une modélisation de la zone locale que l'il intègre lors de l'observation globale de l'image. La forme de la zone est influencée par les caractéristiques générales. La structure d'un texte induit une lecture d'une ligne entière de la gauche vers la droite. L'utilisation successive de plusieurs formes de masques permettra d'obtenir des informations complémentaires.

L'application directe de la méthode qui vient d'être exposée conduit à calculer les couples ordonnés suivant i ($\ln i, \ln N_{\sigma(i)}$). Ne sont évidemment pris en compte que les motifs présents dans l'image et apparaissant effectivement plus d'une fois. En représentant ces données dans le plan, elles constituent un nuage de points. Celui-ci, l'expérience le montre, n'a pas une allure aussi linéaire que celle à laquelle on pouvait s'attendre. On peut le voir sur l'exemple de la FIG. 2 sur laquelle, il semble possible d'identifier plusieurs segments.

Pour obtenir concrètement les points du graphe, il faut, pour chaque motif possible, compter le nombre d'occurrences qui lui correspond dans l'image. L'examen de l'image se fait en considérant des masques recouvrants, ce qui assure aux résultats une indépendance aux translations de faible amplitude. Les nombres de motifs peuvent être très variés suivant les masques

utilisés. Le nombre de motifs différents possibles dans un masque 3x3, dans le cas de 256 niveaux de gris, est de 256^9 , soit de l'ordre de 10^{21} . En présence de ce nombre il n'est pas envisageable de les considérer tous. En revanche, le nombre de motifs distincts ou confondus que peut contenir une image 512x512 n'est que de l'ordre de 10^6 . La solution consiste dans ce cas à ne stocker que les motifs rencontrés au cours du balayage de l'image ce qui permet au programme informatique développé de satisfaire à deux contraintes fondamentales: optimiser l'espace mémoire et réduire le temps de traitement.

De plus, on peut estimer que, même dans cette dernière approche, les motifs différents restent encore bien trop nombreux et ne correspondent pas nécessairement à des perceptions visuelles très différentes, l'oeil étant en effet naturellement plus sensible aux différences relatives qu'aux valeurs absolues. L'impression visuelle laissée par deux images aux contrastes peu différents restera sensiblement la même, alors que les motifs communs aux deux images peuvent être totalement disjoints. Aussi on pourra se contenter pour coder un motif d'utiliser seulement le rang relatif au niveau de gris à l'intérieur du masque. Le nombre (4.10^5) de motifs distincts dans un masque 3x3 est alors très inférieur aux 10^{21} possibilités précédentes. Cette réduction du nombre de motifs permet néanmoins de rendre compte de la diversité des impressions données par une zone locale. L'exemple suivant,

TAB.1. montre un masque 3x3, contenant un motif particulier. Son codage par la méthode dite des rangs distincts est présenté TAB.2.

TAB.1 : motif basé sur les niveaux de gris.

255	255	240
255	0	18
20	0	20

TAB.2 : motif basé sur les rangs distincts.

6	7	5
8	0	2
3	1	4

Néanmoins cette méthode, si elle est avantageuse par son moins grand nombre de motifs possibles, n'est pas, par contre, totalement satisfaisante car, en cas d'égalité des niveaux de gris de plusieurs pixels du masque, elle affecte aux pixels un code quelque peu arbitraire. Il est donc préférable, de manière à limiter la perte et la déformation de l'information, de prendre en compte l'égalité des niveaux de gris. Dans ce cas, le codage précédent du contenu du masque se trouve modifié. Le nouveau codage est indiqué dans le tableau TAB. 4.

TAB. 3 : motif basé sur les niveaux de gris.

255	255	240
255	0	18
20	0	20

TAB. 4 : motif basé sur les rangs généraux.

4	4	3
4	0	1
2	0	2

Cette méthode, dite des rangs généraux [4], permet de garder le profil d'origine du contenu du masque. Elle comporte évidemment un nombre de motifs supérieur au nombre des motifs obtenus par la méthode des rangs distincts mais qui demeure néanmoins en même temps très inférieur au nombre des motifs bruts. Par exemple pour le masque 3x3 on a environ 7.10^6 motifs distincts.

C'est ce codage qui a été retenu pour les images binaires, aussi bien que pour les images en niveaux de gris.

Voici maintenant les résultats que cette technique a permis d'obtenir sur des images variées. Dans les deux exemples proposés nous avons choisi d'utiliser des masques 3x3, car nous n'avons fait aucune hypothèse a priori, sur les images étudiées.

3. Résultats

Cette méthode a été expérimentée sur de nombreuses images aux caractéristiques très variées. L'analyse des résultats issus de ce traitement apporte de multiples renseignements concernant la nature du contenu des images traitées. L'information sous-jacente est exprimée par une courbe traduisant la fréquence d'apparition des principaux motifs qui constituent l'image. Les deux exemples caractéristiques d'images fort différentes qui sont montrées ci-après correspondent à des catégories très générales d'images. La première, que nous pouvons qualifier de naturelle est la photographie de la structure d'un muscle. La deuxième, en revanche est un dessin numérique aux structures bien géométriques. C'est une image artificielle, uvre dessinée par l'homme. L'image est ici très texturée et on y a introduit du bruit pour donner l'impression du naturel. Après traitement mettant en jeu les techniques décrites précédemment, on constate que les courbes de fréquence d'apparition des motifs, obtenues présentent des allures différentes. Les différences sont liées

- soit à la structure de l'image traitée
- soit à la plus ou moins grande hiérarchie de l'information et à la plus ou moins grande richesse et diversité de celle-ci.

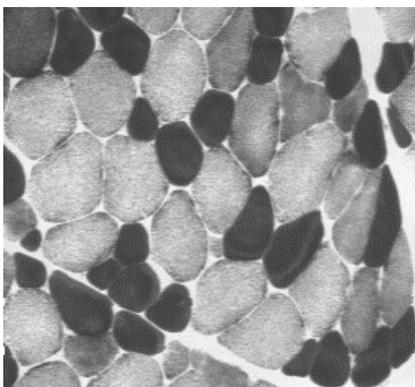


FIG. 1 : image d'un muscle

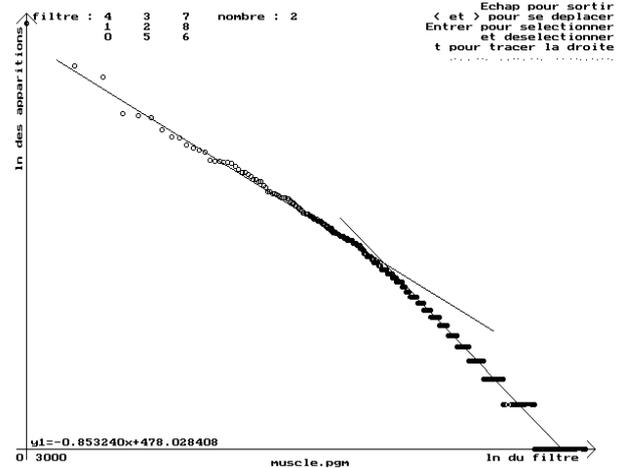


FIG. 2 : courbe de fréquence d'apparition des motifs

L'application du traitement au moyen d'un masque de taille 3x3 et la méthode des rangs généraux, sur le type d'image du muscle qui appartient à la catégorie des images dites "naturelles", fait apparaître des informations très intéressantes.

La courbe de la FIG. 2 met en évidence l'alignement de l'ensemble des motifs présents dans l'image.

La pente de la droite de régression linéaire est assez faible (on l'a estimée à -0,98), et on peut voir qu'un nombre relativement important (3000) de motifs différents sont présents dans l'image.

On peut noter que sur des images présentant un fond uniforme sur lequel se détache une forme, un motif dominant est toujours présent. Il modifie l'allure générale de la courbe de fréquence d'apparition des motifs. Ce qui implique que le premier point ne doit pas être pris en compte dans l'évaluation de la linéarité globale des points du graphe, ni pour le calcul de la pente de la droite.

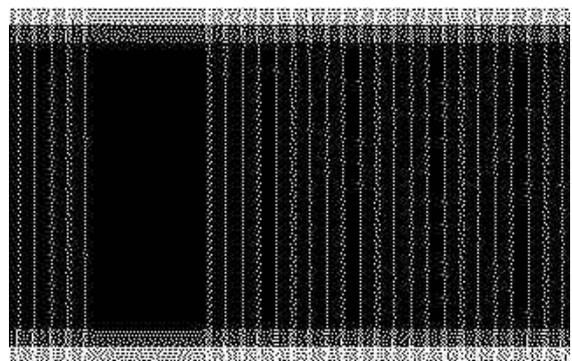


FIG. 3 : image artificielle

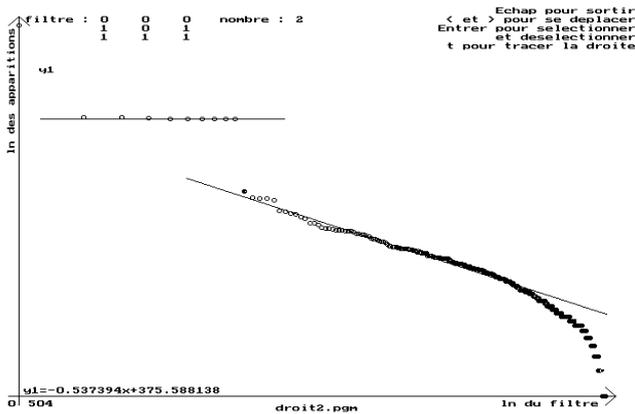


FIG. 4 : courbe de fréquence d'apparition des motifs

Un autre résultat significatif est illustré par l'analyse d'une image texturée bruitée telle que celle de FIG. 3. L'application du traitement au moyen du même motif 3x3 et de la méthode des rangs généraux sur le type d'image texturée et bruitée, dessin numérique réalisé par l'homme, et qui entre dans la catégorie des images artificielles donne un résultat qui n'est pas moins significatif. La courbe de fréquence d'apparition des motifs est fort différente de la précédente. Sa lecture conduit à plusieurs remarques :

D'une part plusieurs motifs qui figurent en grand nombre dans l'image présentent des fréquences d'apparition égales ce qui se traduit par la présence de segments de droite horizontaux. Ce caractère ne se rencontre pas dans les images naturelles où les motifs ont peu de chance d'apparaître en nombre égal.

Les motifs contribuant à la zone linéaire oblique correspondent à la partie bruitée du fond de l'image. D'autre part le nombre de motifs inventoriés est faible (504). Cette caractéristique est commune aux images de synthèse construites au moyen d'un algorithme.

Conclusion et perspectives

L'utilisation de la loi de Zipf dont nous avons montré l'adaptation à la problématique de l'analyse de l'image peut s'étendre à de nombreux domaines. Il permet d'extraire un certain nombre de paramètres qui peuvent caractériser la structure d'une image. Ces paramètres sont : l'existence ou l'absence de zones horizontales, les pentes des droites de régression, le nombre de motifs recensés. L'ensemble de ces données qui traduit en quelque sorte la marque propre à chaque image représente un outil d'un grand intérêt pour procéder à une discrimination des images entre elles ou pour distinguer des zones au sein d'une même image. Parmi les perspectives envisagées, figure la plus attendue, l'authentification des signatures.

Références

[1] S. DJEZIRI, F. NOUBOUD, R. PLAMONDON, Discrimination d'images de documents, 1er Colloque

International Francophone sur l'Ecrit et le Document CIFED'98, Québec, 11-13 mai 1998, pp. 60-65.

- [2] G.K. ZIPF, Human Behavior and the principle of "Least Effort". Addison-Wesley, 1949, New York.
- [3] S. HAVLIN, S.V. BULDYREV, A.L. GOLDBERGER ET AL : Statistical Properties of D.N.A Sequences. Fractal Reviews in the Natural and Applied, Sciences Chapman & Hall, 1995, pp.1-11.
- [4] D. Bi. Segmentation d'images basée sur les statistiques de rangs des niveaux de gris. Thèse de doctorat, Université de Tours, 1997, 180 pages.