

Approches multiéchelles pour l'extraction de structures d'intérêt dans les images naturelles

Multiscale approaches for the extraction of meaningful structures in natural images

A. Turiel[†]

J. Grazzini[‡]

H. Yahia[‡]

I. Herlin[‡]

[†] Departament de Física Fonamental
Universitat de Barcelona
Diagonal, 647. 08028 Barcelona, Espagne.
turiel@ffn.ub.es

[‡] INRIA, Projet AIR
Domaine de Voluceau, Rocquencourt
B.P. 105, 78153 Le Chesnay Cedex, France.
{Jacopo.Grazzini, Hussein.Yahia,
Isabelle.Herlin}@inria.fr

Résumé

Dans cet article, nous nous intéressons à l'extraction de structures géométriques d'une image sur la base d'un critère discriminant qui permettrait de mesurer leur importance relative. Toute la question porte sur la définition de ce critère. L'idée la plus naturelle consiste à utiliser l'entropie définie par Shannon en théorie de l'information, afin d'évaluer la quantité d'information contenue dans les différentes régions de l'image. L'estimation locale de l'entropie fournit une mesure spatialement localisée, mais cependant redondante, de cette quantité : en la généralisant dans un cadre multiéchelle, on respecte les propriétés d'invariance d'échelle des images naturelles et on réduit cette redondance. L'utilisation de cette nouvelle définition permet d'effectuer une décomposition hiérarchique des images suivant la quantité d'information contenue localement. Similairement, l'utilisation d'un modèle d'analyse multifractale permet d'obtenir une segmentation des images en regard d'une mesure qui quantifie la régularité locale du signal. Notre objectif est de mettre en évidence la corrélation entre ces deux approches : l'approche probabiliste de la mesure entropique et l'approche géométrique du formalisme multifractal fournissent une caractérisation locale des structures de l'image qui s'avère similaire. Nous montrons notamment que la hiérarchie des sous-ensembles fractals est reliée à une description entropique de l'image. En particulier, la connaissance de l'un de ces ensembles, proche des maxima locaux de l'entropie multiéchelle, permet la reconstruction de l'image entière. Le formalisme multifractal permet ainsi de remédier aux la-

cunes de l'entropie en prenant en compte l'information topologique à différentes échelles et l'information statistique à différents ordres. Nous sommes finalement en mesure de proposer une interprétation multi-sémantique (en terme de géométrie, de statistique et d'information) des structures et objets détectés dans l'image à l'aide du modèle multifractal.

Mots Clef

segmentation, multiéchelle, information, entropie, multifractal, singularité, reconstruction.

Abstract

In this paper, we are interested with the characterization and the extraction of meaningful structures in images. The more natural way to measure their contribution to the image description consists in using the entropy defined by Shannon which quantifies the information conveyed by subsets. The local estimation of entropy provides a spatially localized, but redundant, measure of the information : we propose to generalize it in a multiscale framework, in order to fit the scale-invariance property of natural images and to reduce this redundancy. This new definition allows to decompose the image according to the local content of information. Similarly, in the context of image segmentation, the use of a multifractal model enables to decompose an image according to the local regularity of the signal. The main purpose of the paper is to show that both approaches are related : the entropy-based concepts and the geometrical approach of the multifractal formalism provide a similar characterization of local structures in images. Namely, the segmentation into fractal subsets is

in good correspondance with the hierarchical decomposition obtained through the entropic description. In particular, the knowledge of one of the fractal subsets, which is quite close to the maxima of the local entropy, allows to reconstruct the image. The use of the multifractal formalism prevents from the limitations of the entropy as it takes into account the topological information at different scales and the statistical information at different orders. These considerations show that the multifractal model enables to supply structures and objects in image with a multi-semantic interpretation (from geometrical, statistical and information theory points of view).

Keywords

segmentation, multiscale, information, entropy, multifractal, singularity, reconstruction.

1 Introduction et motivation

La détection de structures d'intérêt dans une image nécessite tout naturellement de préciser ce que l'on recherche dans cette image et, en conséquence, de définir des quantités mesurant la pertinence des objets à extraire. À ces fins, les concepts issus de la théorie de l'information [1] peuvent être utilisés dans l'analyse et l'interprétation des images [13].

La théorie de l'information est basée sur l'idée qu'un signal peut être vu comme la réalisation d'une variable aléatoire, suivant une certaine distribution de probabilité et apportant une certaine quantité d'information. L'objectif principal de cette théorie est alors d'établir un lien entre la probabilité des événements observés dans le signal et l'information fournie par le signal lui-même [1]. Ainsi, dans un cadre 2D, on considère généralement qu'une image est riche en information si elle met en évidence beaucoup de structures et objets significatifs et variés ; au contraire, elle est pauvre en information si elle contient très peu d'objets ou si elle est uniforme. La théorie de l'information fournit des quantités permettant d'évaluer cette information. En particulier, l'entropie estimée à partir des niveaux de gris [14] : plus celle-ci est grande, plus le nombre de bits nécessaires au codage de l'image est important. Cette mesure trouve des applications non seulement dans la compression et le codage [9] mais aussi le filtrage et l'analyse d'images [20]. Cependant, les méthodes classiques exploitant l'entropie l'interprètent essentiellement comme une mesure globale et ne traitent pas localement la notion d'information [15].

Sporring et Weickert soulignent dans [15] le lien entre les différentes représentations du signal que sont : les entropies généralisées, le spectre multifractal, les moments des niveaux de gris et l'histogramme des niveaux

de gris lui-même. Krim et Brooks [8] relient la notion d'entropie au sens de Shannon avec des concepts multiéchelles en terme d'ondelettes afin de définir un critère de segmentation automatique. Dans [5], Grazzini *et al.* relient les résultats obtenus à l'aide d'une méthode multiéchelle d'estimation de l'entropie locale dans des images météorologiques à un modèle de décomposition multifractale. Ils mettent en évidence l'existence d'un sous-ensemble qui :

- du point de vue de la *théorie de l'information*, est l'ensemble le plus riche en information dans l'image,
- du point de vue de la *théorie multifractale*, autorise la reconstruction de l'image.

Dans la continuité de ces travaux, nous comparons qualitativement les méthodes et nous donnons une justification quantitative des résultats sur des images naturelles. Le formalisme multifractal offre une description géométrique et statistique des images qui conduit à une décomposition hiérarchique. Les mesures entropiques confirment que les structures ainsi extraites sont topologiquement et statistiquement cohérentes au sens où elles sont mises en évidence par des approches différentes.

L'article est organisé comme suit : dans la section 2, nous rappelons les méthodes d'estimation de l'entropie et nous soulignons leurs limitations ; la section 3 présente le modèle d'analyse multifractale et l'algorithme de reconstruction ; dans la section 4, nous comparons et discutons les résultats des deux approches, nous montrons que celles-ci sont liées, mais que l'approche multifractale est davantage adaptée car elle permet d'adopter un point de vue à la fois géométrique et statistique ; la dernière section est consacrée à la conclusion et aux perspectives offertes par ce travail.

2 Mesure de l'information dans les images

2.1 Information et entropie

L'entropie au sens de Shannon [14] permet d'évaluer quantitativement l'information dans une image. La définition classique de l'entropie suppose que l'ensemble de l'information réside dans les distributions marginales des niveaux de gris de l'image (ce qui revient à supposer une invariance translationnelle des images). À partir de l'histogramme des fréquences marginales $p_i = \sum \delta_{y|I(y)=i}$ des niveaux de gris de l'image I (δ est la mesure de comptage), on définit l'entropie globale (EG) s_g :

$$s_g = - \sum_i \frac{p_i}{N} \log_2 \frac{p_i}{N} \quad (1)$$

où N est le nombre total de pixels de l'image. Comme le logarithme est exprimé en base 2, l'**EG** représente le nombre optimal de bits nécessaires au codage des niveaux de gris. Cette construction établit en particulier que l'entropie totale S_g du signal est proportionnelle à s_g : $S_g = N s_g$, i.e. que l'entropie par pixel est égale à l'**EG** ; la quantité obtenue est exprimée en *bits par pixel (bpp)*. Le modèle de probabilité retenu pour l'image est donc celui d'un processus indépendant identiquement distribué : tous les pixels sont obtenus indépendamment les uns des autres à partir de la même distribution. L'**EG** est égale à l'entropie de la variable aléatoire sous-jacente à chacun des pixels de l'image.

2.2 Entropie locale

L'entropie ainsi définie est évaluée globalement sur l'image et ne permet pas de mesurer la quantité d'information contenue dans des structures locales. Elle ne prend en compte aucun facteur contextuel ni aucune connaissance *a priori* sur l'image : les deux images de la figure 1 ne contiennent (de toute évidence) pas la même information, l'image de Lena mettant en évidence des textures et objets définis par leurs contours, elles ont pourtant la même **EG**. On préfère par conséquent au modèle d'**EG** un modèle d'entropie locale, dérivé des méthodes classiques de calcul d'attributs texturels et basé sur les distributions locales des niveaux de gris, dans des voisinages de taille fixe [11]. L'*entropie locale (EL)* est évaluée en un pixel $\mathbf{x}$ par :

$$s_l(\mathbf{x}) = - \sum_i \frac{p_i(\mathbf{x})}{|W|} \log_2 \frac{p_i(\mathbf{x})}{|W|} \quad (2)$$

où $W(\mathbf{x})$ désigne une fenêtre centrée en $\mathbf{x}$, avec $|W| = \sum \delta_{\mathbf{y} \in W(\mathbf{x})}$ le nombre de pixels de cette fenêtre, et où les probabilités empiriques :

$$p_i(\mathbf{x}) = \sum \delta_{\mathbf{y} \in W(\mathbf{x}) | I(\mathbf{y})=i} \quad (3)$$

représentent la distribution marginale locale des niveaux de gris dans la fenêtre $W(\mathbf{x})$ (figure 2). En faisant l'hypothèse que le signal est stationnaire et que les fenêtres analysantes permettent de décomposer l'image en ensembles indépendants, on peut estimer l'entropie d'une région particulière, à partir d'une sommation sur l'ensemble des pixels de cette région :

$$S_\varphi = \sum_{\mathbf{x} \in \varphi} s_l(\mathbf{x}) . \quad (4)$$

En particulier, l'entropie totale s'exprime comme une somme des **EL** sur toute l'image : $S_l = \sum_{\mathbf{x}} s_l(\mathbf{x})$, ce qui signifie, de manière équivalente, que l'entropie par pixel $\bar{s}_l$ est égale à l'**EL** moyenne : $S_l = N \bar{s}_l$. Notons que l'hypothèse de partition en ensembles indépendants



FIG. 1: Deux images contenant une information différente mais ayant la même entropie : $S = 7.40 \text{ bpp}$. L'image de droite est simplement obtenue en réarrangeant spatialement les niveaux de gris de l'image de Lena. Les deux images ont été quantifiées sur 256 niveaux de gris.

est évidemment fausse, d'autant plus que les fenêtres considérées se superposent, mais cette méthode fournit une meilleure estimation de l'entropie que la première.

2.3 Entropie multiéchelle

L'information présente dans une image dépend de l'échelle d'observation. Par conséquent, il est nécessaire d'extraire l'information à différentes échelles. L'approche multiéchelle a donné des résultats probants en compression et codage d'image [9], en analyse et extraction d'attributs [20] ; il apparaît alors naturel d'adopter également un traitement multiéchelle pour calculer l'entropie. L'idée est considérer les variations de l'entropie à travers les échelles pour étudier l'information bas-niveau dans l'image et pour identifier des structures, telles que des formes ou des textures, à des échelles différentes. Jägersand [7] procède ainsi à une analyse de l'information contenue dans les représentations *scale-space* d'une image. L'idée d'appliquer la théorie de l'information aux représentations multirésolutions a également été discutée par Sporring et Weickert, qui s'intéressent à l'évolution des entropies généralisées à travers les échelles [15]. Cependant, ils se concentrent sur des mesures globales de l'information et ne traitent pas de son contenu local. Ferraro *et al.* [3] définissent en revanche une mesure locale qui leur permet de segmenter une image, également à partir de ses représentations *scale-space*, et déterminent la production d'entropie (d'information) à travers les échelles. Cette méthode permet de discriminer différentes régions en fonction de l'information qu'elles contiennent. Cependant, elle nécessite de calculer des représentations multirésolutions des images, ce qui accroît encore la complexité en temps de calcul. L'objection principale à ce type d'approche réside

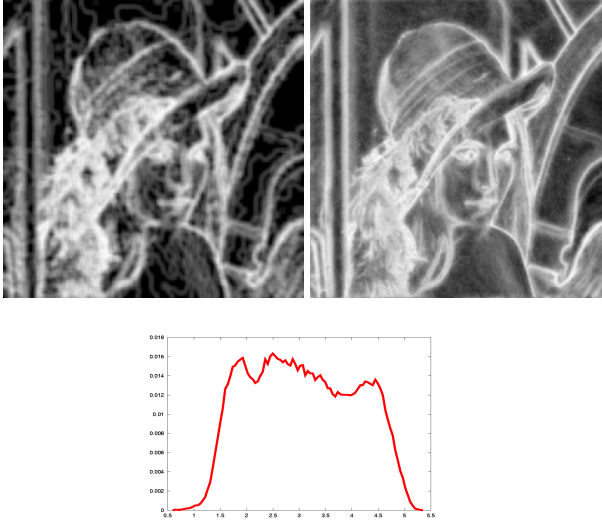


FIG. 2: **En haut** : **EL** et **ELM** calculées sur l'image de Lena ; à gauche, l'entropie est estimée sans pondération sur des fenêtres de taille 10×10 ; à droite, on introduit une pondération multiéchelle sur des fenêtres de taille 21×21 (cf. section 2.3). **En bas** : distribution de l'**ELM** sur l'image.

par ailleurs dans la redondance des représentations *scale-space*. Pour notre part, nous entendons considérer toutes les échelles simultanément dans l'estimation de l'entropie.

On peut affiner la définition (2) pour répondre au paradigme d'échelle grâce à l'introduction d'une fonction de pondération spatiale invariante d'échelle dans l'estimation des distributions locales (3) des niveaux de gris. Au lieu de fenêtres uniformes de taille fixe, les pixels intervenant dans le calcul des distributions locales sont pondérés par une fonction qui décroît en loi de puissance de la distance au pixel d'intérêt :

$$p_i(\mathbf{x}) = \sum_{\mathbf{y} \in W^*(\mathbf{x}) | I(\mathbf{y})=i} |\mathbf{x} - \mathbf{y}|^{-2} \quad (5)$$

où la pondération est effectuée dans la fenêtre $W^*(\mathbf{x})$ privée du pixel d'intérêt $\mathbf{x}$ lui-même (on lui assigne un poids nul). L'*entropie locale multiéchelle (ELM)* est alors calculée comme en (2) avec le facteur de normalisation $|W^*| = \sum \delta_{\mathbf{y} \in W^*(\mathbf{x})} |\mathbf{x} - \mathbf{y}|^{-2}$. De même, l'entropie par pixel est égale à l'**ELM** moyenne.

En introduisant une telle pondération, on fait davantage contribuer les valeurs de I des pixels proches du pixel central $\mathbf{x}$ que les valeurs des pixels éloignés. Afin d'avoir un échantillonnage assez fin du signal et de limiter l'erreur de calcul sur la probabilité estimée, il convient de traiter un nombre suffisamment important de pixels, i.e. un histogramme de taille suffisante.

Expérimentalement, les valeurs de l'**ELM** changent

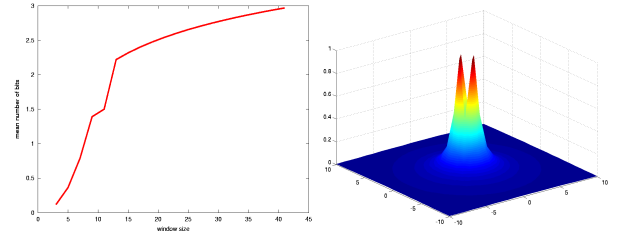


FIG. 3: **À gauche** : évolution de la mesure d'**ELM** (en bits) en fonction de la taille de la fenêtre expérimentale choisie pour l'estimation des probabilités marginales sur l'image de Lena. **À droite** : facteur de pondération multiéchelle finalement choisi sur une fenêtre de rayon 21×21 pixels.

très peu significativement pour des tailles de fenêtres plus grandes que 21×21 du fait de la pondération qui décroît très vite (figure 3). On choisit par ailleurs une discrétisation de I sur $2^6 = 64$ niveaux de gris à l'intérieur de chaque fenêtre. Le choix de ces paramètres n'implique pas de perte en localisation. Il pourrait cependant être optimisé de manière à obtenir le meilleur compromis entre la discrétisation du signal et la taille de la fenêtre [13]. Notons par ailleurs que la discrétisation du signal revient à échantillonner à des valeurs espacées de Δ le signal convolué avec un filtre adapté de largeur de bande Δ . Cette remarque est à l'origine des représentations *scale-space* pour un calcul multiéchelle de l'entropie [15].

L'avantage de notre approche est de pouvoir garantir la stationnarité des données sélectionnées, mêmes dans les zones de l'image comportant des structures très fines ou dans des zones texturées et de considérer un nombre d'échantillons suffisamment grand pour estimer plus précisément les statistiques. Ce modèle permet comme le précédent d'étudier les variations spatiales de l'entropie, mais il atteint en plus un compromis entre localisation et qualité de l'estimation (figure 2).

2.4 Limitations des mesures entropiques

Les statistiques utilisées pour l'estimation de l'entropie sont basées sur des moments d'ordre 1 du signal. Par conséquent, on ne tient pas compte des dépendances mutuelles et de l'éventuelle corrélation entre les pixels. Si l'on désirait le faire, il faudrait évaluer une entropie associée aux distributions conjointes des niveaux de gris : une telle tâche est impossible en terme de capacité mémoire, même pour des images de petite taille. Considérer uniquement les niveaux de gris de l'ensemble n'est cependant pas satisfaisant [16]. En effet, d'une part, l'**EL** est calculée sur des voi-

sinages bidimensionnels (typiquement, les ensembles qui nous intéressent sont en fait uni-dimensionnels ou fractals [19]), d'autre part, il est impossible de connaître l'importance relative d'une structure dans une image sans information supplémentaire sur les autres ensembles. La part d'information fournie par un sous-ensemble à toute l'image dépend également de l'information que lui-même reçoit de tous les autres sous-ensembles. Il faudrait tenir compte de ces informations relatives dans le calcul des informations locales et de l'information totale.

La formule (4) permet d'accéder à la part d'information relative apportée par un sous-ensemble d'une image en évaluant la part d'entropie qu'il contient localement. Cependant, l'**EL** reste une mesure redondante de la quantité d'information : elle est calculée sur des voisinages de pixels qui se superposent. Un ensemble qui contient la totalité de l'information de l'image peut très bien ne pas contribuer majoritairement à l'entropie calculée à partir des **EL** : cet ensemble peut en effet accroître de façon non négligeable l'estimation de l'**EL** de pixels ne lui appartenant pas mais ayant un voisinage en commun. La pondération multiéchelle tend à diminuer cette contribution, mais ce cas de figure est toujours possible. L'**ELM** est toujours une mesure redondante de l'information. Il est en fait nécessaire de prendre en compte l'influence de chaque ensemble séparément pour améliorer le critère de mesure de l'information.

Le caractère fortement non-gaussien des statistiques des images naturelles et leur propriété d'invariance d'échelle [16] justifie que l'on s'intéresse *a priori* aux statistiques d'ordre supérieur du signal et que l'on incorpore de l'information topologique dans l'analyse des structures présentes dans les images.

3 Formalisme multifractal

3.1 Les statistiques des images naturelles

Les images naturelles, *i.e.* sans organisation particulière des objets, ont souvent été décrites comme mettant en évidence un comportement fractal : elles ne possèdent pas d'échelle propre et sont auto-similaires. Cette propriété est principalement mise en évidence et vérifiée sur le seul spectre de puissance ; en effet, expérimentalement, il est statistiquement observé (*i.e.* sur de grandes collections d'images) que la puissance spectrale Γ présente une décroissance en loi de puissance du type [4] :

$$\Gamma(f) \sim |f|^{-(d-\epsilon)} \quad (6)$$

où $d = 2$ est la dimension de l'espace et où f est la fréquence de norme $|f|$. L'exposant ϵ est généralement très petit et dépend des particularités de l'échantillon

d'images considérées [12]. On parle de comportement auto-similaire.

Cependant, cette propriété ne suffit pas à décrire, de manière précise et détaillée, la structure et le contenu local des images naturelles [12]. Ceci provient essentiellement du fait que l'auto-similarité revient à assigner le même exposant d'échelle à tous les pixels de l'image. Il est donc nécessaire, afin d'obtenir une meilleure description des images, de définir localement une variable elle-même capable de décrire des caractéristiques statistiques locales, de la même manière que nous avons dû introduire une mesure locale de l'entropie. L'idée sous-jacente est de pouvoir décrire, à l'aide d'une seule variable définie sur toute l'image, les différents comportements d'échelle des pixels de l'image [19].

3.2 Présentation d'un modèle multifractal

Le formalisme multifractal permet d'établir un lien entre une description géométrique, faisant appel à la notion de dimension fractale d'un ensemble de points, et une description statistique, utilisant les lois d'échelle des moments. Dans [20], Turiel et Parga ont proposé un modèle multifractal de décomposition hiérarchique des images en sous-ensembles fractals. La décomposition obtenue conduit à extraire un ensemble de pseudo-contours et à exhiber également des structures texturées. Ce modèle est basé sur l'analyse d'une mesure

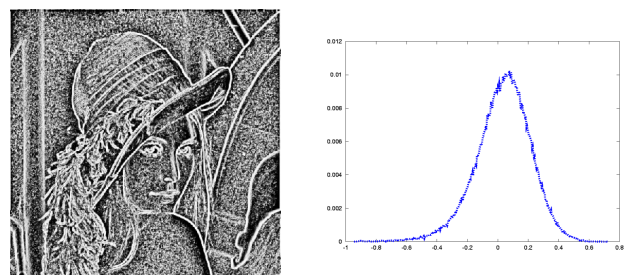


FIG. 4: **À gauche** : image des singularités ; un pixel est d'autant plus clair que son exposant de singularité locale est petit. **À droite** : histogramme de distribution des exposants de singularité.

de la variabilité locale des niveaux de gris. Celle-ci est définie, pour un sous-ensemble quelconque φ de l'image I , par :

$$\mu(\varphi) = \int_{\varphi} d\mathbf{x} |\nabla I|(\mathbf{x}), \quad (7)$$

où $|\nabla I|(\mathbf{x})$ désigne la norme du gradient de l'image. L'évaluation de la mesure $\mu(B_r(\mathbf{x}))$ de boules de rayon r pour différents r permet de mettre en évidence

le comportement en loi d'échelle de la mesure μ :

$$\mu(B_r(\mathbf{x})) \sim r^{h(\mathbf{x})+2} \quad (8)$$

Cette propriété de dépendance, qui caractérise une mesure multifractale, a été démontré pour de grandes collections d'images naturelles [20] et est en fait liée à l'invariance d'échelle des images. L'exposant de singularité local $h(\mathbf{x})$, qui ne dépend que de $\mathbf{x}$, représente le degré de singularité (cas < 0) ou de régularité (cas > 0) du signal $|\nabla I|$. L'estimation des exposants $h(\mathbf{x})$ par des méthodes appropriées [20] permet de décomposer l'image en affectant chaque pixel de l'image au sous-ensemble F_h associé à la valeur h de son exposant de singularité :

$$F_h \equiv \{\mathbf{x} \mid h(\mathbf{x}) \approx h\}.$$

Chacun des sous-ensembles F_h se révèle être un ensemble fractal, *i.e.* il exhibe la même structure géométrique à différentes échelles. L'arrangement des pixels de l'image dans les différentes composantes fractales de l'image ainsi définies est caractéristique de la structure géométrique de l'image et est reflété par les propriétés statistiques du signal.

Parmi tous les ensembles fractals F_h considérés, un sous-ensemble est particulièrement important du point de vue géométrique et statistique. Il s'agit de la **MSM** (pour *Most Singular Manifold*, la variété la plus singulière), la composante F_∞ associée à la singularité la plus forte, *i.e.* celle qui contient les pixels d'exposant h_∞ minimal, ou qui s'en approchent (figure 5). La théorie multifractale nous dit que les caractéristiques statistiques de cette structure lui confèrent des propriétés à la fois en terme de physique et de théorie de l'information [19].

3.3 Reconstruction

Dans son livre sur la vision, Marr [10] a émis la conjecture selon laquelle il est possible de reconstruire une image à partir de ses contours multiéchelles. Turiel et del Pozo présentent dans [18] un algorithme de reconstruction à partir de la **MSM**, considérée pour ses propriétés statistiques, et de la seule connaissance du gradient de l'image sur cet ensemble. Cet algorithme est basé sur des propriétés statistiques et est supposé produire une reconstruction parfaite à partir de cet ensemble. La formule de reconstruction s'exprime ainsi :

$$I(\mathbf{x}) = \vec{g} \star \vec{v}_\infty(\mathbf{x}). \quad (9)$$

où $\vec{g}(\mathbf{f}) = i \mathbf{f}/|\mathbf{f}|^2$ est un noyau de reconstruction universel ayant de "bonnes" propriétés (il est déterministe, linéaire, invariant par translation, isotrope et conserve

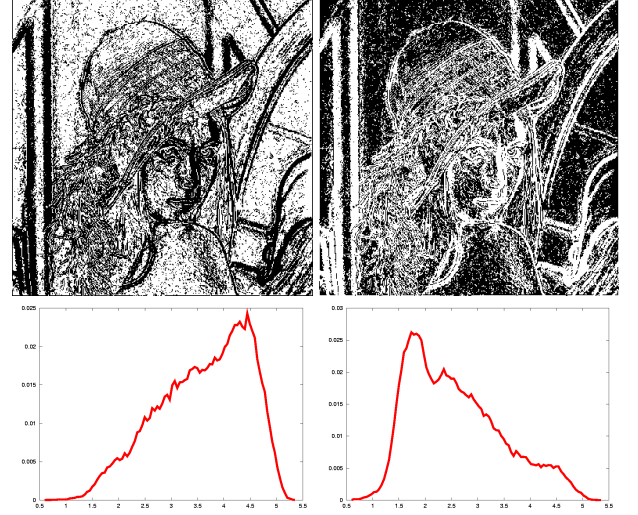


FIG. 5: **En haut** : la **MSM** (42.8% des pixels de l'image) et l'ensemble complémentaire ; **En bas** : distributions de l'ELM sur ces ensembles.

le spectre de puissance). Le champ de vecteurs $\vec{v}_\infty$, appelé *gradient essentiel*, est égal au gradient de l'image originale sur la **MSM** et est nul en dehors :

$$\vec{v}_\infty(\mathbf{x}) = \nabla I(\mathbf{x}) \delta_{F_\infty}(\mathbf{x}). \quad (10)$$

Le principe est en fait celui d'un processus de diffusion du gradient de l'image sur les pixels de la **MSM**, qui permet d'obtenir une estimation du signal original.

3.4 Propriétés

La formule de reconstruction (9) peut en fait s'appliquer à n'importe quel sous-ensemble de l'image, *i.e.* on peut chercher à reconstruire un ensemble à partir d'une région quelconque de l'image. Il suffit en effet de définir un gradient $\vec{v}_\varphi(\mathbf{x}) = \nabla I(\mathbf{x}) \delta_\varphi(\mathbf{x})$ restreint à un sous-ensemble φ et d'appliquer la formule (9) avec le même noyau de propagation :

$$I_\varphi(\mathbf{x}) = \vec{g} \star \vec{v}_\varphi(\mathbf{x}). \quad (11)$$

L'algorithme de reconstruction présente l'avantage d'être linéaire. L'image obtenue par reconstruction à partir de l'ensemble $\varphi_1 \cup \varphi_2$ (avec $\varphi_1 \cap \varphi_2 = \emptyset$) est la somme des reconstructions de chaque sous-ensemble φ_1 et φ_2 : $I_{\varphi_1 \cup \varphi_2} = I_{\varphi_1} + I_{\varphi_2}$. Par ailleurs, la reconstruction est parfaite si l'ensemble considéré est constitué de la totalité des points de l'image. De cette manière, l'algorithme décompose l'image en images non redondantes. Cependant, tous les ensembles utilisés pour la reconstruction n'ont pas la même importance : la **MSM** est l'ensemble le plus pertinent. En théorie, la

reconstruction à partir de la **MSM** est exacte : $I_{F_\infty} = I$; en pratique, la **MSM** est l'ensemble à partir duquel on obtient le meilleur résultat de reconstruction, et celle-ci peut toujours être améliorée en considérant en sus dans le processus de reconstruction, les autres composantes fractales [17]. Notons que cela signifie en particulier que les différentes composantes fractales sont, en un sens, corrélées à la **MSM** puisqu'elles peuvent être prédites par l'application d'un noyau linéaire sur celle-ci.

4 Résultats et discussion

Nous présentons ici les résultats de l'analyse réalisée sur l'image de Lena discrétisée sur 256 niveaux de gris (figure 1). Nous employons l'algorithme multifractal de décomposition et de reconstruction développé dans [18] et présenté dans la section 3 ; nous nous intéresserons à l'extraction de deux ensembles seulement : la **MSM** et son complémentaire. Notons que cette approche peut très facilement être adoptée pour étudier des images de nature quelconque, et notamment des images en couleur : l'analyse et la reconstruction doivent alors être effectuées sur les différentes composantes vectorielles de l'image [21].

L'**EG** calculée sur l'image de Lena vaut 7.40 bpp , ce qui est relativement proche de la valeur maximale possible pour une image sur 256 niveaux de gris (8 bpp). Cependant, l'hypothèse de distribution uniforme est clairement mise en défaut. L'estimation de l'**ELM** constitue un bon moyen de caractériser dans l'image les structures significatives, simplement en calculant la part d'entropie totale qu'elles fournissent à toute l'image. L'**ELM** (figure 2) est maximisée sur les contours de l'image. Cette mesure semble bien plus significative et confirme que l'information n'est pas uniformément distribuée dans l'image. L'**EL** moyenne est notamment de 2.74 bpp , bien moins que l'**EG**, ce qui montre qu'elle est moins redondante. Cependant, cette mesure étant calculée sur des voisinages qui se superposent, elle reste redondante.

Le modèle multifractal permet d'extraire la composante F_∞ associée aux singularités les plus fortes du signal. Expérimentalement, on détermine la valeur h_∞ de l'exposant le plus singulier, à partir de la distribution des singularités, comme la moyenne des exposants aux quantiles à 1 % et 5 % ; on fixe également conventionnellement la dispersion autour de cette valeur afin de déterminer l'ensemble des points appartenant finalement à la **MSM**. Les figures 2 et 5 soulignent la bonne correspondance qu'il existe entre la variété la plus singulière et les maxima locaux de l'**ELM**, tenant compte du fait que les deux techniques employées sont très différentes. Les résultats peuvent être affinés en ajustant



FIG. 6: **En haut** : reconstructions à partir des 2 ensembles de la figure 5 ; ($PSNR = 28.24 \text{ dB}$ pour la reconstruction à partir de la **MSM**, $PSNR = 12.42 \text{ dB}$ lorsqu'elle est effectuée à partir de l'ensemble complémentaire). **En bas** : **ELM** correspondantes.

les paramètres de détermination de la **MSM**. Pour notre exemple, la **MSM** extraite est constituée de 44.15 % des pixels de l'image (figure 5). Cette correspondance confirme le rôle de la **MSM** comme ensemble le plus informatif dans l'image au sens où la reconstruction à partir de ce seul ensemble va être possible. L'algorithme multifractal de reconstruction étant linéaire, la reconstruction à partir de la somme des différents ensembles est égale à la somme des reconstructions à partir des ensembles considérés séparément. La reconstruction à partir du complémentaire de la **MSM** est notamment égale au complémentaire de la reconstruction à partir de la **MSM** (la somme des deux reconstructions est égale à I), et peut ainsi être vue comme une image d'erreur. Les **EG** calculées sur les images reconstruites à partir de ces deux ensembles valent ainsi respectivement 7.42 bpp et 6.74 bpp : on voit là encore les limitations de l'**EG** qui ne permet pas d'appréhender l'organisation des structures dans les images, puisqu'en l'occurrence les deux reconstructions n'apportent pas la même information vis-à-vis de l'image originale. On observe en effet que la reconstruction à partir de la **MSM** est bien meilleure que la reconstruction à partir de l'ensemble complémentaire (image d'erreur sur laquelle on ne peut discerner que des micro-textures

qui ne correspondent pas à des structures significatives de l'image originale), comme le montre le ratio $PSNR$ (figure 6). En revanche, l'**ELM** calculée sur les pixels de la **MSM** vaut 2.44 bpp , et est donc proche de l'**ELM** de l'image originale, alors qu'elle ne vaut que 1.49 bpp sur les pixels de l'ensemble complémentaire. Cette quantité est donc beaucoup plus significative. Cependant, la valeur de l'**ELM** de l'image reconstruite à partir du complémentaire de la **MSM** indiquerait que cet ensemble fournit encore une part d'information non négligeable sur certaines structures locales de l'image. Une mesure plus appropriée de la fidélité de la reconstruction consiste à considérer des probabilités conditionnelles [13]. Une quantité intéressante est notamment fournie par l'**ELM** conditionnelle (**ELMC**) : celle-ci est définie comme la différence entre l'**ELM** des distributions conjointes des niveaux de gris de l'image originale et de l'image reconstruite (à partir de probabilités conjointes estimées similairement à l'équation 5), et l'**ELM** de l'image originale [1]. L'**ELMC** fournit une mesure de la déviation de l'image reconstruite par rapport à l'image originale ; elle s'annule uniquement quand les deux images sont égales. En fait, cette quantité est exactement égale à l'**ELM** de l'erreur conditionnée par l'image originale : elle mesure donc fidèlement la part d'information qui n'est pas apportée par l'image reconstruite. Ainsi, lorsque l'on calcule cette quantité pour la reconstruction à partir de la **MSM** (l'**ELM** vaut alors en moyenne 0.89 bpp seulement), elle apparaît davantage concentrée spatialement sur les régions qui sont effectivement mal reconstruites, et elle semble nettement moins affectées par des perturbations locales (figure 7).

Il apparaît donc que la description géométrique de l'image adoptée par l'algorithme d'analyse multifractale permet d'accéder aux structures et sous-ensembles les plus informatifs dans l'image. L'autre avantage majeur de cet algorithme réside dans le fait qu'il fournit une décomposition non redondante de l'image, dans laquelle la **MSM** est l'ensemble le plus significatif.

5 Conclusion et Perspectives

Nous nous intéressons dans cet article à deux méthodes d'extraction de structures d'intérêt dans une image : une première méthode utilisant une mesure entropique et une autre adoptant une description multifractale. La méthode d'estimation de l'entropie locale que nous présentons est basée sur une approche multiéchelle et utilise les distributions marginales des niveaux de gris dans des voisinages non disjoints de l'image. L'introduction d'un critère local permet d'opérer une classification des pixels de l'image en fonction de la quantité d'information qu'ils lui confèrent et d'extraire locale-

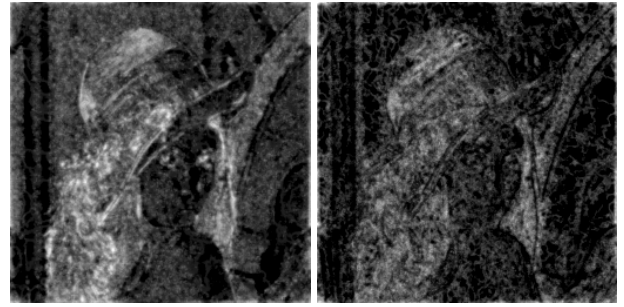


FIG. 7: **ELM** de l'erreur et **ELMC** de l'erreur conditionnellement à l'image originale.

ment les structures les plus informatives du signal : elle apparaît, en ce sens, plus adaptée à l'analyse d'images que les définitions classiques. Cependant, elle constitue une mesure redondante. La méthode multifractale vise quant à elle à vérifier une dépendance en loi de puissance d'une certaine mesure de la régularité locale du signal suivant un paramètre d'échelle. Elle permet ainsi de décomposer l'image en sous-ensembles ayant des propriétés géométriques similaires.

En regard des résultats obtenus, nous avons vu que les maxima de l'entropie locale sont en bonne correspondance avec les points d'une structure particulière dans l'image, la **MSM**. Cette structure, qui correspond principalement aux contours des objets, est fractale et elle constitue l'ensemble le plus informatif de l'image ; il n'est donc pas surprenant, que dans une optique de codage et de compression, elle soit utilisée pour la reconstruction du signal [18]. La méthode multifractale, qui permet d'extraire cette structure particulière avec une bonne localisation, remédie donc aux lacunes de l'entropie. Elle permet notamment de réduire la redondance de la mesure entropique locale car elle tient, en fait, compte à la fois de l'information géométrique à différentes échelles et de l'information statistique à différents ordres. Pour généraliser ces résultats, on cherchera à établir quantitativement la relation entre les deux approches sur de plus grandes collections d'images [6].

Les considérations précédentes sont finalement d'un grand intérêt pour l'interprétation des structures observées dans les images naturelles. En effet, l'acceptation du terme "entropie" est multiple, et les mesures effectuées à l'aide de cette quantité peuvent être appréhendées de diverses façons. En conséquence, l'interprétation des structures détectées à l'aide du modèle multifractal est, elle aussi, multi-sémantique : ces structures traduisent (et caractérisent) des propriétés à la fois en termes statistique et géométrique et en terme de théorie de l'information. Ainsi, dans le contexte de

l'analyse d'images météorologiques [5], cette approche devrait nous permettre d'extraire de l'information de nature thermodynamique sur les systèmes complexes observés. En l'occurrence, on pourra espérer relier l'entropie telle qu'elle est définie ici avec l'entropie réelle des phénomènes atmosphériques sous-jacents.

Remerciements

J. Grazzini est financé par une bourse attribuée par le Conseil Régional d'Île-de-France (accord n° E.1358). A. Turiel est financé par un contrat avec la Generalitat de Catalunya.

Références

- [1] T. M. Cover et J. A. Thomas, "Elements of Information Theory", John Wiley, New York, 1991.
- [2] R. M. Haralick and K. Shanmugam and I. Dinstein, "Textural features for image classification", *IEEE Trans. on Systems, Man & Cybernetics*, 3 :610-621, 1973.
- [3] M. Ferraro, G. Boccignone, T. Caelli, "Entropy-based representation of image information", *Pattern Recognition Letters*, 23 :1391-1398, 2002.
- [4] D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells", *Journal of the Optical Society of America A*, 4 :2379-2394, 1987.
- [5] J. Grazzini, A. Turiel and H. Yahia, "Entropy Estimation and Multiscale Processing in Meteorological Satellite Images", *Proc. of International Conference on Pattern Recognition*, 3 :764-768, 2002.
- [6] J. Grazzini, A. Turiel and H. Yahia, "Information content associated to edges and textures", Soumis à *International Conference on Acoustics, Speech, and Signal Processing*, 2004.
- [7] M. Jägersand, "Saliency maps and attention selection in scale and spatial coordinates : An information theoretic approach" *Proc. of International Conference on Computer Vision*, 195-202, 1995.
- [8] H. Krim et D.H. Brooks, "Feature based segmentation of ECG signals", *IEEE International Symposium on Time/Frequency Analysis*, 1996.
- [9] S. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11 :674-693, 1989.
- [10] D. Marr, "Vision", W.H. Freeman and Co., New York, 1982.
- [11] T. Randen et J.H. Husoy, "Filtering for texture classification : a comparative study", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(4) :291-310, 1999.
- [12] D. Ruderman, "The statistics of natural images", *Network*, 5 :517-548, 1994.
- [13] B. Schiele et J.L. Crowley, "Transformation for Active Object Recognition", *Proc. of International Conference on Computer Vision*, 249-254, 1998.
- [14] C.E. Shannon, "A mathematical theory of communication", *The Bell System Technical Journal*, 27 :379-423, 1948.
- [15] J. Sporring et J. Weickert, "Information measures in scale-space", *IEEE Trans. on Information Theory*, 45(3) :1051-1058, 1999.
- [16] A. Srivastava, A.B. Lee, E.P. Simoncelli et S.-C. Zhu, "On advances in statistical modeling of natural images", *Journal of Mathematical Imaging and Vision*, 18(1) : 17-33, 2003.
- [17] A. Turiel, "Relevance of multifractal textures in static images", *Electronic Letters on Computer Vision and Image Analysis*, 1(1) :35-49, 2003.
- [18] A. Turiel et A. del Pozo, "Reconstructing images from their most singular fractal manifold", *IEEE Trans. on Image Processing*, 11 :345-350, 2002.
- [19] A. Turiel, G. Mato, N. Parga et J.-P. Nadal, "The Self-Similarity properties of natural images resemble those of turbulent flows", *Physical Review Letters*, 80(5) :1098-1101, 1998.
- [20] A. Turiel et N. Parga, "The multi-fractal structure of contrast changes in natural images : from sharp edges to textures", *Neural Computation*, 12 :763-793, 2000.
- [21] A. Turiel, N. Parga, D. Ruderman et T. Cronin, "Multiscaling and information content of natural color images", *Physical Review E*, 62 :1138-1148, 2000.