La segmentation en couleurs au CMM

Jean Serra
Ecole des Mines de Paris, France

Journée ISS, 3 février 2005

Huit thèses. Entre 1998 et 2003, huit thèses impliquant le traitement d'images en couleur se sont déroulées au Centre de Morphologie Mathématique:

Thésard	*Année*	*Directeur*	*Segment*	*Prototype*
Mariusz Młynarczuk [24]	1998	J. Serra	×	
Claire-Hélène Demarty [11]	2000	S. Beucher	×	×
Marcin Iwanowski [23]	2000	J. Serra	×	×
Cristina Gomila [13]	2001	F. Meyer	×	×
Francesca Zanoguera [37]	2001	F. Meyer	×	×
Valéry Risson [26]	2001	S. Beucher	×	×
Allan Hanbury [17]	2002	J. Serra	×	
Jesús Angulo [4]	2003	J. Serra	×	×

Ces thèses ne portent pas toutes exclusivement sur la segmentation en couleur, mais le thème y est suffisamment important pour apparaître une fois au moins dans chaque travail, ainsi que dans les publications associées. La liste qui suit, non exhaustive, concerne cet aspect des études. Quand il y a plusieurs auteurs, ils sont cités par ordre alphabétique. Les thèses de Claire-Hélène Demarty, Marcin Iwanowski, Cristina Gomila et Francesca Zanoguera sont accompagnées d'animations videos, auquelles nous renvoyons le lecteur.

Les synthèses de partitions (M. Młynarczuk, J. Serra) La référence [30] prend pour point de départ la thèse de M. Młynarczuk, qu'elle prolonge. Du point de vue de la couleur, cela correspond à une question très spécifique et dont la généralité, réelle cependant, ne se trouve pas là où on l'attendrait. On veut segmenter des silicates vus en microscopie optique sous quinze orientations de polarisation différentes. La couleur de chaque image est codée sur une palette. Il s'agit donc en fait *d'imagerie multispectrale* sur quinze canaux. Mais comme chaque palette est codée indépendamment des autres, les données sont *hétérogènes* et on ne peut pas construire d'espace vectoriel sur lequel asseoir une analyse en composantes principales (Fig. 1). Il faut trouver autre chose. La méthode de segmentation élaborée ici consiste à segmenter tout d'abord chaque image séparément. D'où quinze partitions, c'est à dire quinze images *homogènes*. Elles sont filtrées et pondérées afin d'engendrer une

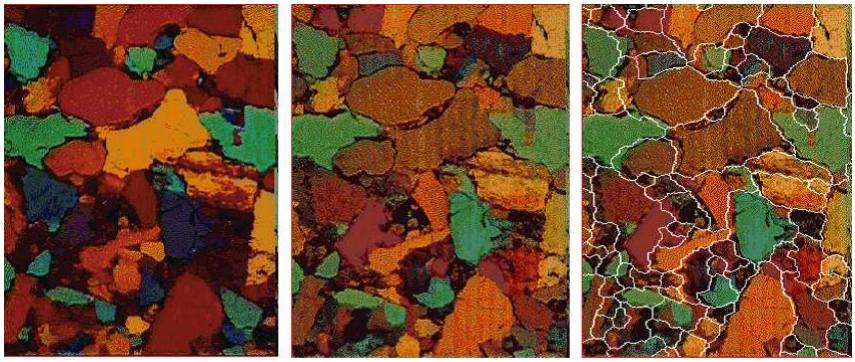

Figure 1: *a) et b) : deux polarisations différentes d'un même grès. c) la segmentation de la série des quinze images du grès.*

fonction numérique qui indique la distance probable à un contour. Cest cette fonction qui est finalement segmentée par ligne de partage des eaux.

L'intérêt de la démarche est double. Il montre qu'on peut traiter quantitativement des images de couleurs par palettes, et conduit à une méthodologie générale aux images multi-dimensionnelles, et située aux antipodes de la démarche par analyse factorielle (celle-ci commence par produire une image moyenne, l'autre finit par là).

Segmentation et indexation par le contenu des nouvelles télévisées (S. Beucher, C-H. Demarty) Le problème traité ici est celui de l'indexation par le contenu de journaux télévisés archivés; elle porte donc sur des séquences d'une demi-heure qu'il s'agit de segmenter temporellement tout d'abord pour en découper les plans successifs (les "shots"), puis de repérer dans chaque plan des images clé, ainsi que diverses caractéristiques, incrustations, caractères α-numériques, etc...de façon à élaborer une sémantique des relations des plans entre eux (par ex. retrouver une interview par le type de succession d'images clé). L'analyse a pour but de résumer la demi-heure des informations en une ou deux pages d'images clés commentées brièvement. Comme la procédure ne présente d'intétêt que si elle est entièrement automatique, elle doit reposer sur une segmentation extrèmement robuste. La référence [9] explicite la méthode de segmentation (voir aussi la Fig.2), la technique d'indexation est présentée dans [10]; toutes deux sont reprises dans la thèse de Claire-Hélène Demarty [11]. On trouvera aussi dans cette thèse le traitement de plusieurs séquences de quelques minutes et d'un joural complet.

L'interpolation des couleurs et des contours (M. Iwanowski, J. Serra) Les références [21] et [22] relatent deux étapes du travail de thèse de M. Iwanowski consacré à l'interpolation [23] des images binaires, numériques et en couleur. La question relève de la segmentation, puisqu'il s'agit de créer des contours. Les algorithmes d'interpolation, affine [22] ou non [22], dérivent de l'étude théorique sur les géodésiques de la distance de Hausdorff [28], ainsi que

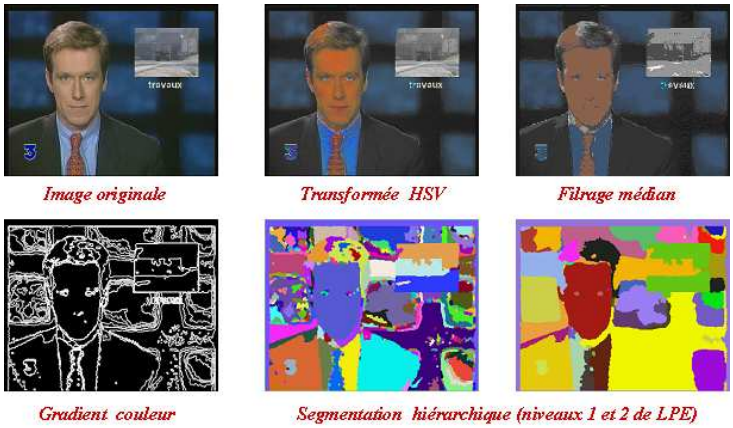

Figure 2: *Les étapes de la segmentation d'une image de journal télévisé.*

des travaux [25] et [8]. Le formalisme établi initialement pour les images numériques est étendu ici au cas de la couleur. L'important n'est pas tant de se référer à des couleurs qui aient une bonne signification perceptuelle ou colorimétrique que de ne pas créer de fausses couleurs. Pour cela, il faut et il suffit que le treillis de travail soit totalement ordonné. On part donc des variables video r', g', b', réduites qui dérivent du R_{709}, G_{709}, B_{709}, de la Rec 709 de la CIE, qu'on modifie par le changement de base suivant

$$\begin{bmatrix} l' & g' & r' \end{bmatrix} = \begin{bmatrix} 0.3 & 0.6 & 0.1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r' \\ g' \\ b' \end{bmatrix}$$

où la luminance L est à peu de choses près celle du standard de la télévision européenne. On évite les fausses couleurs en adoptant pour treillis de travail celui de l'ordre total $l' \rightarrow g' \rightarrow r'$. La Fig. 3 met en évidence l'absence de création de nouvelle couleur dans une situation d'interpolation entre deux couleurs seulement.

Segmentation en videophonie (C. Gomila, F. Meyer) Il s'agit d'un travail à la finalité ciblée, à savoir extraire la tête et le buste de la personne qui s'exprime devant une webcam [12] [13]. La démarche comprend trois volets. Dans un premier temps, chaque image de la séquence est segmentée, par inondation à partir de marqueurs, ce qui conduit à une pyramide de partitions croissantes. On passe ensuite d'un certain niveau hierarchique de segmentation à la reconnaissance des contours de la tête et du buste. Dans un troisème temps enfin, les partitions de deux images successives sont reliées par mise en correpondance de graphes, de façon à en caractériser la compensation de mouvement. L'aspect que nous retenons ici est la seconde étape où co-agissent un critère chromatique et autre, de forme. Une analyse théorique montre qu'ils peuvent conduire à une partition optimale vis à vis de chacun des deux critères [34].

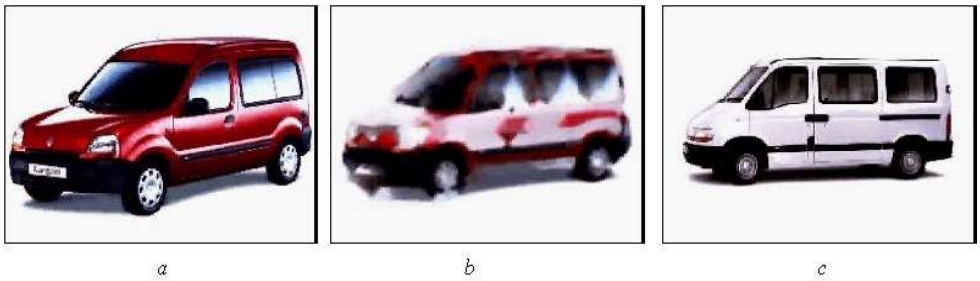

Figure 3: *Exemple d'élément médian entre les deux voitures rouge et blanche.*

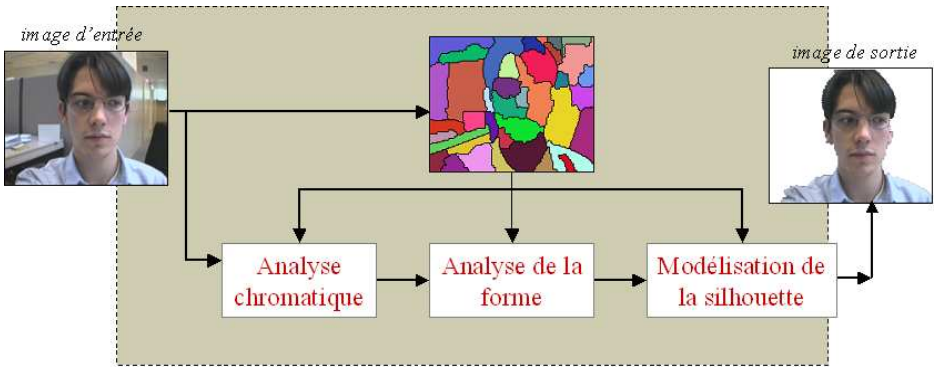

Figure 4: *Schéma d'ensemble d'une double segmentation, par couleur et par forme.*

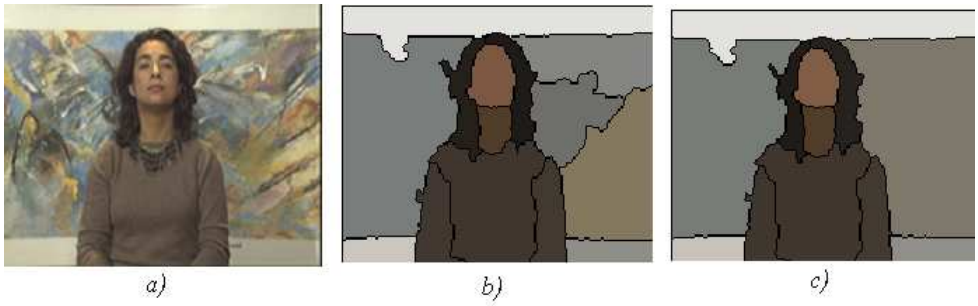

Figure 5: *a) Image initiale; b) niveau hiérarchique correct pour le premier plan, mais qui sursegmente le fond; c) correction manuelle du fond.*

Segmentation interactive à partir de hiérarchies de partitions (F. Meyer, F. Zanoguera) F. Zanoguera a élaboré des pyramides de segmentations d'images fixes et de séquences vidéo, destinées à se prêter commodément à des interactions rapides et efficaces avec un utilisateur. Dans le cas de séquences videos, la segmentation de la seule la première image est remaniée, les autres tenant compte automatiquement du remaniement. L'exemple ci-dessous (Fig. 5) porte sur de l'imagerie fixe, deux niveaux hiérarchiques différents sont combinés manuellement pour obtenir un niveau de détails plus grand pour le premier plan que pour le fond. Cela suppose que lorsqu'on s'élève dans la hiérarchie chaque arc de contours ne peut que rester ou disparaître, mais pas bouger. On montre que seuls les opérateurs connexes offrent cette possibilité [27], [35].

Les représentations polaires à base de normes (A. Hanbury, J. Serra) La référence [19] est l'expression aboutie d'une critique implicite, ou partiellement décrite dans plusieurs études, sur les incohérences des deux représentations polaires courantes HLS et HSV. Si l'on vise des traitements quantitatifs de l'information, il semble indispensable que les grandeurs énergétiques (brillance, saturation) soient des normes, et qu'elles soient indépendantes. On en déduit trois représentations, en normes L_1, L_2 et $max - min$. La plus nouvelle est la représentation L_1, à savoir

$$\begin{cases} l = m = \frac{1}{3}\left(max + med + min\right) \\ s = \begin{cases} \frac{3}{2}\left(max - m\right) & \text{if} \quad m \geq med \\ \frac{3}{2}\left(m - min\right) & \text{if} \quad m \leq med \end{cases} \\ h = \frac{\pi}{3}\left[\lambda + \frac{1}{2} - \frac{1}{2s}(-1)^{\lambda}(max + min - 2med)\right] \end{cases} \tag{1}$$

où max, med et min désignent le maximum, la médiane le minimum des coordonnées (r, g, b) du point point [20]. Le terme λ correspond à six secteurs de 60° du cercle unité (le détail des calculs se trouve dans [31].

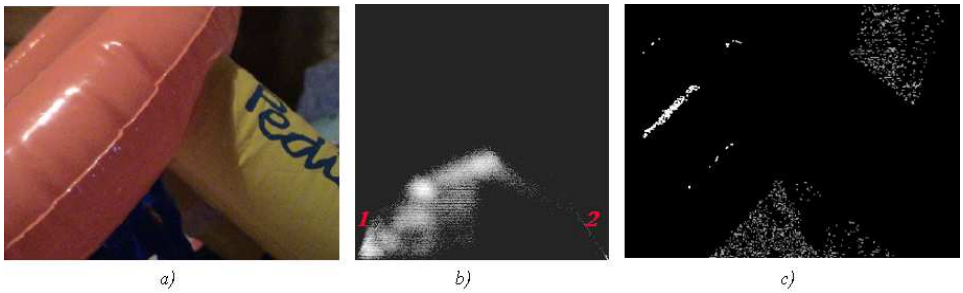

Figure 6: *a) image initale; b) histogramme bi-dimensionnel luminance (axe des x) et saturation (axe des y) en norme L_1; c) points de l'image situés sur le segment 1 (en gris) et sur le segment 2 (en blanc).*

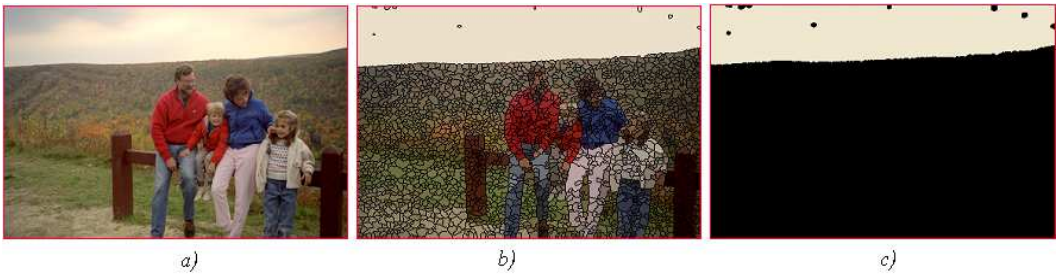

Figure 7: *a) image initiale; b) LPE de son gradient Lab euclidien; c) classe(s) correspondant au ciel.*

Histogrammes bi-dimensionnels en norme L_1 (J. Angulo, J. Serra) Les références [6], [36] et [7] prennent directement le relais de l'analyse précédente, et l'illustrent en comparant les histogrammes bi-dimensionnels *luminance/saturation* selon les trois normes, plus la représentation HLS. Il apparaît que la norme L_1 ci-dessus est la mieux adaptée à la segmentation de reflets, d'ombres et de dégradés de lumière (cf Fig. 6). La référence [34] propose une explication physique fondée sur un modèle di-chromatique de la variation spatiale de la lumière.

Recherche et extraction du ciel sur photographies (S. Beucher,V. Risson) La référence [26] fournit un bel exemple d'intersection de critères connectifs. Il s'agit de détecter la présence du ciel, et de le localiser, dans une base de 487 photographies non orientées, établie par Kodak. Le ciel est caractérisé par les quatre éléments que sont sa couleur, soit bleue, soit achromatique, sa luminance élevée, son absence de texture et sa géométrie: il est rarement entièrement inclus dans le champ. Partant d'une segmentation, ici par LPE du gradient euclidien en représentation Lab, on extrait la ou les classes qui vérifient les quatre critères (Fig.7). Ils suffisent pour caractériser et orienter 90% des images de la base de données.

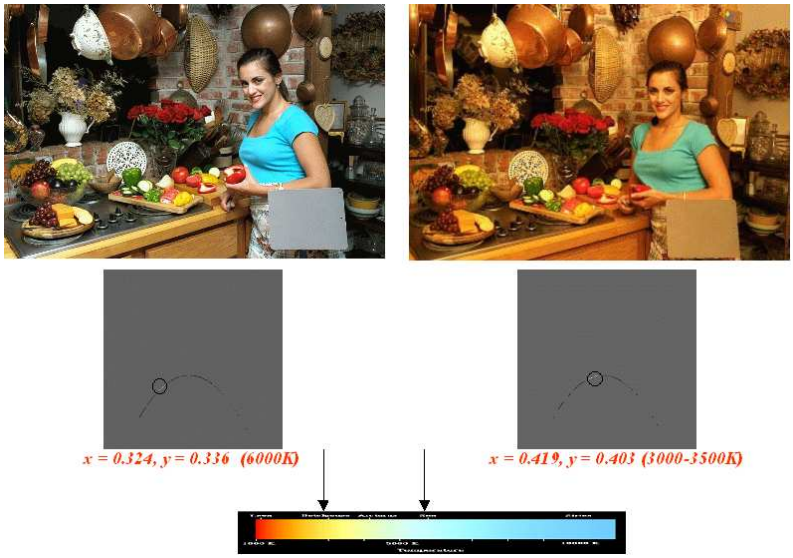

Figure 8: *Estimation de l'illuminant pour chaque photographie.*

Caractérisation de l'illuminant (S. Beucher,V. Risson) Lorsqu'on se place dans le sytème Lab, la constance de couleur permet de déterminer la température de l'illuminant en exploitant le modèle de réflexion dichromatique de Shafer [26]. Il faut dans un premier temps segmenter l'image en régions de réflectance spectrale homogène. On filtre ensuite les régions incompatibles avec le modèle di-chromatique de réflexion. On calcule les droites de convergences par régression linéaire sur les coordonnées chromatiques de chaque région. Faisceau de droites converge alors vers le point du plan chromatique correspondant à l'illuminant sur le diagramme des radiateurs de Planck (Fig. 8).

Cercle unité , Gradients et segmentation (A. Angulo, A. Hanbury, J. Serra) Dans [14], c'est la composante teinte qui est isolée et étudiée en soi, en tant que variable définie sur le cercle unité. Pour répondre à l'objection selon laquelle le cercle unité est tout ce qu'on veut sauf un treillis, quatre démarches sont proposées, dont deux solidement étayées sur de larges bases de données. Lune d'entre elles consiste à étendre au cercle unité des opérateurs des treillis numériques qui ne mettent en jeu que des accroissements. Pour la teinte, qui est un angle, il faut prendre comme accroissement $|t(\alpha) \div t(\beta)|$ entre deux valeurs α et β l'angle aigu d'extremités α et β sur le cercle unité [14]. On pose ensuite

$$\begin{aligned} |\nabla l| &= (\delta(l) - l) \vee (l - \varepsilon(l)) \\ |\nabla t| &= (\delta(t) \div t) \vee (t \div \varepsilon(t)) \end{aligned}$$

où δ et ε sont la dilatation et l'érosion isotropes unité dans l'espace de définition de l'image (i.e. trame carrée ou hexagonale). Le module du gradient couleur de $\mathbf{f}$ est alors introduit par la relation suivante

$$|\nabla \mathbf{f}| = \frac{s}{s_{\max}} \times |\nabla t| + (1 - \frac{s}{s_{\max}}) \times |\nabla l|. \tag{2}$$

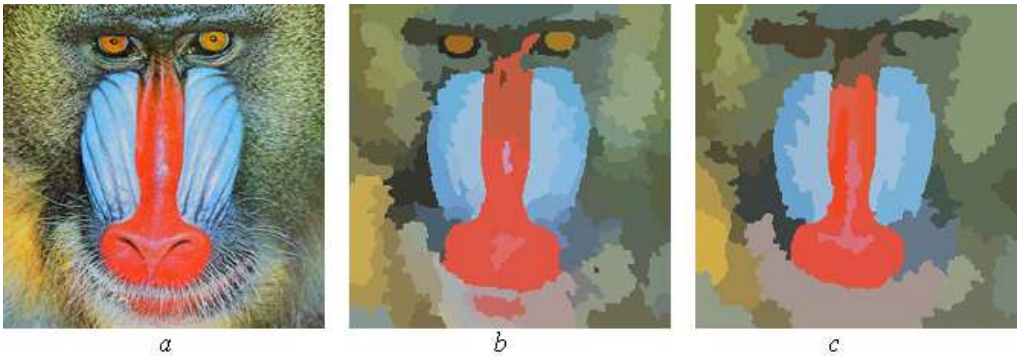

Figure 9: *a) Image initiale, b) et c) segmentations par LPE de gradients, par pondération de saturation (b) puis par gradient euclidien en représentation Lab (c).*

Dans l'étude [5] plusieurs gradients sont comparés sous le critère de la qualité des segmentations qu'ils produisent par l'algorithme des chutes d'eau (arrêté au niveau quatre). Il apparait que les deux gradients les plus pertinents sont celui de la représentation Lab (Fig. 9c), et celui de laRel. (2), (Fig. 9b).

Avec [16], ce sont toutes les composantes de la couleur que l'on cherche à ordonner pour en faire des treillis convenables. On se place dans l'espace Lab, dont les bonnes propriétés colorimétriques et psycho-visuelles sont bien connues, pour y construire une représentation de type polaire. L'ayant obtenue, on classe les points selon leur potentiel newtonien entre les deux extrêmes du gris et de la couleur saturée. Enfin, des treillis totalement ordonnés sont construits sur cette base.

Segmentation pondérée par la Saturation (J. Angulo, J. Serra) La représerntation polaire des couleurs en norme L_1 conduit à de bonnes segmentations fondées sur la démarche suivante en deux étapes [5] :

1. On segmente *séparément* la luminance, et la teinte;
2. les deux partitions qu'on obtient sont alors combinées au moyen de la sauration que l'on prend comme *critère* pour choisir en chaque point x soit la classe de la luminance celle de la teintesoit une classe plus petite, de façon à générer une nouvelle partition synthétique.

   La Fig.10b donne un exemple de cette partition composite.

Segmentation et qualité Chaque exemple, quand on l'analyse de plus près, montre qu'il repose sur une ou plusieurs partitions optimales. Au coeur de toutes ces variations se devine donc un thème commun. Quel est-il?

Figure 10:

En matière de segmentation, la recherche de qualité se traduit par celle d'un optimum. Classiquement, on s'exprime par le biais d'une fonction numérique F qui comprend des variables telles que la longugur des contours segmentés, ou la variance dans les régions intra-contours, etc.... Le calcul variationnel, de type Lagrange, permet de déterminer des segmentations qui optimisent ce nombre.

Dans les exemples que nous venons de voir, l'idée théorique sous-jacente est différente. La qualité ne s'exprime plus par une fonction numérique, mais directement par comparaison des partitions [35]. De fait, comme les partitions du plan forment un treillis complet, on peut parler de *partition optimale* au sens de tel ou tel critère. Seulement, tous les critères ne conduisent pas à un découpage optimal de l'espace: il n'existe pas par exemple, un plus grand découpage en régions où la variation de la fonction numérique f est inférieure ou égale à une constante donnée k. Seuls les critères qui découpent l'espace en classes qui engendrent une *connection* correspondant à des partitions optimales. En revanche, lorsqu'on s'appuie sur de tels critères, on en est récompensé par les bonnes propriétés qu'ils vérifient. Ils garantissent tout d'abord l'existence et l'unicité de la segmentation. De plus quand on compare deux champs Z et Z' qui se recouvrent partiellement, et que les classe $A(x)$ et $A'(x)$ de segmentation au point x sont toutes deux incluses dans $Z \cap Z'$, alors elles sont identiques, i.e. $A(x) = A'(x)$. Une pareille propriété de localisation n'est pas possible pour les segmentations par calcul variationnel (pas plus que, le plus souvent, l'unicité).

La façon connective de concevoir un optimum s'étend à des associations de critères hétérogènes, comme couleur et forme dans l'exemple de la webcam; elle s'applique aussi à des parties du champ, comme dans les histogrammes bi-dimensionnels, et se prête à toute une combinatoire de réunions et d'intersections, comme dans la détection du ciel, ou dans l'utilisation de la saturation comme balancier, ou encore d'érosions et de valeurs probables comme dans l'exemple pétrographique. Par ailleurs, elle s'articule sur la notion d'*opérateur connexe* (i.e. d'opérateur qui élargit les classes connexes), ce qui permet de préserver l'optimalité au travers de filtrages, ainsi que de ce qui en dérive (unicité, localisation, etc..). Enfin, elle se prête à des hiérarchies de partitions, qu'il s'agisse de semi-groupes,

comme dans l'exemple de la segmentation interactive, ou plus généralement de contraintes ordonnées.

Cette démarche théorique est nouvelle, et en cours de publication [35].

References

[1] J. Angulo, J. Serra, G. Flandrin. Quantitative descriptors of the lymphocytes. Analytical Cellular Pathology, 2001. 22(1/2): 69-70.

[2] J. Angulo J. Serra, Morphological colour size distribution for image classification and retrieval *Proc. ACVIS 2002*, Ghent, Belgium, September 9-11, 2002.

[3] J. Angulo J. Serra, Automatic analysis of DNA microarray images using mathematical morphology, *Bioinformatics* vol. 19 n0 5, 2003 pp. 553-562

[4] J. Angulo, *Morphologie mathématique et indexation d'images couleur. Application à la microscopie en biomédecine.* Thèse de Doctorat en Morphologie Mathématique, ENSMP, Paris, Décembre 2003.

[5] J. Angulo, J. Serra, Color segmentation by ordered mergings, in *Proc. of IEEE International Conference on Image Processing (ICIP'03),* IEEE, Vol. 2, Barcelona, Spain, Sept. 2003, p. 125–128.

[6] J. Angulo J. Serra, Histogrammes luminance/saturation pour les images couleur en représentation par norme L_1 *Ecole des Mines*, N-01/03/MM, mars 2003, 25p.

[7] J. Angulo, J. Serra, Traitement des images de couleur en représentation luminance/saturation/teinte par norme L_1, à paraitre dans *Traitement du Signal* en 2005.

[8] S. Beucher , Interpolation of sets, of partitions and of functions, in Mathematical Morphology, and its applications to image and signal processing, H.Heijmans and J. Roerdink eds.Kluwer, 1998

[9] C.-H. Demarty, S. Beucher, Color segmentation algorithm using an HLS transformation. In *Proceedings of the International Symposium on Mathematical Morphology (ISMM '98)*, p. 231-238.

[10] C-H Demarty, S. Beucher, Morphological tools for indexing video documents, in Proceedings IEEE International Conference on Multimedia Computing and Systems - ICMCS'99, June 7-11, 1999, Florence, Italy. p. 991-992..

[11] C-H Demarty, *Segmentation et structuration d'un document vidéo pour la caractérisation et l'indexation de son contenu sémantique : Application aux journaux télévisés.* Thèse de Doctorat en Morphologie Mathématique, ENSMP, 25 Janvier 2000. 232 p.

[12] C. Gomila, F. Meyer, Automatic video object generation tool : segmentation and tracking of persons in real time. Annales des Télécommunications, 2000. 55(3-4): p. 172-183.

[13] C. Gomila, *Mise en correspondance de partitions en vue du suivi d'objets.* Thèse de Doctorat en Morphologie Mathématique, ENSMP, 12 septembre 2001, 242 p

[14] A.Hanbury et J. Serra, Morphological operators on the unit circle. *IEEE Transactions on Image Processing* 10 (12) pp. 1842-1850, 2001.

[15] A. Hanbury, J. Serra, Mathematical morphology in the HLS colour space, in *Proceedings of British Machine Vision Conference*, Tim Cootes and Chris Taylor, Eds BMVA Manchester, Sept. 2001 pp. 451-460

[16] A. Hanbury, J. Serra, Mathematical Morphology in the L*a*b* space. 8th ECS and Image Analysis. *Image Anal. Stereol.*, 2001. 20(Supplt.1): 48-53.

[17] A. Hanbury *Morphologie mathématique sur le cercle unité, avec application aux teintes et aux textures orientées.* Thèse Ecole des Mines de Paris, Mars 2002.

[18] A. Hanbury, The taming of the hue, saturation and brightness colour space, in *CVWW'02, Computer Vision Winter Workshop* H. Wildenauer and W.G. Kropatsch (eds.), Bad Aussee, Autriche, february 2002, pp. 234-243.

[19] A. Hanbury J. Serra, A 3D-polar Coordinate Colour Representation Suitable for Image Analysis, Ecole des Mines, N-37/ 02/MM, nov. 2002, 39p (soumis à JVCIR)

[20] A. Hanbury, J.Serra, Colour Image Analysis in 3D-polar coordinates, in *Proc. of DAGM symposium,* Vienna, April 2003.

[21] M. Iwanowski, J. Serra, Morphological Interpolation and Colour Images, *Proc. International Conference on Image Processing* ICIAP'99 Venice, Sept.1999.

[22] M. Iwanowski, J. Serra, The Morphological-affine object deformation, *Mathematical Morphology and its Applications to Image and Signal Processing*, J. Goutsias, L. Vincent, D.S. Bloomberg (Eds.) Kluwer Ac.Publ. 2000, pp.81-90.

[23] M. Iwanowski, *Application de la morphologie mathématique pour l'interpolation d'images numériques.* Thèse de Doctorat en Morphologie Mathématique, ENSMP // Ecole Polytechnique de Varsovie, 15 Novembre 2000. 137 p.

[24] M. Młynarczuk, *Opis wybranych struktur skalnych przy użyciu metod morfologii matematycznej i analizy obrazów*, PhD thesis, IMG PAN, (1998)

[25] F.Meyer, A morphological interpolation method for mosaic images, in Mathematical Morphology and its applications to image and signal processing, Maragos P. et al. eds.Kluwer, 1996.

[26] V. Risson, *Application de la Morphologie Mathématique à l'analyse des conditions d'éclairage des images couleur.* Thèse de Doctorat en Morphologie Mathématique, ENSMP, 17 déc. 2001, 203 p.

[27] P. Salembier and J. Serra, Flat Zones Filtering, Connected Operators, and Filters by Reconstruction *IEEE Transactions on Image Processing*. August 1995, vol. 4, n° 8, 1153-1160 .

[28] J.Serra, Hausdorff distance and Interpolations, in *Mathematical Morphology and its applications to image and signal processing*, (Proc. ISMM'98) H.Heijmans and J. Roerdink eds Kluwer,1998, pp.107-114.

[29] M. Iwanowski , J. Serra, The Morphological-affine object deformation, in *Mathematical Morphology and its Applications to Image and Signal Processing*, J. Goutsias, L. Vincent, D.S. Bloomberg (Eds.) Kluwer Ac. Publ. 2000, pp.81-90.

[30] J. Serra, M.Mlynarczuk, Morphological merging of multidimensional data. in *STERMAT'2000*, 6th Int. Conf. on Stereology and Image Analysis in Materials Science. Cracow, 20 Sept. 2000. p. 385-390.

[31] J. Serra, Espaces couleur et traitement d'images, rap. tech. EMP N-34/02/MM Octobre 2002 13p

[32] J.Serra, J. Angulo, Aplicación de la morfología matemática a la telemedicina y la biotecnología : caracterización morfológica de células de la sangre y análisis de cDNA microarrays. In *Reconocimiento de patrones* CIARP'02 J.L Diaz de Leon Santiago y C. Yàñes Marquez (Eds)„ México, nov. 2002. pp. 39-50.

[33] J. Serra, Connection, image segmentation and filtering. in *Avances en Ciencias de la computacion e Ingenieria de Computo* J.H. Sossa Azuela, et al., Eds. , CIC'02 Invited lecture, México, nov. 2002. pp. 37-50.

[34] J. Serra Morphological Segmentation of Colour Images, *ISMM'05*, Paris, April 2005.

[35] J. Serra, A lattice approach to Image segmentation, 75 p. (to be published by JMIV)

[36] F. Torres, J. Angulo, F. Ortiz, "Automatic detection of specular reflectance in colour images using the MS diagram", Dans (Petkov et Westenberg Ed.) *Computer Analysis of Images and Patterns (Proceedings CAIP 2003),* Lecture Notes in Computer Science, Vol. 2756, 2003, p. 132–139.

[37] M-F. Zanoguera, *Segmentation interactive d'images fixes et de séquences video basée sur des hiérarchies de partitions*, thèse Ecole des Mines de Paris, décembre 2001, 200

[38] M-F. Zanoguera, On the implementation of non-separable vector levelings. In Mathematical Morphology VI, Proc. ISMM'02, Sydney, Australia, April 2002, H. Talbot and R. Beare, eds. CSIRO. pp. 369-377.