Mophological Operators for the Segmentation of Colour Images

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

August 2004

Abstract

A panel of several recent studies on colour processing at CMM is presented, and referred to G. Matheron activities. A first section introduces new representations of the triplet "luminance, saturation, hue" that are consistent with a few basic quantization prerequisites. One of these representations, which derives from the L_1 norm, produces alignments in luminance/saturation histograms. A physical model of light propagation provides an interpretation of such alignments. The next section starts from the particular role of the saturation, which balances the opposition between luminance and hue, and uses it to combine the segmentations of the three components into synthetic colour segmentations. A new section deals with mathematical morphology applied to image data distributed on the unit circle, such as the hue. It is followed by another section where the question of colour images interpolation is treated by means of Hausdorff geodesics. The last section proposes a morphological approach to multivariate analysis of heterogeneous data, and applies it to the segmentation of colour images coded by palettes.

Key words : *colour, segmentation, saturation, norms, light propagation, unit circle, interpolation, multivariate analysis*

1 Introduction

During his long and fruitful career, Georges Matheron never tackled the two themes of segmentation and colour processing of images. At the beginning of the nineties, when I was entering the first theme via connected filters, I asked him why he had avoided them. He told me that he could not take up working on all possible problems, and after a silence, added with his marvellous tactful smile: "besides, I guess it is already left to you".

Indeed, G. Matheron was partly right. The work on connected filters contained the seed of a morphological theory of segmentation that arose ten years later [31]. But by this time, my knowledge about colour processing was just a matter of two ideas. First, one represents colour by three grey images, in RGB or HLS modes, second, the compressed bit streams of grey tone images have to be multiplied by two (and not by three) when passing to colour.

I entered the field of colour processing quite accidentally. In 1999, Allan Hanbury, who was one of my Phd students, was developing a methodology for an automatic classification of pieces of woods according to their textures (veins, knots, ...). The directional features were, of course, preponderant and led us to elaborate morphological tools for circular data [11]. Now, in imagery, the two major situations where such data appear, are orientations and hues. Therefore, we decided to try also our algorithms on colour representations involving hue, i.e. on the two HSV and HLS systems. It rapidly led us to a critical analysis of both systems, and to propose more consistent ones [30][15]. In the meantime, A. Hanbury defended his thesis and left the Centre de Morphologie Mathématique (CMM, in brief) in 2002. I carried on the approach with Jesus Angulo, another Phd student at CMM, whose thesis subject did not really involve colour processing, but who was enthusiastic about the question. Together, we started from the new colour systems and discovered that, in the 3-D histograms generated by the L_1 representations, the pixels of an arbitrary image used to present specific alignments. What did it mean ?

The new representations had particularly modified the definition of the saturation parameter, allowing us to perceive his meaning more clearly. With J. Angulo, we decided to exploit this parameter to balance luminance versus hue, when segmenting colour images [2].

This thread of ideas is not the only one we developed during these recent years. Since I was thrown into colour problems, it was an excellent opportunity to have a look at RGB representations and to those based upon palettes. I developed the first theme jointly with Marcin Iwanowski [12], and the second with Mariuscz Mlynarczuk [29], both polish Phd students at CMM.

The text which follows surveys these studies on colour image processing. I shall try and show how my way of thinking was pervaded by G. Matheron's one; and since Matheron used to say that the pedagogical order must be the reverse of that of the discovery, we will begin by the most recent developments.

2 The 3-D polar representations of the colour

2.1 Light intensities and gamma correction

Consider a television receiver. It uses three different colour representations. On the one side, the input Hertzian signal is coded as one grey image plus two other ones, associated to green-red and blue-yellow contrasts (i.e. one luminance and two chrominances). On the other side, the image on the monitor is obtained from three electrical signals, which excite three layers of green, red and blue photo-receivers. These two representations are quite different, although technically sound for their respective purposes. However, the manufacturers take none of them for the user's interface, and prefer human adjustments based on light (luminance), contrast (saturation), and, in case of an old receiver, from hue. Hence, this last triplet turns out to be the simplest one for human vision.

What are the relationships between these various representations? Do the technological steps modify the initial light that enters a device? Colour image processing rests on a few basic operations (addition, comparison,..) and properties (increasingness, distances..). Have

these tools a physical meaning? In colour imagery the basic notion is the spectral power distribution (SPD) of the light radiating from or incident on a surface. This intensity has the dimension of an energy per unit area, such as watt per m^2. When the light arrives at a photo-receiver, this sensor filters the intensities of each frequency by weighting them according to fixed values. The sum of the resulting intensities generates a signal that exhibits a certain "colour". The CIE (Commission Internationale de l'Eclairage), in its *Rec 709,* has standardized the weights which yield the triplet $R_{709}, G_{709,} B_{709}$ [8]. As energies, the intensities are additive, so that all colours accessible from an RGB basis are obtain by sums of the primary colours R, G, and B and by multiplications by non negative constants.

The exploration of the spectrum is lower bounded by $R = G = B = 0$ (zero energy) higher bounded by a maximum red R_0, green G_0 and blue B_0 that are given by the context (illumination, technological limits of the sensors, or of the eye, etc.) in which we work. Generally, each technology fixes the three bounds, which therefore define the reference white, and then introduces the *reduced variables*

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

The digital sensitive layers of cameras transform the light intensities into voltages; conversely, the cathodic tubes (CRT) and the flat screens that display images return photons from the electrical current. Now, their response is not linear, but a power function of the input voltage whose exponent γ, (gamma), varies around 2.5 according to the technologies. If we want the light intensities of the CRT to be proportional to those of the scene itself, the gamma effect has to be compensated. In video systems, this *gamma correction* is universally at the camera. The *Rec. 709* of CIE proposes the following correction function

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

that we write here for the reduced red intensity r, and where$1/\gamma = 0.45$. The same transfer function is applied to both green and blue bands.

Fig. 1, drawn from [24] depicts the graph of Rel.(2). The variation domain $[0, 1]$ is the same for the reduced intensities (r) as for the video colours (r'), which implies that the white point $R_0\ G_0\ B_0$ is left invariant. The linear beginning in Rel.(2) minimizes the effect of the sensor noise. An ideal monitor should invert the transform Rel.(2). Indeed, they generally have neither linear segment, nor gamma exponent equal to$1/0{,}45$ [24].

Figure (1) shows that for r closed to1, the graph looks like a straight line. More precisely, the limited expansion

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

for small u, leads us to replace the second equation (2) by

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

i.e., numerically

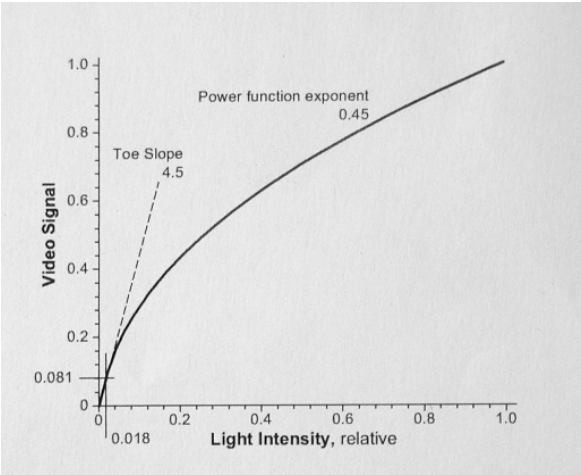


Figure 1: *Gamma correction function.*

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%

In comparison with the noise of the video systems, we can consider the approximation r'^* is perfect for $r \geq 0.8$ and excellent for $0.6 \leq r \leq 0.8$.

2.2 Colour Vector Spaces

Their linearity provide the intensities r, g, b with the structure of a 3 dimensions vector space, or rather of the part E which is limited to the unit cube $[0,1] \times [0,1] \times [0,1]$ of $\mathbb{R}^3$. For colour image processing purposes, it would be wise to go back from the video bands (r', g', b') to the reduced intensities (r, g, b) by the inverse transform of Rel.(2). When starting from the usual 3×8 bits (r', g', b') images, the best should probably be to code in 3×16 bits for computation (or in floating variables). But as a matter of fact, people keeps the (r', g', b') video space, which is implicitly modelled as a part of a vector space, from which one builds arithmetic means, projections, histograms, Fourier transforms, etc., which often gives significant results.

What are the real consequences of the gamma correction Rel.(2) on the processing of colour data? Formally speaking, one can always consider the unit video cube (r', g', b') as a part, E' say, of a 3-dimensions vector space. This allows us to formulate operations, but their physical interpretations demand we come back to the intensities (r, g, b).

Fig. 2 depicts the unit cube E'. The vector $\vec{x}'$, of coordinates (r', g', b') can also be decomposed into two orthogonal vectors $\vec{c}'$ and $\vec{l}'$ of the chromatic plane and the a-chromatic (or gray) axis respectively. The latter is the main diagonal of the cube going through the origin O and the chromatic plane is perpendicular to the gray axis in O. The two vectors $\vec{c}'$

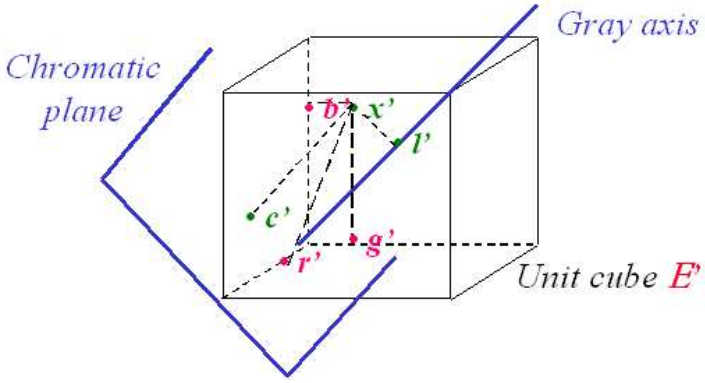


Figure 2: *Chromatic plane and a-chromatic axis.*

and $\vec{l'}$ have the following coordinates

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

Consider the red band $r'(z)$ over a zone Z in a colour image. What meaning can we give to the average red in Z ? As we just saw, the only average that has a physical meaning is the quantity $\bar{r} = \frac{1}{Z}\int (r'(z))^{\gamma}\, dz$, which needs to be corrected into $\bar{r}^{1/\gamma}$ for display purposes (for the moment we neglect the constants 1,099 and 0,099 in Rel.(2)). On the other hand, the usual segmentations aim to split the space into regions Z where the colour is nearly constant. Then at each point $z \in Z$, we can approximate $r(z)$ by the limited expansion

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

where $\bar{r}' = \frac{1}{Z}\int_Z r'(z)dz$. Under averaging in Z, the coefficient of the γ term in the right member becomes zero, so that

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

Therefore, the arithmetic mean of the video red r' equals, at the second order, the mean of the red intensity r followed by the gamma correction. The result remains true when the coefficients of Rel.(2) are added, when the the average is weighted, and also for the dark zones Z where the first Rel.(2) applies. It extends to the greens and blues. Rel.(6) turns out to be a theoretical justification of the "mosaic" based image segmentations (e.g. waterfall algorithm).

2.3 Brightness

From the point of view of physics, brightness is nothing but the integral of the power spectrum, i.e., here, the sum of the three components r, g, and b, that stand for this spectrum. For colorimetric purposes, this sum has to be weighted relatively to the spectral sensitivity of the

eye. The CIE *Rec. 709* defines a white point and three weighting functions of the spectrum which lead to the variables R_{709}, G_{709} and B_{709}, then to the *luminance*

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

and to the luminance Y_W of the associated white point. The three coefficients of Rel.(7) are related to the brightness sensitivity of the human vision and have been estimated by colorimetric measurements on a comprehensive population. The luminance Y_{709}, as a linear function of intensities, is an energy $(watts/m^2)$.

Human vision responds to intensities in a logarithmic way, according to laws of the type $di/i = constant$. Just as we took into account the spectral sensitivity of the eye, we should not ignore its energetic sensitivity. Now, by an amazing coincidence vision response to intensity is closed to the gamma correction of Rel.(2) : for example, when the luminance of a source is reduced to 20%, the eye perceives an intensity reduction of 50% only. Therefore, following many authors, we can consider the transforms

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

for $\gamma \simeq 2.2$ as generating *perceptual intensities* . For example, the *Rec. BT 601-E* proposes the *luma* y'_{601} as a perceptual brightness measurement

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

However, this luma, as established from video values has not an energy dimension, and not any more the deriving additivity properties. The CIE follows the same direction, but defines the *lightness* l^* by taking a slightly different exponent

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

As regards the operations of segmentation in image processing, the situation is different. They do not hold on a *perceived* brightness, but on that of the *object under study.* In microscopy, the histological stainings usually range from blue to violet; the spectrum of a sunset, or that of a human face have nothing to do with the weights given to r, g, and b in Rel.(7) or (9). Thus in the absence of *a priori* information on the spectra of the objects under study, the purpose of segmentation leads us to take as brightness a *symmetrical function* of primary colours.

As regards the perceived energies now, consider, in the intensity space E, a vector x whose direction is given by $x_o = r_o$, g_o, b_o but whose intensity varies, i.e.

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

The point x describes the segment S which begins in O, goes through (r_o, g_o, b_o) and ends on the edge of cube E. In the video space E' there corresponds to x the point x' :

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

with $x'_0 = r_0^{1/\gamma}$, $g_0^{1/\gamma}$, $b_0^{1/\gamma}$. Similarly, the point x' describes a segment S' in E'. When x varies, if we want its perceptual brightness to seem additive, then Rel.(10) implies that the corresponding brightness of x' is a linear function of the three primary components. Finally, since this "image processing brightness" has to vary from 0 to 1, as r and r' do, the only possibility is to take for it the arithmetic mean m' of the primary colours :

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

Put $\lambda' = \lambda^{1/\gamma}$. The two expressions

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

turn out to be different distances in segments S and S' respectively. The exponent $1/\gamma$ provides the second one with a meaning of *perceptual homogeneity*. But image processing is more demanding, as we must be able to express that a colour point E' (or more generally a set of points) gets closer to another even when these two points are not aligned with the origin. Now, the mean (11) is nothing but the restriction to the cube E' of the L_1 norm, which is defined in the whole space $\mathbb{R}^3$ (i.e. for $r', g', h' \in [-\infty, +\infty]$) by taking $\alpha = 1$ in the relation

$$n\left(x'\right) = \left(\left|r'(x)\right|^\alpha + \left|g'(x)\right|^\alpha + \left|b'(x)\right|^\alpha\right)^{1/\alpha} \qquad \alpha \geq 1 \tag{12}$$

(This Rel.(12) introduces indeed a family of norms as soon as $\alpha \geq 1$. For $\alpha = 2$, we obtain the Euclidean norm L_2, and for $\alpha = \infty$, the "max" norm). In a vector space V, any norm n generates a distance d_n (see [9], section VII-1-4) by the relation

$$d_n\left(x'_1, x'_2\right) = n\left(x'_1 - x'_2\right) \qquad x'_1, x'_2 \in V \tag{13}$$

Therefore L_1 is a distance, as well, of course, as its restriction to the unit cube E'.

Thus, for $\alpha = 1$, both brightness $m'(x')$ and distance $d\left(x'_1, x'_2\right) = m'\left(|x'_1 - x'_2|\right)$ in E' derive from a unique concept. This latter relation is important, as in segmentation a number of algorithms which were established for numerical functions extend to vector functions when a distance is provided (e.g. watershed).

2.4 Saturation

The CIE was more interested in various formulations of the brightness (luminance, lightness ...) than in saturation, that it defines as "*the colourfulness of an area judged in proportion to its brightness*". In other words, it is the concern of the part of uniform spectrum (i.e. of gray) in a colour spectrum, so that any maximal monochromatic colour has a unit saturation and so that any triplet $r = g = b$ has a zero saturation.

Intuitively, what the CIE means here is clear, but its definition of the saturation lends itself to various interpretations. From a given point $x \in E$, one can draw several paths along which the colour density varies in proportion to brightness. For example, in Fig. 2,

supposed to represent cube E, we can take the perpendicular xc to the chromatic plane, or the perpendicular xl to the gray axis, or again the axis Ox, etc.. Which path to choose?

Indeed, these ambiguities vanish as soon as we set the context in the chromatic plane. The cube E is projected according to Hexagon H centered in O. Consider a point $x_o \in E$, of projection c_0 in H, and such that $c_0 \neq O$. Following the CIE, we define as a saturation any non negative function along the axis Oc_0 that *increases* from O; in O, it equals zero (pure gray) and has its maximum value when the edge of Hexagon H is reached, in $c_{\max}$ say (saturated colour). The hue remains constant along the segment $[0, c_{\max}]$, and the hue of the opposite segment $[0, \bar{c}_{\max}]$ is said to be *complementary* of that of segment $[0, c_{\max}]$. For a point $c \in [0, c_{\max}]$, we have $c = \lambda c_0$, $0 \leq \lambda \leq 1$. Thus, given $c_0 \in H$, the saturation $s(c) = s(\lambda c_0)$ is a function of λ only, and this function is increasing.

We have to go back to the 3-D cube E, as point c_0, projection of x_0, is just an intermediary step (moreover $c_0 \notin E$).The saturation $s(x_0)$ of point $x_0 \in E$ is then defined by

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

Note that when a point $x \in E$ moves away from the chromatic plane along the perpendicular $c_0 x_0$ to this plane, its gray proportion increases, but its saturation $s(x)$ does not change: it is indeed a matter of *chromatism* and not of *energy* of the light intensity.

As point c describes the radius $[0, \bar{c}_{\max}]$ which is at the opposite of $[0, c_{\max}]$ in the chromatic plane, we have

$$c \in [0, \bar{c}_{\max}] \qquad \Longleftrightarrow \qquad c = \lambda c_0 \qquad \lambda(\bar{c}_{\max}) \leq \lambda \leq 0$$

where λ indicates the proportionality ratio, now negative, between c and $\bar{c}_0$. This purely vector equivalence admits a physical interpretation if we extend the definition of the saturation to all diameters $D(c_0) = [0, c_{\max}] \cup [0, \bar{c}_{\max}]$, $c_0 \in H$, of the hexagon H (saturation was previously introduced for radii only). This can be done by putting $c \in D(c_0)$, $s(c) = s(\lambda c_0) = s(|\lambda| c_0)$. Two opposite points have the same saturation, and more generally if $c_1 \in [0, c_{\max}]$ and $c_2 \in [0, \bar{c}_{\max}]$, then $c_1 + c_2 = (\lambda_1 + \lambda_2) c_0$, with $\lambda_1 \geq 0$ and $\lambda_2 \leq 0$. As s is increasing we have

$$c_1 \in [0, c_{\max}], \qquad c_2 \in [0, \bar{c}_{\max}] \qquad \Longrightarrow \qquad s(c_1 + c_2) \leq s(c_1) + s(c_2). \tag{14}$$

When $c_1 = c_{\max}$ and $c_2 = \bar{c}_{\max}$ we find in particular Newton's disc experiment, reduced to two complementary colours.

When considering the saturation in the video cube E', the conditions of increasingness of s' along the radii (now of H') and of its nullity on the gray axis are still valid. They must be completed by the two requirements of image processing, namely the symmetry w.r.t. r', g', b' and the fact that $s'(x'_1 - x'_2)$ must be a distance in E'.

We saw that the mean m', in Rel.(11), was the L_1 norm expressed in the unit cube E', and that $3m'(x')$ was both the norm of x' and of its projection l' on the gray axis, i.e.

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

It is tempting to keep the same norm for the hexagon H' of the chromatic plane. By using Rel.(5) we find

$$s'\left(x'\right) = L_1\left(c'\right) = \frac{1}{3}\left[\left|2r' - g' - b'\right| + \left|2g' - b' - r'\right| + \left|2b' - r' - g'\right|\right]. \tag{15}$$

By symmetry, $s'\left(x'\right)$ depends on the three functions $max' = max(r', g', b,')$, $min' = min(r', g', b,')$, and $med' = mediane(r', g', b,')$ only, which gives

$$s' = \begin{cases} \frac{3}{2}\left(max' - m'\right) & \text{if} \quad m' \geq med' \\ \\ \frac{3}{2}\left(m' - min'\right) & \text{if} \quad m' \leq med' \end{cases} \tag{16}$$

On can find in [30] the derivation yielding s', and that of the following expression h' of the hue (which avoids to bring trigonometric terms into play) ,

$$h' = \frac{\pi}{3}\left[\lambda + \frac{1}{2} - (-1)^{\lambda}\frac{max' + min' - 2med'}{2s'}\right] \tag{17}$$

with λ equals

$$\begin{array}{lll} 0 \text{ if } r > g \geq b & 1 \text{ if } g \geq r > b & 2 \text{ if } g > b \geq r \\ 3 \text{ if } b \geq g > r & 4 \text{ if } b > r \geq g & 5 \text{ if } r \geq b > g. \end{array}$$

The hue h', as a coordinate on the unit circle, is defined modulo 2π. The value $h' = 0$ in Eq.(17) corresponds to the red. For $s' = 0$, colour point lies on the gray axis, so that its hue is meaningless. The inverse formulae are given in [15], and the detailed proofs may be found in [30].

The relations (15) and (13) entail that $s\left(c'_1 - c'_2\right) = L_1\left(c'_1 - c'_2\right)$ is a distance in the chromatic plane, which therefore brings into play both saturation and hue. On the other hand, as L_1 is a norm, Rel.(14) becomes true for all triplets c'_1, c'_2 and $c'_1 + c'_2$ that are on a same diameter of H'. Remark that here the L_1 norm is the concern of the *projections* c', the norm of the vectors x' themselves being their arithmetic mean. Finally, the above comments apply also to the Euclidean norm and to the max-min, which both induce distances in the chromatic hexagon H'.

When passing from the video variables to the intensities, a first result is obtained by observing that the averaging of the saturation s' follows the same law than that of the brightness m', namely Rel.(6), in the zones Z where the colour varies weakly. Moreover, the mapping $x'_0 = (r'_0, g'_0, b'_0) \rightarrow x_0 = \left(r_0'^{\gamma}, g_0'^{\gamma}, b_0'^{\gamma}\right)$ shows that $x' = \lambda x'_0$ becomes $x = \lambda^{\gamma} x_0$, hence

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

In other words, the L_1 norm is increasing on the radii, and is zero for the grays, on both chromatic hexagons H of the intensities and H' of the video variables. Thus it represents a saturation *in both spaces*. It seems difficult to go further, as two points $x'_0, x'_1 \in E'$ whose projections c'_0 and c'_1 lie on a same radius of H' may have homolog points x_0 and $x_1 \in E$ whose projections are not always aligned with O.

2.5 Two other norms

How to build polar representations which be not contradictory with the previous requirements? Besides the L_1 norm, we can think of two ones. Firstly, the Euclidean norm L_2. In practical image processing, it turns to be less convenient than the L_1 norm, which suits particularly well to linear and morphological operations, and provides nice inverses. In addition, the associated 2-D histograms are rather unclear (see Fig. 4).

Another possibility is to correct the classical HLS system [16], by replacing its saturation by $max(r,g,b) - min(r,g,b)$. In the whole space, the quantity $max - min$ is a semi-norm only: two distinct vectors c and c', whose difference c - c' is a gray have the same $max - min$ [15]. However, in the chromatic plane, $max - min$ becomes a norm. It can be used for the saturation in parallel with m' for the brightness. This is what we will do below each time $max - min$ norm is introduced.

Finally, the norm and distance based approach presents the significant advantage that it *separates the variables* : two points x'_1 and $x'_2 \in E'$ which have the same projection on the chromatic plane (resp. on the gray axis) have the same saturation (resp. the same brightness). However, the last property, on brightness, vanishes when the three bands are given different weights in the means m or m'.

2.6 The classical polar representations

Even though the transformation from RGB to hue, saturation and brightness coordinates is simply a transformation from a rectangular colour coordinate system (RGB) to a three-dimensional polar (cylindrical) coordinate system, one is faced with a bewildering array of such transformations described in the literature (HSI, HSB,HSV, HLS, etc.). Most of them date from the end of the seventies [33], and were conceived neither for processing purposes, nor for the current computing facilities. This results in a confusing choice between models which essentially all offer the same representation. The most popular one is the HLS triplet of System (18), which appears in many software packages. The comments which follow hold on this particular model, but they apply to the other ones. The HLS triplet derives from RGB by the following system

$$\left\{ \begin{array}{l} l'_{HLS} = \frac{\max(r',g',b')+\min(r',g',b')}{2} \\ \\ s'_{HLS} = \left\{ \begin{array}{ll} \frac{\max(r',g',b')-\min(r',g',b')}{\max(r',g',b')+\min(r',g',b')} & \text{if} \quad l'_{HLS} \leq 0.5 \\ \\ \frac{\max(r',g',b')-\min(r',g',b')}{2-\max(r',g',b')-\min(r',g',b')} & \text{if} \quad l'_{HLS} \geq 0.5 \end{array} \right. \end{array} \right. \tag{18}$$

One easily checks that the HLS expressions do not preserve the above requirements of linearity (for the brightness), of increasingness (for the saturation) and of variables separation. The HLS luminance both RGB triplets $(1/2, 1/2, 0)$ and $(0, 1/2, 1/2)$ equals $1/4$, whereas that of their mean equals $3/8$, i.e. it is lighter than both terms of the mean. The HLS saturations of the RGB triplets $(4/6, 1/6, 1/6)$ and $(2/6, 3/6, 3/6)$ equals $3/5$ and $1/5$ respectively, whereas that of their sum is 1: it is just Newton's experiment denial! Finally the independence

a) Rubber ring *b) Ana Blanco*

Figure 3: *Two test images.*

property is no more satisfied. Take the two RGB triplets $(1/2, 1/2, 0)$ and $(3/4, 3/4, 1/4)$. One passes from the first to the second by adding the gray $r' = g' = b' = 1/4$. Hence both triplets have the same projection on the chromatic plane. However, the HLS saturation of the first one equals 1 and that of the second 1/2.

3 2-D histograms and linearly regionalized spectra

In practice, is it really worth deviating from beaten tracks, and lengthening the polar triplets list? What for? We may answer the question by comparing the *luminance/saturation* bi-dimensional histograms for HLS system and for L_1, L_2 and $max - min$ norms. J. Angulo and J. Serra did so on a dozen images [3] [4]. Two of them are depicted below, in Fig.3.

3.1 Bi-dimensional histograms

In the first image, we observe strong reflections on the rubber ring, and various types of shadows. The corresponding histograms are reported in Fig.4, with luminance on the x axis and saturation on y axis. No information can be drawn from HLS histogram, some structures begin to appear in L_2 and $max - min$ norms, but the most visible ones come from L_1 norm.

By coming back to the images, we can localize the pixels which give alignments, as depicted in Fig.5. They correspond to three types of areas :

- shadows with steady hue,
- graduated shading on a plane,
- reflections with a partial saturation.

Consider now the more complex image of "Ana Blanco", in Fig.3b. It includes various sources light (television monitor, alpha-numerical incrustations), and the light diffused by the

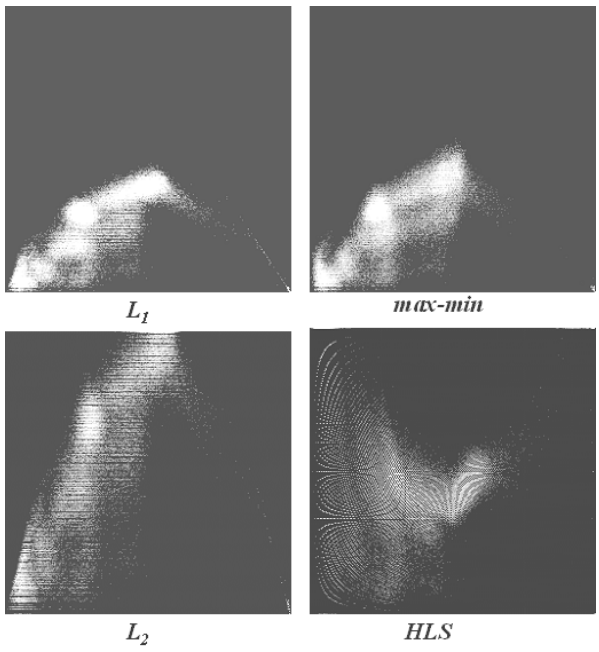


Figure 4: *Bi-dimensional histograms of the "rubber ring" image. The x-axis corresponds the luminance and the y-axis to the saturation.*

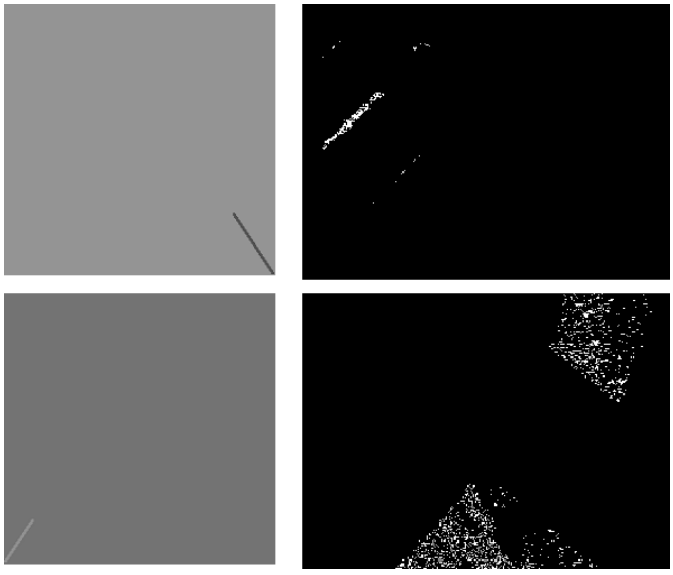
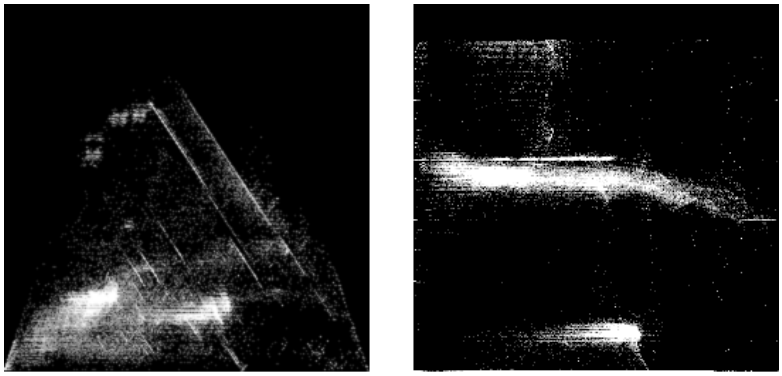


Figure 5: *Zones of "Rubber ring" associated with alignments. The left two images show the supports of the alignments detected in Fig.4 for the L_1 norm, and the right two images indicate the locations of the aligned pixels in the space of the initial picture.*



a *b*

Figure 6: *(a) and (b) the two histograms of "Ana Blanco", in the* luminance/saturation *and the* luminance/hue *plane respectively, both in* L_1 *norm.*

background is piecewise uniform over the space. However, there are still alignments, which do not always go through points $(0,0)$, or $(1,0)$, and are sometimes parallel. In the *lum/hue* plane of the L_1 norm representation, several horizontal lines (constant hue) are located at different hue levels, and alternate with elongated clouds of points (Fig.6b).

All in all, we draw from the above histograms four main informations.

1. In the *lum/sat* histogram, there is no accumulation of pixels at point (1,0). It means that the sensors we use are not physically saturated, which make realistic the proposed linear approach;
2. Still in the *lum/sat* histogram, some well drawn alignments can be extrapolated to point (0,0) or point (1,0). The others are parallels to the first ones;
3. However, most of the pixels form clouds in both *lum/sat* and *lum/hue* histograms are not aligned at all, whether the model does not apply, or the homogeneous zones are too small;
4. In the *lum/hue* histogram, most often the aligned pixels exhibit a (quasi) constant hue, i.e. draw horizontal lines. But sometimes, these "lines" turn out to be a narrow horizontal stripe.

Such characteristic structures, such distinct lines suggest we seek a physical explanation of the phenomenon. This is what we will do now. But besides any physical model, a first point is worth to be noticed: the only norm that enables us the extraction of reflection areas, of shadows and gradations is L_1. No other polar model results in such an achievement.

3.2 Linearly regionalized spectra (LR model)

If we assume that the alignments are a property of the spectrum, and not an artefact due to some particular representation, we have to express the spectrum in such a way that the

sequence

$$(spectrum) \rightarrow (r'g'b') \rightarrow (m's'h') \rightarrow (m' = \alpha s' + \beta)$$

be true (in the alignments) whatever the weights generating r, g and b are, and also whatever the spectrum itself is. Consider a zone Z of the space whose all pixels yield an alignment in the L_1 histogram. Denote by $sp(\nu; z)$ the spectrum of the light intensity at point $z \in Z$. We will say that this spectrum is *linearly regionalized* in Z when for each point $z \in Z$ one can decompose $sp(\nu; z)$ into the sum of a first spectrum $sp_0(\nu)$, independent of point z, and of a second one, $\overline{\omega}(z)sp_1(\nu)$, which proportionally varies in Z from one point to another. For all $z \in Z$, we have

$$sp(\nu; z) = sp_0(\nu) + \overline{\omega}(z)sp_1(\nu) \tag{19}$$

where $\overline{\omega}(z)$ is a numerical function which depends on z only, and where sp_0 and sp_1 are two fixed spectra.

In the spectrum $sp(\nu; z)$, though sp_0 usually corresponds to diffuse light and sp_1 to specular one, we do not need to distinguish between the emitted and reflected components of the light. It can be the concern of the light transmitted through a net curtain, for example, or of that of a TV monitor; but it can also come from passive reflectance, such as those described by Shafer's di-chromatic model [32], or by Obein's model of glossiness [22]. But unlike these two models, the term $\overline{\omega}(z)sp_1$ may also represent an absorption, when it is negative. Similarly, we do not need to distinguish between diffuse and specular lights. The term sp_0 may describe a diffuse source over the zone Z, as well as a constant specular reflection stemming from the same zone. But above all, the emphasis is put here on the *space variation* of the spectrum. It is introduced by the weight $\overline{\omega}(z)$, that depends on point z, but not on spectrum sp_1. This weight may bring into play cosines, when the angle of the incident beam varies, or the normal to a glossy surface, etc...

The three spectra sp, sp_0 and sp_1 are known only through the weighting functions that generate a (R, G, B) triplet. We use here the notation (R, G, B) in a canonical manner, i.e. it may designate the (X, Y, Z) coordinates of the CIE, or the perceptual system (L, M, S) [35], as well as the (Y, U, V) and (Y, I, Q) TV standards. In all cases it is a matter of *scalar products* of the spectra by such or such frequency weight. In particular, the white colour given by $r = g = b = 1$ can be obtained from a spectrum which is far from being uniform. We write

$$r(z) = \int sp(\nu; z)\,\xi(\nu)d\nu = \int [sp_0(\nu) + \overline{\omega}(z)sp_1(\nu)]\,\xi(\nu)\,d\nu = r_0 + r_1\overline{\omega}(z) \tag{20}$$

$$g(z) = \int sp(\nu; z)\,\chi(\nu)\,d\nu = g_0 + g_1\overline{\omega}(z) \tag{21}$$

and

$$b(z) = \int s(\nu; z)\,\psi(\nu)\,d\nu = b_0 + b_1\overline{\omega}(z) \tag{22}$$

where ξ, χ and ψ are the three weighting functions that generate the primary colours r, g and b.

As sp_0 and sp_1 are power spectra, they induce *intensities* r, g, and b. Now, in the above histograms, the L_1 norm applies to the *video variables* $r' = r^{1/\gamma}$, $g' = g^{1/\gamma}$, and $b' = b^{1/\gamma}$ (if we neglect the behaviour near the origin). Then we draw from Rel.(20)

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

with similar derivations for the video green and blue bands.

Is the linearly regionalized model able to explain the alignments in video histograms, despite the gamma correction? For the sake of simplicity, we will tackle this question by fixing the order of the video bands as $r' \geq g' \geq b'$, and $m' \geq g'$. Then we have

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

Alignments with the dark point

In the *luminance/saturation* histograms in L_1 norm, several alignments are in the prolongation of the point $(0, 0)$, of zero luminance and saturation. The shadow regions of the "rubber ring" image illustrate this situation.

Suppose that, in the relation (19) which defines the LR spectrum, the term $sp_0(\nu; z)$ is identically zero. Then $r(z)$ reduces to $\overline{\omega}(z)r_1$, which gives

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

with similar derivations for two other bands. Therefore we have

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

and

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

hence $m'(z)s_1' = m_1's'(z)$. In the space E of the intensities, we find in the same way that $m(z)s_1 = m_1 s(z)$. Therefore the nullity of the constant spectrum $sp_0(\nu)$ entails that both m' and s' on the one hand, and m and s on the other one, are proportional. Each *video* alignment indicates a zone where the *intensities* spectrum varies proportionally from one point to another.

Alignments with the white point

The "rubber ring" image generates also an alignment along a line going through the point $(1, 0)$, i.e. the point with maximum luminance and zero saturation. That suggests to suppose the spectrum $sp_0(\nu; z)$ constant and equal to 1, and in addition that the three colors r_1, g_1, b_1 are not identical (if not, the saturation s' should be zero). We have

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

and the two sister relations for $g(z)$ and $b(z)$. Under gamma correction, $r(z)$ becomes

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

Now, to say that the alignment is closed to a point of maximum luminance comes down to saying that r_1, g_1, and b_1 are small with respect to 1, or again that

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

hence $m'(z) = 1 + \frac{\overline{\omega}(z)}{\gamma} m_1$ and $s'(z) = \frac{\overline{\omega}(z)}{\gamma}[2r_1 - g_1 - b_1]$. We observe that the two conditions $r_1 \geq 0$ and $r'_1 \leq 1$, jointly with Rel.(24) imply that the coefficient $\overline{\omega}(z)$ is negative. Moreover, as the three colours r_1, g_1, b_1 are distinct, the condition $s'(z) \geq 0$ implies in turn that the quantity $2r_1 - g_1 - b_1$ is strictly negative. By putting $\sigma_1 = -(2r_1 - g_1 - b_1) > 0$ (σ_1 is not the saturation at point z_1), we obtain the following linear relation with positive coefficients

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

As in the previous case, but without approximations, the mean $m(z)$ and the saturation $s(z)$ of the intensities are linked by the same equation (25): it is a direct consequence of Eq.(23). Again, both video and intensity histograms carry the same information, and indicate the zones of almost white reflections.

Alignments with a gray point

There appears in some images, as "AnaBlanco", series of parallel alignments. Their supports go through points of (quasi) zero saturation but their luminance is strictly comprised between 0 and 1. The interpretation we just gave for the case of reflections extends to such a situation. It is still assumed that $r_0 = g_0 = b_0$, but with $0 < r_0 \leq 1$, and that the terms $\overline{\omega}(z) r_1$, $\overline{\omega}(z) g_1$, and $\overline{\omega}(z) b_1$ are small with respect to r_0. Then we have,

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

and the two sister relations for g' and b'. Hence

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

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

so that, finally

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

When the colour component (r_1, g_1, b_1) remains unchanged, but that the gray component (r_0, g_0, b_0) takes successively various values, then each of them induces an alignment of the same slope $\frac{m_1}{s_1}$. The property extends to the histograms of the intensities themselves. Finally, we derive from Eq.(17) that, in the three cases, the hue remains *constant* in each of these zones.

4 Saturation weighted segmentations

The most radical change between the classical HLS system and those based on norms holds on the saturation equation. In system (18), when $\min(r, g, b) = 0$, (with $l \leq 0.5$) or when $\max(r, g, b) = 1$ (with $l \geq 0.5$), then the saturation equals 1. Now for human vision, the most significant parameter is the hue in high saturated areas, and it turns to luminance when saturation decreases. Any person whose reaction to colours is normal can easily check it. In the darkness, or, at the opposite, in white scenes (e.g. a landscape of snowy mountains), the eye grasps the contours by scrutinizing all small grey variations, whereas when the scene juxtaposes spots of saturated colours, then the eye localizes the frontiers at the changes of the hue. But how to transcribe quantitatively such a remark by a saturation function that takes its maxima precisely when the colours loose their saturation, as the classical HLS system does?

The norms based representations correct this drawback, so that their saturations may serve to split the space into hue-dominant versus grey-dominant regions. This very convenient key to entering the segmentation of colour images was initially proposed by C.Demarty and S.Beucher [10]. They introduce the function $max - min$ on the image under study, and threshold it at a level s_0 that depends on the context. Then they adopt the HSV representation, but they replace its saturation by 1 in the regions above s_0 and by 0 elsewhere. Their downstream segmentations become easier and more robust.

However, they did not take the plunge of a new representation, and they worked at the pixel level, which is not the most informative. In order to go further in the same way of thinking, J. Angulo and J. Serra propose, in [2], the following two steps segmentation procedure:

1. to *separately* segment the luminance, the saturation and the hue in a correct Newtonian representation;

2. to combine the obtained partitions of the luminance and of the hue by means of that of the saturation: the later is taken as a *criterion* for choosing at each place either the luminance class, or the hue one.

The three bands of the "parrots" image of Fig.13, in L_1 representation, are depicted in Fig.7(a-c). Each band is segmented by the *jump connection* algorithm [31] (one groups in same classes all points x where $f(x)$ differs by less than k of an extremum in the same connected component, these classes are then withdrawn from the image, and one iterates). The method depends only on the jump positive value k.

As the parameter k increases, the over-segmentations reduce, but in compensation heterogeneous regions appear. A satisfactory balance seems to be reached for $k = 20$ (for 8-bits images), up to the filtering of a few very small regions. We obtain the two segmentations depicted in Fig.8.

4.1 Synthetic partition

How to combine the two partitions of images 8a and 8b? The idea consists in splitting the saturation image into two sets X_s and X_s^c of high and low saturations respectively, and in

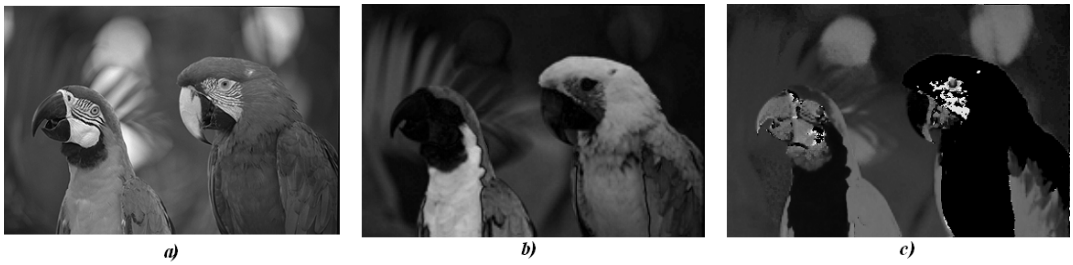


Figure 7: *Representation of the"Parrots" image 13 in L_1 norm : a) luminance, b) saturation, c) hue.*

Figure 8: *Grey segmentations of the luminance (a) and the hue (b). Both are depicted in false colour.*

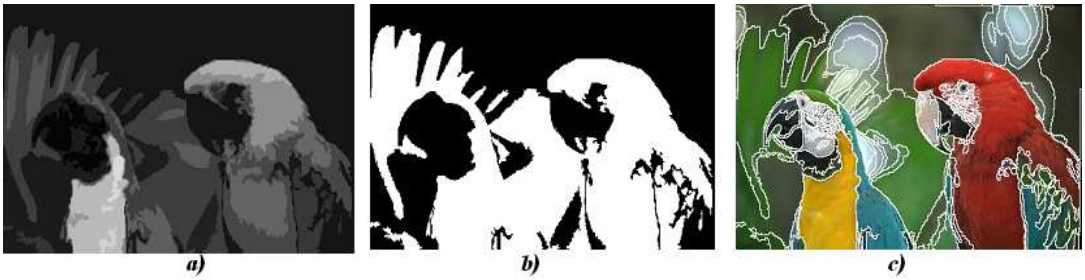


Figure 9: *a) Segmentation of the saturation (presented in grey tones); b)optimal threshold of a); c) final synthetic partition, superimposed to the initial image.*

assigning the hue partition to the first set, and the luminance one to the second. A class of the synthetic partition is either the intersection of a luminance class with the low saturation zone X_s^c, or the intersection of a hue class with the high saturation zone X_s. If the classes of the luminance, the hue, and the synthetic partition at point x are denoted by $A_m(x)$, $A_h(x)$, and $A(x)$ respectively, we have

$$\begin{aligned} A(x) &= A_m(x) \cap X_s^c \quad \text{when} \quad x \in X_s^c \\ A(x) &= A_h(x) \cap X_s \quad \text{when} \quad x \in X_s. \end{aligned} \tag{26}$$

The simplest way to generate the set X_s consists, of course, in thresholding the saturation image. But this risks to result in an irregular set X_s, with holes, small particles, etc. Preferably, one can start from the mosaic image of the saturation provided by the same segmentation algorithm as for the the hue and the luminance (Fig.9a). An optimal threshold on the saturation histogram determines the value for the a-chromatic/chromatic separation (Fig.9b). By applying Rel.(26) we finally obtain the composite partition depicted in Fig.9c, which is excellent.

5 The unit circle and the hue

The unit circle, like the round table of King Arthur's knights, has no order of importance, and no dominant position. In mathematical terms, this signifies that we cannot construct a lattice on the unit circle, unless assigning it an arbitrary origin. This is a severe verdict against morphological treatments (i.e. operators relying on lattices) when we use them on the unit circle.

However, is it really impossible to bypass this interdiction? If we consider the standard morphological operators, three paths at least seem possible, that A. Hanbury and I investigated in [11]. Only he first path is presented here, because it was designed it by transposing Matheron's idea of working on increments, that underlies the whole linear geostatistics. He did it because the range of fluctuations of the grades, in some mineral deposits, seems practically infinite, although the increments at finite distances remain finite. In the case of the hue, we are not disturbed by the infinity, but by the choice of an origin, which arbitrarily

forces the reds to be the smallest values and the purples largest ones (see for example the histogram of Fig.11). Can we transfer to the circular case the class of operators which bring into play differences only, such as gradients, top-hats, medians, etc.?

5.1 Circular centered operators

We fix an origin a_0 on the unit circle C with centre o by, for example, choosing the topmost point, and indicate the points a_i on the circle by their curvilinear coordinate in the trigonometric sense between 0 and 2π from a_0. Given two points a and a', we use the notation $a \div a'$ to indicate the value of the acute angle aoa', i.e.

$$a \div a' = \begin{cases} |a - a'| & \text{if} \quad |a - a'| \leq \pi \\ 2\pi - |a - a'| & \text{if} \quad |a - a'| \geq \pi \end{cases} \tag{27}$$

If the a_i are digital values between 0 and 255 (for example), the expression "$\leq \pi$" becomes "≤ 127", and "2π" becomes "255". However, we continue using the notation in terms of π, as it is more enlightening. Rel.(27) appears in [23] applied to the treatment of the hue band of colour images.

5.2 Circular hue gradient

We know that in the Euclidean space $\mathbb{R}^n$, to determine the modulus of the gradient, at point x, of a numerical differentiable function f, one considers a small sphere $S(x, r)$ centered on x with radius r. Then one takes the supremum minus the infimum of the increments $|f(x) - f(y)|$, where y describes the small sphere $S(x, r)$, i.e.

$$2g(x, r) = [\vee \{|f(x) - f(y)|, y \in S(x, r)\}] \bigvee [\wedge \{|f(x) - f(y)|, y \in S(x, r)\}] \tag{28}$$

Finally, one determines the limit of the function $g(x, r)$ as r tends to zero. This limit exists as the function f is differentiable in x. In the two-dimensional digital case, it is sufficient to apply Rel.(28), taking for $S(x, r)$ the unit circle centered on x (square or hexagon). This is the classic Beucher algorithm for the gradient.

Consider now an image of hues or of directions, i.e. a function $h : E \to C$, where E is an Euclidean or digital space, and C is the unit circle. As the previous development only involves increments, we can transpose Rel.(28) to the circular function h by replacing all the $|h(x) - h(y)|$ by $|h(x) \div h(y)|$. This transposition then defines the modulus of the gradient of the circular distribution. For example, in $\mathbb{R}^n$, $K(x)$ indicates the set of neighbours at distance one from point x, hence

$$2\,|\text{grad h}|(x) = [\vee \{|h(x) \div h(y)|, y \in K(x)\}] \bigvee [\wedge \{|h(x) \div h(y)|, y \in K(x)\}] \tag{29}$$

As an illustration, consider the hue component of Fig.10a, shown in Fig.10b. This image was chosen as it is mostly red in colour, and in the angular hue encoding, the red has value zero. This means that pixels which appear red could have low hue values (for example, 0° to 30°) and high hue values (330° to 360°). A large discontinuity is therefore visible in the hue image, with red pixels appearing at the extremities of the histogram (Fig.11). A classical

Figure 10: *a) Pere Serra's Painting of the Virgen (detail, St Cugat Monastery, Barcelona); b) corresponding image of the hue in HLS system (image size* 352×334 *pixels).*

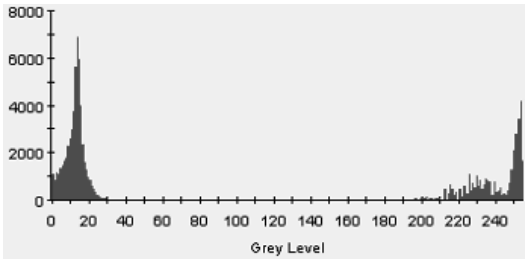


Figure 11: *Histogram of the hue band.*

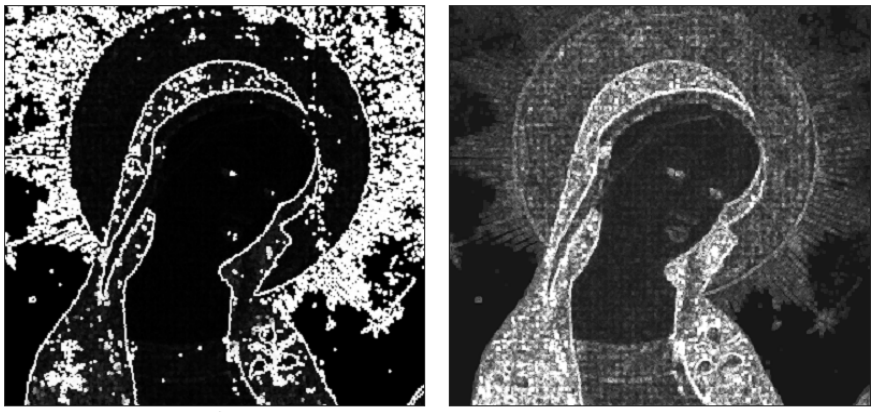


Figure 12: *a) Classical morphological gradient on the hue band; b) Circular centered gradient on the hue band. The gradients were calculated using a* 3×3 *structuring element*

gradient on this hue band produces a large number of spurious high-valued pixels, as shown in Fig.12a.

These high-values are present even though the neighbouring pixels appear very similar in colour, and are due to the discontinuity in the hue encoding. A good illustration of this is the outer part of the halo, which appears smooth in Fig.10a, but results in very high gradients in Fig.12a. The gradient calculated using Eq.(29), shown in Fig.12b, overcomes this problem. As its range is relatively small, we reinforced its contrast for the display, and the texture of the underlying canvas becomes visible. Note that if we rotate the hue band pixel values by π, the classical gradient will be the same as the angular gradient. The angular gradient is, however, invariant to rotations of the pixel values.

5.3 A use of the circular hue gradient

Consider a colour function f, and its L_1 norm representation. The circular gradient of the hue of f can be inserted in a global gradient |grad f| which takes the three components of f into account. There are number of such gradients, but the above comments on the saturation in L_1 norm representation suggests us to construct a *barycentric* gradient, where the L_1 saturation balances the effects of luminance and hue gradients, i. e.

$$|\text{grad f}| = \frac{s}{s_{\max}} \times |\text{grad h}| + (1 - \frac{s}{s_{\max}}) \times |\text{grad l}| \,. \tag{30}$$

It is well known that the watershed of the gradient |grad f| of a numerical function f provides us with the contours of f [19]. This property extends to the vector functions of the space, because then the term |grad f| is still a numerical function, which is sufficient for the theory. We can then use Rel.(30) to segment the parrots image, and compare the result with that we obtained by combined partitions in Fig.9c .

Figure 13: *a) Initial "Parrots" image; b) watershed partition at level 4 of the waterfall pyramid (superimposed to the initial image).*

Under iteration, the watershed operator turns out to generate the so-called waterfall pyramid, which is non parametric, and increases the partitions [19]. For the parrots, the best fit occurs at the fourth level of the hierarchy, see Fig.13b. The obtained segmentation is fair, but less convincing than that by combined partitions of Fig.9c : in the watershed process, the reduction (colour)→(numerical) by Rel.(30) arises too early.

6 Colour Interpolation

6.1 Morphological interpolators

During the nineties, the need for image interpolation arose with the development of video coding. Several proposed algorithms were efficient, though no theory justified them; one could not say whether they were optimal in some sense, just as kriging derives from a variance minimization for example. I asked G.Matheron which concept could serve as a substitute for variance in morphological interpolation, and he suggested me to see whether Haudorff metric admitted geodesics. The Hausdorff distance ρ is a metric defined on the class of the non empty compact sets of $\mathbb{R}^n$ by the relation

$$\rho(X,Y) = \inf\{\lambda : X \subseteq \delta_\lambda(Y)\ ;\ Y \subseteq \delta_\lambda(X)\} \tag{31}$$

where δ_λ stands for the Minkowski dilation by the ball of size λ. The value ρ indicates a degree of similarity between the two sets. A geodesic (if it exists) between sets X and Y from distance μ apart is an ordered family $\{Z_\lambda,\ 0 \leq \lambda \leq \mu\}$ of sets. Each Z_λ is at distance λ from X and $\rho - \lambda$ from Y. They are optimal interpolators in the sense that for any λ the sum $\rho(X, Z_\lambda) + \rho(Z_\lambda, Y) = \lambda + (\mu - \lambda)$ is minimum.

Matheron's intuition was right: the Hausdorff distance admits geodesics, and more than one [28]. However:

-1 the intermediary sets Z_λ are always inflated, but one geodesic swells less than the others;

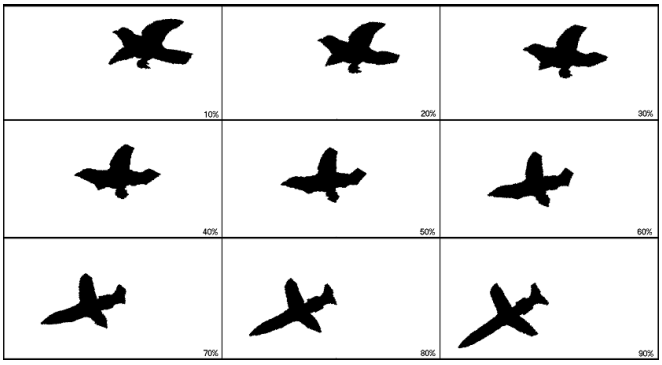


Figure 14: Morphological interpolations based on the Hausdorff distance.

-2 they are not self dual, as the interpolator of X^c and Y^c differs from the complement of the interpolator between X and Y ;

-3 when one of the two sets, X say, is shifted or rotated, the Z_λ's are modified.

Moreover, the duality between dilation and erosion w.r.t. complementation suggests to play with both Hausdorff distances by dilations, on $X \cap Y$, and by erosions, on $X \cup Y$. Then, the less inflating geodesic permits a self dual variant which gives, for $\lambda = 0.5\rho$, the so called *morphological median*

$$M(X,Y) = \bigcup_{\sigma>0} [(X \cap Y) \oplus \sigma B] \cap [(X \cup Y) \ominus \sigma B] \tag{32}$$

(assuming $X \cap Y \neq \emptyset$). Concerning the third point, one can always find a displacement that minimizes the distance $\rho(X,Y)$, hence optimizes the median [28]. Rel.(32) establishes a common theoretical basis to the previous works of F. Meyer [20], and S. Beucher [5]. Nice developments based on distance function have also been proposed by P. Soille [34], J.R. Casas [7] and by P. Moreau and Ch. Ronse [18].

M. Iwanowski found a method to determine the best displacement, and generated sequences of geodesics by subdividing morphological medians [14] [13]. A binary example is depicted in Fig.14.

6.2 A false colour case

In Rel.(32), the passage to the corresponding numerical version is straightforward. As the operator increases with both arguments, it suffices to replace all $\cup$ and $\cap$ by $\vee$ and $\wedge$ respectively. The extension to colour images is more subtle. For the first time in this paper, we have to produce a *colour* image, and neither a partition, nor a modulus of gradient. By treating separately the three bands, of any representation, we would obtain new colours that risk to parasite the quality of the result. The only way to be sure that every colour vector of the transform is already present in the initial image consists in *totally* ordering the set of the colour vectors. There are many ways to generate such a lattice. For the scenes of the

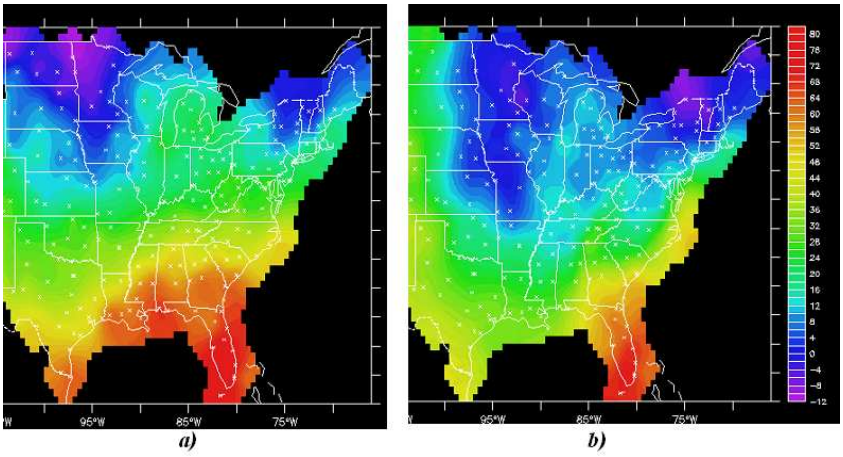


Figure 15: *Maps of the temperatures at two consecutive days.*

everyday life M. Iwanowski tried several possibilities and finally adopted an ordering with the luminance $l = 0.3r + 0.6g + 0.1b$ first, then g and finally r [12].

But the reference to G. Matheron leads me propose an example in the themes of geostatistics. The two maps depicted in Fig.15 come from the web-site of the American Meteorological Institute, and indicate the daily maximum temperature on January the 5th and the 6th, 1996. They are themselves interpolations from point measurements, but we do not mind about it and take them as an input. The complete series comprises seven maps, for one week, and which cover the whole country. The final interpolated sequence contains 102 images. Here we just calculate the morphological median between the two maps of Fig.15.

The gamut of colours, artificial, has been established in such a way that the temperatures linearly decrease with the hue. Moreover, for graphic reasons, the saturation is maximum everywhere. Not only the circularity argument against the hue vanishes, but this parameter is the most representative of the physics of the phenomenon. Therefore the convenient ordering is as follows

$$
\begin{array}{lll}
f'(h,l,s,) \geq f(h,l,s) \text{ when} & \text{either } h' < h & \\
& \text{or } h' = h \quad \text{and } l' > l & \\
& \text{or } h' = h \quad l' = l \quad \text{and } s' > s &
\end{array}
$$

It results in the morphological median depicted in Fig.16a. The arithmetic mean of the same maps is placed at the right (Fig.16b) for the sake of comparison. The morphological operator better follows the fronts because it keeps unchanged the number of gradations of the temperatures scale.

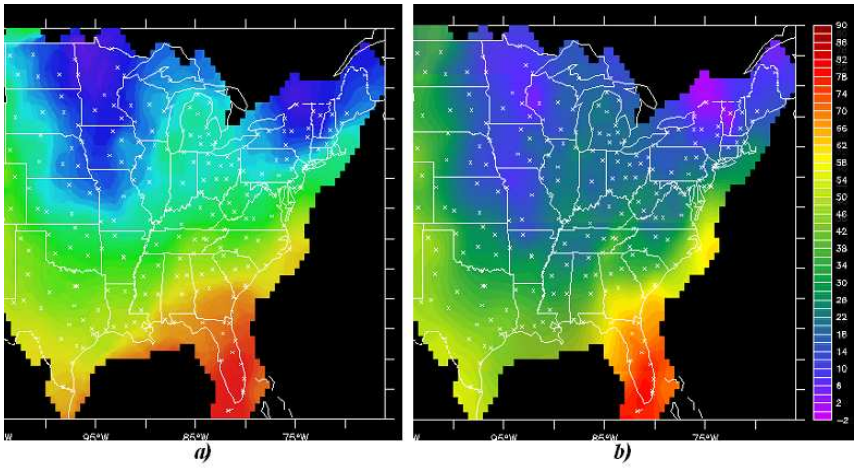


Figure 16: *a) Morphological median between the two maps of figure 15; b) linear interpolation between the same maps.*

7 A morphological approach to multivariate analysis

This last study on colour imagery commemorates G. Matheron less by the approach than by the theme of the application. Indeed mathematical morphology is born on the occasion petrographical problems [17], and it is my pleasure to conclude the survey by an example which recalls our beginnings.

In digital processing, it occurs that a series of different descriptors be attached to each pixel, and that one wishes to segment the space into the homogenous regions on the basis of this information. The descriptors may combine data that are physically heterogeneous. For example, in geographic information systems (G.I.S.), each pixel is sometimes assigned a slope and a population density. These two quantities, although connected, are physically

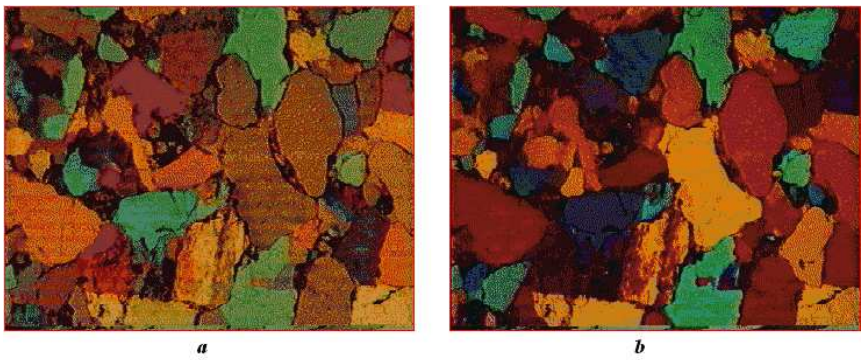


Figure 17: *a) and b) two polarized views of the Lipka sandstone thin section.*

non comparable, and their mean, hence a linear descriptor, is just a physical nonsense. Such an heterogeneity becomes more complicated again when some of the quantities vary over the unit circle (as the colours obtained by various polarization angles), and when the colour information has been coded by palettes. It will be the case here, but our approach, though designed for this current case, applies to any type of heterogeneous multivariate data. The underlying idea is the following. By segmenting separately each of the n variables, we obtain n *partitions*, which are data of the same type, and without the initial heterogeneity. The goal then consists in combining these partitions for producing a synthetic one. One possible solution is developed below: a numerical function is generated by adding indicator functions associated with each partition, then this numerical function is segmented.

7.1 Polarized thin sections

Polarized light is a useful tool for examination of rock thin sections. It assigns a specific hue to each rock component and allows distinguishing and separating the grains according to the crystalline orientations [21]. But it is generally used in a qualitative way.

In this section we purpose to use a partition based approach for reaching the same goal, but in a quantitative manner. In other words, starting from a sequence of polarized images of a same microscopic field, we propose to determine automatically the contours of the grains.

J. Serra and M.Mlynarczuk have treated four rocks of different microscopic appearance, in order to base the proposed method on a significant spectrum of rock structures [29]. They are two sandstones from Tumlin and Lipka, one quartzite from Wiśniówka and one dolomite from Rędziny. From each selected rock a thin section was prepared. Then a chosen field was observed in a polarized microscopy, in such a way that the polarization prism was turned 15 times by 12 degrees. The obtained sequence (15 images of 8 bits depth and of size 352 X. 268), was digitized by means of a CCD camera. The colour of each image has been coded on 256 levels by using a specific palette for each image. Fig.17 gives an outline of the obtained images. By observing the images, one can notice that:

- for the same thin section, the variations of colours between the images observed in different polarizations are considerable;
- for a same grain, and under a given polarization, the hue remains more or less uniform;
- the crystals are placed side by side, without any visible border;
- finally, we can notice that the images are of poor quality.

7.2 Proposed approach

A palette is a look-up table, i.e. a mapping $[0, 255] \times [0, 255] \times [0, 255] \rightarrow [0, 255]$. Therefore, the 256 output values do not constitute a perceptual grey-tone axis. In a same image the luminance associated with the level 120, say, may be higher or lower than that of level 140. Moreover, if at point x, two different images of a same sequence have the same numerical value, this does not mean that the two images have the same colour or the same luminescence at point x, because the palette changes from one image to another. But, fortunately, in each

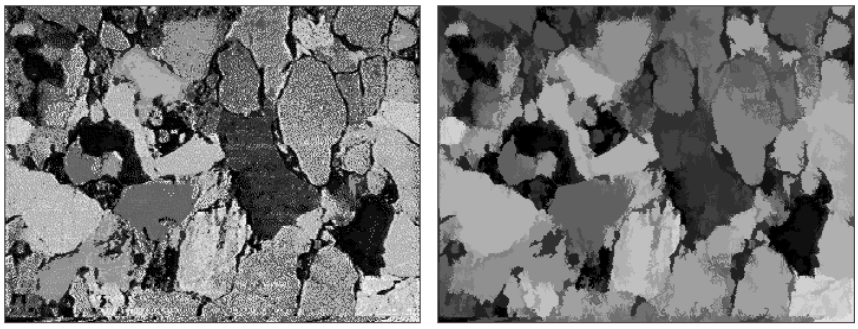


Figure 18: *a) artificial grey tone image which is associated with the palette of Fig. 17a; b) flat zones filtering of image a), namely image fi.*

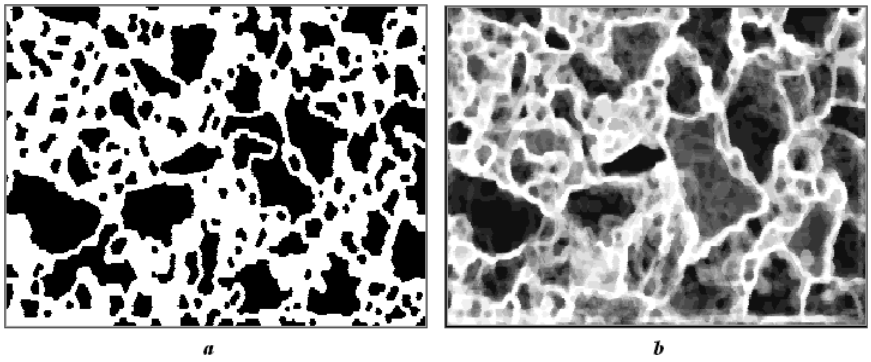


Figure 19: *a) Black and white image m_i obtained by partition erosion ; b) Sum of all m_i images of the sequence, resulting in a probability map.*

image, the majority of points of the same crystals admit the same numerical values. This happens in most cases, but not always, because the crystals are not absolutely pure.

Classically, a processing of multidimensional data begins by reducing the number of dimensions by means of linear multivariate analysis techniques, such as principal component analysis or the Karunen-Loève transformation. When there are no more than one or two major variables, different filters and segmentations conclude the processing. Here linearity is inadvisable and we have to invent another approach [29].

Our procedure turns out to be the exact opposite of multivariate analysis. If we want an approach to be valid for dimensionally heterogeneous data, it is recommended not to try and mix them in a first stage, which should be more or less linear. On the contrary, we must:

i) first associate a partition of the space with each individual image. A partition $\mathcal{D}$ of a space E is a segmentation of E into classes $D(x)$ that do not overlap, and that cover the whole space E. For this stage, we will have to *erode* partitions, in order to enhance boundaries, by applying the following result

Proposition 1 *Given an arbitrary set E, every set erosion $\varepsilon^* : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ for which $\varepsilon^*(\emptyset) = \emptyset$ induces, on the lattice $\mathcal{D}$ of the partitions of E, a unique erosion ε defined by:*

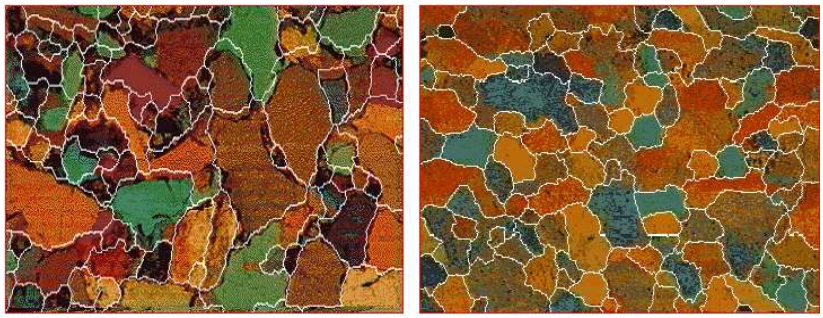
$(\varepsilon D)(x) = \varepsilon^*[D(x)]$ *when* $x \in \varepsilon^*[D(x)]$

$(\varepsilon D)(x) = \{x\}$ *when not .*

ii) and on the second stage only, group the obtained partitions into a synthetic one. To do this, one considers them as different realizations of a random partition, and one estimates the probability for each pixel to be at a given distance d from the border of its class. In this probability map, the closer to the border a pixel is, the higher is its numerical value. Finally, the watershed of this probability map provides the wanted segmentation.

7.3 An example

We will illustrate the algorithm by means of the Lipka sandstone sequence.



a *b*

Figure 20: *a) Watershed of the probability map Fig. 19 b), in superimposition with one of the polarized views. Same procedure, and same parameters applied to a quartzite from Wisniowka.*

First stage: Filter each image of the sequence by flat structuring elements. The optimal filter is the smaller that eliminates the various defects, since the defects are always smaller than the grains. The solution adopted here consists an alternating sequential filter by reconstruction of size two [26]. We call f_i the images after filtering, where i indicates the label of the image in its sequence (Fig.18b). The obtained flat zones are then eroded by applying Proposition 1 for an hexagonal erosion of size two. In such a process, each class of the partition, i.e. each flat zone, is narrowed independently of the others, and the rest of the space is occupied by classes reduced to points. In Fig.19a, the non point classes are given value zero, and the others value one. Let m_i denote the obtained image.

Second stage The border probability map is estimated by the estimate $m = 255 - 15(\sum m_i)$. Fig.19b depicts the image m associated with Lipka sequence. The final detection of the contours is provided by watersheding image m. Since we are interested in large regions without internal details, we take as a marker, for watershed, the set of the maxima of

function m, after an opening of size 3. On Fig.20a the final watershed lines are superimposed on one of the polarized images of the initial sequence. Fig.20b depicts the result of the same algorithm, with the same parameters, for Wiśniówka quartzite sequence.

Acknowledgements *The author gratefully thanks Prof. F.Meyer, Dr P.Dokladal and Dr J.Angulo for their valuable comments and the improvements they suggested.*

References

[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*, Jan. 2003.

[2] 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.

[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, Dec. 2003.

[4] J. Angulo, J. Serra, Traitements des images de couleur en représentation luminance/saturation/teinte par norme L_1, *technical report Ecole des mines de Paris* April 2004, 26p.

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

[6] T. Carron, *Segmentations d'images couleur dans la base Teinte-Luminance-Saturation: approche numérique et symbolique*,Thèse doctorale, Univ. de Savoie, 1995.

[7] J.R. Casas**,** *Image compression based on perceptual coding techniques*, PhD thesis, UPC, Barcelona, March 1996.

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

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

[10] Claire-Hélène Demarty and Serge Beucher, Color segmentation algorithm using an HLS transformation. In *Proceedings of the International Symposium on Mathematical Morphology (ISMM '98)*, p. 231-238, 1998.

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

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

[13] M. Iwanowski and 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.

[14] M. Iwanowski *Application de la Morphologie Mathématique pour l'interpolation d'images numériques* Phd thesis Ecoles des Mines de Paris- Ecole Polytechique de Varsovie, 15 nov. 2000.

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

[16] H. Levkowitz, G.T. Herman, GLHS : a generalized lightness, hue and saturation color model, *Graphical Models and Image Processing,* Vol. 55, No. 4, 1993, p.271–285.

[17] G. Matheron and J. Serra, The birth of Mathematical Morphology, *in Mathematical Morphology*, H. Talbot, R. Beare, Edts Proc. ISMM 2002, CSIRO Sydney 2002 pp.1-16.

[18] Moreau P., and Ronse Ch., Generation of shading-off on images by extrapolations of Lipschitz functions, *Graph. Models and Image Processing*, **58**(6), July 1996, pp. 314-333.

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

[20] 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.

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

[22] 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.

[23] R. A. Peters II. Mathematical morphology for angle-valued images, in *Non-Linear Image Processing VIII.* SPIE volume 3026, 1997.

[24] 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

[25] J. Pokorny, V.C. Smith, S.J. Starr, Variability of color mixture data I. The effect of viewing field size on the unit coordinates, *Vision Research*, Vol. 16, 1976, p.1095–1098.

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

[27] J. Serra Connectivity for sets and functions, *Fundamenta Informaticae, 41 (2000) 147-186*

[28] J. Serra, Hausdorff distance and Interpolations, *Mathematical Morphology and its applications to image and signal processing*, H.Heijmans and J. Roerdink eds Kluwer,1998, pp.107-115

[29] J. Serra and M.Mlynarczuk, Morphological merging of multidimensional data, *Proc. STERMAT'2000*, Cracow, Sept. 2000, pp.385-390 and 455.

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

[31] J. Serra, A lattice approach to Image segmentation, *Rapport Technique CMM-Ecole des Mines de Paris, N-02/04/MM,* 87 p. Janv. 2004 (to be published by JMIV)

[32] 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.

[33] A.R. Smith, Color gammet transform pairs, *Computer Graphics,* Vol. 12, No. 3, 1978, p.12–19.

[34] P. Soille, Spatial Distribution from Contour Lines: An efficient Methodology based on Distance Transformations, *J. Vis. Com. and Im. Under.* Vol.2 n°2 , June 1991, pp.128-150.

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