A paraître dans la revue "Traitement du Signal"

Traitements des images de couleur en représentation luminance/saturation/teinte par norme L_1

Colour image processing from luminance/saturation/hue in L_1 representation

Jesús Angulo et Jean Serra
Centre de Morphologie Mathématique - Ecole des Mines de Paris
35 rue Saint Honoré - 77300 Fontainebleau - France
`Jesus.Angulo@ensmp.fr` - `Jean.Serra@ensmp.fr`

25 novembre 2004

Résumé

Les représentations polaires classiques de la couleur, du type (HSV, HLS,...) conduisent à des brillances et des saturations aux propriétés incohérentes. Il leur est substitué des paramètres d'intensité (brillance et saturation) qui sont des normes, dont L_1 et $max - min$.

Deux exemples illustrent la pertinence d'une norme convenable. Le premier porte sur les histogrammes luminance/saturation. Construits dans le cadre des diverses normes introduites plus haut, ils mettent en évidence des alignements dans le cas de la norme L_1, alignements dont les points caractérisent des zones d'ombres ou de reflets dans l'image. Un modèle de comportement régionalisé linéairement des flux lumineux est proposé pour l'interprétation physique de ce phénomène.

Le second exemple porte sur la segmentation d'images en couleur, à partir de représentations polaires. L'idée qui préside ici est de faire jouer à la saturation un rôle d'équilibrage entre luminance et teinte. Chacune des trois composantes luminance, saturation et teinte est traitée indépendamment des autres selon une technique classique de segmentation pour les fonctions numériques. Dans les classes de haute valeur de la saturation on retient alors la teinte comme paramètre significatif, et dans celles de basses saturation, la luminance. A nouveau, la norme L_1 donne des résultats meilleurs que les autres représentations.

Mots clés : couleur, segmentation, saturation, normes, histogrammes bi-dimensionnels, HLS, morphologie mathématique.

Abstract

The classical colour polar-based representations, of type (HSV, HLS, etc.), lead to brightness and saturation with non consistent properties. In this paper they are replaced by means of norms, in particular L_1 and $max - min$.

Two examples illustrate the suitable relevance of L_1 norm with regard to segmentation issues. The first issue deals with the histograms luminance/saturation which are plotted

for four different representations. They highlight alignments in the case of the norm L_1 and these alignments characterise zones of shadows or of reflections in the image. A model of linearly regionalised spectrum for light flux is proposed for a physical interpretation of this proportionality.

The second example relates to morphological colour image segmentation, starting from polar representations. The main underlying idea here is to fully exploit the nature of the saturation component as a balance between the luminance and the hue. More precisely, each of the three components luminance, saturation and hue is processed independently from the others, according to a traditional segmentation technique for numerical functions. In the regions of high saturations, the hue is retained as the significant parameter, and in those of low saturations, the luminance replaces the hue. Again the norm L_1 presents interesting advantages with respect to the other representations.

Keywords : colour, segmentation, saturation, norms, bi-variate histograms, HLS, mathematical morphology.

1 Introduction

La segmentation des images en couleur est un thème d'études actuellement fécond (cf. par exemple [10]). Nous nous proposons de l'aborder au moyen de représentations du type "luminance, saturation, teinte", et en particulier d'étudier le rôle joué par la variable "saturation". Comment cette variable aide-t-elle à repérer des zones uniformes, comment l'utiliser dans la construction de gradients, ou dans celle de partitions qui segmentent l'image ? Voilà les trois questions qui seront abordées ci-dessous.

L'observation suivante peut nous guider : pour la vision humaine, il apparaît que le paramètre le plus significatif est la teinte dans les zones de grande saturation, et devient de plus en plus la luminance quand la saturation diminue. Toute personne dotée d'une sensibilité normale aux couleurs vérifie facilement le phénomène sur elle-même. Cette idée de rôle de balancier de la saturation a déjà été exploitée par P. Lambert et T. Carron [9].

Or, pour présenter ce fil directeur, nous venons d'utiliser les trois termes de *luminance*, de *saturation* et de *teinte*. Avant d'aller plus loin, il convient sans doute d'en préciser les sens qualificatifs, et pour ce qui est de l'imagerie digitale, quantitatifs.

Intuitivement, on peut définir le triplet "luminance, saturation, teinte" en disant que c'est celui sur lequel nous agissons quand nous voulons régler une image en couleur. Prenons, par exemple, un récepteur de télévision. Il met en jeu trois représentations distinctes de la couleur. D'une part, le signal reçu par l'antenne est codé sous forme d'une image de gris et de deux autres associées aux contrastes vert-rouge et bleu-jaune (on parle d'une luminance et de deux chrominances). D'autre part, l'image vue à l'écran s'obtient par trois intensités électriques qui excitent chacune une couche de photo-récepteurs, verts, rouges ou bleus. Ces deux représentations sont donc différentes bien que techniquement indispensables l'une de l'autre. Pourtant, les constructeurs n'en veulent pas pour l'interface avec l'utilisateur, et leur préfèrent des réglages manuels à base de lumière (luminance), de contraste (saturation), et si l'on possède un très vieux poste, de teinte. C'est donc bien ce dernier triplet qui est le plus accessible à la perception humaine.

Lorsqu'on se propose de construire une représentation de la couleur de type brillance-

saturation-teinte, en vue de segmenter des images vidéo, trois ordres de considération différents doivent être pris en compte. Viennent en premier lieu les technologies d'acquisition et de visualisation (l'expression "image vidéo" recouvre ici toutes les images digitalisées afin d'apparaître sur un écran d'ordinateur ou de récepteur de télévision). A ces considérations s'ajoutent ensuite celles relatives à la perception humaine. Elles font l'objet des travaux de la commission internationale de l'éclairage (CIE) dont les recommandations colorimétriques conduisent à définir des standards. Enfin, le traitement quantitatif des images possède ses propres exigences, telles que pouvoir manipuler des distances, dire que deux fonctions sont indépendantes, dire qu'une mesure est symétrique , qu'elle est homogène (i.e. qu'elle a une dimension, au sens physique du terme), etc ... Ces exigences ne sont d'ailleurs pas les mêmes selon qu'il s'agit de filtrage ou de segmentation. Dans le premier cas, on transforme une image en une autre, dans le second en une partition de l'espace ; la contrainte de ne pas créer de fausse couleur (par exemple), qui est primordiale dans le premier cas, n'apparaît pas dans le second.

Les motivations qui ont poussé A. Hanbury et J. Serra, à proposer la norme L_1 pour la brillance et la saturation [8] étaient surtout liées à la cohérence interne du traitement d'image. Aussi convient-il de mieux situer cette démarche par rapport aux deux autres points de vue, de la technologie vidéo et de la perception. Toutefois, indépendamment des justifications qui vont suivre, la norme L_1 s'est déjà montrée bien adaptée à la détection des zones d'ombres et de reflets. Une première étude [19], relayée par des travaux plus systématiques de J. Angulo [1] [3], ont mis en évidence des alignements caractéristiques dans les histogrammes bi-dimensionnels $luminance \times saturation$ en norme L_1 (cf figures 4-8 ci-dessous).

La section 2 est un bref rappel sur la correction gamma, nécessaire à notre propos. On trouvera une excellente présentation du sujet dans le livre de Ch. Poynton [14], et pour les lecteurs francophones dans l'ouvrage de C. Fernandez et Al [20]. Concernant les notations, nous suivons Ch. Poynton, pour qui les couleurs électroniques sont marquées de l'indice "prime" (e.g. R') alors que les intensités physiques correspondantes n'en ont pas (e.g. R) et nous étendons cette règle aux opérateurs, e.g. la moyenne arithmétique s'écrit m pour les intensités, et m' pour les variables video. Nous faisons nôtre aussi la convention de la CIE de désigner les quantités absolues par des majuscules (e.g. X, Z) et les valeurs relatives correspondantes par des minuscules (e.g. x, z).

La section 2 introduit la correction gamma, elle est suivie de l'examen des notions de brillance et de saturation dans les sections 3 à 5 qui reprennent la présentation [17]. Après avoir présenté en section 6 les histogrammes bidimensionnels en norme L_1, nous présentons en section 7 un modèle régionalisé de la lumière, pour conclure en sections 8 et 9 sur les segmentations à base de saturation.

2 Intensités lumineuses et correction gamma

En imagerie couleur, la notion première est le spectre *sp* (SPD en anglais) du flux de puissance de la lumière qui irradie d'une surface ou qui arrive sur elle, et plus précisément, dans ce spectre électromagnétique, la gamme des fréquences du domaine visible. Ces intensités lumineuses ont donc pour dimension une énergie par unité de surface, telle que des watts par m^2.

Lorsqu'un spectre est reçu par un photo-récepteur, celui-ci filtre les intensités de chaque fréquence en les pondérant par des poids fixes. La somme des intensités résultantes produit un signal qui représente une certaine "couleur". Ainsi, la *Rec 709* de la CIE a défini les poids qui conduisent au triplet $R_{709}, G_{709,} B_{709}$ [4]. Ce ne sont pas les seuls possibles, mais ils permettent de fixer les idées. Les intensités se comportent de manière additive à la fois pour des raisons physiques (il s'agit d'énergies) et aussi pour la vision de la couleur. Ce principe de superposition fait que toutes les couleurs accessibles à partir d'un système RGB sont construites par sommation et multiplication par des constantes positives.

L'exploration du spectre est limité inférieurement par $R = G = B = 0$ (énergie nulle) et supérieurement par un rouge R_0, un vert G_0 et un bleu B_0 maximaux pour les conditions d'expérience (spectre illuminant, limites technologiques des capteurs, ou de l'oeil, etc...) dans lesquelles on se place. En général, chaque technologie fixe les trois bornes une fois pour toutes et elles définissent le blanc de référence. On exprime alors les intensités en *variables réduites*

$$r = \frac{R}{R_0}, \qquad g = \frac{G}{G_0}, \qquad b = \frac{B}{B_0}. \tag{1}$$

Les photo-récepteurs des caméras et des appareils de photos numériques transforment l'intensité lumineuse en courant électrique ; inversement, les tubes cathodiques ou les écrans plats restituent des photons à partir d'un courant électrique. Or, la réponse de ces derniers n'est pas une fonction linéaire du voltage d'entrée, mais une fonction puissance, dont l'exposant γ, "gamma", est de l'ordre de 2.5 et varie selon les technologies. Si l'on veut que les intensités lumineuses produites par le tube soient proportionnelles à celles de la scène captée par le photo-récepteur, il faut corriger cet effet gamma, ce que les constructeurs font au niveau de l'électronique du photo-récepteur. La *Rec. 709* de la CIE propose la fonction de transfert suivante

$$\begin{array}{ll} r' = 4.5r & r \leq 0.018 \\ r' = 1.099r^{0.45} - 0.099 & r > 0.018 \end{array} \tag{2}$$

écrite ici pour l'intensité lumineuse réduite rouge, et où l'on a $1/\gamma = 0.45$. La même fonction de transfert est appliquée aux composantes verte et bleue.

La figure (1), reprise de [14] montre le graphe de la Rel. (2). Le domaine de variation $[0, 1]$ est le même pour les intensités réduites (r) et pour les couleurs vidéo (r'), ce qui implique que le point blanc RGB de la Rel. (1) reste inchangé. Le changement de définition aux petites valeurs dans la Rel. (2), non seulement évite de créer des couleurs vidéo négatives, mais corrige du bruit de fond des capteurs. Un tube cathodique idéal devrait avoir pour fonction de transfert l'inverse de la Rel. (2). En fait, ils n'ont en général pas de partie linéaire [14], ni d'exposant gamma exactement égal à 1/0,45.

La figure (1) montre que pour r proche de 1, le graphe ressemble à une droite. Plus précisément, le développement limité

$$(1 - u)^{1/\gamma} = 1 - \frac{u}{\gamma} + \epsilon\,(u) \tag{3}$$

valable pour u petit, conduit à remplacer la seconde équation (2) par

$$r'^{*} = (0.55 + 0.45r)1.099 - 0.099 \tag{4}$$

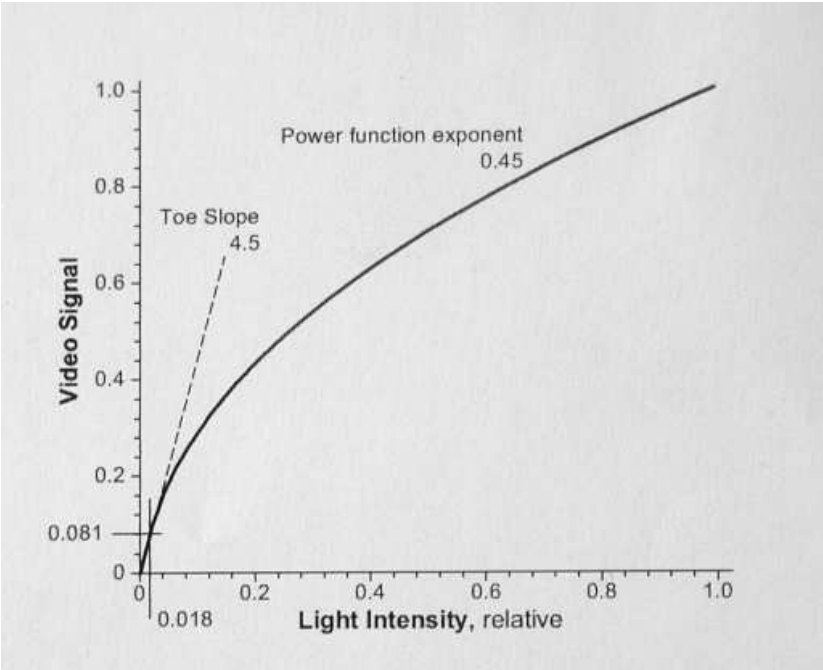


FIG. 1 – Fonction de correction Gamma.

soit, numériquement

r	0.9	0.8	0.7	0.6	0.5
r'	0.949	0.895	0.837	0.774	0.705
r'^*	0.950	0.901	0.851	0.802	0.753
$\frac{r'-r'^*}{r'}$	0.1%	0.6%	1.4%	2.8%	4.8%

Par rapport aux bruits de fond des signaux vidéo, on peut considérer l'approximation r'^* comme parfaite pour $r \geq 0.8$ et excellente pour $0.6 \leq r \leq 0.8$.

3 Les espaces vectoriels de la couleur

La linéarité des intensités r, g, b leur confère une structure d'espace vectoriel à 3 dimensions, ou plutôt de la partie E de cet espace limitée au cube unité $[0,1] \times [0,1] \times [0,1]$ de $\mathbb{R}^3$. Il serait donc sage, quand on veut faire du traitement d'images vidéo en couleur de remonter des couleurs videos (r', g', b') aux intensités réduites (r, g, b) par la transformation inverse de la Rel. (2). Si l'on part d'images de 3×8 bits pour (r', g', b'), il faudrait sans doute passer en 3×16 bits pour les calculs proprement dits (ou en flottant), mais cela n'a rien d'exhorbitant. Or, c'est un fait, les gens travaillent en pratique dans l'espace (r', g', b'), qu'ils modélisent comme une partie d'espace vectoriel à partir duquel ils construisent des moyennes arithmétiques, des projections, des histogrammes, des transformations de Fourier, etc... ce

qui donne souvent des résultats sensés.

La question se pose alors de savoir quelles sont les conséquences *réelles* de la correction gamma (Rel. (2)). Formellement parlant, on peut toujours considérer le cube unité (r', g', b') comme partie E' d'un espace vectoriel à trois dimensions. Cela permet d'exprimer des opérations, mais leur interprétation physique demandera qu'on remonte aux intensités (r, g, b).

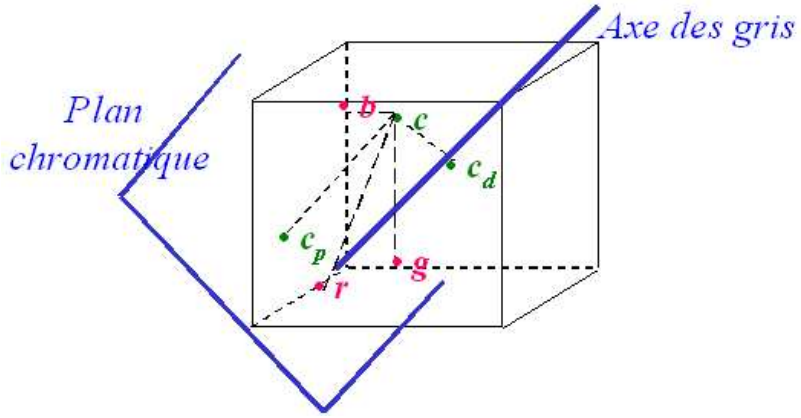


FIG. 2 – Cube E des intensités réduites, avec le plan chromatique et l'axe a-chromatique des gris. La figure décrit aussi bien le cube E' des couleurs videos réduites.

La figure 2 visualise ce cube unité. Le vecteur courant $\vec{x}'$, de coordonnées (r', g', b') se décompose aussi en deux vecteurs orthogonaux $\vec{c}'$ et $\vec{l}'$ portés par le plan chromatique et par l'axe a-chromatique. Ce dernier est la diagonale principale du cube passant par l'origine O et le plan chromatique lui est perpendiculaire en O. Les vecteurs $\vec{c}'$ et $\vec{l}'$ ont pour coordonnées

$$\begin{aligned} 3\vec{c}' &= (2r' - g' - b', 2g' - b' - r', 2b' - r' - g') \\ 3\vec{l}' &= (r' + g' + b', r' + g' + b', r' + g' + b') \end{aligned} \tag{5}$$

Considérons la composante rouge $r'(z)$ à travers une zone Z d'une image en couleur. Que signifie la moyenne de ce rouge dans Z ? Comme nous venons de le voir, seule la quantité $\bar{r} = \frac{1}{Z}\int (r'(z))^\gamma \, dz$ possède un sens physique, et sa visualisation demande qu'on lui affecte la correction γ en prenant $\bar{r}^{1/\gamma}$ (on fait abstraction pour le moment des constantes 1,099 et 0,099 de la Rel. (2)). Par ailleurs, les segmentations que nous avons en vue visent à découper l'espace en régions Z de couleur presque constante. En tout point $z \in Z$, on peut alors écrire, en appliquant le développement limité

$$r(z) = r'(z)^\gamma = \overline{r}'^\gamma \left[1 - \frac{\overline{r}' - r'(z)}{\overline{r}'}\right]^\gamma = \overline{r}'^\gamma \left[1 - \gamma\left(\frac{\overline{r}' - r'(z)}{\overline{r}'}\right) + \varepsilon(r)\right]$$

où $\bar{r}' = \frac{1}{Z}\int_Z r'(z)dz$. Lorsqu'on passe à la moyenne dans Z, le coefficient du terme en γ dans

le membre de droite s'annule et l'on trouve après correction γ :

$$(\bar{r})^{1/\gamma} = \bar{r}' + \bar{\epsilon}\left(r'\right) \tag{6}$$

Ainsi, la moyenne arithmétique du rouge vidéo r' est égale au second ordre près à la moyenne de l'intensité r suivie d'une correction γ. Le résultat reste vrai si l'on rajoute les coefficients de la Rel. (2), ainsi d'ailleurs que pour les zones Z faiblement rouges, i.e., régies par la première des Rel. (2), et il s'étend aux deux autres couleurs primaires. La Rel. (6) fournit donc une justification théorique aux images mosaïques qui apparaissent fréquemment dans les techniques de segmentation.

4 Brillance

Du point de vue de la physique, la brillance n'est autre que l'intégrale du spectre de puissance, i.e., ici, la somme des trois composantes r, g, b, qui le représentent. Pour la colorimétrie, cette somme doit être pondérée relativement à la sensibilité spectrale de l'oeil.

La *Rec. 709* de la CIE définit un point blanc et trois fonctions de pondération du spectre qui conduisent à R_{709}, G_{709} et B_{709}, puis la luminance

$$Y_{709} = 0.212R_{709} \ + 0.715G_{709} \ + 0.072B_{709} \tag{7}$$

ainsi que la luminance Y_W du point blanc associé. Les trois coefficients de la Rel. (7) tiennent compte de la sensibilité spectrale de la vision humaine, estimée par des mesures colorimétriques. Cette luminance Y_{709}, fonction linéaire d'intensités lumineuses a donc la même dimension énergétique que celles-ci ($watts/m^2$).

La vision humaine répond de manière non linéaire aux intensités, mais par des lois de type $di/i = constante$, qui engendrent des réponses logarithmiques. De même qu'on a tenu compte de la sensibilité spectrale de l'oeil, de même il convient de ne pas ignorer sa sensibilité énergétique. Or, par chance, cette dernière se modélise de façon à peu près satisfaisante par la Rel. (2) : par exemple, une réduction de 20% de l'intensité est perçue comme une diminution de moitié. A la suite de nombreux auteurs, on peut donc considérer les transformations

$$r' = r^{1/\gamma} \qquad g' = g^{1/\gamma} \qquad b' = b^{1/\gamma} \tag{8}$$

pour $\gamma \simeq 2.2$ comme engendrant des *intensités perceptuelles*. C'est dans cet esprit que la *Rec. BT 601-E* propose comme mesure de brillance le *luma*

$$y'_{601} = 0.299r' + 0.587g' + 0.144b'. \tag{9}$$

Toutefois ce luma, établi à partir des grandeurs videos, n'a pas la dimension d'énergie de Y_{709}, donc pas davantage les propriétés d'additivité qui en découlent. De son côté la CIE prend un exposant légèrement différent pour définir la luminosité (*lightness*) comme

$$l^* \ = 116(\frac{Y_{709}}{Y_W})^{1/3} \ - 16 \qquad\qquad Y_{709} \geq 0.0089Y_W.$$

Les opérations de segmentation en traitement d'image, quant à elles, ne portent pas sur une brillance *perçue*, mais sur celle de *l'objet analysé*. En microscopie, les colorations histologiques vont le plus souvent du bleu au violet, le spectre d'un coucher de soleil, ou d'un visage ne correspondent en rien aux poids donnés aux rouge, vert et bleu des Rel. (7) ou (9). En l'absence d'informations a priori sur les spectres des objets d'étude, l'objectif de segmentation conduit donc à prendre pour brillance une fonction *symétrique* des couleurs primaires.

Si sur le plan spectral la contrainte de symétrie nous écarte du point de vue colorimétrique, on peut en compensation s'en rapprocher sur le plan des énergies perçues. Considérons dans l'espace E des intensités un vecteur x de direction donnée par $x_o = r_o$, g_o, b_o mais dont l'intensité varie, i.e.

$$x = (\lambda r_0, \lambda g_0, \lambda b_0) \qquad \lambda \in [0, \lambda_{\max}]$$

Le point x décrit le segment S partant de O, passant par (r_o, g_o, b_o) et se terminant sur les bords du cube unité. Il lui correspond dans l'espace E' des couleurs vidéo le point x' :

$$x' = \left((\lambda r_0)^{1/\gamma}, (\lambda g_0)^{1/\gamma}, (\lambda b_0)^{1/\gamma}\right) = \lambda^{1/\gamma} x'_0 \tag{10}$$

avec $x'_0 = r_0^{1/\gamma}$, $g_0^{1/\gamma}$, $b_0^{1/\gamma}$. Le point x' décrit donc lui aussi un segment S' dans E'. Si nous voulons que quand x se déplace, sa brillance perceptuelle (dans E) semble additive, il faut d'après la relation (10) que la brillance correspondante de x' soit fonction linéaire de ses trois composantes primaires. Enfin si l'on veut que ces deux valeurs varient de 0 à 1, la seule possibilité consiste alors à prendre pour brillance vidéo la moyenne arithmétique m' des couleurs primaires :

$$m' = \frac{1}{3}(r' + g' + b') \tag{11}$$

Posons $\lambda' = \lambda^{1/\gamma}$. Les deux expressions

$$\begin{aligned} |m(x_1) - m(x_2)| &= |\lambda_1 - \lambda_2|\, m(x_0) \\ \left|m'(x'_1) - m'(x'_2)\right| &= \left|\lambda_1^{1/\gamma} - \lambda_2^{1/\gamma}\right| m'\left(x'_0\right) \end{aligned}$$

sont des distances, différentes, dans S et S' respectivement. La présence de l'exposant $1/\gamma$ donne à la seconde une signification d'homogénétité perceptuelle. Mais le traitement d'image demande davantage, car on doit pouvoir exprimer qu'un point coloré de E' (ou plus généralement un ensemble de points) se rapproche d'un autre, même lorsque les deux points ne sont pas alignés avec l'origine. Or, la moyenne (11) n'est autre que la restriction au cube E'de la norme L_1, définie, elle, dans l'espace $\mathbb{R}^3$ tout entier (i.e. $r', g', h' \in [-\infty, +\infty]$). D'ailleurs la norme L_1 est elle-même l'élément pour $\alpha = 1$ de la famille des normes dites de Minkowski

$$n\left(x'\right) = \left(\left|r'(x)\right|^\alpha + \left|g'(x)\right|^\alpha + \left|b'(x)\right|^\alpha\right)^{1/\alpha} \qquad \alpha \geq 1$$

(Pour $\alpha = 2$, on trouve la norme euclidienne, et pour $\alpha = \infty$, la norme "max"). Par conséquent ([5] section VII-1-4), pout tout $\alpha \geq 1$, la fonction

$$d\left(x'_1, x'_2\right) = n\left(x'_1 - x'_2\right) \qquad x'_1, x'_2 \in \mathbb{R}^3$$

est une distance, ainsi, bien sûr, que sa restriction au cube unité E'.

Pour $\alpha = 1$, on dispose donc d'un même concept pour exprimer la brillance $m'(x')$ et la distance $d(x'_1, x'_2) = m'(|x'_1 - x'_2|)$ entre les points colorés de E'. Cette dernière est importante, car en segmentation, beaucoup d'algorithmes qui ont été établis pour les fonctions numériques s'étendent aux fonctions vectorielles lorsqu'on dispose d'une distance (e.g. les lignes de partage des eaux). Qui plus est, dans le cube E, la brillance $m(x)$ résulte elle-même d'une pondération du spectre de la lumière, par la moyenne des pondérateurs du rouge, du vert et du bleu. Cette propriété n'est plus vérifiée en norme L_2.

5 Saturation

La CIE s'est davantage intéressée aux diverses formulations de la brillance (luminance, luminosité ...) qu'à la saturation. Elle définit cette dernière comme "*the colourfulness of an area judged in proportion to its brightness*"[4]. Il s'agit donc de la portion de spectre uniforme (i.e. de gris) dans le spectre d'une couleur, toute couleur monochromatique maximale étant de saturation unité, et tout triplet $r = g = b$ de saturation nulle.

Intuitivement, on comprend bien ce que la CIE veut dire, mais sa définitition de la saturation se prête à diverses interprétations. Lorsqu'on part d'un point $x \in E$, on peut construire beaucoup de chemins le long desquels la densité de couleur varie par rapport à la brillance. Par exemple, dans la Fig. (2), supposée représenter le cube E, on peut prendre la perpendiculaire xc au plan chromatique, ou la perpendiculaire xl à l'axe des gris, ou encore l'axe Ox, etc... Quel chemin choisir ?

En fait, ces ambiguités disparaissent dès qu'on se place dans le plan chromatique. Le cube E s'y projette selon l'hexagone H centré en O. Considérons un point $x_o \in E$, de projection c_0 dans H, avec c_0 distinct de l'origine. Suivant la CIE, définissons comme saturation toute fonction non négative sur l'axe Oc qui *augmente* à partir de O ; en O, elle vaut zéro (gris pur) et il prend sa valeur maximale quand on atteint le bord de l'hexagone en $c_{\max}$ (couleur saturée). Le long du segment $[0, c_{\max}]$, la teinte reste constante, et celle du segment $[0, \bar{c}_{\max}]$ de direction opposée est dite *complémentaire* de la précédente. Pour tout point $c \in [0, c_{\max}]$, on a $c = \lambda c_0$, $\lambda \geq 0$. Ainsi, à $c_0 \in H$ fixé, la saturation $s(c) = s(\lambda c_0)$ est une fonction de λ seulement, et cette fonction est croissante.

Il faut ensuite remonter au cube tridimensionnel E, car le point c_0, projection de x_0, n'est qu'un intermédiaire de calcul (d'ailleurs $c_0 \notin E$).On définit alors la saturation $s(x_0)$ du point $x_0 \in E$ par

$$s(x_0) = s(c_0)$$

Quand un point $x \in E$ s'écarte du plan chromatique le long de la perpendiculaire $c_0 x_0$ à ce plan, sa proportion de gris augmente, mais sa saturation $s(x)$ ne change pas : il s'agit bien d'une donnée *chromatique* et non pas d'*énergie* de l'intensité lumineuse.

Faisons varier maintenant le point c dans le segment $[0, \bar{c}_{\max}]$ opposé à $[0, c_{\max}]$ dans le plan chromatique. Nous pouvons toujours écrire

$$c \in D(c_0) \qquad \Longleftrightarrow \qquad c = \lambda c_0 \qquad \lambda(\bar{c}_{\max}) \leq \lambda \leq \lambda(c_{\max})$$

où $\lambda(\bar{c}_{\max})$ indique le rapport de proportionalité, négatif, entre c_0 et c. Cette équivalence purement vectorielle s'interprète physiquement si l'on étend aux diamètres $D(c_0) = [0, c_{\max}] \cup [0, \bar{c}_{\max}]$, $c_0 \in H$, de l'hexagone H la saturation définie jusqu'ici le long de ses rayons. On le fait en posant pour $c \in D(c_0)$, $s(c) = s(\lambda c_0) = s(|\lambda| c_0)$. Deux points diamétralement opposés ont alors la même saturation ; plus généralement, si $c_1 \in [0, c_{\max}]$ et $c_2 \in [0, \bar{c}_{\max}]$, alors $c_1 + c_2 = (\lambda_1 + \lambda_2) c_0$ avec $\lambda_1 \geq 0$ et $\lambda_2 \leq 0$. Comme s est une fonction croissante, il vient

$$c_1 \in [0, c_{\max}], \qquad c_2 \in [0, \bar{c}_{\max}] \qquad \Longrightarrow \qquad s(c_1 + c_2) \leq s(c_1) + s(c_2). \tag{12}$$

On trouve en particulier l'expérience du disque de Newton, mais réduit à deux couleurs complémentaires, en prenant $c_1 = c_{\max}$ et $c_2 = \bar{c}_{\max}$.

Quand l'on part de l'espace vidéo, on a toujours la condition de croissance de s' sur les rayons de H' et de sa nullité sur l'axe des gris. A ces deux conditions s'ajoutent les deux contraintes, déjà énoncées, du traitement d'images, à savoir une saturation symétrique en r', g', b' et telle que $s'(x'_1 - x'_2)$ soit une distance dans E'.

Nous avons vu que la moyenne m' Rel.(11), définie dans le cube unité E' s'étendait à tout l'espace sous forme de norme L_1, et que $3m'(x')$ était à la fois la norme de x' et de sa projection l' sur la première diagonale (l'axe des gris), i.e.

$$L_1(x') = L_1(l') = 3m'(x')$$

Il est tentant dans ces conditions de ne pas changer de norme et de voir comment celle-ci s'exprime dans l'hexagone H' du plan chromatique. Compte tenu de la Rel. (5) il vient

$$s'(x') = L_1(c') = \frac{1}{3}\left[\left|2r' - g' - b'\right| + \left|2g' - b' - r'\right| + \left|2b' - r' - g'\right|\right]. \tag{13}$$

Par symétrie, $s'(x')$ ne dépend que des trois fonctions $max' = max(r', g', b,')$, $min' = min(r', g', b,')$, et $med' = mediane(r', g', b,')$. En les substituant à $r', g', b,'$ on obtient

$$s' = \begin{cases} \frac{3}{2}(max' - m') & \text{si} \quad m' \geq med' \\ \\ \frac{3}{2}(m' - min') & \text{si} \quad m' \leq med' \end{cases} \tag{14}$$

On trouvera dans [16] le détail du calcul conduisant à s', ainsi que celui de l'expression suivante de la teinte t', qui évite de mettre en jeu des lignes trigonométriques

$$t' = \frac{\pi}{3}\left[\lambda + \frac{1}{2} - (-1)^{\lambda}\frac{max' + min' - 2med'}{2s'}\right] \tag{15}$$

avec

$$\lambda = \begin{cases} 0 \text{ si } r > g \geq b \\ 1 \text{ si } g \geq r > b \\ 2 \text{ si } g > b \geq r \\ 3 \text{ si } b \geq g > r \\ 4 \text{ si } b > r \geq g \\ 5 \text{ si } r \geq b > g \end{cases} \tag{16}$$

La teinte t', coordonnée curviligne du cercle unité, est donc définie à 2π près. Son zéro t' dans l'équation (15) correspond au vecteur purement rouge. Pour $s' = 0$, le point coloré est sur l'axe des gris, et sa teinte n'a plus de sens.

Dans le cas de la norme L_1, la relation (13) entraîne que $s'\left(c'_1 - c'_2\right) = L_1\left(c'_1 - c'_2\right)$ est une distance dans le plan chromatique, mettant en jeu par conséquent à la fois la saturation et la teinte. D'autre part, comme L_1 est une norme, la relation (12) devient vraie pour tous les triplets c'_1, c'_2 et $c'_1 + c'_2$ qui sont sur un même diamètre de H'. On notera qu'il est question ici de la norme L_1 des projections c', la norme L_1 des vecteurs x' eux-mêmes étant la moyenne arithmétique. Enfin, les commentaires précédents s'appliquent également à la norme euclidienne et à la quantité max-min, qui toutes deux engendrent des distances dans l'hexagone chromatique H'.

Concernant le passage des variables vidéo aux variables d'intensité, notons tout d'abord que nous retrouvons la même relation entre les saturations video et celle des intensités qu'avec les moyennes, à savoir la Rel (6), dans les zones où la couleur varie peu. Par ailleurs, l'application $x'_0 = (r'_0, g'_0, b'_0) \to x_0\left(r'^{\gamma}_0, g'^{\gamma}_0, b'^{\gamma}_0\right)$ montre que $c' = \lambda c'_0$ devient $c = \lambda^{\gamma} c_0$, d'où

$$s'(x') = L_1\left(c'\right) = \lambda L_1\left(c'_0\right) = \lambda s'(x'_0) \Leftrightarrow s(x) = L_1\left(c\right) = \lambda^{\gamma} L_1\left(c_0\right) = \lambda s(x_0).$$

La norme L_1 est croissante sur les rayons et nulle pour les points gris *à la fois* dans l'hexagone chromatique H des intensités et dans celui, H', des variables video. Elle représente donc une saturation dans ces deux espaces. Elle vérifie l'inégalité (12) de type Newton le long de chaque diamètre de H comme de H'. Il est difficile d'aller plus loin, car deux points $x'_0, x'_1 \in E'$ dont les projections c'_0 et c'_1 sont sur un même rayon de H' n'ont pas nécessairement d'homologues x_0 et $x_1 \in E$ dont les projections soient alignées avec O.

Nous remarquerons, pour conclure, que la démarche à base de normes et de distances L_1 ou L_2 présente l'avantage considérable de *séparer les variables* quand on choisit des axes orthogonaux. Deux points x'_1 et $x'_2 \in E'$ qui se projettent au même point sur le plan chromatique ont même saturation, et s'ils se projettent au même point sur l'axe des gris, ils ont même brillance (mais cette dernière propriété disparait si l'on donne des poids différents aux couleurs dans la moyenne de brillance, puisqu'alors les axes ne sont plus orthogonaux).

6 Histogrammes bi-dimensionnels

On peut se demander si des normes aux propriété mathématiques identiques, comme L_1 ou L_2, décrivent aussi bien le comportement physique de la lumière dans l'espace. Pour juger de la pertinence des diverses normes, nous nous proposons de tester sur quelques images les histogrammes bi-dimensionnels *luminance/saturation* qu'elle induisent, et de les comparer au système HLS. Nous avons effectué ces comparaisons sur une douzaine d'images [3], dont trois, aux comportements bien significatifs, sont reportées ici. Dans tous les histogrammes l'axe des x représent toujours la brillance et l'axe des y la saturation ou la teinte selon le cas.

Rappelons que dans le système HLS la luminance l et la saturation s ont pour expressions

$$l = \frac{\max(r,g,b)+\min(r,g,b)}{2}$$

$$s = \begin{cases} \frac{\max(r,g,b)-\min(r,g,b)}{\max(r,g,b)+min(r,g,b)} & \text{si} \quad l \leq 0.5 \\ \frac{\max(r,g,b)-\min(r,g,b)}{2-\max(r,g,b)-\min(r,g,b)} & \text{si} \quad l \geq 0.5 \end{cases} \tag{17}$$

et que, comme sa saturation donne des résultats particulièrement aberrants (cf. [8]), on peut remplacer cette dernière par $max(x') - min(x')$. La quantité $max - min$ est une norme dans le plan chromatique, que l'on complète dans les histogrammes ci-dessous, dits "max-min", en prenant le luma (9) comme valeur de la brillance.

Les deux premières images, que l'on voit en figure (3), sont assez proches en ce sens qu'elles ne font que ré-émettre des lumières incidentes. Les histogrammes qui leur correspondent sont reportés ci-dessous, en plaçant la luminance en axe des x et la saturation en axe des y. L'histogramme HLS ne véhicule aucune information, quelques structures se dessinent en L_2 et $max - min$, mais les plus nettes viennent de la norme L_1. Les histogrammes de la seconde image ("bouée") corroborent le fait.

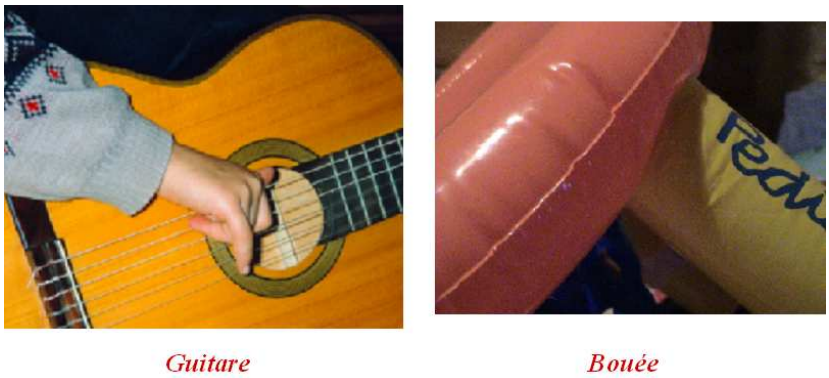


FIG. 3 – Deux images test.

Retournant aux images, voyons quels sont les pixels dont les couples ($luminance/saturation$) forment des alignements.

- Dans chaque masque dont les pixels forment des alignements, la teinte garde une valeur constante (ex. main de "guitare" : $t = 24 \pm 2$; ombres portées de "bouée" $t = 165 \pm 3$) ;
- Dans toutes les images, les droites d'alignement passent toutes soit par l'origine $(0, 0)$ du plan $lumminance/saturation$ (pentes > 0) soit par le point $(255, 0)$ (pentes < 0) ;
- Les modules des pentes sont toujours élevés.

Les alignements peuvent correspondre à trois types de régions, à savoir :

- ombres de teinte fixe, (ombres portées de "bouée"),

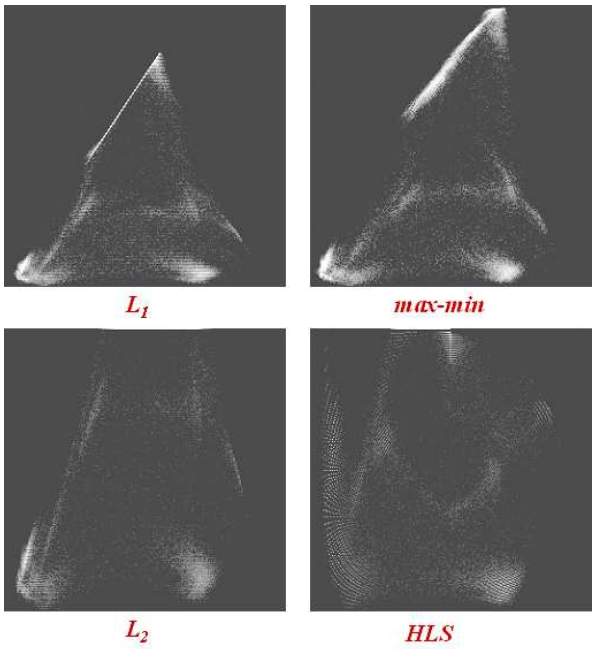


FIG. 4 – Histogrammes bi-dimensionnels de l'image "guitare".

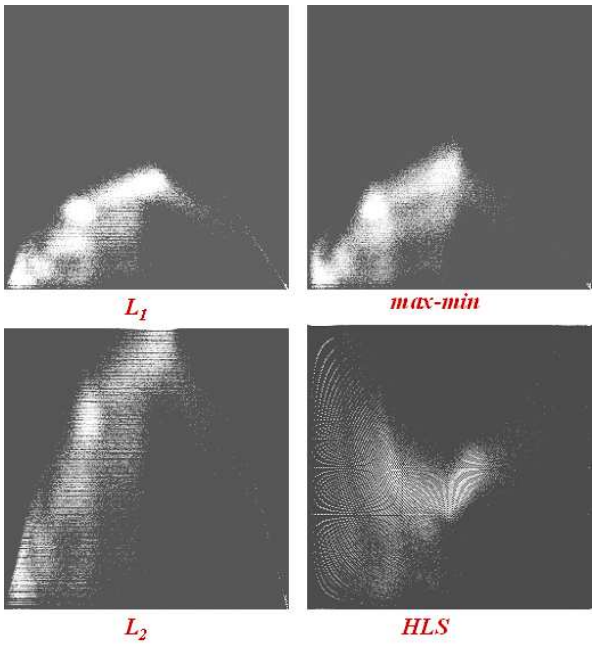


FIG. 5 – Histogrammes bi-dimensionnels de l'image "bouée".

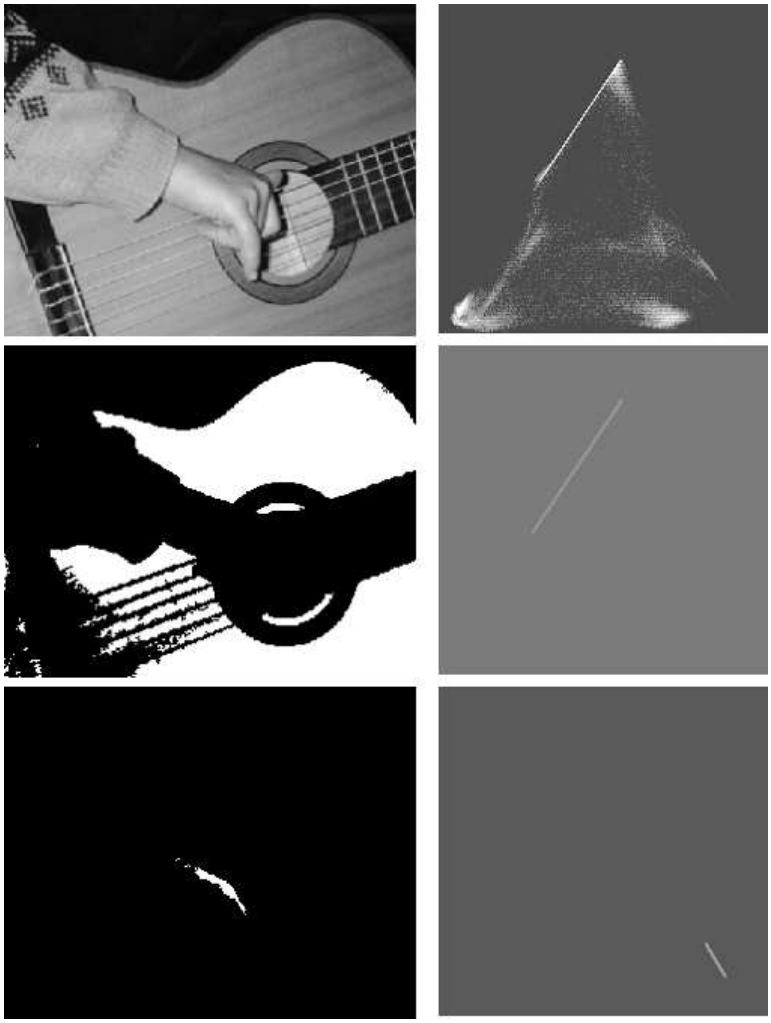


FIG. 6 – Zones de l'image "guitare" associées à des alignements.

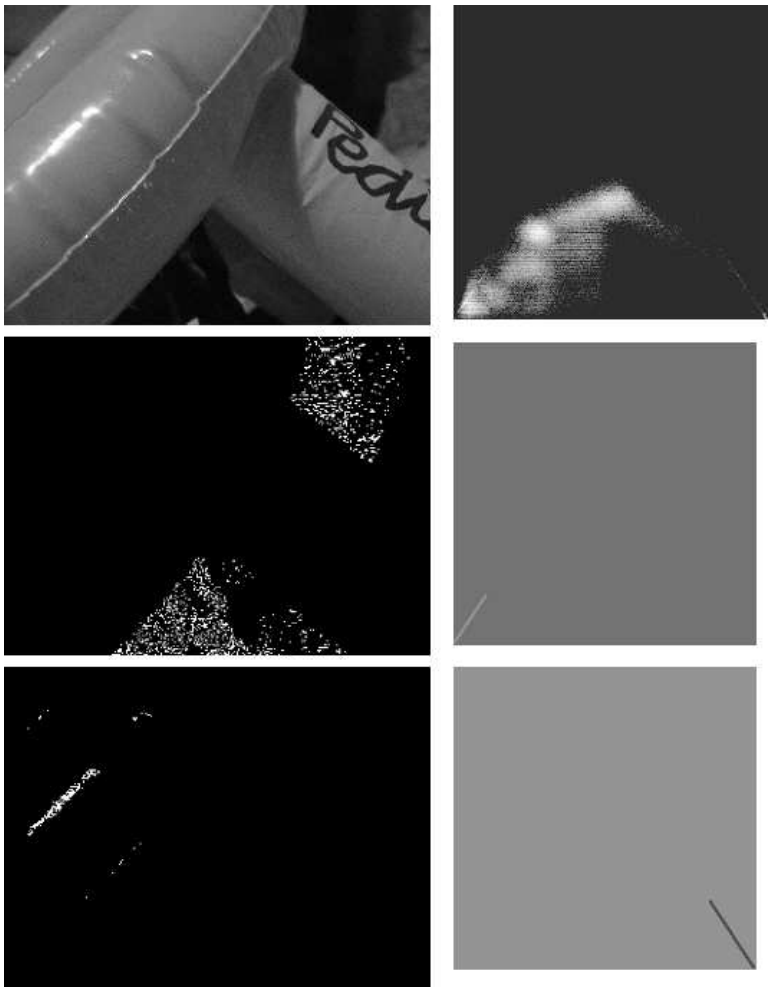


FIG. 7 – Zones de l'image "bouée" associées à des alignements.

– dégradés de la lumière sur un plan (“guitare”),
– reflets non complètement saturés (“guitare”, “bouée”).

Passons maintenant à une image plus complexe, figure (8), car elle comprend des sources émettrices de lumière (moniteurs de télévision, bandeau alpha-numérique). De plus, la lumière diffuse ambiante n’y est pas uniforme. On retrouve cependant des alignements, mais qui cette fois ne passant pas tous par les points $(0, 0)$, ou $(1, 0)$, et sont parfois parallèles. Dans le plan $luminance/teinte$ les horizontales (teinte constante) alternent avec des nuages de points.

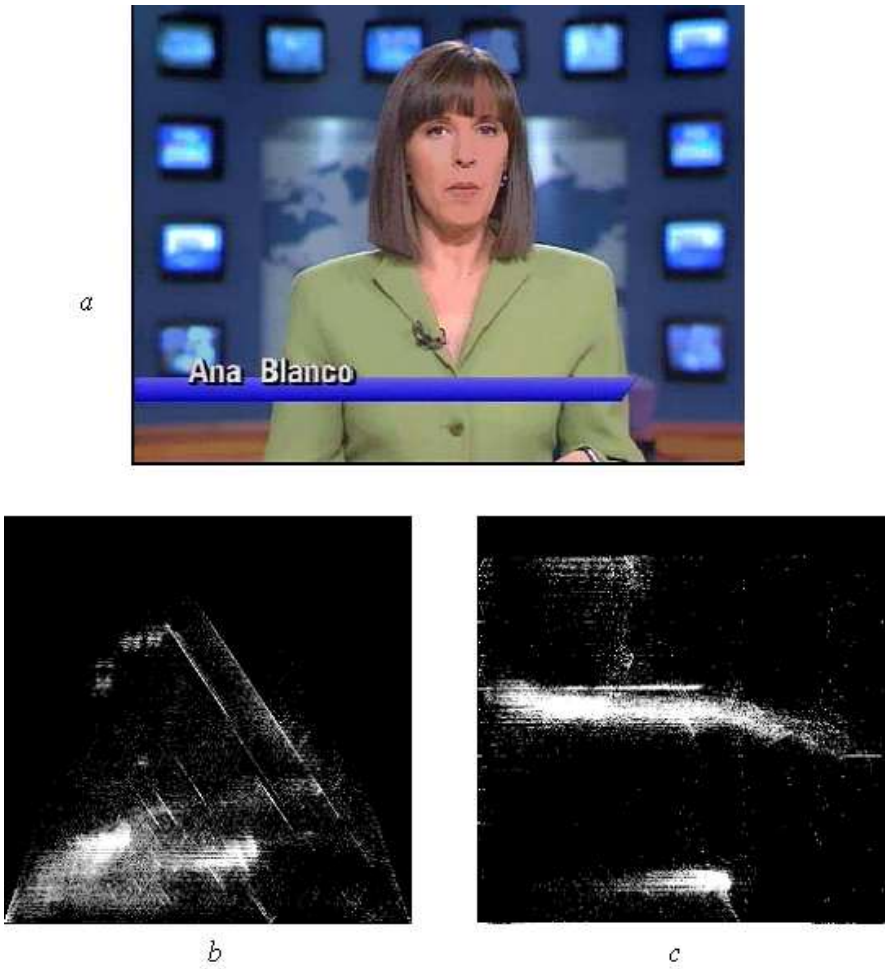


FIG. 8 – (a) Image “AnaBlanco”, (b) ses deux histogrammes $x = luminance/y = saturation$, (c) et $x = luminance/y = teinte$, tous deux en norme L_1. Le premier présente une série de segments obliques parallèles, le second des horizontales dont la plus importante est portée par l’axe des x, i.e. correspond à la teinte rouge.

Comme on peut le constater [fig.(9)], des alignements dans le plan $luminance/saturation$,

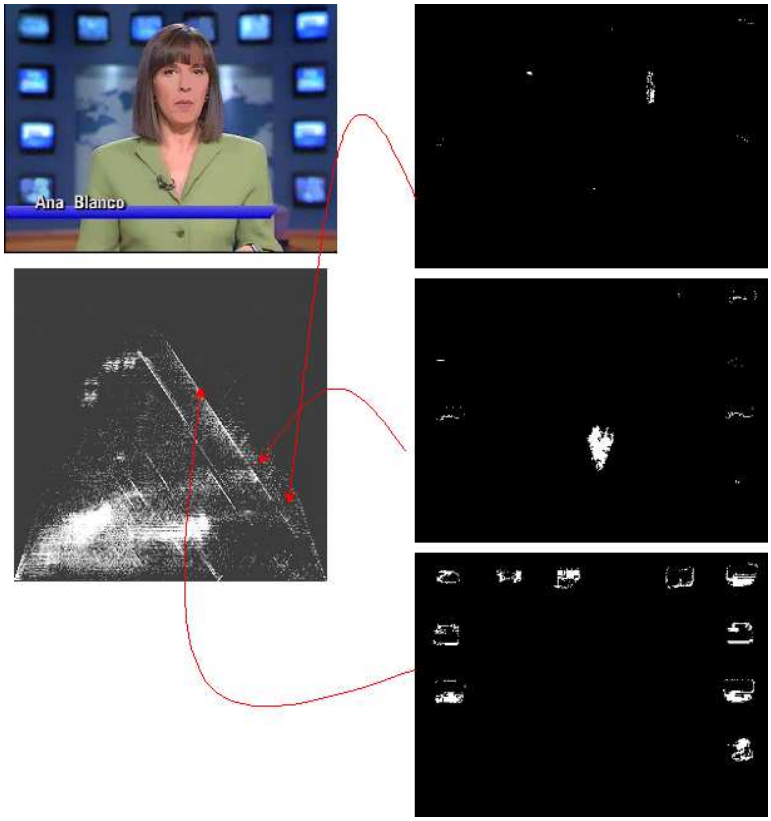


FIG. 9 – Zones de l'image "AnaBlanco" associées à des alignements.

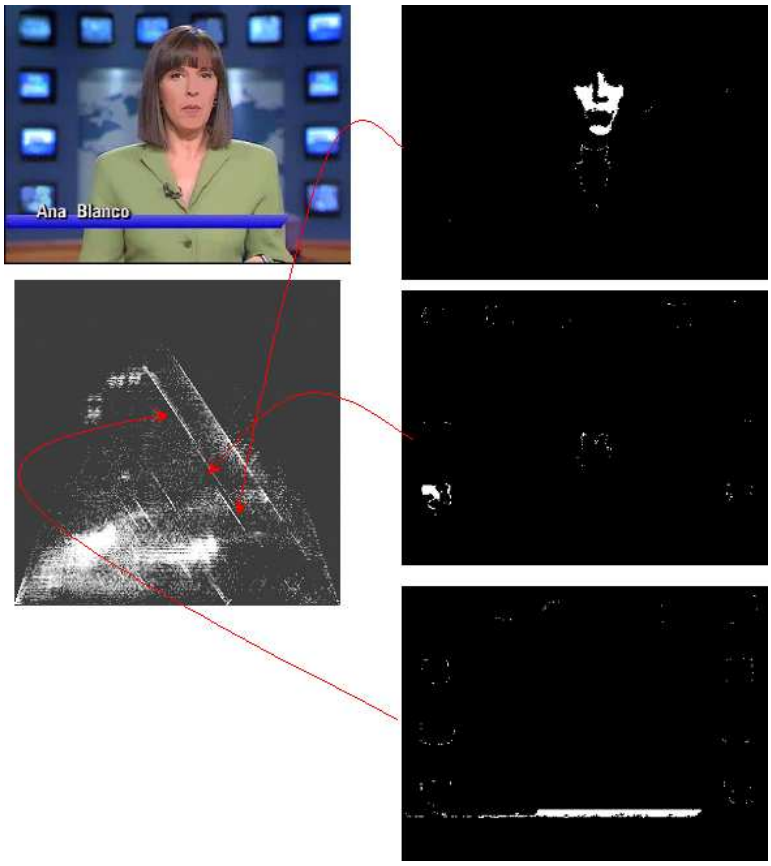


Fig. 10 – Autres zones de l'image "AnaBlanco" associées à des alignements, parallèles aux précédents.

i.e. en projection, peuvent correspondre à deux horizontales parallèles, mais de hauteur différentes dans l'espace tridimensionnel TSL.

Des structures aussi typées, des droites aussi marquées poussent à rechercher une interprétation physique du phénomène. C'est ce que nous allons faire maintenant. Mais hors tout modèle physique, une constatation s'impose : la norme L_1 a permis d'extraitre des images des régions de reflets, des régions d'ombres et d'autres de dégradés de lumières. Aucune autre norme, ni l'analyse de la seule image de la luminance, n'auraient conduit à pareil résultat.

7 Spectres régionalisés linéairement (Modèle SRL)

Les histogrammes bidimensionnels précédents suggèrent un modèle qui combine de façon linéaire les spectres de lumière issus d'un point. Le modèle est présenté avec plus de détails, et surtout plus de justifications dans [17]. Nous en reprenons l'essentiel ci-dessous. Considérons la lumière émanant d'une zone Z de l'objet d'étude, et désignons par $sp\,(\nu; z)$ le spectre des inensités au point $z \in Z$. Nous dirons que ce spectre est *régionalisé linéairement dans* Z quand pour tout $z \in Z$ on peut décomposer $sp\,(\nu; z)$ en somme d'un premier spectre $sp_0\,(\nu)$, indépendant du point z, plus d'un second, $\overline{\omega}(z)sp_1(\nu)$, qui varie proportionnellement d'un point à l'autre de Z. En tout point $z \in Z$, on a donc

$$sp\,(\nu; z) = sp_0\,(\nu) + \overline{\omega}(z)sp_1(\nu) \tag{18}$$

où $\overline{\omega}\,(z)$ est un réel qui ne dépend que de z, et où sp_0 et sp_1 sont deux spectres fixes. Le modèle SRL est donc di-chromatique, tout comme celui, très connu, de S.A. Shafer [18], mais dans un sens bien différent. En effet,

- le modèle de Shafer décrit la reflexion de la lumière sur un matériau homogène, et la décompose en la somme de deux termes, à savoir les spectres de corps et de surface. Ici, la décomposition (18) porte autant sur l'émission que sur la réflexion de la lumière au point z. Elle s'applique par exemple à la lumière transmise par une fenêtre avec un voilage, ou à celle d'un récepteur de télévision en fonctionnement, autant qu'à des réflectances passives ou au brillant de Obein *et al.* [13] ;
- de plus, dans le modèle SRL l'accent est mis sur la *variation spatiale* du spectre. Celle-ci est introduite par le poids $\overline{\omega}(z)$, qui est une fonction du point z, mais pas du spectre sp_1. Ce poids peut mettre en jeu des cosinus, quand une zone réfléchissante est arrondie par exemple, ou s'il y a variation d'angle d'un faisceau incident, etc...
- enfin, les deux composantes de la reflexion de Shafer sont nécessairement positives, alors que dans le modèle SRL la fonction d'espace $\overline{\omega}(z)$ peut être négative. Le terme $\overline{\omega}(z)sp_1$ décrit alors une absorption.

En somme il s'agit dans un cas (Shafer) de l'interaction lumière/matière, et dans l'autre (SRL) de l'interaction lumière/espace. Les trois spectres sp, sp_0 et sp_1 ne nous sont connus qu'à travers des fonctions d'absorption engendrant le triplet $R_{709}, G_{709,}B_{709}$. Il en va de même pour le triplet (X, Y, Z) ou encore (L, M, S), selon l'espace de travail choisi. Il s'agit dans tous les cas de *produits scalaires* des spectres par telle ou telle fonction d'absorption des fréquences

ν. En fait, le modèle régionalisé linéaire est indépendant du triplet de couleurs de base, et dans ce qui suit les symboles r, g et b doivent être entendus dans un sens générique. Posons

$$r(z) = \int sp(\nu; z)\,\xi(\nu)d\nu = \int [sp_0(\nu) + \overline{\omega}(z)sp_1(\nu)]\,\xi(\nu)\,d\nu = r_0 + r_1\overline{\omega}(z) \tag{19}$$

ainsi que

$$g(z) = \int sp(\nu; z)\,\chi(\nu)\,d\nu = g_0 + g_1\overline{\omega}(z) \tag{20}$$

et

$$b(z) = \int sp(\nu; z)\,\psi(\nu)\,d\nu = b_0 + b_1\overline{\omega}(z) \tag{21}$$

où ξ, χ et ψ sont les trois fonctions d'absoption qui génèrent les couleurs primaires r, g et b.

Comme sp_0 and sp_1 sont des spectres de puissance, ils engendrent des intensités r, g, b. Or dans les histogrammes ci-dessus la norme L_1 est appliquée aux variables vidéo $r' = r^{1/\gamma}$, $g' = g^{1/\gamma}$, $b' = b^{1/\gamma}$ (si l'on néglige la zone au voisinage de l'origine). La relation (19) dévient par conséquent

$$r'(z) = [r(z)]^{1/\gamma} = [r_0 + \overline{\omega}(z)\,r_1]^{1/\gamma}, \tag{22}$$

le vert et le bleu vidéo g' et b' s'établissant de manière similaire.

Le modèle régionalisé linéaire est-il capable expliquer les alignements dus à des variables video, et ce malgré les corrections gamma ? Pour étudier cette question, nous nous placerons, par souci de simplicité, dans le cas où $r' \geq g' \geq b'$, et $m' \geq g'$. Il vient alors

$$\begin{aligned} 3m'(z) &= r'(z) + g'(z) + b'(z) \\ 2s'(z) &= 2r'(z) - g'(z) - b'(z) \end{aligned}$$

Points alignés avec le point noir

Dans les histogrammes *luminance/saturation* en norme L_1, on trouve souvent des alignements qui prolongés, passent par le point $(0, 0)$, de luminance et de saturation nulles. La table de "guitare" ou la zone d'ombre de "bouée" en sont deux exemples.

Supposons que dans la Rel. (18) qui définit le spectre régionalisé linéaire le terme $sp_0(\nu; z)$ soit identiquement nul. On a alors $r(z) = \overline{\omega}(z)r_1$, puis

$$r'(z) = r^{1/\gamma} = \overline{\omega}^{1/\gamma} r_1^{1/\gamma} = \overline{\omega}^{1/\gamma} r_1'(z),$$

ainsi que les relations homologues pour les verts et les bleus. On en déduit

$$3m'(z) = \overline{\omega}^{1/\gamma}(z)\left[r_1^{1/\gamma} + g_1^{1/\gamma} + b_1^{1/\gamma}\right] = 3\overline{\omega}^{1/\gamma}(z)m_1'$$

et

$$2s'(z) = 2r'(z) - g'(z) - b'(z) = \overline{\omega}^{1/\gamma}(z)\left[2r_1^{1/\gamma} - g_1^{1/\gamma} - b_1^{1/\gamma}\right]$$

d'où $m'(z)s'_1 = m'_1 s'(z)$. De plus, si l'on se place dans l'espace E des intensités, on montre de la même manière que $m(z)s_1 = m_1 s(z)$. La nullité du spectre constant $sp_0(\nu)$ entraîne donc à la fois que m' et s' d'une part et m et s de l'autre soient proportionnels. L'alignement *vidéo* correspond à une zone où le spectre des *intensités* varie proportionnellement d'un point à l'autre.

Points alignés avec le point blanc

Passons aux reflets. Les images de "guitare" et de "bouée", engendrent des alignements vidéos le long de droites passant par $(1, 0)$, i.e. le point de luminance maximale et de saturation nulle. Supposons cette fois que le terme spectral $sp_0(\nu; z)$ soit constant et égal a 1, et de plus que les trois couleurs r_1, g_1, b_1 ne soient pas identiques. Il vient alors

$$r(z) = 1 + \overline{\omega}(z) r_1 \tag{23}$$

ainsi que les deux autres relations pour $g(z)$ et $b(z)$. Après correction gamma, on a

$$r'(z) = (1 + \overline{\omega}(z) r_1)^{1/\gamma}.$$

Or dire qu'on est au voisinage du point de luminance maximale revient à dire que r_1, g_1 et b_1, sont petits vis-à-vis de 1, ou encore que

$$r'(z) = \left(1 + \frac{\overline{\omega}(z)}{\gamma} r_1 + \varepsilon(r_1)\right) \tag{24}$$

d'où $m'(z) = 1 + \frac{\overline{\omega}(z)}{\gamma} m_1$ et $s'(z) = \frac{\overline{\omega}(z)}{\gamma}[2r_1 - g_1 - b_1]$. Observons que les deux conditions $r_1 \geq 0$ et $r'_1 \leq 1$, jointes à la Rel. (24) impliquent que coefficient $\overline{\omega}(z)$ soit négatif. De plus, comme les trois couleurs r_1, g_1, b_1 sont distinctes, la condition $s'(z) \geq 0$ implique à son tour que la quantité $2r_1 - g_1 - b_1$ soit strictement négative. Posant $\sigma_1 = -(2r_1 - g_1 - b_1) > 0$ (σ_1 n'est pas la saturation du point z_1), on obtient donc la relation linéaire à coefficients positifs

$$m'(z) = 1 - \frac{m_1}{\sigma_1} s'(z). \tag{25}$$

Comme dans le cas précédent, la moyenne $m(z)$ et la saturation $s(z)$ des intensités sont reliés elles aussi par l'équation (25) : c'est une conséquence directe de l'équation (23). Cette fois encore l'histogramme vidéo et celui des intensités véhiculent la même information, qui correspond aux zones de reflets à la limite du blanc.

Points alignés avec un point gris

On trouve dans certaines images, comme "AnaBlanco" des séries d'alignements parallèles dont les droites passent par un point de saturation nulle (ou très faible), mais dont la luminance est strictement comprise entre 0 et 1. L'interprétation que nous venons de donner pour la situation des reflets s'étend à ce cas de figure. On suppose toujours que $r_0 = g_0 = b_0$, mais

avec $0 < r_0 \leq 1$, et l'on considère que les termes $\overline{\omega}(z)r_1$, $\overline{\omega}(z)g_1$, et $\overline{\omega}(z)b_1$ sont assez petits vis-à-vis de r_0. Ils vient alors,

$$r'(z) = (r_0 + \overline{\omega}(z)r_1)^{1/\gamma} = r_0^{1/\gamma} + r_0^{1/\gamma-1}\frac{\overline{\omega}(z)}{\gamma}r_1,$$

ainsi que les relations homologues pour g' et b'. D'où

$$m'(z) = r_0^{1/\gamma} + r_0^{1/\gamma-1}\frac{\overline{\omega}(z)}{\gamma}m_1,$$

$$s'(z) = -r_0^{1/\gamma-1}\frac{\overline{\omega}(z)}{\gamma}\sigma_1,$$

et finalement

$$m'(z) = r_0^{1/\gamma} - \frac{m_1}{\sigma_1}s'(z).$$

Lorsque la composante colorée (r_1, g_1, b_1) reste la même, mais que la composante grise (r_0, g_0, b_0) prend successivement plusieurs intensités, il correspond à chacune d'elles un alignement qui garde la même pente $\frac{m_1}{s_1}$. La propriété s'étend aux histogrammes des intensités elles-mêmes. Enfin, on déduit de l'équation (15) que la teinte reste *constante* dans chacune de ces zones.

7.1 Comparaison avec d'autres représentations

Les modèles précédents, et en particulier les cas de la *lumière grise* et des *effets proportionnels purs* s'appliquent aussi à la norme $max - min$. L'on doit donc s'attendre à trouver des alignements dans le plan (l, s) pour ces deux cas. On en trouve effectivement, mais ils sont nettement moins affirmés qu'en norme L_1. Cela vient de ce que la saturation L_1 est plus précise que $max - min$, puisqu'elle vaut $1/2\,[(\max - med) + (\max - \min)]$ ou $1/2\,[(\max - \min) + (med - \min)]$ selon les cas. Si par conséquent l'hypothèse de lumière grise pour s_0 est valable à une erreur aléatoire près, ou si cette hypothèse signifie plutôt que seule la partie centrale du spectre est constante, alors le terme de saturation $u_0\,(x) - \frac{v_0(x)-w_0(x)}{2}$ (norme L_1) se rapproche davantage de zero en moyenne que $u_0\,(x) - w_0\,(x)$ (norme $\max - \min$).

Revenons sur les raisons qui nous ont poussés à ne pas poursuivre la piste ouverte par Lambert et Carron [9]. Ces auteurs construisent une base orthogonale formée de la diagonale principale du cube r, g, b (ou r', g', b') puis de deux axes orthogonaux dans le plan chromatique, donnant les composantes Ch_1 et Ch_2. Ils définissent ensuite le chroma $C = \sqrt{Ch_1^2 + Ch_2^2}$, de sorte que finalement, la brillance est prise en norme L_1 (c'est la même que la nôtre), et le chroma (ou la saturation) en norme L_2 . Si l'on recalcule le modèle SRL dans ce cadre, on continue de trouver des relations linéaires entre brillance et chroma, mais nous préférons nous fixer sur une seule norme plutôt que de mêler L_1 et L_2 sans raisons particulières. D'autant que les figures 4 et 5 n'engagent pas tellement à aller chercher la norme L_2. La suite de l'article de Lambert et Carron développe une méthode de segmentation par croissance de régions dont l'esprit est complètement différent de ce que nous proposons ci-dessous.

8 Segmentation par saturation pondératrice I : gradient couleur

Après l'examen des histogrammes bi-dimensionnels, nous orientons maintenant le projecteur vers la segmentation, en nous demandant comment formaliser l'observation psychovisuelle de l'introduction, selon laquelle l'oeil segmente à l'aide des variations de teinte dans les régions saturées, et des variations de luminance dans les regions grises. Nous examinerons successivement deux méthodes [2], toutes deux initialement dans le cadre de la norme L_1 pour ensuite les comparer avec les autres normes.

L'algorithme de la ligne de partage des eaux, ou LPE, construit les lignes de crêtes d'une fonction numérique f [2], [3]. Quand on l'applique au module du gradient de f, les lignes de crêtes du gradient délimitent les contours des points d'inflexion, que l'on interprète perceptuellement comme des contours d'objets. La démarche s'étend facilement aux images en couleur à cause du passage obligé par un gradient. Il existe en effet un assez grand nombre de modules de gradients, plus ou moins classiques, pour les images couleur. Six d'entre eux (HLS, LAB, ...) ont été utilisés et comparés dans [3]. Celui qui donne les meilleurs résultats visuels se trouve être, précisément, celui où la saturation joue un rôle de balancier entre luminosité et teinte, en norme L_1. C'est lui que nous reportons ici.

Notons $\mathbf{f} = (l, s, h)$ la fonction vectorielle associée à une image en couleurs en représentation TLS, et par $|\nabla \mathbf{f}|$ le module de son gradient. 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 [7]. On pose donc

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

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

$$|\nabla \mathbf{f}| = \frac{s}{s_{\max}} \times |\nabla t| + (1 - \frac{s}{s_{\max}}) \times |\nabla l|. \tag{26}$$

Dans cette expression barycentrique, plus la saturation est élevée, plus la variation de teinte est prise en compte, ce qui était recherché.

On peut alors appliquer à la fonction $\mathbf{f}$, et à son gradient au sens de la rel. (26) la méthode de *segmentation* dite "*des chutes d'eau*", généralisation hiérarchique de la ligne de partage des eaux [12], qui procède comme suit :

1. on calcule $|\nabla \mathbf{f}|$,
2. puis la ligne de partage des eaux de $|\nabla \mathbf{f}|$, ce qui conduit à une première segmentation du plan,
3. on affecte à chacune des classes segmentées la médiane des rouges, des verts et des bleus de la fonction $\mathbf{f}$ dans la classe en question ; ce qui produit une nouvelle image en couleur $\mathbf{f}_1$,
4. on itère les trois premières étapes en partant de la fonction $\mathbf{f}_1$.

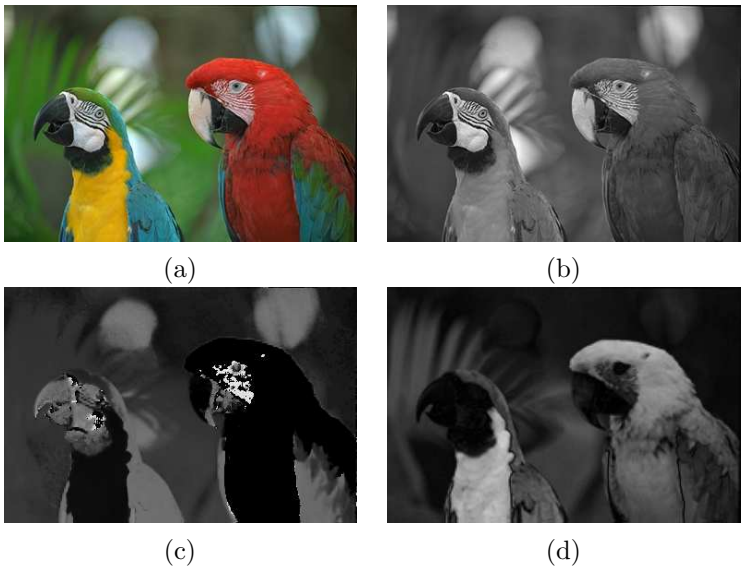


FIG. 11 – Image initiale de “parrots” (a), et décomposition de l'image couleur selon la représentation en norme L_1 : (b) luminance, (c) teinte (d) saturation.

Nous prendrons pour illustrer cette démarche l'image standard “parrots” de la figure 11, en procédant à quatre itérations de l'opérateur. Cela conduit aux mosaïques de la figure 12. A première vue les formes majeures semblent préservées. Pourtant une perte réelle d'information se manifeste quand on superpose les partitions à l'image initiale, en figure 13. Les trois premiers niveaux contiennent beaucoup trop de régions non significatives, et l'on doit attendre le quatrième pour une meilleure partition, mais pas vraiment convainquante. Des petites structures comme les yeux ont disparu, des frontières irréalistes apparaissent sur les plumes, etc...

9 Segmentation par saturation pondératrice II : combinaison de partitions

En fait, la démarche précédente pèche par une réduction trop précoce de l'information, qui devient scalaire dès le départ, quand on calcule le premier $|\nabla \mathbf{f}|$. On préservera mieux l'information pertinente en

1. segmentant *séparément* la luminance, la saturation et la teinte
2. et en combinant les partitions de la luminance et de la teinte à l'aide de la saturation, qu'on prendra comme *critère* pour choisir en chaque endroit l'une ou l'autre classe des segmentations de la luminance et de la teinte.

Il y a ici, comme précédemment avec le gradient couleur, plusieurs façons de construire les segmentations séparées, ainsi que de les mixer (en seuillant ou non la saturation, par

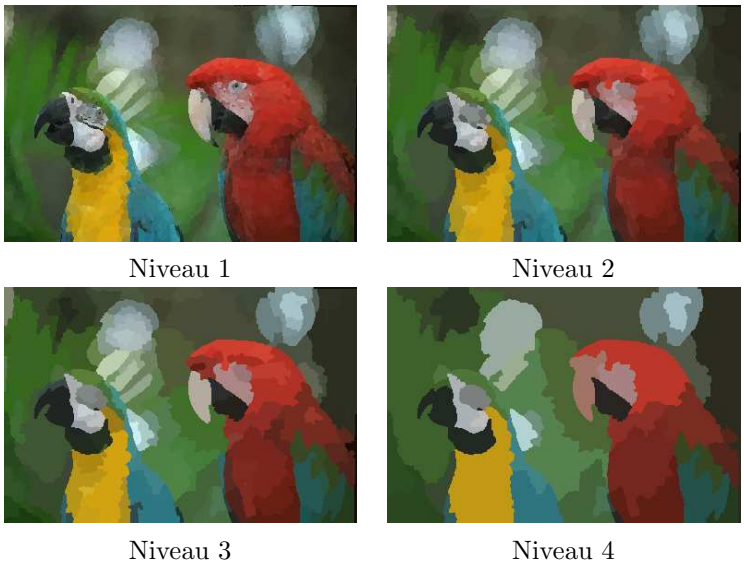


FIG. 12 – Quatre itérations de la hiérarchie des “chutes d’eau” qui met en jeu des images mosaïques.

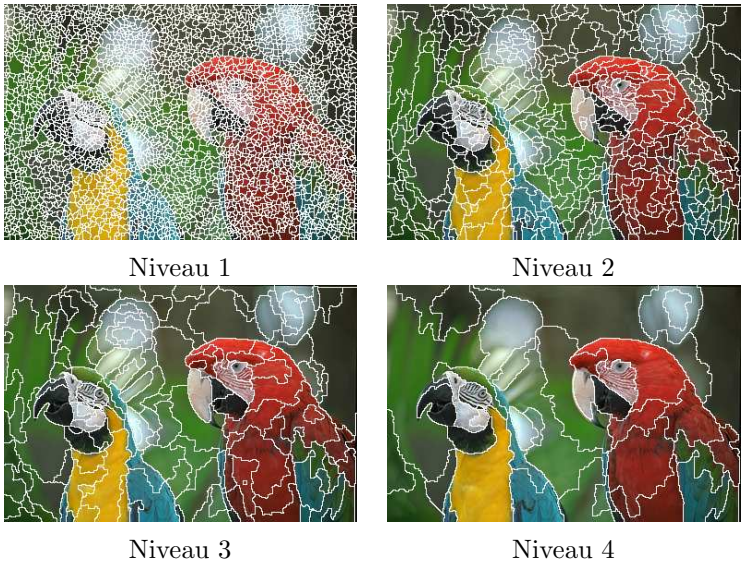
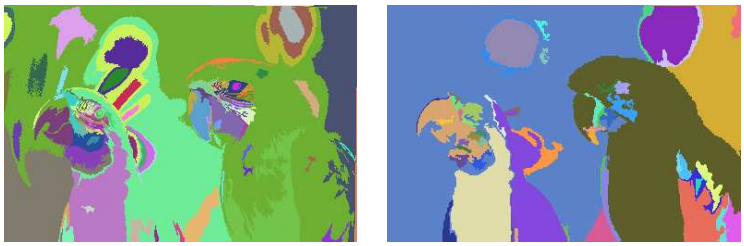


FIG. 13 – Partitions associées aux “chutes d’eau”, superposées à l’image initiale.

exemple). Nous renvoyons encore à la référence [3] pour une étude comparative, et nous focalisons sur la démarche qui a très sensiblement donné les meilleurs résultats. Elle repose sur la *segmentation par sauts* [15], définie pour les fonctions numériques, mais qui ne s'exprime pas par l'intermédiaire de leur gradient (de ce fait, elle n'aurait pas pu servir dans la démarche précédente).

Les trois canaux en norme L_1 de l'image "parrots" sont présentés en figure 11(b)-(d). Segmentons par sauts les deux images scalaires de la luminance et de la saturation (on met dans une même classe tous les points x où $f(x)$ diffère de moins de k d'un extremum, on retire ces classes du plan de l'image, et on itère). La méthode dépend du seul paramètre positif k des sauts.



(a) (b)

FIG. 14 – Segmentations selon la connection par sauts de valeur $k = 20$: (a) luminance, de fontières A_l, (b) teinte, de fontières A_t. Les deux partitions sont présentées en fausse couleur.

Quand ce paramètre croit, les sur-segmentations diminuent, mais en compensation des régions hétérogènes apparaissent. Un bon équilibre semble atteint pour $k = 20$ (pour des images scalaires à 256 niveaux de gris). Toutefois, il reste encore quelques petites régions à supprimer. On peut ne conserver que les classes d'une surface 50, et reconstituer la partition par squelette par zones d'influence (ou aussi bien par un algorithme de croissance de régions [6]). Les deux méthodes conduisent aux résultats à peu près identiques de la figure 14(a).

La démarche que nous venons de proposer pour la luminance s'applique aussi à la teinte. La seule différence vient de ce que l'on doit chosir une origine, nécessairement arbitraire, des teintes. Ce sera le rouge, selon l'usage (d'autres méthodes permettent un choix adaptatif pour chaque image de la valeur à prendre comme origine [3]). Les deux segmentations finales sont indiquées en figure 14.

9.1 Saturation et partition composite

Comment combiner ces deux partitions ? L'idée ici consiste à réduire l'image de saturation à un ensemble X_s, qui corresponde aux pixels de saturation élevée, de restreindre la partition de la teinte à X_s, et celle de la luminance à X_s^c, puis de réunir les deux résultats. Plus précisément, une classe de la partition synthétique est soit l'intersection d'une classe de la luminance avec la zone X_s^c de faible saturation, soit l'intersection d'une classe de la partition

de la teinte avec l'ensemble X_s de haute saturation. Si $A_m(x), A_h(x)$, et $A(x)$ désignent les classes de la brillance, de la teinte et de la synthèse au point x, il vient

$$\begin{aligned} A(x) &= A_m(x) \cap X_s^c \quad \text{si} \quad x \in X_s^c \\ A(x) &= A_h(x) \cap X_s \quad \text{si} \quad x \in X_s. \end{aligned} \tag{27}$$

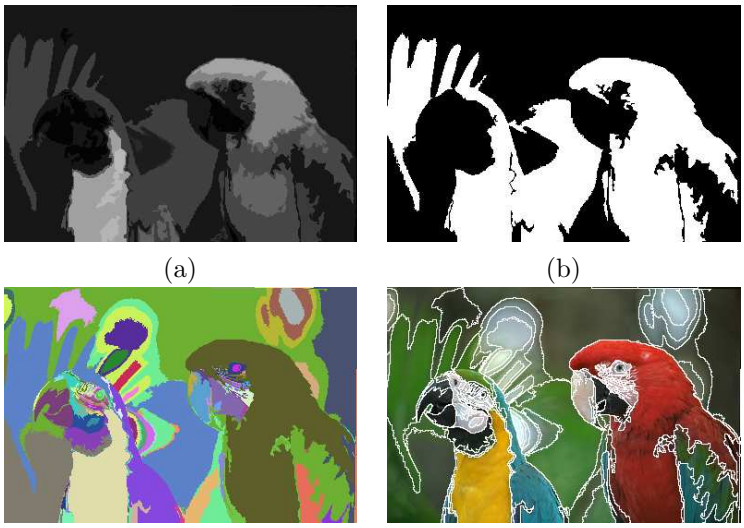
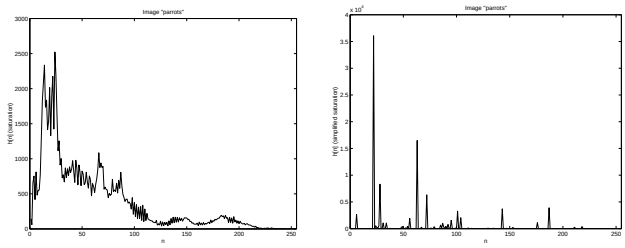


FIG. 15 – (a) Image mosaïque de "parrots" pour la saturation produite par la même séquence d'opérations que pour la figure 14, (b) ensemble des pixels de l'image (a) de valeur ≥ 45, (c) partition composite finale et (d) contours superposés à l'image initiale de cette partition.

Le procédé le plus simple pour déterminer l'ensemble X_s consiste, bien sûr, à seuiller la saturation à un certain niveau h. Mais ce type de sélection ne permet pas de contrôler la forme de l'ensemble $X_s(h)$, qui risque d'être irrégulier, avec des trous, des petites particules, etc... Il est préférable de travailler sur une image lissée, ou, puisqu'il s'agit de partitions, sur une image mosaïque de la saturation. Nous prendrons le même algorithme par sauts que pour la luminance et la teinte, en affectant ensuite à chaque classe la valeur moyenne de la saturation de ses pixels, ce qui donne la figure 15(a), plus régulière, donc plus robuste que l'image initiale de saturation de la figure 11(d). Cela apparait aussi sur les histogrammes correspondants, où le minimum d'abcisse 45 est plus affirmé. Le seuillage de l'image mosaïque pour $h = 45$ engendre l'ensemble X_s de la figure 15(b). Un algorithme de seuillage automatique optimal à partir de l'histogramme peut permettre d'obtenir la valeur optimale de séparation achromatique/chromatique pour chaque image [3], p.314. Cet algorithme recherche les deux modes les plus significatifs, et optimise un critère de variance par rapport à eux.

Il ne reste plus qu'à appliquer la rel.(27) aux deux partitions A_l et A_t de la figure 14. La partition composite correspondante est présentée en figure 15(c), avec les contours super-



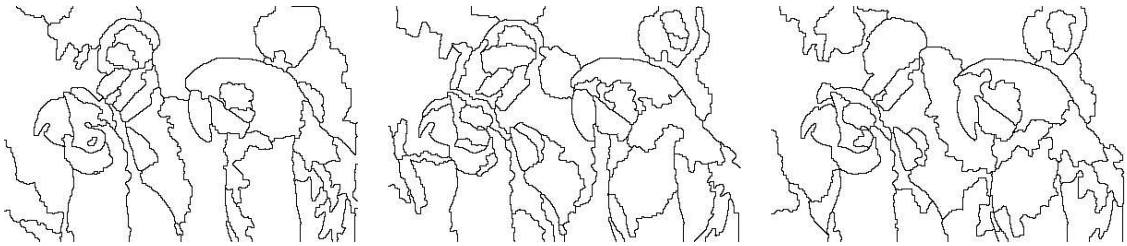
(a) (b)

FIG. 16 – Histogrammes de la saturation initiale (a) et de son image mosaïque (b). Le minimum aux environ de 45 est plus affirmé dans (b).

posés à l'image initiale en figure 15(d). Comme on pouvait s'y attendre, cette approche plus multidimensionelle délimite de meilleurs contours que la méthode scalaire des chutes d'eau (figure 13, niveau 4).

9.2 Comparaison des segmentations selon la norme utilisée

Nous avons illustré les deux méthodes de segmentation par gradient et par combinaison de partitions en prenant par triplet TLS celui de la norme L_1. Si, toute chose égale par ailleurs, on remplace ce triplet par le système HLS, ou par la norme $max - min$, les résultats deviennent-ils meilleurs, ou pires ? Considérons les contours de la partition pour l'image "parrots", en figure 17. On constate qu'en règle générale, la norme L_1 donne les meilleures segmentations (il faut rappeler de nouveau que la définition de la saturation en norme L_1 est plus fine que celle plus simple de la norme $max - min$).

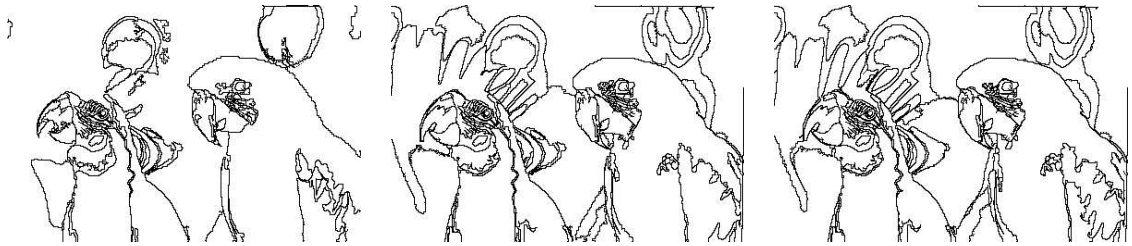


(a) (b) (c)

FIG. 17 – Comparaison des segmentations par "chutes d'eau" pour l'image couleur "parrots" selon la representation TLS utilisée : (a) HLS classique, (b) Y_{709} pour la brillance et norme $max - min$ pour la saturation, et (c) norme L_1 pour les trois composantes. Les contours correspondent au niveau 4 de la pyramide non-paramétrique.

Si maintenant nous comparons les trois représentations TLS en utilisant l'algorithme de

fusion ordonnée de partitions, la même conclusion s'affirme davantage : la saturation en norme L_1 fournit des segmentations bien plus précises. D'une part la norme L_1 récupère le fond (la palme de gauche par exemple, que la représentation HLS occulte complètement), et d'autre part les contours sont moins tortueux, [cf fig. 18]. L'amélioration sensible mais relativement légère par rapport à la norme $max - min$, est considérable par rapport au système HLS. Ce résultat ne fait que confirmer les multiples exemples de [3], pages 131-133.



(a) (b) (c)

FIG. 18 – Comparaison des segmentations par fusion ordonnée de partitions pour l'image couleur "parrots" la representation TLS utilisée : (a) HLS classique, (b) L de la télévision et norme $max - min$ pour S, et (c) norme L_1 pour les trois composantes. Les mêmes paramètres ont été utilisés pour la connexion par sauts et la mosaïque de la saturation a été seuillée de manière optimale pour chaque cas.

9.3 Berkeley Segmentation Dataset and Benchmark (BSDB)

Pour compléter la partie résultats de cet article, nous avons appliqué les deux algorithmes de segmentation couleur présentés sur une base d'images couleur de référence, la "Berkeley Segmentation Dataset and Benchmark" (BSDB) [11], disponible en ligne. Il s'agit d'une base de données avec 12,000 images naturelles à niveaux de gris et en couleur (300 disponibles publiquement pour des tests) qui ont été segmentées manuellement. Le fait d'évaluer nos résultats sur la BSDB nous permet de comparer la performance de l'approche développée avec une segmentation humaine de référence et par ailleurs, une comparaison parmi les différentes techniques de segmentation proposées par la communauté scientifique.

Les résultats sont montrés dans les figures 19, 20 et 21 (le numéro correspond à celui assigné dans la BSDB). En ce qui concerne la représentation couleur, nous avons utilisé la norme L_1 pour toutes les composantes. Pour les deux méthodes, toutes les images ont été segmentées avec les mêmes paramètres : 1) itération 4 de la "chute d'eau" et 2) connexion par saut avec $k = 20$ et une surface > 50 (évidemment, la saturation a été seuillée de manière optimale pour chaque cas). Comme nous pouvons constater, pour la plupart des images montrées, les résultats obtenus avec les deux méthodes morphologiques sont satisfaisants par rapport à la référence manuelle : les régions chromatiques/achromatiques les plus importantes sont bien détectées (les mosaïques associées nous donnent une représentation couleur minimale acceptable de l'image). En général, l'itération de la LPE sur le gradient couleur donne des

régions assez larges mais qui peuvent découper des régions homogènes ou bien perdre des petits détails. Tandis que la fusion de connexions par sauts fournit des résultats plus équilibrés entre les grandes régions homogènes et les petits détails significatifs. Comme il était prévisible, les performances les plus médiocres correspondent toujours aux images les plus texturées (il faut rappeler que les images sont traitées sans aucun pré-filtrage, qui pourrait réduire notablement les textures).

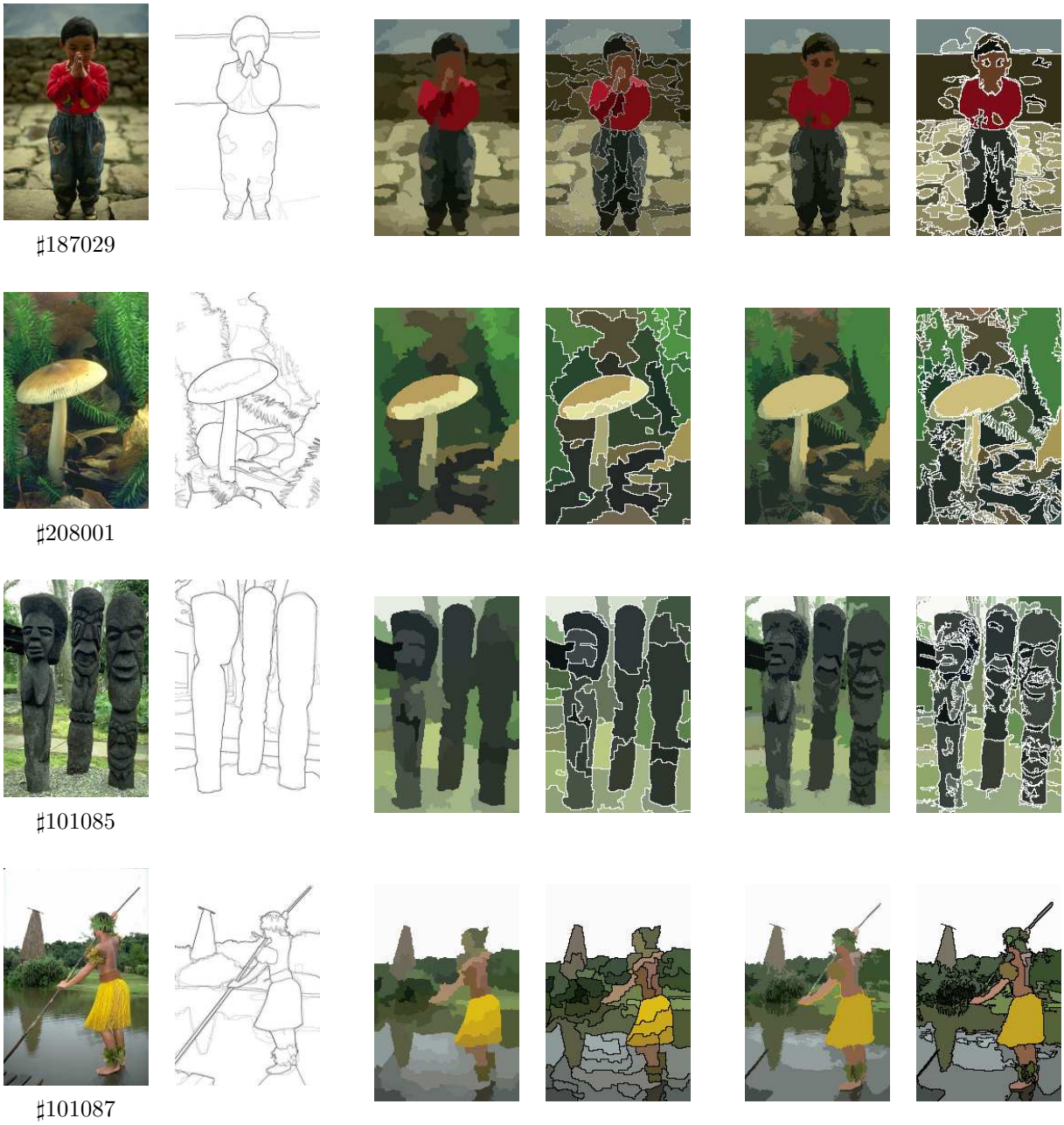


Fig. 19 – Résultats de segmentation avec les deux méthodes présentées sur des images de la BSDB. De gauche à droite : images originales, segmentations humaines de référence (BSDB), segmentations par “chutes d’eau” (images mosaïque et contours associés), segmentations par fusion ordonnée de partitions, (images mosaïque et contours associés).

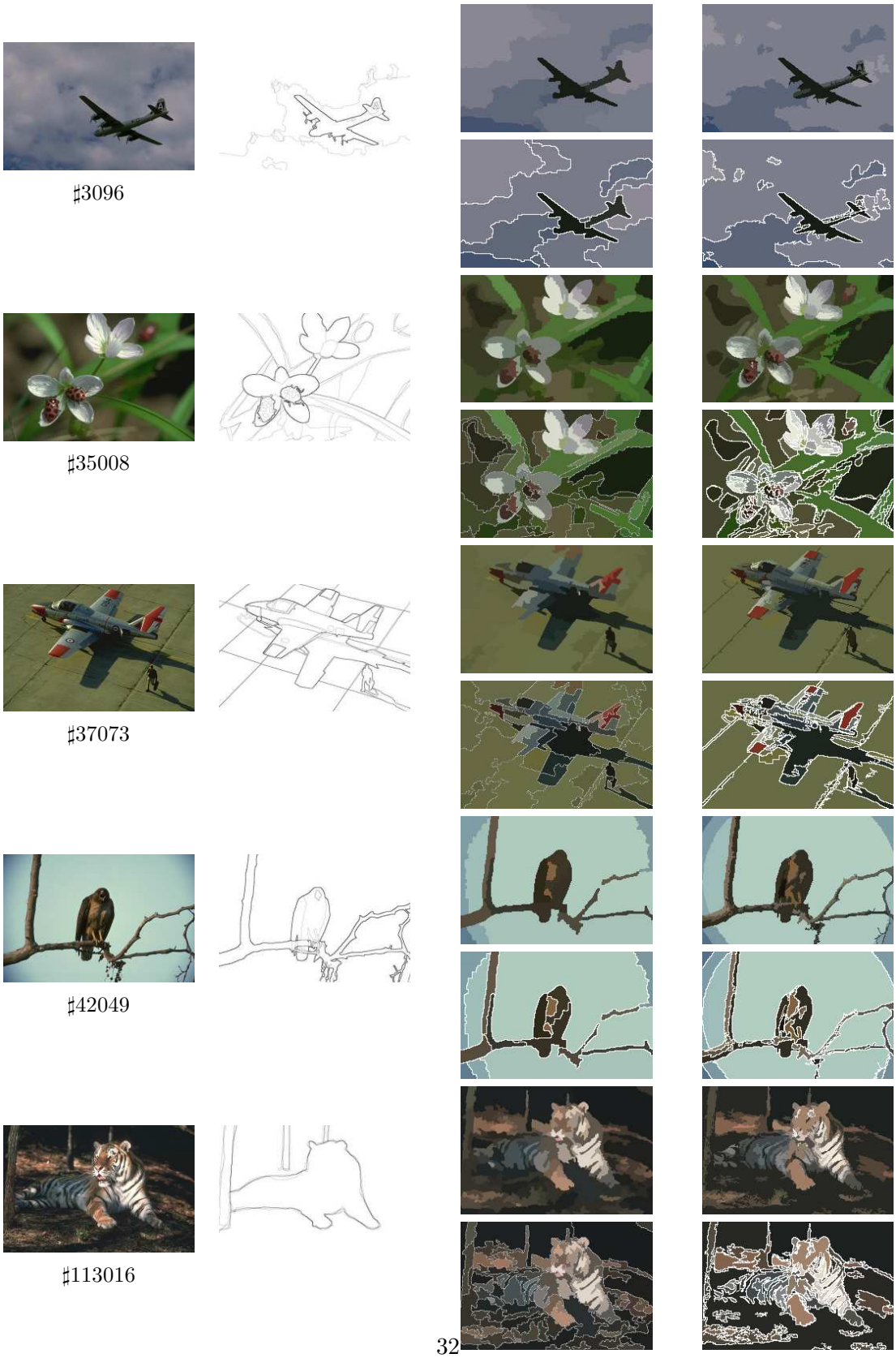


Fig. 20 – Résultats de segmentation avec les deux méthodes présentées sur des images de la BSDB. Idem. figure 19.

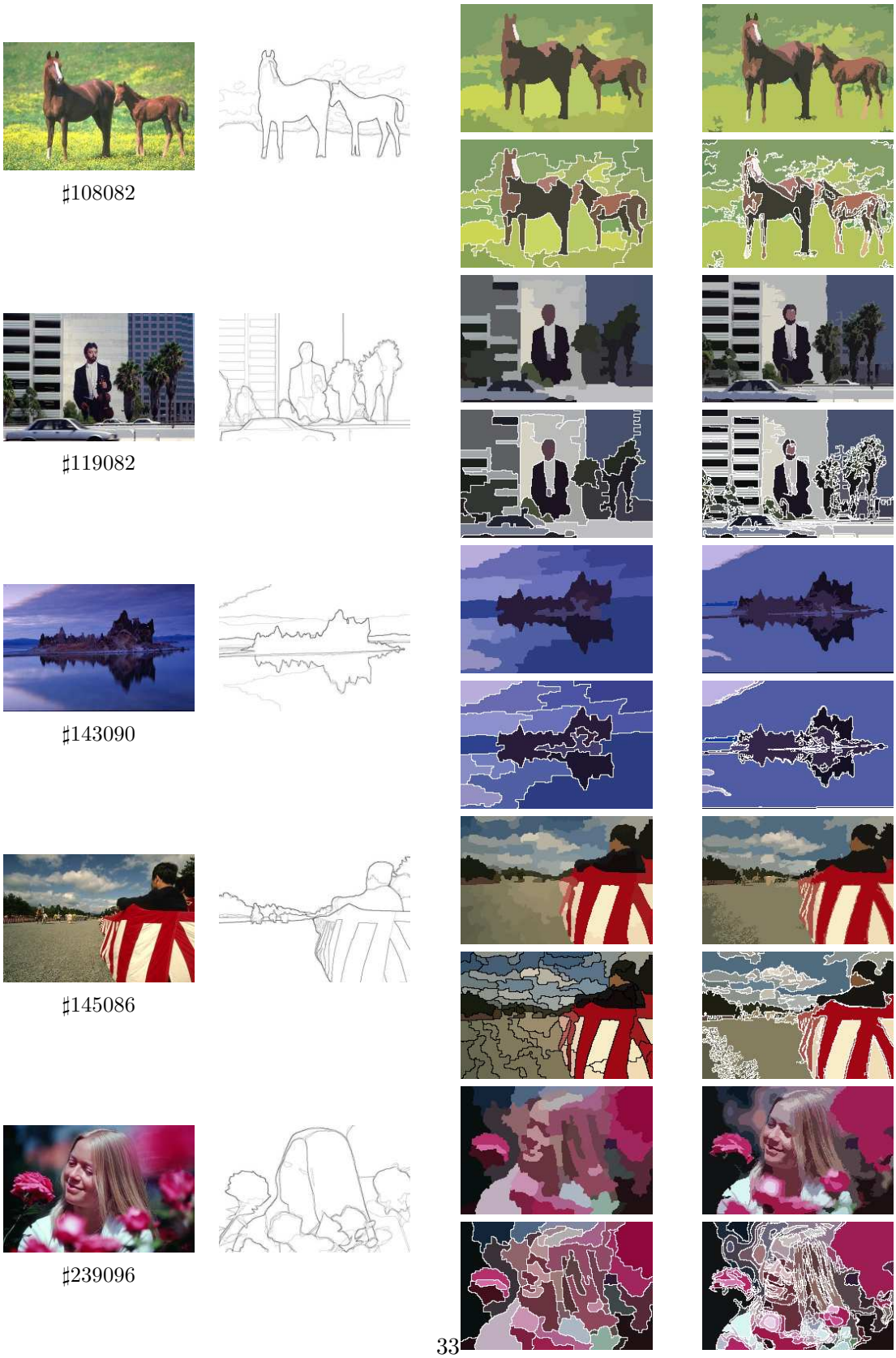


Fig. 21 – Résultats de segmentation avec les deux méthodes présentées sur des images de la BSDB. Idem. figure 19.

10 Conclusions et perspectives

Le but premier de cet article a été de proposer un cadre pour les représentations couleur du type "luminance, saturation, teinte" qui soit adapté au traitement quantitatif des images. Dans ce triplet vectoriel, la saturation joue un rôle fondamental de pondération entre la luminance et la teinte. Par ailleurs, bien que cette variable ne soit pas une grandeur physique mesurée, au sens de la colorimétrie, l'interprétation visuelle très intuitive de son rôle ainsi que son importance dans la mise en œuvre des opérateurs de traitement de l'image nous laisse penser que des études plus approfondies sur la perception et la quantification de la saturation vue comme norme devraient être menées.

En plus de donner des définitions mathématiques des différentes représentations valides pour la luminance, la saturation et la teinte, nous avons largement illustré les avantages de la représentation en norme L_1 ; aussi bien lorsqu'on utilise les histogrammes bi-dimensionnels *luminance/saturation* et *luminance/teinte* pour l'étude des phénomènes du type reflets, ombres, dégradés d'illumination, etc... selon un modèle linéaire d'émission-reflectance, que lorsqu'on veut segmenter des images au moyen d'une saturation pondératrice.

Références

[1] J. Angulo, "Colour segmentation using bivariate histograms in 3D-polar colour spaces", *Rapport Technique CMM-Ecole des Mines de Paris, N-03/03/MM*, Janvier 2003.

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

[3] J. Angulo, *Morphologie mathématique et indexation d'images couleur. Application à la microscopie en biomédecine.* Thèse doctorale, Centre de Morphologie Mathématique, Ecole des Mines, Paris, Décembre 2003.

[4] Commission Internationale de l'Eclairage (CIE), Colorimetry, Second Edition. *CIE Publication No. 15.2,* Vienna, 1986.

[5] G. Choquet, "Topologie", Academic Press, N.Y., 1966.

[6] L. Garrido, P. Salembier, D. Garcia, "Extensive operators in partition lattices for image sequence analysis", *Signal Processing,* Vol. 66, No. 2, 1998, p. 157–180.

[7] A. Hanbury, J. Serra, "Morphological operators on the unit circle", *IEEE Transactions on Image Processing,* Vol. 10, No. 12, 2001, p. 1842–1850.

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

[9] P. Lambert, T. Carron, "Symbolic fusion of luminance-hue-chroma features for region segmentation", *Pattern Recognition,* Vol. 32, 1999, p. 1857–1872.

[10] L. Macaire *Exploitation de la couleur pour la segmentation et l'analyse d'image* Thèse doctorale, Univ. de Lille Nov. 2004, 274 p.

[11] D. Martin, C. Fowlkes, D. Tal, J. Malik,"A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics", Dans *Proc. 8th Int. Conf. Computer Vision (ICCV'01)*, Vol. 2, p. 416–423, Vancouver, Canada, 2001.
http ://www.cs.berkeley.edu/projects/vision/grouping/segbench/

[12] F. Meyer, S. Beucher, "Morphological Segmentation", *J.of Visual Communication and Image Representation*, Vol. 1, No. 1, 1990, p. 21–46.

[13] G. Obein, K. Knoblauch, F. Viénot, "Perceptual scaling of the gloss of a one-dimensional series of painted black samples", *Perception*, Vol. 31, Suppl., 2002, p. 65.

[14] Ch. Poynton *A technical Introduction to Digital Video.* New York : Wiley, 1996. Chapter 6, "Gamma" is available online at
http ://www.poyton.com/PDFs/TIDV/Gamma.pdf

[15] J. Serra, "Set connections and discrete filtering", Dans (G. Bertrand, M. Couprie and L. Perroton Eds.), *Proc. DGCI 1999 - Lecture Notes in Computer Science,* Springer, Vol. 1568, 1999, p. 191–206.

[16] J. Serra, "Espaces couleur et traitement d'images", *Rapport Technique CMM-Ecole des Mines de Paris, N-34/02/MM*, Octobre 2002.

[17] J. Serra, "Morphological Segmentation of Colour Images by Merging of Partitions," Proc. of the International Symposium on Mathematical Morphology (ISMM'05), Kluwer, 2005.

[18] S.A. Shafer, "Using color to separate reflection components from a color image", *Color Research and Applications,* Vol. 10, No. 4, 1985, p. 210–218.

[19] 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.

[20] A.Tremeau, Ch. Fernandez-Maloigne, P. Bonton, *Image numérique couleur*, Ed. Dunod, Paris, 2004.

Brève biographie des auteurs

Jesús Angulo

Dr Jesús Angulo est né à Cuenca, Espagne, en 1975. En 1999, il a obtenu son diplôme d'Ingénieur en Télécommunications de l'Université Polytechnique de Valence (Espagne), dont le projet de fin d'études portait sur le Traitement des Images et de la Vidéo. En 2003, il a passé sa thèse de Doctorat en Morphologie Mathématique et Traitement d'Image à l'Ecole des Mines de Paris (France), encadré par le Professeur Jean Serra. Il est actuellement chercheur (post-doc) au Centre de Morphologie Mathématique de l'Ecole des Mines de Paris. Ses thèmes de recherche sont le traitement des images couleur et la morphologie mathématique, ainsi que leurs applications à la biomédecine et aux biotechnologies.

Jean Serra

Jean Serra est né en 1940, de nationalité française. En coopération avec G. Matheron, il a créé la morphologie mathématique dans les années 60, dans le but de décrire les formes et les textures de phénomènes naturels, à diverses échelles. Il est directeur de recherches à l'Ecole des Mines de Paris et Docteur honoris causa de l'Université Autonome de Barcelone. Son ouvrage majeur est un traité en deux volumes intitulé "Image Analysis and Mathematical Morphology" (1982, 1988). Ces cinq dernières années, ses principales contributions portent sur les connexions, la segmentation et la couleur. Il a créé en 1993 la société internationale de morphologie mathématique.