Marble classification using scale spaces

G.Dislaire & E.Pirard
Université de Liège, GeomaC, Géoressources Minérales, Liège, Belgium.

M.Vanrell
Universitat Autònoma de Barcelona, Spain.

ABSTRACT: Marble texture classification is an error prone undertaking when performed by humans. Therefore normalized methods are needed in order to obtain reproducible results. Technological advances in digital image acquisition and computing allows for the building of systems based on such methods.
The classification will be represented here by dyadic scale-space models (powers of 2). We will take into account the functioning of the human visual system in reproducing its natural ability to extract the features of textures: opponent-colors space is used as well as dyadic approaches for both light-dark multi-scale feature detection and inter-pathway resolution ratios.
The spatial organization will be captured with the use of features from the statistical sum and difference histograms, from a model-based blob-oriented morphological scale-space and from statistics of wavelet coefficients.
Features will then be classified with a common method, a Learning Vector Quantization network.

1 INTRODUCTION

The requirement for aesthetic appearance constancy in marble products is essential to certify that slabs sold to the client are alike. In this way, a robust texture definition is important for the classification of slabs into homogeneous classes. Visual discrimination of the human expert can be translated into algorithms in order to reduce subjectivity of the human classification. Classification methods will be presented to illustrate the benefit of using scale-based models to improve the classification.

2 TEXTURE

It is always important to specify what we mean when discussing textures. Generally speaking, a texture is a repetition of pattern(s) with possible random variations in the primitive placement rules. To be more precise, we have to say that unlike structures or other organization types, the texture is strongly linked to the visual perception of this order. This explains the importance of the psychophysiology and of the translation into algorithms of the multi-channel frequency and orientation analysis performed by retinal and cortical neurons. In addition, a texture definition also depends on the observation scale. Therefore we will build a marble classification system based on a *visual perception* model of *spatial organization* of light intensities on a given *scale range*.

2.1 Texture definitions

In Tuceryan and Jain (1998) several texture definitions are proposed, definitions intrinsically linked to feature extraction methods chosen to identify the texture. They group methods into geometrical, structural, statistical, model-based and signal processing-based.

Structural methods assume that textures are composed of primitives - as textiles are composed of threads. Texture elements are first extracted and then the placement rule is analyzed. Elements can be blobs. Lindeberg's scale-space researches can be used to extract them at different scales (Lindeberg, 1994). Morphology can be used to analyze them and placement rule can be defined as a tree grammar using symbols - the primitives - to form strings - the textures.

Texture is related to the spatial distribution of light intensities, so statistical methods such as the co-occurrence matrices are reasonable texture analysis tools. Autocorrelation captures repetitive placements and drops slowly or quickly depending on whether the texture is coarse or not. This last property can be linked to the power spectrum in signal processing models.

Model-based methods take out a set of parameters defining a model generally used as a constraint for synthesizing similar texture. Known models are based on the Markov random fields or on the fractals. Portilla and Simoncelli introduce "A parametric texture model based on joint statistics of wavelet coefficients" (Simoncelli & Portilla, 2000) that seems to capture the nature of the texture - its essential features. It binds model-based methods to signal processing ones.

2.2 Classification comparison

Randen and Husøy (1999) compare texture classification using statistical and signal processing approaches [2]. For multi-textured images the best classification performance is achieved with the highly complex Quadratic Mirror Filter f16b filter bank. Nevertheless the computationally more efficient DCT approaches or QMF critically sampled filter are of interest because, as the feature dimensionality decreases, the classifier complexity decreases too. They also conclude that the co-occurrence and the popular Gabor filter are not superior.

The processing time must not be forgotten in comparative methods. A classification system has to be viewed as a whole: the complexity of the feature model extracted will require an efficient classifier. Therefore some of the co-occurrence features such as mean, energy and contrast could already be enough in some cases and lighter for a classifier.

2.3 Visual perception – Color and scale ranges

Psychophysiology has provided a very useful model for color reproduction. We are now aware of trichromacy, the ability to arrange a color match using three primary colors. Human oriented color spaces have also been constructed to reflect the opponent-colors pathways: the light-dark, the red-green and the blue-yellow channels. We will use the opponent-colors space presented in (Zhang. & Wandell, 1996) as a preliminary step to feature extraction:

$$X = 0.6067\,R + 0.1736\,G + 0.2001\,B$$
$$Y = 0.2988\,R + 0.5868\,G + 0.1143\,B$$
$$Z = 0.0000\,R + 0.0661\,G + 1.1149\,B$$

$$O_1 = 0.279\,X + 0.72\,Y - 0.107\,Z$$
$$O_2 = -\,0.449\,X + 0.29\,Y - 0.077\,Z$$
$$O_3 = 0.086\,X - 0.59\,Y + 0.501\,Z$$

Less known but also important is the difference of resolution within opponent-color spaces: the light-dark contrast achieves a maximum at 10 cycles per degree.

[1 degree corresponds to 0.89 centimeters on a screen viewed at 50 centimeters. In this configuration, 10 cycles per degree corresponds to 10/0.89 = 11.2 cycles per centimeters.]

We will see that, more than the mean, especially when searching color-similar sub-classis of a given

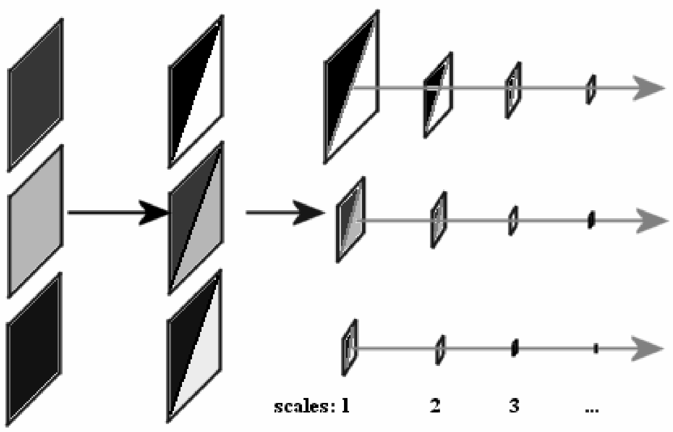

Figure 1. Using opponent color space model, we can link to a given scale different resolutions according to color pathways.

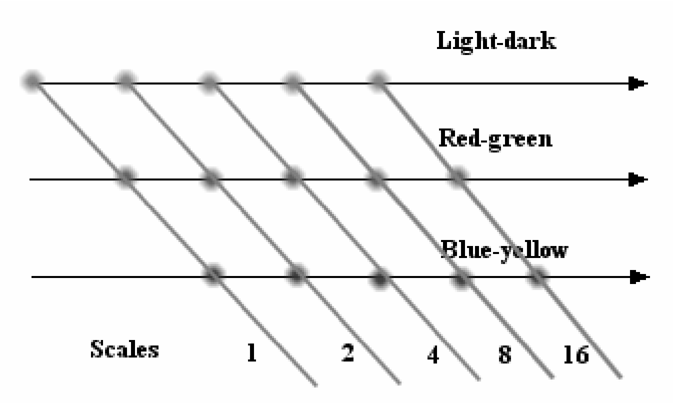

Figure 2. In a dyadic scale space decomposition of a color opponent model-based image, a given scale s_0 will be characterized by feature of the scale s for the light-dark channel, 2 x s_0 for the red-green channel, 4 x s_0 for the blue-yellow one.

slab type, the contrast seems to be an effective discriminating factor in the human classification criteria. Actually this is due to an induction phenomenon intrinsically rooted in the filtering process performed by the retinal neurons: they compress the information by reacting especially to transitions.

Induction is the source of many illusions with two opposite forms: contrast for low frequency (0.7 cpd for blue-yellow), assimilation for high frequency (9 cpd for blue-yellow) (Vanrell & Baldrich, 2003). Note that for 10 cpd, we have assimilation in blue-yellow but contrast in light-dark.

Light-dark contrast falls for resolutions above 30 cpd, red-green for 10-20 cpd when blue-yellow for 5-6 cpd (Wandell, 1999). Undertaking a practical approach allows us to imagine dyadic scales (powers of 2) depending on the pathways: given a scale s_0 for the light-dark pathway we use a scale s_1 = 2 x s_0 for the red-green and s2 = 4 x s0 for the blue-yellow one.

Let's conclude the present contrast presentation with a numerical example: a 30 centimeter tile seen from 50 a centimeters distance will give a maximum light-dark contrast for variations of 0.89 millimeters - corresponding to 30x11.2 = 337 variations in the width of the field of view - and a resolution of 0.30 millimeters - corresponding to 1011 variations. A dyadic scale will give resolution of 0.30/2 = 0.15 millimeters in the red-green pathway, of 0.075 millimeters for the blue-yellow one (see Fig. 1).

When we move away from a surface, we gain the larger scales - limited by the field of view - as we lose the smaller ones - limited by the retinal resolution.

3 EXPERIMENTAL MATERIAL

A set of Marfil slabs have been acquired using a tri-CCDs linear camera to obtain color images of the diffuse component of the light reflected by the slabs. Images have been pre-processed to remove shading effects due to non-homogeneity of the lighting. This set of images has been used as the basis of the scale-based models discussed in the followings paragraphs.

4 METHODS

Dyadic scales have been used to classify the Marfil slabs using different image analysis methods. For a given acquisition we train color features at different scales by shifting dyadic pathway ratios from high to low resolution (see Fig. 2).

4.1 Spatial organization

In order to feed a vector of features to the classifier we have to capture the nature of the spatial organization in a digital form. In this way statistical models capture mean, energy, entropy, contrast and homogeneity. Other features, such as uniformity, density, coarseness, regularity, linearity, directionality… have various implementations. A 'texture browsing descriptor' is considered by the MPEG-7 compression format.

But the most important is to keep relevant descriptors; thus depending on the texture, only certain ones are retained for defining a given model. We will compare the use of features from the statistical Sum and Difference Histograms (SDH), from a model-based blob-oriented morphological scale-space and from statistics of wavelet coefficients.

4.2 Marble classification

Ornamental stone textures are so varied that it is difficult to build a model classifying all the possible varieties found on the market. When granites seems to be easily classified due to a certain homogeneity of the repeated pattern at a given scale, marble often are characterized by the presence of veins that will produce a texture on a scale higher than the scale of the marble slab. Such slabs are evaluated by human experts with subjectivity and fatigue giving inconsistent results. Automatic classifications have been introduced by Martinez & Tomás (1999) to solve this problem using the SDH method which computes features on small neighborhoods.

Our experiments will focus on slabs of the type 'Crema Marfil' coming from Murcia.

4.3 Scale-space variations

To improve results based on statistical methods, we will study the texture on different scales. Three methods are presented. For each one we work in the opponent-color space model as described in Figure 2. A Gaussian kernel is each time applied at a given scale and subsampled to produce the larger scale.

4.4 Sum and Difference Histograms

The SDH algorithm is a powerful alternative to the usual spatial grey level dependence method or co-occurrence matrices: for a distance vector (d1,d2), the combination of two pixels zx,y and zx+d1,y+d2 forms the sum and difference vectors:

sx,y = zx,y + zx+d1,y+d2
dx,y = zx,y - zx+d1,y+d2

The normalized histograms are:

ps (i) = hs(i)/N = #(sx,y = i) / N
pd (i) = hd(i)/N = #(dx,y = i) / N

The statistical features used are:

$$\mu = \frac{1}{2}\sum_i iP_s(i) \qquad energy = \sum_i P_s(i)^2 \sum_j P_d(j)^2$$

$$entropy = -\sum_i P_s(i)\log(P_s(i)) - \sum_j P_d(j)\log(P_d(j))$$

$$contrast = \sum_j j^2 P_d(j) \qquad \mathrm{hom}\,og = \sum_j \frac{1}{1+j^2} P_d(j)$$

The Figure 3 illustrates the mean and two contrasts at different scales. The use of a scale factor for this last feature will improve the classification from 75% (1 scale) to 88% (6 scales) – an improvement factor of 17%.

We have to notice the poor initial result for 1 scale related to the 90% with the same method used by Martinez & Tomás (1999). This is likely due to a different set of images and only the improvement factor should be retained.

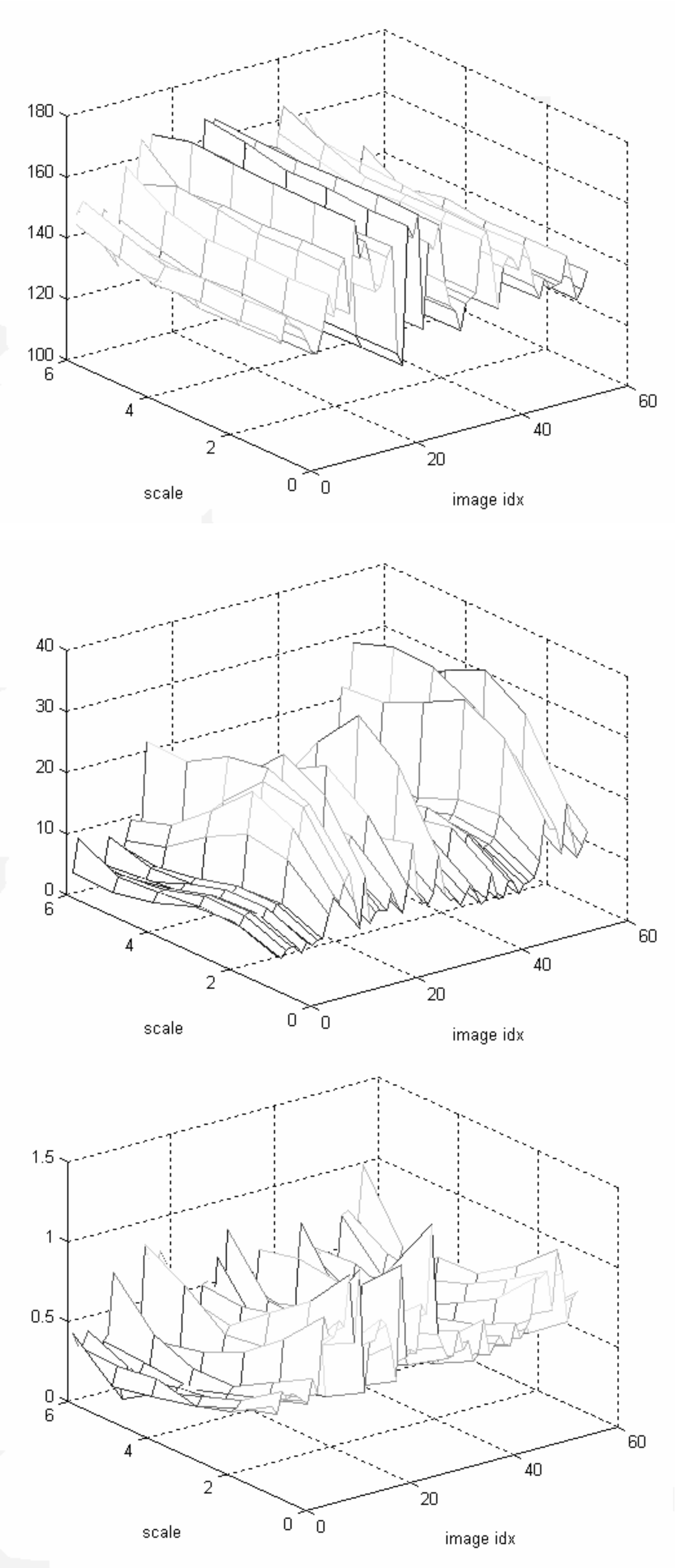

Figure 3. Light-dark mean, light-dark contrast and blue-yellow contrast for an image set and for 6 scales. The mean remains constant whatever the scale (what is expected) but the contrast present discriminent profiles.

4.5 Blob analysis in the Lindeberg's Scale-space

We explained that the contrast is a very discriminating factor in the human classification criteria. But how to implement it to reflect this specificity to an image acquired by a static vision system? The answer is the Laplacian of Gaussian.

The Gaussian kernel and its derivates are one of the most precious tools in image analysis. For instance, filtering with a Gaussian kernel simulates the assimilation as a perceptual blurring; filtering with the second derivate is named the Laplacian of Gaussian and it simulates the contrast as a perceptual sharpening (Vanrell&Baldrich, 2003).

The Laplacian filter gives a strong response in blob detection but is too sensitive to the noise, so a first Gaussian filtering has to be applied. Practically, it is the same to filter directly by the Laplacian of Gaussian:

$$I' = \nabla^2 (I * G_\sigma) = I * (\nabla^2 G_\sigma)$$

The Lindeberg's scale-space theory introduces normalization to allow comparison of blobs responses between scales. It automatically selects the scale at which local image structures are better detected by differential operators (Salvatella, 2002).

With a normalized Laplacian,

$$\nabla^2_{norm} = \sigma^2 \nabla^2$$

blobs responses are computed for all scales and the maximum over all scales gives all the image blobs no matter their size.

Basing ourselves on that fact, we will present a sharpening operator to not only find black and white blobs - by getting the minimas and the maximas over all scales but will produce an image with a flat background and contrasted response (see Fig. 4 & 5).

We have classified the Marfil slabs using Blob analysis on the segmented blobs with extraction of features like the mean and maximum area, the mean and maximum ellipsoid major axes, the mean and maximum of the eccentricity weighted by the corresponding diameter. Results of classification of 82% are promising because the blobs features are still not fully exploited. More detailed feature distributions analysis would give better results. Luengo (2004) uses the size distribution to characterize structures; this distribution is estimated using a so called granulometry, which is the projection of a morphological scale-space on the scale axis.

The major advantage of this technique is its ability to extract veins. Indeed, statistical methods do not find such 'non-textural' feature. Actually, veins

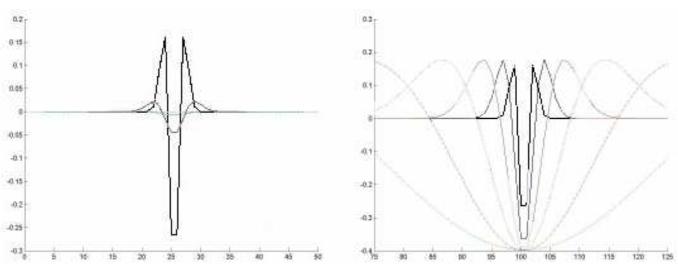

Figure 4. For σ = 1, 2, 4, 8, 16, profiles of Laplacian of Gaussian and normalized Laplacian of Gaussian.

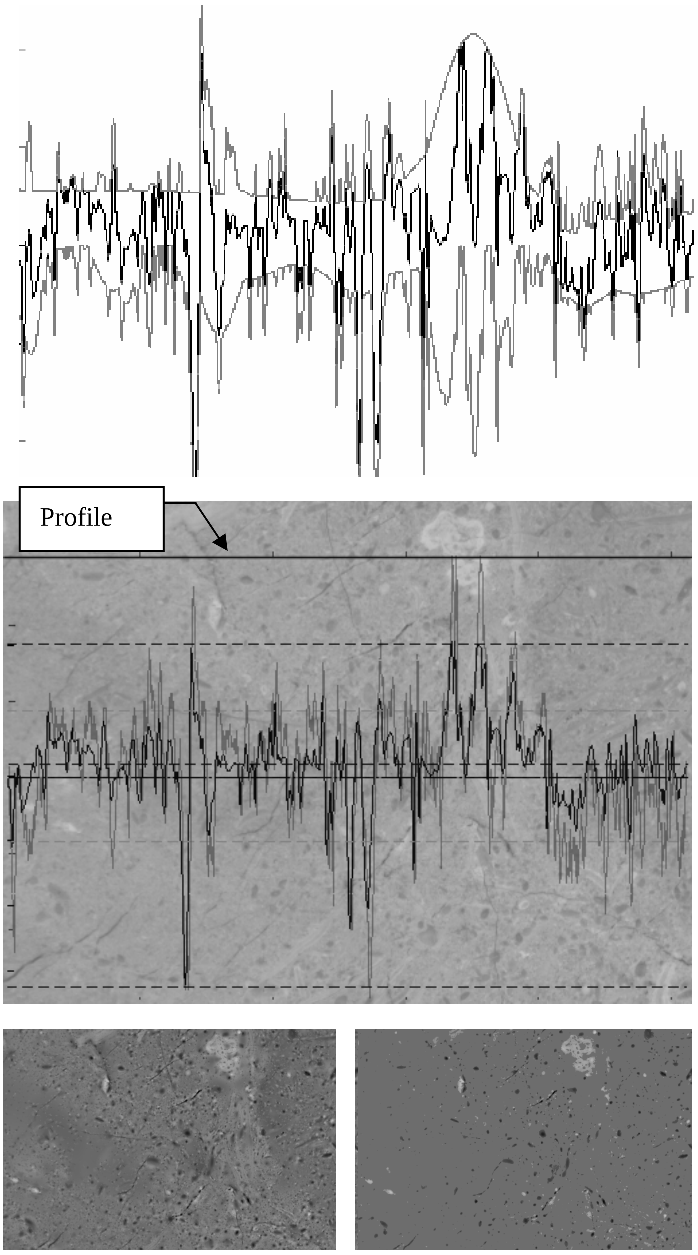

Figure 5. For given profile, the top picture illustrates the profiles of the maximums and minimums over all scales of LoG-normalized transformations giving top and bottom covers over these scales. Their difference gives the (LoG-normalized) sharpen profile. The middle picture shows, in overlay on a Marfil original image, the original profile (grey) and the sharpen one (dark). This sharpen operation flats the background and allows an easy segmentation by threshold. The bottom images illustrate a segmentation of the transformed original image.

are not repeated patterns producing a texture but produce a texture on a scale wider than the scale of the slab: that of the tiling.

4.6 Parametric Texture Model based on Statistics of Wavelets coefficients

Portilla and Simoncelli (2000) propose a universal model to capture important features of various texture images. It can serve as a high-level texture representation for characterization and segmentation applications. It uses a pyramidal approach similar to the Laplacian pyramid but capturing orientations: a steerable pyramid. From this representation, key features are extracted to define four statistical constraints: capturing the pixel intensity distribution (marginal statistics), the periodic or globally oriented structure (raw coefficient correlation), the structural information such as edges, corners…(coefficient magnitude statistics) and illumination gradients due to 3D appearance (cross-scale phase statistics).

This complex representation summarizes the nature of the texture in 710 feature values that can serve for classification. An important property of this representation is the ability to synthesize texture from these features to verify their validity.

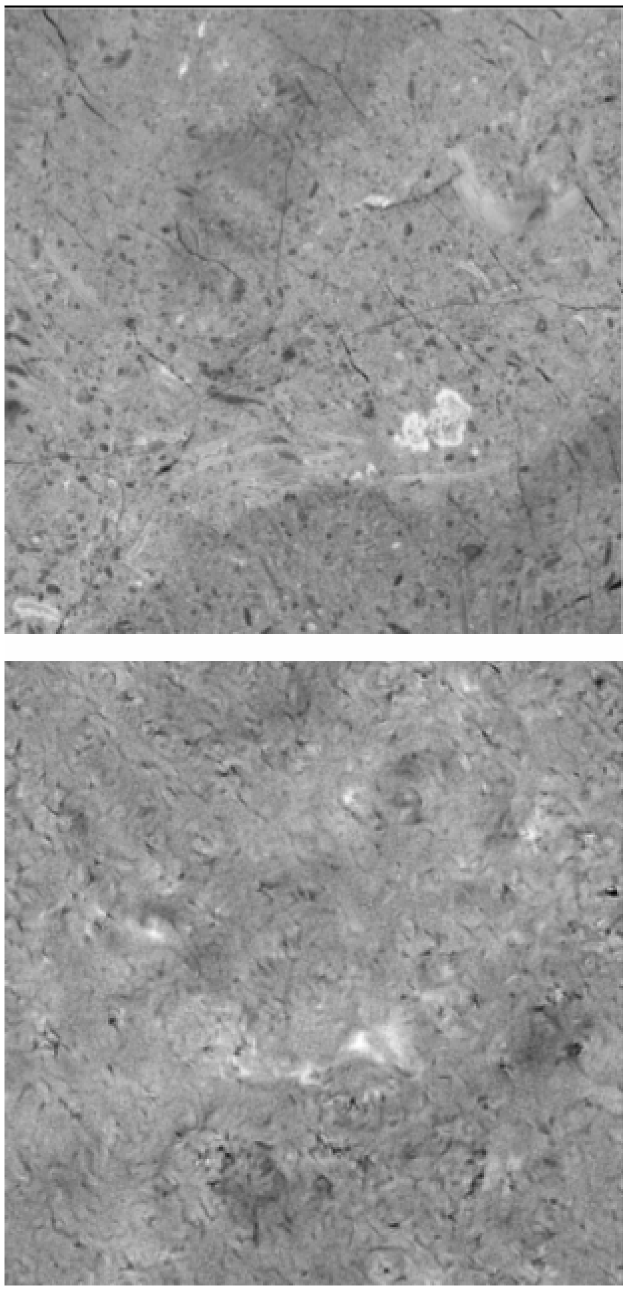

Figure 6. A Marfil slab and a synthesized image from the set of statistics of wavelet coefficients.

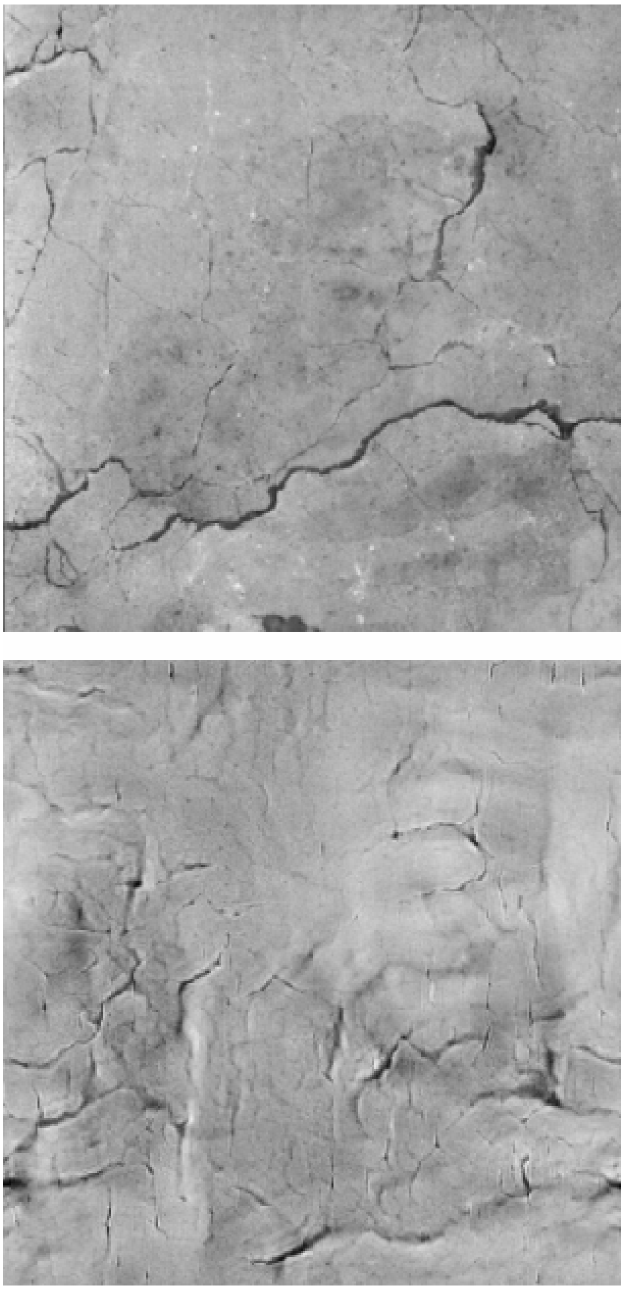

Figure 7. An other Marfil slab and its synthesized image illustrating a non-working case due to the veins extension not captured by the model.

5 CONCLUSION

Three scale-based models have been presented.

The SDH model gives interesting classification results on the Marfil slabs and we showed an improvement by introducing scale. This improvement seems to be essentially due to the contrast evolution over scales. This probably captures features of different sizes.

The Parametric Texture Model based on Statistics of Wavelets coefficients (PTMSWC) is very complete. It seems to be the "holy grail" of the statistic-based models due to its fully automatic research of inter- and intra-scale features. Nevertheless the classification of the 710 parameters has to be improved by a reduction of these parameters to the relevant ones. In this way, principal component analysis or manual discriminating feature evaluation will be studied.

The blob-oriented classification is the only one that is able to identify veins. Indeed the statistical models fail to capture such features due to the 'non-textural' nature of a vein. Therefore this method seems to be the more appropriated for veins analysis. Nevertheless, for background description the other approaches are more complete. Another problem of the blob-oriented model is its sensibility to the choice of the threshold used to extract blobs.

These three scale-based methods demonstrated the interest of scale in texture analysis. SDH over scales is an automatic and simple model that provides significant results. PTMSWC is more sophisticated and implies other studies to use the parameters as input to a classifier. The blob model catches the veins but is sensitive to the blob extraction threshold level.

In the three cases the scale is the only way to capture features of different sizes.

6 REFERENCES

Lindeberg, T. 1994. Scale-space theory: A basic tool for analyzing structures at different scales. *Journal of Applied Statistics, vol. 21, no. 2, 225-270, 1994.*

Martínez, J. & Tomás, L.-M. 1999. Marble Slabs Quality Classification System using Texture Recognition and Neural Networks Methodology. *European Symposium an Artificial Neural Networks 1999 proceedings*. 75-80.

Randen, T. & J.H. HusØy. 1999. Filtering for Texture Classification: A Comparative Study, *IEEE Trans.Pattern Analysis and Machine Intelligence vol. 21, no. 4, 291-310, April 1999*.

Salvatella, A. 2002. Texture Description based on Subtexture Components.

Simoncelli, P. & Portilla, J. 2000. A Parametric Model Based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision 40(1), 49-71, 2000.*

Tuceryan, M. & Jain, A.K. 1998. Texture Analysis *Handbook of Pattern Recognition and Computer Vision (2nd Edition)*: 207-248. World Scientific Publishing Co.

Vanrell, M; & Baldrich, R. & Salvatella, A. 2002. Induction operator for computational Color Texture Description. *Elsevier Science.*

Wandell, B. 1999. Computational Neuroimaging: Color Representations and Processing, *Cognitive Neuroscience edited by Michael Gazzaniga.*

Zhang, X. & Wandell, B. 1996. A Spatial Extension of CIELab for Digital Color Image Reproduction. *Proceedings of the SID Symposiums*, 1996.