

# Analyse et représentation ensembles-échelle d'une image

Laurent GUIGUES<sup>1</sup>, Hervé LE MEN<sup>1</sup>, Jean-Pierre COCQUEREZ<sup>2</sup>

<sup>1</sup>Institut Géographique National / Laboratoire MATIS  
2 Avenue Pasteur, 94165 Saint Mandé Cedex, France

<sup>2</sup>Université de Technologie de Compiègne / Heudiasyc UMR CNRS 6599  
BP 20529 Compiègne Cedex, 60205, France

{laurent.guigues,herve.le-men}@ign.fr, cocquerez@hds.utc.fr

**Résumé** – Cet article propose une approche multi-échelles de la segmentation d'image qui s'appuie sur un principe d'optimisation. Notre objectif est de calculer une séquence complète de partitions monotones qui approxime les minima d'une fonctionnelle dépendant d'un paramètre réel qui se comporte comme un paramètre d'échelle. Pour pouvoir manipuler toutes ces solutions, nous introduisons la représentation *ensembles-échelles*, qui constitue le pendant orienté régions des représentations scale-space issues du filtrage. Nous aboutissons à un algorithme efficace et sans paramètre, baptisé *escalade*, qui produit une description ensembles-échelle d'une image dont les coupes correspondent à toute une famille de partitions de finesse variable.

**Abstract** – We propose a multiscale approach to energy minimization-based image segmentation. Our goal is to compute a complete series of ordered partitions which approximates the minima of a functional depending on a real parameter which acts as a scale parameter. In order to capture all these solutions, we introduce the *scale-sets* representation, which is the region-oriented equivalent of the scale-space representation. We end up with an efficient parameter-free algorithm, called *scale climbing*, which produces a scale-sets description of an image whose cuts correspond to a whole family of solutions, from fine to coarse.

## 1 Introduction

Délimiter des unités «cohérentes» dans une image, ou segmenter l'image, est un problème fondamental d'analyse de bas niveau d'information visuelle. Ce problème est habituellement envisagé comme un problème de partitionnement : découper le domaine de l'image en régions mutuellement exclusives. Il est cependant évident que des unités cohérentes apparaissent à différentes échelles : texton/texture, feuille/arbre/forêt... et qu'aucune échelle d'analyse n'est intrinsèquement meilleure qu'une autre. Le choix d'une échelle appropriée dépend fondamentalement de la tâche de haut niveau à accomplir (interprétation, reconnaissance...). S'appuyant sur ce constat simple, la théorie scale-space propose qu'une analyse de bas niveau, pour pouvoir constituer une couche robuste de structuration d'information visuelle, doit rester indifférente à la position, à l'orientation et à l'échelle des phénomènes. Elle doit donc produire une analyse *continue* en échelle [7]. Cette nécessité est étroitement liée à des problèmes d'invariance des systèmes de vision et à la délicate question du paramétrage des algorithmes. Ce dernier point, aboutir à des méthodes automatiques d'analyse d'image, i.e. sans paramètre, a été présenté comme un «challenge» de la discipline en séance plénière du colloque Gretsi'01 [1].

Cet article propose une méthodologie de segmentation multi-échelles qui constitue le pendant orienté régions des approches scale-space du filtrage d'image. L'approche comporte deux aspects : un aspect *structurel* et un aspect *énergétique*, la résolution du problème d'optimisation que nous proposons dépend fondamentalement de la manière dont on peut représenter ses solutions.

## 2 Formalisation du problème

Considérons un algorithme de partitionnement d'image  $A$  dépendant d'un paramètre réel  $\lambda$ . En suivant Koepfler et al. [5], disons que  $A$  est un algorithme *multi-échelles* et que  $\lambda$  est un *paramètre d'échelle* si, pour une image donnée, les partitions  $A(\lambda)$  renvoyées par  $A$  sont *monotones* en fonction de  $\lambda$ , c'est-à-dire constituent une séquence de descriptions à finesse variable de l'image. Cette propriété de structure pyramidale constitue l'analogie orientée régions de l'*axiome de causalité* de la théorie scale-space [7]. Fixons-nous pour objectif d'obtenir un tel algorithme et d'en calculer *toutes les solutions* en fonction de  $\lambda$ .

Dans la lignée des approches variationnelles [8], Bayésiennes [2] ou par codage minimal [6], envisageons la segmentation comme un problème de modélisation (déterministe ou probabiliste) des observations. Pour la segmentation, le modèle recherché est défini par morceaux, i.e. s'appuie sur une partition de l'image en régions. Ces approches aboutissent toutes à l'idée que le problème est «mal posé» et doit être régularisé. Les formulations énergétiques obtenues prennent alors globalement la forme d'une opposition entre deux termes, un terme  $D$  d'attache aux données (ou vraisemblance/longueur de description des erreurs) et un terme  $C$  de régularisation (ou potentiel a priori/complexité de modèle). L'énergie d'une partition  $P$  s'écrit alors

$$E_\lambda(P) = \lambda C(P) + D(P). \quad (1)$$

Le paramètre  $\lambda$  correspond intrinsèquement à un réglage d'échelle. On peut en effet envisager l'énergie  $E_\lambda$  comme le Lagrangien du problème d'optimisation sous contrainte

$$\text{minimiser } D(P) \text{ sous la contrainte } C(P) \leq \gamma \quad (2)$$

Il est bien connu que les minima du Lagrangien 1 coïncident avec les solutions du problème 2, chaque valeur de contrainte  $\gamma$  correspondant à une valeur spécifique du paramètre de Lagrange  $\lambda$ . La formulation 2 montre qu'augmenter  $\lambda$  correspond à renforcer une contrainte sur la complexité  $\mathbf{C}$  de la solution recherchée. Donc, si  $\mathbf{C}$  attribue une plus forte «complexité» aux partitions les plus fines, augmenter  $\lambda$  doit conduire à des solutions de plus en plus grossières. On reconnaît également en 2 l'expression débit/distorsion classique d'un problème de compression avec perte. Les énergies considérées permettent donc de formuler des questions de compression d'image orientée régions ou de simplification/caricature optimale d'image [9].

Nous nous intéressons ici à la classe des énergies *séparables*, i.e. qui s'écrivent

$$E_\lambda(P) = \lambda \sum_{x \in P} \mathbf{C}(x) + \sum_{x \in P} \mathbf{D}(x). \quad (3)$$

De nombreuses énergies classiques sont séparables (fonctionnelle de Mumford-Shah [8], Champs de Markov utilisant un potentiel de Potts [2], critères MDL s'appuyant sur un codage de Freeman des partitions [6], etc.).

### 3 Représentation ensembles-échelle

Le premier problème que nous traitons est un problème de *représentation* : si l'ensemble des solutions  $\{A(\lambda)\}_{\lambda \in \mathbb{R}^+}$  renvoyées l'algorithme  $A$  que nous recherchons forme effectivement une séquence de partitions monotones, quelle structure permet de les représenter *toutes*? La question n'est pas triviale car  $\lambda$  est un paramètre *continu*. Cependant, l'ensemble  $H$  des parties produites par  $A$  quand  $\lambda$  balaye  $\mathbb{R}^+$  est fini, lui, et peut donc être explicité.

Nous proposons donc de représenter  $\{A(\lambda)\}_{\lambda \in \mathbb{R}^+}$  par la donnée de  $H$  et de la fonction  $\Lambda : H \rightarrow \mathcal{P}(\mathbb{R}^+)$  qui à chaque région  $x \in H$  associe

$$\Lambda(x) = \{\lambda \in \mathbb{R}^+ \mid x \in A(\lambda)\} \quad (4)$$

En français,  $\Lambda(x)$  représente l'ensemble des valeurs du paramètre d'échelle  $\lambda$  pour lesquelles  $x$  appartient à une partition renvoyée par l'algorithme  $A$ . On vérifie alors aisément que comme  $A$  est un algorithme multi-échelles :

i)  $H$  est une **hiérarchie de régions**, i.e.  $\forall (x, y) \in H^2 : x \cap y \in \{x, y, \emptyset\}$ . On peut donc la représenter par un arbre d'inclusions ;

ii) Les  $\Lambda(x)$  sont des **intervalles** de la forme

$$\Lambda(x) = [\lambda^+(x), \lambda^-(x)]. \quad (5)$$

Nous appelons  $\lambda^+(x)$  l'*échelle d'apparition* de  $x$ ,  $\lambda^-(x)$  son *échelle de disparition*<sup>1</sup> et  $\Lambda(x)$  son intervalle de *persistance* dans une solution de l'algorithme.

On voit alors que si  $f(x)$  désigne le père d'une région  $x$  dans la hiérarchie  $H$ , alors  $\forall x \in H : \lambda^-(x) = \lambda^+(f(x))$ . Les solutions de  $A$  sont donc totalement déterminées par le couple  $(H, \lambda^+)$ , où  $\lambda^+$  est la fonction qui renvoie l'échelle d'apparition des régions de  $H$ .  $\lambda^+$  est une fonction croissante :  $x \subset y \Rightarrow \lambda^+(x) < \lambda^+(y)$ . En classification hiérarchique, le couple  $(H, \lambda^+)$  s'appelle une *hiérarchie indicée*. Dans le

1. On les appelle inner et outer scale - échelle interne et externe - en théorie scale-space [7].

cadre de l'image<sup>2</sup>, nous dirons que  $(H, \lambda^+)$  est la représentation **ensembles-échelle (scale-sets)** des solutions de  $A$  (cf. fig. 1). Son axe *continu* en  $\lambda$  la pose en effet comme l'analogue orienté région des représentations scale-space issues du filtrage et des théories de la diffusion.

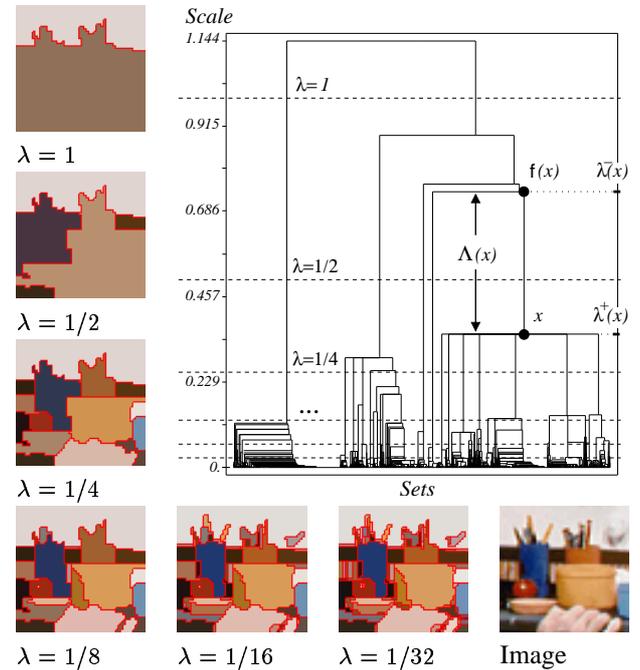


FIG. 1: Description ensembles-échelle d'une image et certaines de ses coupes.

Les partitions  $A(\lambda)$  prennent alors une forme *implicite* dans  $(H, \lambda^+)$  : chaque  $A(\lambda)$  correspond à une **coupe horizontale** du scale-sets. Les éléments de  $A(\lambda)$  sont les plus hautes régions de la hiérarchie situées sous la droite  $y = \lambda$ .

Si les solutions de  $A$  sont monotones, elles forment donc une séquence de coupes horizontales d'une hiérarchie indicée  $(H, \lambda^+)$ . Notre démarche est alors la suivante : nous supposons dans un premier temps que la hiérarchie  $H$  a déjà été construite, par un moyen quelconque, et nous étudions le problème de coupe minimale dans  $H$  pour une énergie  $E_\lambda$ , c'est-à-dire la manière de construire l'indice sur  $H$ . Cette étude nous conduit alors à un principe naturel pour construire la hiérarchie elle-même.

### 4 Coupes minimales

En général, une *coupe* d'une hiérarchie  $H$  est définie comme une partition  $C = \{x_1 \dots x_p\}$  uniquement constituée d'éléments de  $H$ . Du point de vue de l'arbre  $T$  représentant  $H$ ,  $C$  est un ensemble de nœuds que toute branche de  $T$  intersecte une et une seule fois (cf. fig. 2). Nous notons  $\mathcal{C}(H)$  l'ensemble des coupes de  $H^3$ .

Soit  $H$  une hiérarchie et  $E_\lambda$  une énergie séparable (eqn. 3). Pour  $\lambda \in \mathbb{R}^+$ , notons  $C_\lambda^*$  la coupe qui minimise  $E_\lambda$  sur  $\mathcal{C}(H)$ .

2. ou du signal, le modèle est valable en dimension arbitraire.

3. Notons que le nombre de coupes possibles d'une hiérarchie est très important. Par exemple, si  $H$  est un arbre binaire équilibré s'appuyant sur un ensemble de taille  $n$  alors  $|\mathcal{C}(H)| > \sqrt{2}^n$ , cf. [3]. Deux coupes sont a priori non ordonnées, comme les coupes  $C_1$  et  $C_2$  de la figure 2.

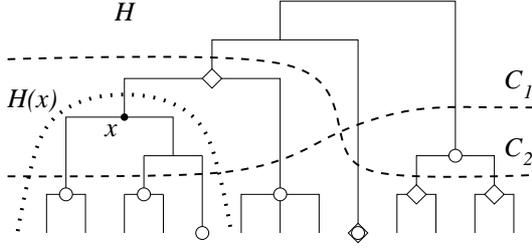


FIG. 2: Une hiérarchie  $H$  et deux de ses coupes  $C_1$  et  $C_2$ , constituées respectivement des nœuds ronds et losanges.  $H(x)$  est la hiérarchie partielle de sommet  $x$

Nous montrons alors dans [3] le

**Théorème 1 (coupes minimales causales) :**

Si  $C$  est décroissante dans le treillis  $\mathbb{P}(\mathcal{D})$  des partitions du domaine  $\mathcal{D}$ , i.e. vérifie  $\forall(P, Q) \in \mathbb{P}^2(\mathcal{D})$

$$P > Q \Rightarrow C(P) \leq C(Q)$$

alors l'ensemble  $\{C_\lambda^*\}_{\lambda \in \mathbb{R}^+}$  est causal, i.e.

$$\lambda_2 \geq \lambda_1 \Rightarrow C_{\lambda_2}^* \geq C_{\lambda_1}^*.$$

La condition de décroissance du régulariseur  $C$  est relativement naturelle, la requérir revient à juger qu'une partition grossière est plus «simple» qu'une partition plus fine. Toutes les énergies séparables mentionnées ci-dessus emploient des régulariseurs décroissants. Une classe importante de problèmes de partitionnement par minimisation d'énergie admet donc des solutions monotones quand l'espace des solutions est restreint aux coupes d'une hiérarchie. Comme nous l'avons vu, si ces partitions sont monotones on peut en obtenir une représentation exhaustive en déterminant les intervalles de persistance des nœuds de  $H$ . Voyons maintenant que ces intervalles peuvent être calculés exactement et efficacement. La méthode de calcul est une généralisation de la méthode de programmation dynamique que nous avons proposée dans [4].

$\forall x \in H$ , le sous-ensemble de  $H$  défini par  $H(x) = \{y \in H \mid y \subseteq x\}$  est également une hiérarchie, de sommet  $x$ , que nous appelons la *hiérarchie partielle* issue de  $x$  (cf. fig. 2). Remarquons alors que pour un nœud  $x$ , l'évolution de son énergie en fonction de  $\lambda$  est une fonction *affine*

$$E_x : \lambda \rightarrow E_x(\lambda) = \lambda C(x) + D(x) \quad (6)$$

que nous appelons l'*énergie propre* de  $x$ . Considérons également pour chaque hiérarchie partielle  $H(x)$  la fonction  $E_x^*$  qui donne l'énergie de ses coupes optimales en fonction de  $\lambda$ . Appelons cette fonction l'*énergie partielle* de  $x$ . Nous montrons alors dans [3] le

**Théorème 2 :** Si  $C$  est décroissante alors  $\forall x \in H$

i) Les énergies partielles des nœuds de  $H$  sont reliées par la relation de programmation dynamique fonctionnelle

$$E_x^* = \inf \left\{ E_x, \sum_{f \in \mathcal{F}(x)} E_f^* \right\} \quad (7)$$

où  $\mathcal{F}(x)$  désigne l'ensemble des fils de  $x$  dans  $H$ .

ii)  $E_x^*$  est une fonction affine par morceaux continue, non décroissante et concave.

iii)  $\lambda^+(x)$  est l'unique solution de

$$E_x(\lambda) = \sum_{f \in \mathcal{F}(x)} E_f^*(\lambda). \quad (8)$$

iv)  $\lambda^-(x)$  est donnée par

$$\lambda^-(x) = \min \{ \lambda^+(y) \mid y \in H, x \subset y \} \quad (9)$$

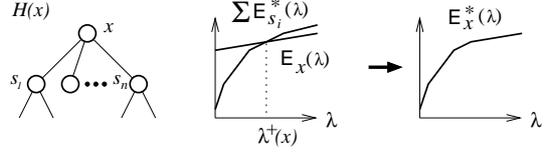


FIG. 3: Un pas de programmation dynamique fonctionnelle.

Ce théorème permet de calculer exactement  $\lambda^+$  et  $\lambda^-$  en deux passes dans  $H$  : ascendante puis descendante.

Les énergies partielles sont des fonctions affines par morceaux (cf. fig. 3) que l'on peut mémoriser explicitement pour chaque nœud. La passe ascendante permet de les calculer toutes par deux opérations formelles sur des affines par morceaux : addition et intersection avec une fonction affine. Pour chaque nœud  $x$ , on additionne les énergies partielles de ses fils. L'intersection de cette somme avec l'énergie propre de  $x$  permet de résoudre l'équation 8 et donne  $\lambda^+(x)$ . L'équation 7 donne alors  $E_x^*$ .

Dans un second temps, une passe descendante permet de calculer  $\lambda^+$  grâce à l'équation 9.

Cette optimisation fait apparaître des nœuds qui n'appartiennent à aucune coupe optimale pour  $E_\lambda$ . Ils vérifient  $\Lambda(x) = \emptyset$ , i.e.  $\lambda^-(x) \leq \lambda^+(x)$ . Nous les supprimons donc de  $H$  pour obtenir ce que nous appelons la **hiérarchie persistante** pour  $E_\lambda$ , notée  $H^*$ .  $\lambda^+$  est alors une fonction croissante dans  $H^*$  et  $(H^*, \lambda^+)$  est la représentation ensembles-échelle des coupes optimales de  $H$  pour  $E_\lambda$ .

## 5 Escalade

La monotonie des coupes minimales d'une hiérarchie suggère alors un principe naturel de construction de la hiérarchie elle-même, par fusion de régions, que nous appelons le principe d'**escalade**. La hiérarchie est construite de manière ascendante, en partant de sa base - une partition fine de l'image, éventuellement en pixels - et en ajoutant itérativement l'union des deux régions voisines qui possède la *plus faible échelle d'apparition*  $\lambda^+$ . L'escalade consiste donc à progressivement régulariser la fonctionnelle - i.e. renforcer la contrainte  $\gamma$  - et à collecter les régions au fur et à mesure où elles deviennent optimales. L'escalade peut être vue comme une procédure d'optimisation par continuation [6] contrainte à suivre une séquence ordonnée de partitions. On montre également que la stratégie d'escalade correspond à un choix localement optimal d'un point de vue débit/distorsion. En outre, la méthode est *exacte* : l'algorithme de programmation dynamique permet de calculer exactement l'échelle  $\lambda^+$  à laquelle chaque hypothèse de fusion apparaîtrait comme solution.

Nous obtenons donc un algorithme *totalemment exempt de paramètre* étant donné un couple  $(C, D)$  d'énergies antagonistes. Les scale-sets calculés par escalade vérifient plusieurs propriétés, la plus importante étant une *invariance d'échelle* :

**Théorème 3 :** Si  $(H^*, \lambda^+)$  est le scale-sets d'escalade associé à l'énergie  $(C, D)$  alors celui associé à  $(\nu C, \mu D)$  est  $(H^*, \frac{\mu}{\nu} \lambda^+)$ .

Cette propriété fondamentale indique que la méthode permet d'éliminer le paramètre  $\lambda$  de l'étape d'analyse de bas niveau : régler le poids relatif de  $C$  et  $D$  *a priori*, dans la formulation des énergies, ou *a posteriori*, par coupe dans le scale-sets est strictement équivalent. L'invariance d'échelle conduit en outre, pour certaines énergies comme celle de Mumford-Shah, à des scale-sets *invariants par homothétie*.

## 6 Exemples et conclusion

Le cadre théorique étant relativement général, il permet de modéliser toutes sortes de problèmes de segmentation : images naturelles, de profondeur, 2D/3D, multispectrales, etc. Le tout est de définir l'énergie adaptée à chaque problème. Nous illustrons ici la pertinence de l'approche pour la segmentation d'images naturelles en utilisant la fonctionnelle constante par morceaux de Mumford-shah [8]. La sur-segmentation initiale est obtenue par ligne de partage des eaux (LPE) mais on peut partir des pixels individuels. L'algorithme est en pratique quasi-linéaire en fonction de la taille de la sur-segmentation et rapide (2 à 3 s. de CPU à 1,2 GHz depuis une LPE sur une image  $256 \times 256$ ). Nous sommes évidemment obligés de réaliser des coupes dans les scale-sets pour pouvoir présenter des résultats. La figure 4 montre des séries de coupes selon une progression géométrique en  $\lambda$ . On observera que la méthode permet de délimiter des unités globales, éventuellement très texturées (e.g. panier, ananas dans l'image (b)), suivre aussi la simplification progressive du miroir dans l'image (c).

Les représentations proposées sont susceptibles de multiples applications : réglage interactif de la finesse de segmentation (à  $\lambda$  fixe ou réglé localement) et segmentation par marqueurs [10], séparation objet/fond, recherche active de modèles d'objets, caricature/compression d'image [9], etc. Par ailleurs, en explicitant les durées de vie des régions dans une solution, les scale-sets ouvrent plusieurs voies nouvelles pour l'analyse d'image orientée régions : caractérisation des régions pertinentes en tant qu'étapes *stables* du processus de simplification [6, 7], recherche *a posteriori* des «bonnes» échelles d'analyse de l'image, i.e. celles où apparaissent des formes stables, caractérisation de la structure des images par la structure de leur scale-sets, etc. Ces différentes idées semblent particulièrement prometteuses pour les problèmes d'indexation d'images par le contenu et de fouille de données.

## Références

[1] A. Desolneux, L. Moisan, and J.-M. Morel. Automatic image analysis : a challenge for computer vision. In *GRETSI'01, Toulouse, France, 2001*.

[2] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE trans. PAMI*, 6(6):721–741, 1984.

[3] L. Guigues. Modèles multi-échelles pour la segmentation d'images. Thèse de doctorat de l'Université de Cergy-Pontoise - en préparation.

[4] L. Guigues, H. Le Men, and J.-P. Cocquerez. Segmentation d'image par minimisation d'un critère MDL dans

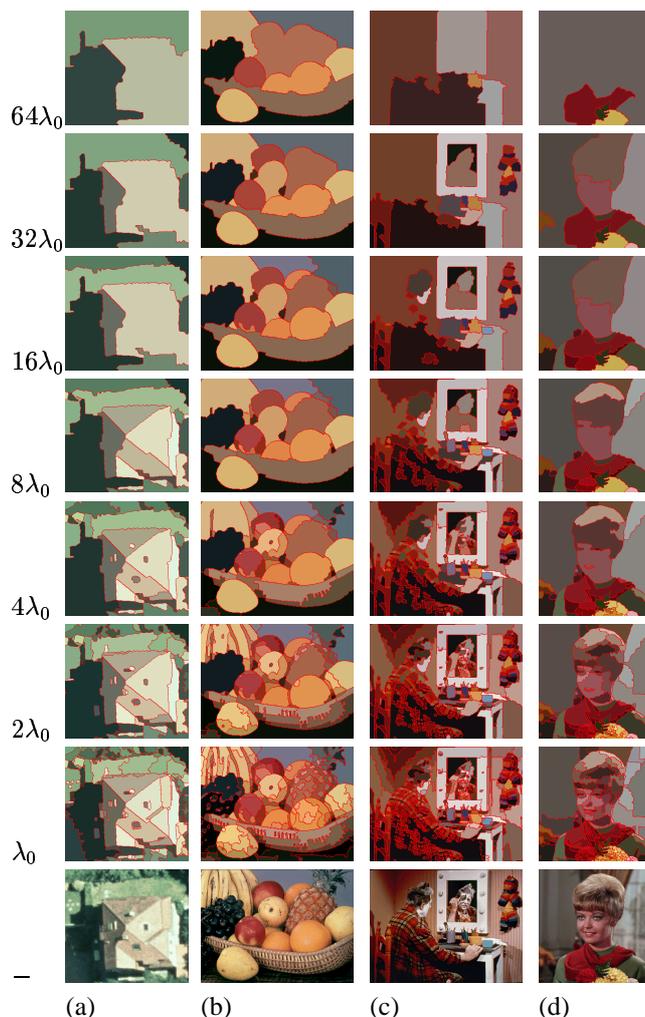


FIG. 4: Exemples de coupes dans des scale-sets.

une pyramide de segmentations. In *GRETSI'01, Toulouse, France, 2001*.

[5] G. Koepfler, C. Lopez, and J.-M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM Journal on Numerical Analysis*, 31(1):282–299, 1994.

[6] Y. G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3(1):73–102, 1989.

[7] T. Lindeberg. *Scale-space theory in computer vision*. Kluwer Academics, 1994.

[8] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. on Pure and Applied Math.*, 17(4):577–685, 1989.

[9] P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE trans. IP*, 9(4):561–576, April 2000.

[10] F. Zanoguera. *Segmentation interactive d'images fixes et de séquences vidéo basée sur des hiérarchies de partitions*. PhD thesis, Ecole des Mines de Paris, Centre de Morphologie Mathématique, 2001.