

Une mesure de complexité pour la détection de la zone d'intérêt d'une image

A complexity measure to detect the region of interest of an image

Y. Caron¹

H. Charpentier¹

P. Makris¹

N. Vincent²

¹ Laboratoire d'Informatique
Ecole Polytechnique de l'Université de Tours
64 avenue Jean Portalis
37200 TOURS

² Laboratoire CRIPS-SIP
Université Paris 5
45, rue des Saints Pères
75270 PARIS CEDEX 06

Yves.Caron@etu.univ-tours.fr, makris@univ-tours.fr, nicole.vincent@math-info.univ-paris5.fr

Résumé

Cet article présente une méthode permettant de détecter une zone d'intérêt dans une image numérique en utilisant une mesure de complexité basée sur la modélisation de la répartition des motifs de l'image à l'aide de lois puissance, en particulier la loi de Zipf et la loi de Zipf inverse. Différents codages peuvent être utilisés pour définir les motifs. L'image est segmentée en imagerie et on effectue une classification des imagerie en fonction des propriétés des lois puissance associées. Le résultat de cette classification permet de détecter une zone d'intérêt dans l'image. La loi de Zipf et la loi de Zipf inverse peuvent toutes deux être utilisées pour cette analyse. On se réfère également à la loi de Zipf en considérant différents codages des images. Les régions d'intérêt obtenues avec ces méthodes correspondent à celles qui sont reconnues par un observateur humain.

Mots Clef

Zone d'intérêt, lois puissance, loi de Zipf

Abstract

This paper presents a method allowing the detection of a region of interest in a digital image using a complexity measure based on the analysis of the distribution of image patterns using power laws, namely Zipf law and inverse Zipf law. Different codings can be used to define the patterns. The image is segmented into sub-images and a classification of the sub-frames is performed using the properties of the associated power law models. The result of this classification allows the detection of a region of interest in the image. Both Zipf law and inverse Zipf law can be used for this analysis. Zipf law can also be referred to with different image codings. The regions of interest detected with these methods match the regions found by a human observer.

Keywords

Region of interest, power laws, Zipf law

1 Introduction

La détection de zones d'intérêt dans des images est un des problèmes les plus difficiles à résoudre en vision par ordinateur. La difficulté vient du fait que, dans le cas général, il est difficile de déterminer a priori, sans connaissance préalable sur la nature de la scène, la position dans l'image et le contenu de la zone d'intérêt. Une zone d'intérêt étant une partie de l'image qui intéresse l'utilisateur, cette notion est subjective et dépend de l'application envisagée. On s'intéressera essentiellement ici aux images photographiques numériques en niveaux de gris. Dans ce type d'image, cette zone d'intérêt peut être par exemple un objet que l'on cherche à détecter. En général, une zone d'intérêt présente des caractéristiques visuelles particulières qui permettent de la distinguer du reste de l'image. Si l'on admet cette hypothèse, il est possible de chercher à définir une mesure permettant de caractériser la présence d'une zone qui se distingue nettement du reste de l'image. Evidemment, les critères de distinction sont très nombreux et nous avons choisi d'utiliser des critères de complexité de l'image, complexité de la répartition des niveaux de gris des pixels d'un point de vue spatial.

Différentes méthodes ont été proposées pour détecter des zones d'intérêt dans les images. La méthode proposée par Di Gesu et al. [8] utilise des mesures locales des moments et un estimateur de symétrie pour détecter une zone d'intérêt. Wang et al.[22] utilisent une transformée en ondelettes pour détecter des objets dans une image à faible profondeur de champ. Kadir et Brady [12] utilisent une approche multirésolution et une mesure d'entropie

afin de détecter des régions saillantes. Beaver et al. [1] emploient une méthode fractale pour détecter des objets dans des vues aériennes. Enfin, Stentiford [20] propose un estimateur de complexité locale de l'image qui s'inspire des mécanismes de la vision humaine de bas niveau.

Parmi les applications possibles pour la recherche automatique d'une zone d'intérêt, on peut notamment citer l'option d'optimisation de la compression présente dans la norme JPEG 2000.

La méthode présentée dans cet article utilise une approche issue de l'analyse linguistique mais rarement utilisée en analyse d'image : l'analyse statistique des motifs de l'image à l'aide de modèles de loi puissance. L'analyse statistique basée sur les propriétés des lois puissances, et en particulier de la loi de Zipf, a été utilisée dans différents domaines comme l'analyse linguistique par Zipf [23], Miller et Newman [16] ou Cohen et al. [6], l'étude de la répartition géographique des populations par Hill [10] et Makse et al. [13], le séquençage du génome humain par Mantegna et al. [15], l'analyse du trafic sur Internet par Breslau et al. [3], ou l'étude des bruits de la déglutition par Dellandrea et al. [7]. Dans le domaine de l'analyse d'images, la loi de Zipf a déjà été utilisée avec succès par Vincent et al. [21] pour la mesure de la qualité des images compressées et par Caron et al. [4] pour la détection des objets artificiels dans des milieux naturels. La détection d'objets artificiels dans des milieux naturels peut être considérée comme un cas particulier de la recherche de zone d'intérêt, et la méthode présentée ici est une généralisation de cette approche.

Après avoir présenté les modèles de loi puissance utilisés, à savoir les lois de Zipf et de Zipf inverse ainsi que quelques paramètres qui leur sont liés, nous allons étudier leur application au cas des images, notamment en ce qui concerne le codage des motifs. Ensuite nous présenterons les méthodes utilisées pour la détection de zones d'intérêt à l'aide de la loi de Zipf et de la loi de Zipf inverse, ainsi que l'utilisation simultanée de plusieurs codages des motifs. Enfin, nous discuterons de l'influence de la taille des imagerie sur le résultat, et on comparera entre elles les différentes méthodes de détection.

2 Le principe des mesures

Depuis les travaux de Mandelbrot [14], la mesure de complexité d'un phénomène est souvent liée à la dimension fractale des mesures d'observation. Elle repose sur la modélisation de l'évolution d'un paramètre par une loi puissance. Ici, nous avons choisi une approche qui repose aussi sur la loi puissance, outil bien adapté à la description de la nature, la loi de Zipf. Plus précisément, la méthode de détection de zones d'intérêt présentée dans cet article est basée sur deux modèles de loi puissance : la loi de Zipf et la loi de Zipf inverse.

Nous avons aussi utilisé les notions d'entropie associée à ces deux lois.

2.1 La loi de Zipf

La loi de Zipf est une loi empirique qui a été mise en évidence et énoncée en 1949 par G.K. Zipf [23] à partir de ses travaux sur l'étude de la fréquence des différents mots de la langue anglaise. Selon Zipf, dans un ensemble de symboles organisés topologiquement, par exemple un texte, les n-uplets de symboles ne se répartissent pas de manière aléatoire, mais la fréquence d'apparition des différents n-uplets, c'est à dire dans cet exemple les mots du texte, s'organise suivant une loi puissance, dite loi de Zipf. Selon Zipf, une telle loi s'observe dans les textes écrits en langage naturel suffisamment longs. L'interprétation donnée par Zipf de cette loi était liée au principe du moindre effort. Si l'on classe les différents n-uplets présents dans l'ensemble étudié suivant l'ordre décroissant de leur fréquence d'apparition, la fréquence d'apparition $N_{s(i)}$ du n-uplet de rang i dans la suite est donnée par la formule suivante :

$$N_{s(i)} = k \cdot i^{-\alpha} \quad (1)$$

Dans cette formule, k et α sont des constantes positives et la valeur de l'exposant α caractérise la loi puissance. Dans le cas des textes en langage naturel, la valeur de l'exposant α est proche de 1. Cette distribution en loi puissance peut se représenter graphiquement en échelle bi-logarithmique, avec en abscisse le rang des motifs et en ordonnée leur fréquence d'apparition. On appellera courbe de Zipf une telle représentation graphique.

2.2 La loi de Zipf inverse

La loi de Zipf inverse est un autre modèle de loi puissance, lui aussi mis en évidence par Zipf [24] à partir de ses travaux sur les textes. Contrairement à la loi de Zipf définie précédemment qui met en évidence les motifs les plus fréquents, la loi de Zipf inverse s'intéresse aux n-uplets dont la fréquence d'apparition est faible. La loi de Zipf inverse donne pour chaque fréquence d'apparition f , le nombre I de n-uplets différents ayant cette fréquence d'apparition, d'après la formule suivante :

$$I(f) = a f^{-b} \quad (2)$$

Dans cette formule, les coefficients a et b sont des constantes et la valeur de l'exposant b est proche de 2 dans le cas des textes étudiés par Zipf. La loi de Zipf inverse a notamment été utilisée dans les travaux de Cohen et al. [6] sur l'analyse de textes naturels et artificiels.

2.3 L'entropie

La notion d'entropie peut également être utilisée pour caractériser la complexité d'un ensemble de symboles comme un texte ou une image. L'entropie au sens de la

théorie de l'information a été définie par Shannon [14]. Pour un texte de longueur T contenant R mots distincts, l'entropie peut être définie par la formule suivante :

$$\tilde{H}_w = - \sum_{r=1}^R \frac{f(r)}{T} \log_R \frac{f(r)}{T} \quad (3)$$

Dans cette formule, on utilise un logarithme à base R , de manière à permettre de comparer les résultats obtenus avec des textes différents, il permet de donner une valeur de l'entropie comprise entre 0 et 1 quelle que soit la valeur de R . Cette entropie est maximale quand tous les mots du texte ont la même fréquence d'apparition, elle est minimale quand la fréquence relative d'un des différents mots atteint 1. L'entropie permet donc de mesurer l'uniformité de la distribution des différents mots. Une autre formulation de l'entropie a été utilisée par Cohen et al. [6] dans leurs travaux sur les textes. L'entropie est ici définie relativement à la fréquence d'apparition des différents mots du texte. L'entropie relative à la fréquence est définie par la formule suivante :

$$\tilde{H}_f = - \sum_{f=1}^F \frac{I(f)}{R} \log_F \frac{I(f)}{R} \quad (4)$$

Dans cette formule, $I(f)$ représente le nombre de mots distincts ayant une fréquence d'apparition égale à f et F représente le nombre total d'occurrences des mots dans le texte. L'utilisation d'un logarithme à base F assure que la valeur de l'entropie soit comprise entre 0 et 1. Par rapport à la précédente, cette définition de l'entropie permet de donner un poids plus important aux mots les moins fréquents. Or, ce sont les mots peu fréquents qui sont souvent les plus discriminants quant au contenu du texte. Plusieurs travaux ont été effectués sur l'utilisation de l'entropie pour la détection de régions saillantes dans une image, notamment ceux de Jägersand [11], ou ceux de Kadir et Brady [12].

3 Application aux images

Après avoir défini les modèles mathématiques utilisés, il convient maintenant de préciser comment on peut les appliquer au domaine de l'image. Tout d'abord, nous définirons les codages utilisés pour pouvoir les utiliser pour l'analyse d'images. Nous montrerons ensuite quelques résultats.

3.1 Codage de l'image

Les n -uplets de symboles seront les motifs de l'image. Chaque motif est constitué d'un bloc de pixels adjacents de taille 3×3 . D'autres tailles de motifs seraient possibles, mais une taille plus grande des motifs aurait pour conséquence que chaque motif n'aurait qu'une faible probabilité de se retrouver plusieurs fois dans l'image, et la distribution des fréquences des motifs ne serait pas véritablement significative. De plus, une taille de motifs trop importante aurait pour conséquence d'augmenter considérablement le temps de calcul. C'est pourquoi une

taille des motifs de 3×3 nous a semblé un bon choix pour notre application. Une première idée de codage serait d'utiliser directement les niveaux de gris de l'image. Cependant, le nombre de motifs possibles serait dans ce cas de 256^9 , soit $4,7 \times 10^{21}$. Avec un tel nombre de motifs possibles, la probabilité d'apparition d'un même motif plusieurs fois dans l'image serait également très faible. Le nombre de pixels dans une image est négligeable devant ce nombre de motifs possibles, c'est pourquoi on a choisi d'utiliser d'autres codages pour les motifs afin de réduire le nombre de motifs différents de l'image.

Un premier type de codage consiste à partitionner l'échelle des niveaux gris en un nombre réduit de classes. Nous désignerons cette méthode par méthode des n classes. L'échelle des niveaux de gris est divisée en plusieurs intervalles ou classes numérotées de 0 à $n - 1$ et on affecte à chaque pixel le numéro de la classe qui correspond à son niveau de gris. Le nombre de classes a été fixé à 9 pour notre application, ce nombre permet de réduire significativement le nombre de motifs différents, qui n'est plus alors que de $3,874 \times 10^8$ tout en préservant une certaine complexité de l'image, nécessaire afin de pouvoir distinguer la zone d'intérêt.

Un autre codage possible est celui des rangs généraux, utilisé par Bi et al. [2]. Contrairement au codage précédent qui distingue entre les grandes zones de l'image, le codage des rangs généraux s'intéresse aux faibles variations de niveaux de gris. La méthode des rangs généraux consiste, à l'intérieur du bloc 3×3 , à numérotter les pixels en fonction des niveaux de gris classés dans l'ordre croissant de leur valeur en affectant le même rang quand si les niveaux de gris ont même valeur. On affecte la valeur 0 au niveau de gris le plus bas, et on incrémente la valeur d'une unité quel que soit l'écart relatif entre deux niveaux de gris consécutifs.

Afin de diminuer les influences des faibles variations de niveau de gris, il est possible d'utiliser un seuil pour le codage des rangs généraux. Dans ce cas, on affecte la même valeur lors du codage aux pixels dont la différence de niveaux de gris est inférieure au seuil choisi. Une autre manière de diminuer l'influence des faibles variations de niveaux de gris est d'appliquer la méthode des rangs généraux non pas sur les niveaux de gris d'origine, mais sur les motifs obtenus à l'aide de la méthode des 9 classes. La figure 1 montre la représentation d'un motif par les différents codages. Il est également possible d'utiliser l'écart type des niveaux de gris pour coder les pixels de l'image. Cet écart type est calculé par rapport à la moyenne des niveaux de gris du motif centré sur un pixel. Les valeurs obtenues sont ensuite normalisées par rapport à la valeur maximale de l'écart type dans l'image. L'échelle des valeurs est ensuite partitionnée en 9 classes et on affecte à chaque pixel la valeur de la classe. On voit que des caractéristiques variées peuvent être ainsi mises en évidence dans les images.

255	210	210	a	8	7	7	b	5	4	4	c
25	2	34		0	0	1		1	0	2	
40	2	40		1	0	1		3	0	3	

4	3	3	d	3	2	2	e
1	0	2		0	0	1	
2	0	2		1	0	1	

Fig. 1 : motif original (a), codé par la méthode des 9 classes (b), par la méthode des rangs généraux (c), par les rangs généraux avec un seuil de 10 (d), par les rangs généraux appliqués sur un motif codé par les 9 classes (e)

3.2 Application des modèles de loi puissance

L’analyse d’une image en utilisant la loi de Zipf consiste donc à balayer l’image par un masque 3x3, à coder les motifs en utilisant l’un des codages définis précédemment puis à affecter à chaque motif distinct sa fréquence d’apparition dans l’image. Les motifs sont ensuite classés dans l’ordre décroissant de leur fréquence d’apparition. Le résultat de l’étude peut être représenté graphiquement dans un repère bi-logarithmique, cette représentation est la courbe de Zipf associée à l’image. D’un point de vue théorique, si l’image respecte la loi de Zipf, on devrait obtenir un ensemble de point alignés sur une droite dont la pente serait égale à - α selon la formule (1). Les courbes de Zipf associées à une même image sont différentes suivant le codage utilisé, comme le montre la Fig. 2.

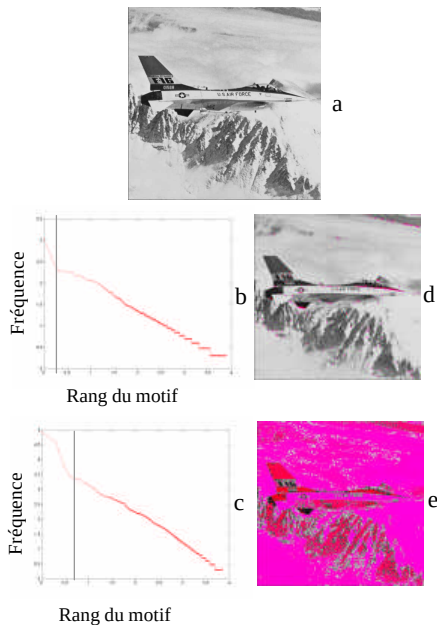


Fig. 2 : Courbes de Zipf associées à une image (a) obtenues avec les rangs généraux (b) et le codage des 9 classes (c). Les points colorés sur les images (d) et (e) représentent les motifs de la partie gauche de la courbe.

On remarque que le plus souvent la courbe de Zipf d’une image en niveau de gris peut être divisée en deux parties linéaires. Dans le cas du codage des 9 classes, la partie gauche de la courbe correspond aux zones uniformes de l’image c’est à dire les régions et sa partie droite correspond aux zones non uniformes, comme les contours et les détails fins de l’image. Sur les courebes de Zipf de la figure 2, la séparation entre les deux parties de la courbe est représentée par un trait vertical et les motifs situés dans la partie gauche de la courbe sont représentés par des points colorés sur les images de droite (d) et (e). On peut également remarquer que plus le nombre de pixels de l’image est grand, plus la courbe obtenue se rapproche d’une droite. On peut donc dire que la répartition des fréquences d’apparition des motifs de l’image suit bien une distribution en loi puissance. La forme de la courbe dépend également de la complexité de l’image, comme on peut le voir sur la Fig. 3, qui montre les courbes de Zipf associées à deux zones extraites de l’image précédente, codées ici à l’aide des 9 classes. On remarque que les courbes de Zipf associées à des zones uniformes n’ont pas la même allure que celles associées à des zones comportant de nombreux détails. Un léger décalage de l’image n’a qu’une influence négligeable sur l’allure de la courbe de Zipf obtenue. Il est donc possible d’utiliser la courbe de Zipf pour caractériser la complexité d’une image afin de détecter une zone d’intérêt dans celle-ci.

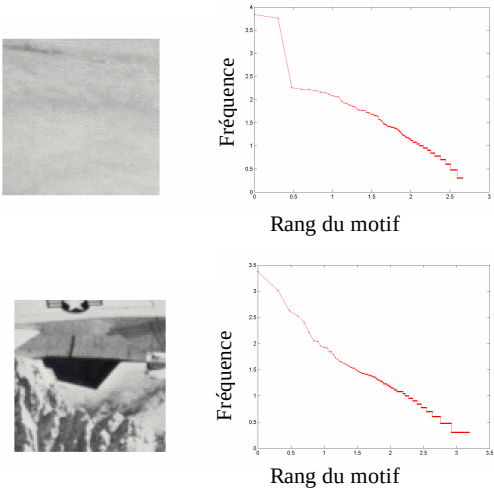


Fig. 3 : Courbes de Zipf associées à une zone de l’image uniforme (en haut) et à une zone comportant de nombreux détails (en bas).

Pour utiliser la loi de Zipf inverse afin d’analyser une image, on procède de la même manière qu’avec la loi de Zipf pour extraire les motifs de l’image et les coder. On représente ensuite la courbe de Zipf inverse associée à l’image dans un repère bi-logarithmique. La Fig. 4 montre la courbe de Zipf inverse associée à l’image précédente, obtenue ici avec le codage des 9 classes.

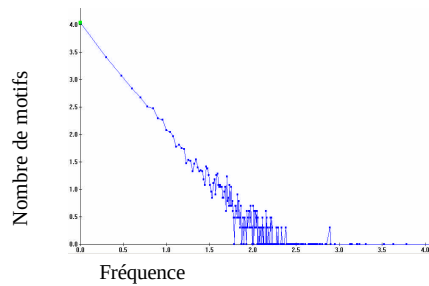


Fig. 4 : exemple de courbe de Zipf inverse pour l'image présentée en 2 (a).

Sur cette courbe, on remarque que les points indiquant la répartition du nombre de motifs en fonction de leur fréquence sont alignés sur une droite pour les motifs les moins fréquents dans l'image. L'hypothèse de Zipf est donc vérifiée dans le cas des images. De même qu'avec la loi de Zipf, on peut remarquer que l'allure de la courbe de Zipf inverse est différente selon le contenu de l'image, comme on peut le voir sur la figure 5. Dans le cas d'une image complexe, la pente de la courbe et l'ordonnée à l'origine sont plus importantes que dans le cas d'une image uniforme, en effet plus la complexité de l'image est grande, plus il y a de nombreux motifs n'apparaissant qu'une fois dans l'image. La loi de Zipf inverse permet donc de caractériser la quantité de détails présente dans l'image.

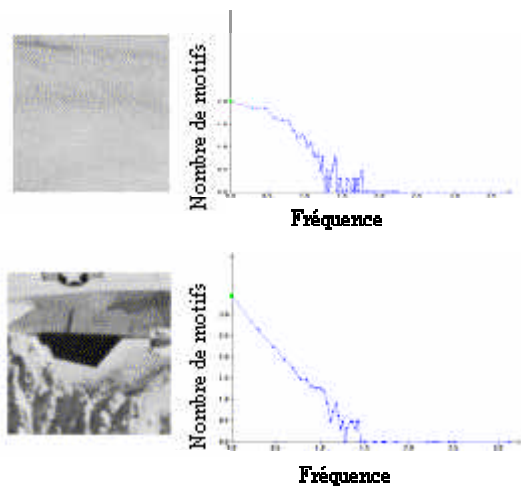


Fig. 5 : Courbes de Zipf inverse associées à deux zones différentes de la même image.

4 Détection de zones d'intérêt

Si les modèles de loi puissance permettent de caractériser la complexité des textures d'une l'image, ils ne donnent par eux-mêmes aucune information spatiale quant à la position d'une éventuelle zone d'intérêt dans l'image. C'est pourquoi il est nécessaire de segmenter l'image en imagerie afin de pouvoir détecter une zone d'intérêt. La

segmentation n'ayant pas besoin d'être très précise, on pourra se contenter d'un découpage en imagerie rectangulaires de taille fixe. La méthode de détection consiste alors, après avoir découpé l'image en imagerie, à appliquer la loi de Zipf ou de Zipf inverse sur chacune des imagerie, en utilisant un ou plusieurs codages des motifs et à effectuer une classification en fonction des propriétés des courbes de Zipf ou de Zipf inverse obtenues. La méthode de classification utilisée est une méthode non supervisée qui consiste à représenter les courbes de Zipf associées à chaque imagerie par un nuage de points dont les coordonnées représentent les caractéristiques des courbes obtenues, et à classer les points du nuage en deux classes correspondant respectivement à la zone d'intérêt et au fond. Contrairement à des méthodes basées sur les matrices de cooccurrences comme celle de Haddon et Boyce [9] ou aux réseaux de neurones comme ceux employés notamment par Carpenter et Grossberg [5], par Srinivasan et Kanal [19] ou par Rogers et al. [17], cette méthode permet de se passer de phase d'apprentissage. La Fig. 6 montre l'organigramme de la méthode de détection.

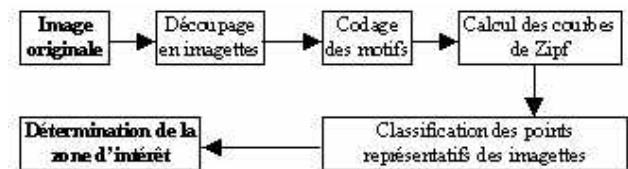


Fig. 6 : Organigramme de la méthode de détection de zones d'intérêt

4.1 Utilisation de la loi de Zipf

Pour détecter les zones d'intérêt à l'aide de la loi de Zipf, on utilisera le codage des 9 classes. En effet, ce type de codage donne une information qui est davantage conforme à la perception visuelle humaine que celle obtenue à l'aide du codage des rangs généraux, car il s'agit en fait d'un simple changement d'échelle des niveaux de gris. Nous avons vu que la courbe de Zipf associée à une image codée avec la méthode des 9 classes peut être divisée en deux parties. La séparation entre les deux parties de la courbe est définie par le point le plus éloigné de la droite joignant les deux points extrêmes de la courbe. La courbe peut alors être caractérisée par les deux pentes des droites de régression linéaire des deux parties de la courbe. Ces deux pentes seront utilisées comme paramètres de classification des imagerie. La courbe de Zipf associée à chacune des imagerie sera représentée par un point dans un plan, l'abscisse de ce point représentant la pente de la partie gauche de la courbe et l'ordonnée celle de la partie droite. Pour la plupart des images, la zone d'intérêt est constituée d'imagerie pour lesquelles la pente de la partie gauche de la courbe qui contient les zones uniformes est supérieure à celle de la partie droite qui correspond aux

détails. Cela correspond sur le graphe, dont un exemple est donné sur la Fig. 7, à la partie du nuage de points situés sous la première bissectrice. Plus spécifiquement, il a été déterminé expérimentalement que la zone d'intérêt correspond aux points situés dans la partie du nuage de points délimitée par les droites d'équation $y=x$ et $y=1,2Gx$, où Gx est l'abscisse du centre de gravité du nuage de points. Cependant, tous les points situés dans cette partie du nuage ne correspondent pas à des imagerie appartenant à la zone d'intérêt de l'image. Dans ce cas, on ne considère que la plus grande composante connexe des imagerie vérifiant le critère de classification. La Fig. 6 montre un exemple de zone d'intérêt obtenu avec cette méthode.

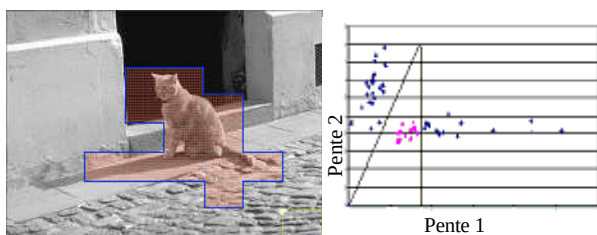


Fig. 7 : zone d'intérêt détectée avec la loi de Zipf et le codage des 9 classes.

4.2 Utilisation de la loi de Zipf inverse

De même que la loi de Zipf, la loi de Zipf inverse peut être utilisée pour détecter des zones d'intérêt dans une image. Comme avec la loi de Zipf, on représente la courbe de Zipf inverse associée à chaque imagerie par un point dans un graphe. La courbe de Zipf inverse est caractérisée par la pente des points situés à gauche et par son ordonnée à l'origine. Une image peut alors être représentée par un point dans un espace de dimension 2 dont l'abscisse est la pente et l'ordonnée est l'ordonnée à l'origine de la courbe de Zipf inverse. Sur la plupart des images, la zone d'intérêt comporte davantage de détails que le fond, elle sera donc moins uniforme. Cela se traduit par le fait que les courbes de Zipf inverse associées aux imagerie de la zone d'intérêt ont une ordonnée à l'origine élevée et se situent donc dans la partie supérieure du nuage de points. La zone d'intérêt est donc constituée des imagerie situées au-dessus du centre de gravité du nuage de points, comme on peut le voir sur la Fig. 8. Cependant, toutes les imagerie vérifiant ce critère ne font pas partie de la zone d'intérêt principale de l'image, c'est pourquoi on ne retiendra que la plus grande composante connexe des imagerie situées au-dessus du centre de gravité du nuage de points. Il se peut qu'il y ait des trous dans cette zone connexe, dans ce cas ils peuvent être comblés en considérant que les imagerie dont toutes les voisines appartiennent à la zone d'intérêt appartiennent elles aussi à la zone d'intérêt.

Afin de toujours trouver une zone d'intérêt de taille suffisante, on peut imposer que la taille de la zone d'intérêt soit comprise entre 20 et 50% de la taille de l'image. Si après l'extraction de la plus grande composante connexe et le rebouchage éventuel des trous, la zone d'intérêt comprend plus de 50% des imagerie, on peut modifier la position de la droite de séparation jusqu'à ce que la taille de la zone d'intérêt trouvée occupe moins de 50% de la taille de l'image. À l'inverse, si la zone d'intérêt a une surface inférieure à 20% de celle de l'image, on peut diminuer l'ordonnée de la droite de séparation jusqu'à obtenir une zone d'intérêt de la taille voulue. La méthode de Zipf inverse donne souvent de meilleurs résultats que la méthode de Zipf, elle fournit des zones d'intérêt plus précises et elle permet de détecter des zones d'intérêt dans des images où la méthode de Zipf échoue.

Il existe cependant des images pour lesquelles la zone d'intérêt est plus uniforme que le fond. C'est le cas en particulier des images d'objets artificiels dans des milieux naturels. Ces images se caractérisent par le fait que plus de 50% des motifs n'apparaissent qu'une seule fois dans l'image. Dans ce cas, la zone d'intérêt est constituée par la partie du nuage de points dont l'ordonnée est inférieure à celle du centre de gravité du nuage de points. Ce cas peut être automatiquement révélé par la position globale du nuage de points dans le plan. De même que dans le cas précédent, on extrait la plus grande composante connexe et on ajuste la taille de la zone d'intérêt entre 20 et 50% de la taille de l'image.

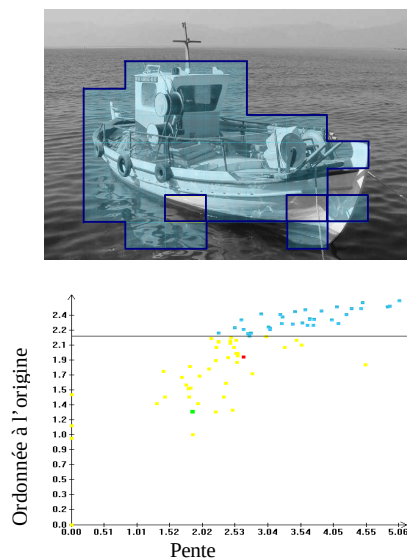


Fig. 8 : Détection de zone d'intérêt avec la méthode de Zipf inverse

4.3 Utilisation de l'écart type

Il est également possible d'utiliser l'écart type des niveaux de gris pour détecter des zones d'intérêt dans l'image. La méthode proposée commence par la construction d'une image de l'écart type. Pour cela, on

balaie l'image par un masque 3x3 dans, et on calcule l'écart type des niveaux de gris des pixels du masque. On associe la valeur de l'écart au pixel situé au centre du masque. On partitionne ensuite en 9 classes l'échelle des valeurs de l'écart type trouvées sur l'ensemble de l'image et on affecte à chaque pixel la valeur de la classe. La Fig. 9 montre un exemple d'image de l'écart type réalisée avec cette méthode.

L'image est ensuite partitionnée en imagerie de taille 32x32 et l'on construit la courbe de Zipf associée à chaque imagerie en utilisant un masque 3x3 et les classes d'écart type définies précédemment. On calcule la pente de la droite de régression de chacune des courbes de Zipf. On considère que l'imagerie fait partie de la zone d'intérêt si la pente de sa courbe de Zipf est supérieure à 0,8 fois la valeur maximale de la pente sur l'ensemble des imageries. Cette méthode donne de bons résultats dans le cas des images à faible profondeur de champ, notamment pour les objets artificiels dans des milieux naturels. Elle permet également de repérer des surfaces construites par l'homme dans des photographies aériennes ou satellites. Elle est en revanche assez mal adaptée aux autres types d'images. De plus, le temps de calcul est plus long du fait de la construction préalable d'une image de l'écart type.

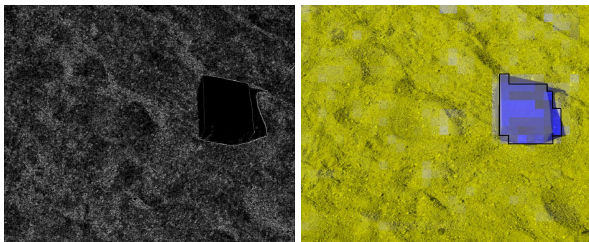


Fig. 9 : Image de l'écart type des niveaux de gris (à gauche) et zone d'intérêt obtenue avec cette méthode (à droite)

4.4 Utilisation de l'entropie

L'entropie permettant de donner une mesure de la quantité de détails présents dans une image, on peut envisager de l'utiliser afin de détecter des régions d'intérêt, étant donné que ces régions comportent en général davantage de détails que le reste de l'image. Pour détecter des zones d'intérêt avec l'entropie, on découpe l'image en imagerie, et pour chacune des imageries on calcule l'entropie relative à la fréquence des motifs en utilisant la formule de Cohen rappelée en (3) et (4). Le calcul de l'entropie permet de vérifier que les imageries dont l'entropie est la plus élevée sont effectivement celles qui contiennent le plus de détails. Cependant, cette méthode ne donne pas toujours des résultats satisfaisants en ce qui concerne la détection de zones d'intérêt. En effet, les imageries d'entropie élevée correspondent souvent à des détails isolés plutôt qu'à la zone d'intérêt principale de l'image, dont l'entropie se situe autour de

la moyenne et ne se distingue pas véritablement de celle des imageries de l'arrière-plan. La Fig. 10 montre un exemple d'utilisation de l'entropie pour la détection de zones d'intérêt.

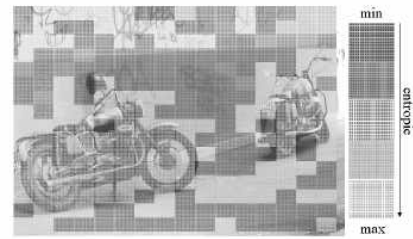


Fig. 10 : Evaluation de l'entropie sur une image découpée en imagerie.

5 Utilisation simultanée de plusieurs codages

Afin d'améliorer les résultats de la détection, on peut envisager de combiner les résultats provenant de plusieurs calculs de courbe de Zipf effectués avec des codages différents et de fusionner les résultats obtenus. Parmi les combinaisons possibles de codages, deux ont donné des résultats particulièrement prometteurs, à savoir le codage des 9 classes associé à l'écart type, et la méthode des rangs généraux appliquée sur des motifs préalablement codés à l'aide de la méthode des 9 classes, elle aussi associée à une mesure de l'écart type.

5.1 Ecart type et méthode des 9 classes

Après segmentation de l'image en imagerie, on construit deux courbes de Zipf pour chacune des imageries. La première courbe est calculée à l'aide de la méthode des 9 classes déjà évoquée. Pour la seconde, on calcule une mesure de l'écart type des niveaux de gris sur un motif de taille 3x3. Afin de diminuer le temps de calcul, on n'a pas construit préalablement d'image de l'écart type, mais on s'est contenté de calculer l'écart type sur les valeurs des pixels codés par la méthode des 9 classes pour construire la courbe de Zipf qui représente les fréquences d'apparition des différentes valeurs. On calcule ensuite la pente de chaque courbe de Zipf et pour chacune des imageries, on représente en abscisse la pente de la courbe obtenue à l'aide de la méthode des 9 classes et en ordonnée la pente de la courbe obtenue avec l'écart type. On peut remarquer que les imageries contenant des zones totalement uniformes se caractérisent par une pente nulle pour la courbe obtenue avec la méthode des 9 classes. Les autres imageries contenant davantage de détails se concentrent autour du centre de gravité du nuage de points. Or, sur la plupart des images, ce sont souvent les imageries qui comportent le plus de détails qui constituent la zone d'intérêt. Il a été montré expérimentalement que la zone d'intérêt correspond, sur la plupart des images, aux imageries représentées par des points dont la distance au centre de gravité est inférieure

à 0,8 fois l'abscisse de celui-ci. La Fig. 11 montre un exemple de résultat obtenu avec cette méthode.

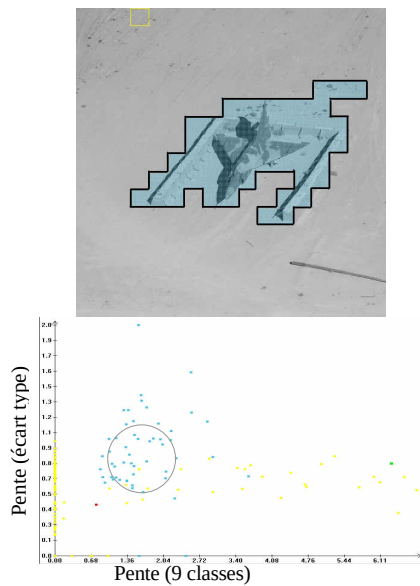


Fig. 11: Exemple de zone d'intérêt détectée avec les 9 classes et l'écart type

5.2 Ecart type, rangs généraux et méthode des 9 classes

Une autre méthode dérivée de la précédente a également été testée. Elle consiste à coder l'image avec la méthode des 9 classes, puis à calculer la mesure de l'écart type définie précédemment, ainsi qu'à appliquer la méthode des rangs généraux sur les motifs codés avec les 9 classes. On construit ensuite deux courbes de Zipf pour chacune des imagerie, et on représente chaque imagerie par un point dans un nuage de points, où l'abscisse représente la pente de la courbe obtenue à partir des motifs codés avec les rangs généraux et les 9 classes et l'ordonnée, la pente de la courbe de Zipf obtenue à partir des valeurs de l'écart type. On peut ici remarquer que les imagerie représentant l'arrière-plan de l'image tendent à se regrouper autour de la première bissectrice du graphe, tandis que les imagerie comportant des détails plus complexes sont situées sous celle-ci. Avec cette méthode, la zone d'intérêt a été définie à partir des résultats expérimentaux comme la plus grande composante connexe des imagerie situées sous la première bissectrice et dont l'abscisse est inférieure à 2,2 fois celle du centre de gravité du nuage de points. Cette méthode donne de bons résultats surtout pour les images dont la zone d'intérêt comporte de nombreux détails fins. La Fig. 12 montre un exemple de zone d'intérêt obtenue avec cette méthode. On remarque que la zone d'intérêt obtenue par cette méthode est plus précise qu'avec la précédente.

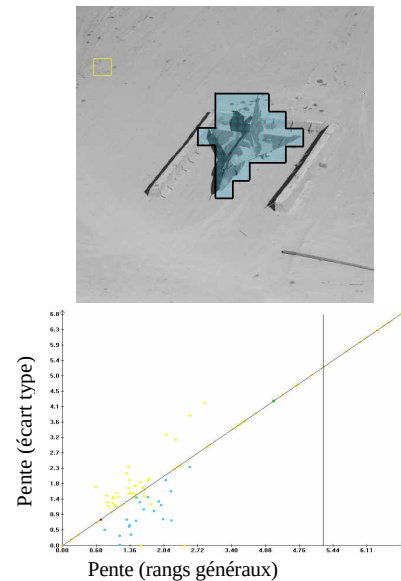


Fig. 12 : Exemple de zone d'intérêt détectée avec l'écart type et les rangs généraux appliqués sur les 9 classes.

6 Influence de la taille des imagerie

La taille des imagerie utilisées dans le découpage de l'image a évidemment une influence sur la qualité de la détection des zones d'intérêt. Notre approche reposant sur des techniques statistiques, les imagerie ne peuvent être trop petites. Nous montrons sur la figure 13, pour une même image avec différents découpages en imagerie, les zones d'intérêts détectées par la méthode de Zipf inverse pour trois découpages différents. Avec un découpage en 8x8 imagerie, la détection échoue, la zone d'intérêt détectée est ici le sol au premier plan de l'image. Avec un découpage en 19x19 imagerie, la zone d'intérêt, c'est à dire la moto, est correctement détectée. Par contre, avec un découpage en 32x32 imagerie, la zone d'intérêt ne correspond plus que partiellement à l'objet à détecter, certaines parties uniformes de l'objet étant considérées comme faisant partie du fond et des zones texturées du fond comme appartenant à la zone d'intérêt. Dans le cas des méthodes de Zipf et de Zipf inverse utilisant seulement le codage des 9 classes, les meilleurs résultats ont été obtenus avec des imagerie dont la surface est de l'ordre de 5000 pixels, ce qui correspond, sur l'exemple ci-dessus, au découpage en 19x19 imagerie.

De la même manière, la taille optimale des imagerie a été déterminée pour les méthodes utilisant l'écart type associé à un autre codage, décrites aux paragraphes 5.1 et 5.2. Pour ces deux méthodes, la taille optimale des imagerie est ici de l'ordre de 1000 pixels. Ces deux méthodes permettent de travailler avec des imagerie plus petites, ce qui doit en principe permettre d'obtenir des zones d'intérêt plus précises. Le temps de calcul est fonction de la taille de l'image et des performances de la

machine utilisée, il est de l'ordre de 2 à 3 secondes pour une image de 1024x768 sur un Pentium III 1,2 Ghz.

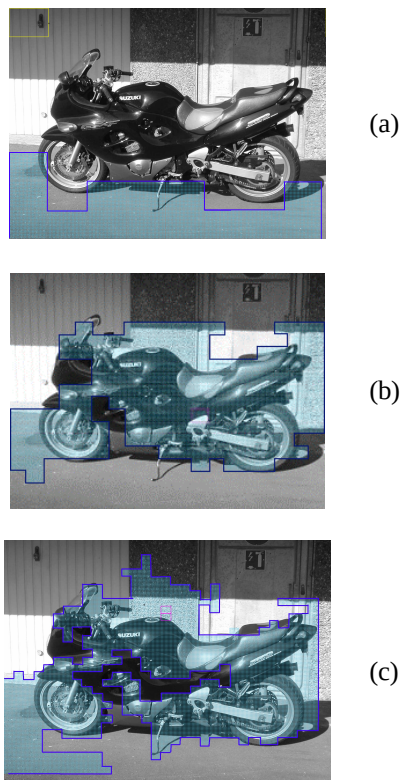


Fig. 13 : Zones d'intérêt détectées à l'aide de Zipf inverse avec un découpage en 8x8 (a), 19x19 (b) et 32x32 (c) imageries.

7 Comparaison des différentes méthodes

Les différentes méthodes présentées dans cet article vont maintenant être comparées entre elles relativement à un objectif de détection d'une zone d'intérêt dans une image. Une base de test de 100 images a été constituée, chacune de ces images présentant une zone d'intérêt nettement visible pour un observateur humain. Ces images sont des photographies numériques de sujets divers, et elles sont de résolution variable. De telles images sont représentatives de ce que l'on pourrait par exemple avoir à traiter avec un logiciel de compression. Les différentes méthodes ont été testées sur chacune des images. On considère que la détection est réussie lorsque la région d'intérêt détectée par le programme contient l'objet que l'on cherche à détecter, c'est à dire celui qui serait défini comme zone d'intérêt par l'observateur humain. Les résultats sont présentés dans le tableau 1. La méthode basée sur l'écart type seul présentée au paragraphe 4.3 n'a pas été évaluée, en effet, elle n'est efficace que sur certains types d'images particuliers, comprenant des objets uniformes dans un fond texturé, qui ne sont pas représentatifs des images testées ici.

D'après le tableau, on remarque que la méthode de Zipf inverse est de loin la plus efficace. Cependant, elle a l'inconvénient de détecter souvent des zones d'intérêt trop grandes, qui associent des portions du fond à la véritable zone d'intérêt que l'on voudrait détecter. Les méthodes utilisant l'écart type ont moins cette tendance. On voit également que la loi de Zipf avec le codage des 9 classes donne des résultats médiocres et que l'entropie est peu efficace pour détecter des zones d'intérêt. Il faut cependant noter que toutes les méthodes n'échouent pas sur les mêmes images, et qu'une méthode ayant un moins bon taux de réussite peut s'avérer meilleure sur un type d'image particulier.

Méthode	Taux de détection réussie
Zipf (9 classes)	52 %
Zipf inverse	82 %
Entropie	38 %
Ecart type - 9 classes	60 %
Ecart type - 9 classes – Rangs généraux	69 %

Tableau 1 : Taux de détection réussie en fonction de la méthode.

8 Conclusion

L'application aux images des modèles de loi puissance permet la détection de zones d'intérêt dans une image conformes à celles que verrait un observateur humain. La loi de Zipf, comme la loi de Zipf inverse, peuvent être utilisées pour effectuer cette détection. Les motifs de l'image peuvent être codés de plusieurs manières, et les différents codages peuvent être utilisés simultanément ou non. La taille optimale des imageries a pu être déterminée expérimentalement, elle dépend du codage utilisé pour les motifs. La méthode basée sur la loi de Zipf inverse donne les meilleurs résultats, de plus elle est facilement adaptable à différents types d'images, cependant elle a tendance à détecter des zones d'intérêt plus grandes que prévu.

Afin d'améliorer les résultats, on peut envisager d'utiliser simultanément les méthodes de Zipf et de Zipf inverse. Nous comptons développer notre étude en utilisant un plus grand nombre de paramètres ainsi qu'une méthode de classification des résultats plus élaborée.

Bibliographie

[1] P. Beaver, S.M. Quirk, J.P. Sattler, Object Identification in Greyscale Imagery using Fractal Dimension, in M Novak, *Fractal Reviews in the Natural and Applied Science*, Chapman & Hall, London, pp. 63-73, 1995

- [2] D. Bi, J.P. Asselin de Beauville, M.Mraghni, Spatial gray levels distribution based unsupervised texture segmentation, *Proceedings of 3rd International Conference of Signal Processing (ICSP96)*, Pékin (Chine), 1996
- [3] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, Web caching and Zipf-like distributions, Evidence and implications, *Proceedings of IEEE Infocom 99*, New York, pp. 126-134, 1999
- [4] Y. Caron, P. Makris, N. Vincent, A method for detecting artificial objects in natural environments, *International Conference on Pattern recognition (ICPR - IAPR)*, Québec (Canada) pp. 600-603, 2002
- [5] G.A. Carpenter, S. Grossberg, *Neural networks for vision and image processing*, MIT Press, Cambridge, MA, 1992
- [6] A. Cohen, R.N. Mantegna, S. Havlin, Numerical analysis of word frequencies in artificial and natural language texts, *Fractals*, Vol. 5 No.1, pp. 95-104, 1997
- [7] E. Dellandrea, P. Makris, M. Boiron, N. Vincent, A medical acoustic signal analysis method based on Zipf law, *International Conference on Digital Signal Processing (DSP 2002)*. Santorin (Grèce). Vol. 2, pp. 615-618. Juillet 2002.
- [8] V. Di Gesu, C. Valenti, L. Strinati, Local Operators to Detect Regions of Interest, *Pattern Recognition Letters*, Vol.18, No. 11-13, pp. 1077-1081, 1997
- [9] J.F. Haddon and J.F. Boyce, Co-occurrence matrices for image analysis, *Electronics and Communication Engineering Journal*, pp. 71-83, 1993.
- [10] B.M. Hill, Zipf's law and prior distributions for the composition of a population, *Journal of the American Statistical Association*, 65, pp.1220-1232, 1970
- [11] M. Jägersand, Saliency Maps and Attention Selection in Scale and Spatial Coordinates: An Information Theoretic Approach. *International Conference on Computer Vision (ICCV 95)*, Boston, MA. June 20-23, 1995 , pp. 195-202
- [12] T. Kadir, M. Brady, Scale, Saliency and Image Description, *International Journal of Computer Vision*, Vol. 45 No. 2, pp. 83-105, 2001
- [13] H.A. Makse, S. Havlin, H.E. Stanley, Modelling urban growth patterns, *Nature*, 377, pp. 608-612, 1995
- [14] B. Mandelbrot, *The Fractal Geometry of Nature*, W.H Freeman, New York, 1983
- [15] R.N. Mantegna, S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.K. Peng, M. Simons, H.E. Stanley, Linguistic Features of Noncoding DNA Sequences, *Phys. Rev. Lett*, 73,3169, 1994
- [16] G.A. Miller, E.B. Newman, Test of a Statistical Explanation of the Rank-Frequency Relation for Words in Written English, *American Journal of Psychology*, No. 71, pp. 209-218, 1958
- [17] S.K. Rogers, J.M. Colombi, C.E. Martin, J.C. Gaaney, K.H. Fielding, T.J. Burns, D.W. Ruck, M.Kabrisky, M.Oxley, Neural Networks for Automatic Target Recognition, *Neural Networks*, Volume 8, Issues 7-8, 1995, pp. 1153-1184
- [18] C.E. Shannon, A Mathematical Theory of Communication, *Bell System Tech. Journal*, No. 27, pp. 79-423, 1948
- [19] S. Srinivasan, L. Kanal, Qualitative landmark recognition using visual cues, *Pattern Recognition Letters*, Volume 18, Issues 11-13, November 1997, pp. 1405-1414
- [20] F. Stentiford, An Estimator for Visual Attention through Competitive Novelty with Application to Image Compression, *Picture Coding Symposium*, Seoul (Corée du Sud) , pp. 25-27, April 2001
- [21] N.Vincent, P.Makris, J.Brodier. Compressed Image Quality and Zipf's Law, *Proceedings of International Conference on Signal Processing (ICSP - IFIC-IAPRWCC2000)* , pp. 1077-1084, Pékin (Chine), 2000
- [22] J.Z. Wang, J.Li, R.M. Gray, G. Wiederhold, Unsupervised Multiresolution Segmentation for Images with Low Depth of Field, *IEEE PAMI*, Vol. 23 No.1, pp. 85-90, 2001
- [23] Zipf, G.K, *Human Behavior and the Principle of Least Effort*, Addison-Wesley, New York, 1949
- [24] Zipf, G.K, *The Psychology of Language, an Introduction to Dynamic Philology*, M.I.T. Press, Cambridge, Massachusetts, 1965