

# Color texture recognition through multiresolution features

Felipe Lumbreras, Joan Serrat, Ramon Baldrich, Maria Vanrell

Juan José Villanueva

Computer Vision Center & Departament d'Informàtica

Edifici O, Universitat Autònoma de Barcelona

08193 Cerdanyola, Spain.

Tel. +34 93 5811828, Fax +34 93 5811670

[felipe@cvc.uab.es](mailto:felipe@cvc.uab.es)

Preference : poster

## Abstract

In this paper we address the problem of color texture classification and present results on two practical problems. The central idea is to combine color and texture information through the multiresolution decomposition of each channel in order to take as classification vector the energies and cross correlations of the coefficient images. However, this simple approach can be materialized in many different ways, as a several decisions have to be taken, each one allowing multiple choices : the multiresolution decomposition scheme (for instance, Mallat's, *à trous*, wavelet packets), the subspaces base family (and within it, which specific base), number of decomposition levels, space for color representation and finally, the classification features to be computed from the decomposition. Instead of simply trying some possibilities and take the best one, we have assessed a very large number of combinations, trying to find out which are the important and the non-relevant issues with regard the classifier performance. In addition, we propose three image models as a framework for color texture classification, depending on how texture is combined with color. This allows us not only to initially select the appropriate types of features but also to reduce the number of classification parameters so that the training set does not need to be large. This framework has been successfully applied to two specific machine vision problems, namely, the sorting of ceramic tiles into perceptually homogeneous classes and the recognition of metalized paints for car refinishing.

*Keywords:* color, texture, wavelets, recognition, sorting, ceramic tile, refinishing, paint.

## 1 Introduction

In this work we present a study on the wavelet decomposition and classification of color textures. It was prompted to solve an industrial machine vision problem, namely, the on-line sorting of polished ceramic tiles. Later, we approached another application, paint identification from microscopy images for car refinishing. We will see that the solution to the first application fits very well also to the second one.

From a computer vision point of view, we are addressing in both cases a problem of color texture representation and classification. The objective is to devise a numerical representation of images that captures both the color and texture features. In the present case, we are interested in the benefits of this representation for classification purposes, but it may also be useful in other contexts like color texture synthesis and database image indexing.

The whole application process consists of the following steps, being the slanted items those directly associated with the method of color texture analysis we propose in this article :

- image acquisition
- *change of color space representation*
- *multiresolution decomposition*
- *feature extraction*
- supervised classification

We will focus on the justification and performance of different choices for color representation spaces and multiresolution decomposition schemes. Our aim is to assess all possible combinations in terms of minimum classification error over a relatively large set of samples. Furthermore, we want to provide a sound explanation in terms of *why* each choice achieves

its result. This is done in the context of three color texture analysis models we propose, depending on how the image content within scales and channels is related.

This paper is organized as follows. Section 2 reviews previous work on joint computational representations of color and texture visual cues, including wavelet transforms of multichannel images. Next section describes the planned experiments, namely, the choices for color space, wavelet transform scheme and the classification features derived from them, which we are going to combine and assess. In section 4 we introduce the problem of ceramic tile classification and the main results obtained. Section 5 briefly deals with the same issues for the second case study of paint recognition. Finally, section 6 contains the conclusions and future work.

## 2 Related work

Color texture representation is a current topic in computer vision. Although both are properties of a surface, these two visual cues have been usually studied separately. One reason is that while color is a point feature given by the value of a pixel in several bands or channels, texture has been modeled as a spatial relationship of the point with its neighbours within each channel. An excellent review of approaches used in computer vision to deal with the texture representation problem can be found in [1], whereas an introduction to color representation is given in [2].

The study of color texture representations has received increasing attention in the last years. The objective of many researchers is to find co-joint representations of spatial and chromatic information which capture the spatial dependence (in particular, correlation) *within and among* spectral bands [3, 4, 5]. One of the most frequent approaches is the construction of a feature vector mixing grey level texture features and color features [4]. Another one is to extend classical texture models, such as Markov Random fields and the autocorrelation function, in order to deal with multichannel images [5, 6]. Other works, like [7], convert RGB values into a single code from which texture measurements are computed as if it were a grey scale image. Spatio-chromatic representations are computed in [3, 8] over the smoothed Laplacian of the image. Other works have been influenced by known perceptual mechanisms of the human visual system like Gabor filters [9, 11].

In parallel, multiresolution texture analysis has come to age thanks to the setting of a sound theoretical basis for wavelet transforms and filter banks. Recent works on texture incorporate color as an additional image dimension [9, 12, 13]. This has been applied to analysis but also to synthesis [14, 15] and texture classification [9, 11].

A color texture analysis based on a multiresolution decomposition representation normally involves to make up two decisions: the selection of the decomposition scheme to perform the texture analysis and the definition of a space to represent color. A general framework for image decomposition is to apply a bank of filters. Gabor filter banks and wavelet transforms are two common approaches found in the literature.

The simplest way to extend them to cope with color images is to filter or transform each channel (RGB for instance) independently. However, some authors propose to represent color in other spaces such as the opponent color space [9, 11], inspired in biological evidences of the human visual system. Both works start from similar color representations, followed by different texture analysis methods. We are going to devote some attention to them, as

they are closely related to our study.

The first one [9] uses the orthogonal wavelet decomposition and calculates the energy,  $e_i^k$ , of each detail level and the cross terms between different channels at the same detail level,  $c_i^{kl}$ :

$$e_i^k = \int (d_i^k(u))^2 du \quad (1)$$

$$c_i^{kl} = \int d_i^k(u) d_i^l(u) du \quad , \quad (2)$$

where  $u$  denotes spatial coordinates,  $i$  the decomposition level,  $k$  and  $l$  are channel indexes. Thus,  $d_i^k$  is the detail at level  $i$  of the channel  $k$ . In this specific case,  $d$  is an image of detail coefficients of a orthogonal wavelet decomposition, but it can be seen also as one of the outputs of a filter bank.

The second work [11] uses a set of Gabor filters where the response at different levels and channels is analysed. A biological model is implicit in this scheme due to the use of Gabor filters and to the extraction of the information between channels following the opponent color model. Energies at each level of every channel (terms  $e_i^k$  of equation (1) for all  $i$  and  $k$ ) are calculated, but also the energies associated to the inhibition between channels at different levels

$$I_{ij}^{kl} = \int (d_i^k(u) - d_j^l(u))^2 du \quad , \quad (3)$$

where  $d_i^k$  are now the responses of a Gabor filter bank. If we expand the inhibition terms of equation (3) we obtain the energies  $e_i^k$ ,  $e_j^l$  and a cross term that could be expressed as  $-2c_{ij}^{kl}$ , using the notation of equation (2). Therefore, both papers are using a similar representation.

To end this review, we want to mention a sound comparative study on the performance of texture classification algorithms by Randen and Husøy [10]. Like us, they want to assess combinations of wavelet decompositions and features, including additional filter bank schemes. However they test them only on graylevel images. But the main shortcoming of their study with regard ours is that they work with clearly distinct textures, that is, a subset of the Brodatz, Meastex and Vistex collections. Conversely, we are trying to differentiate among textures much more visually similar, as they come from the same industrial process (at least in the tiles case), this being a much tougher problem, as real problems usually are.

There are a few previous works to be considered in the specific subject of tile inspection. Some research effort has been devoted to the detection of other kinds of defects like cracks and spots. Only in [16, 17] the same problem of tile color texture classification is addressed. The authors try to solve it taking as features only statistical measurements on the color histogram. Therefore, results are poor in the event of overall similar color but different textural aspect, as it happens in our samples. Better results were obtained by some authors of the present paper by performing a color segmentation prior to an analysis of blob features [18].

### 3 Multiresolution color texture classification

#### 3.1 Color spaces

It is a common practice to use color representation that try to decorrelate information across channels, thus reducing the number of meaningful classification features. However, we will consider other choices, such as conversion from color to intensity and no transformation at all, in order to compare them with the decorrelation transforms. Therefore, the envisaged color spaces/transforms are:

**C.a** Color to grey level conversion by simple averaging of the R, G and B channels. Hence, only intensity is taken into account. This would make sense in images where texture is the only relevant feature for classification.

**C.b** Raw RGB values. That is, no transformation is applied to the image provided by the camera and frame grabber. In many applications this is sufficient to introduce the color information and classify successfully.

**C.c** Generic Karhunen–Loève transform. This is obtained through the base that best decorrelates the spectral information of a large set of color images. According to [9], it is given by the following fixed linear transform:

$$\begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.5 & 0.0 & -0.5 \\ -0.5 & 1.0 & -0.5 \end{bmatrix} \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix}. \quad (4)$$

It does not decorrelate spatially but somehow gets three new channels weighting each one by its real contribution when describing the input data with the new base.

**C.d** Specific Karhunen–Loève transform. Now, it is sought the base which achieves the maximum spectral decorrelation but over the training set of the application, In our case, it is the set of images for each class and model.

To fix ideas and for the sake of simplicity, we will illustrate concepts of this section with figures of the tile problem. Figure 1 shows the former four transforms for a  $128 \times 128$  region.

#### 3.2 Decomposition scheme and bases

The decomposition scheme is “application dependent”. Thus, for time critical applications, an orthogonal scheme like wavelets or wavelet packets with a reduced number of levels is generally preferred. Conversely, in images with high frequency content in the middle zone of the spectrum, a wavelet packet scheme should be better *a priori* because it allows to focus the analysis on the levels where the relevant information is. Likewise, in images which exhibit a regular behaviour and without a privileged direction, an isotropic and symmetric decomposition makes more sense, but then it must be non-orthogonal and redundant like the *à trous* one. Therefore, the following four wavelet transforms have been considered :

**D.a** Multiresolution analysis with Mallat’s algorithm [12].

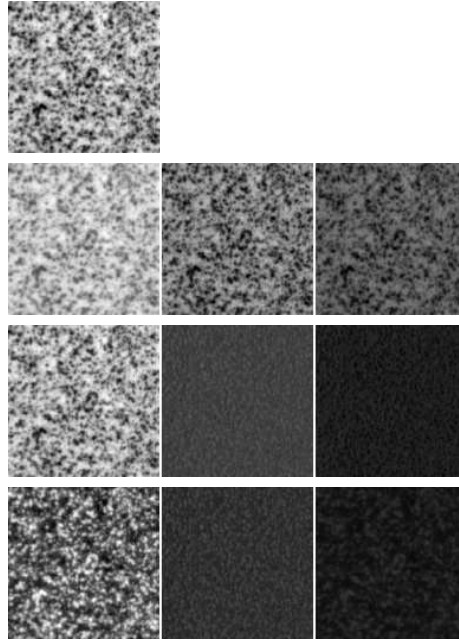


Figure 1: From top to bottom: grey level, RGB, general and specific Karhunen–Lòève transform. Images have been linearly contrast enhanced for the sake of visualization.

**D.b** *À trous* algorithm [19]. Opposite to D.a and D.c, it is a non-orthogonal and hence redundant transform.

**D.c** Wavelet packets transform [20] using a few fixed tree structure patterns.

In addition to the wavelet transform scheme, a suitable base must be selected. There are many families of bases, each having different properties like symmetry, orthogonality and regularity (related to the number of vanishing moments). This adds still a new dimension to the search space in which we want to minimize the classification error. In order to cut down the number of tests, we have fixed the base family for each scheme after a number of trials which we will not report here. Accordingly, Mallat’s multiresolution analysis and wavelet packets transform are performed with Daubechies orthogonal bases and the *à trous* decomposition uses a first order B-spline base. Figure 2 summarizes the decomposition scheme followed by each transform, and figure 3 shows an example over the R channel of a tile.

### 3.3 Feature extraction

Once the decomposition has been performed, we need to compute a vector of feature measures. In the literature of wavelet texture analysis two types of features are mostly used: energy and entropy. They are applied to the coefficients of the approximation and details at each level, though in some works cross energies (correlation signatures in our terminology) of details at different levels are also computed. Joint entropy [21] of couples of details or approximations at different levels and/or channels could also be computed and assessed.

In our application both had a similar performance. Actually, energy attained less than 1% improvement on the classification error over entropy, at least when features were re-

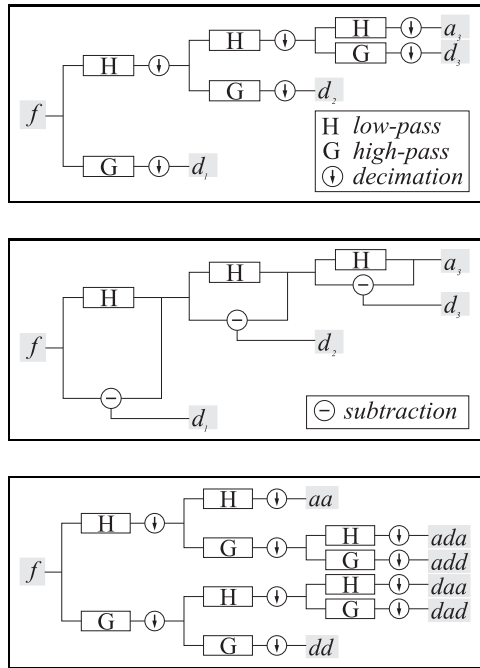


Figure 2: Decomposition schemes of a 1D signal  $f$  into detail and approximation coefficients for the Mallat's, *à trous* and a wavelet packet transforms.

stricted to be the energies of details and approximation for each channel (terms  $e_i^k$ , see below). For this reason we have restricted our study to energy related features.

The terms we will compute for the analysis stage are the energy and the cross correlation between levels and channels. We call all of them *correlation signatures* like in [9] :

$$c_{ij}^{kl} = \int d_i^k(u) d_j^l(u) du . \quad (5)$$

Note that  $c_{ij}^{kl}$  include the energy terms because  $e_i^k = c_{ii}^{kk}$ .

The number of features provided by the former three decompositions grows rapidly as the number of levels increases. For instance, a three-levels Mallat's wavelet transform of a RGB image gives rise to 30 images (1 approximation plus 29 detail images) on which 306 correlation signatures of images of the same size are possible. In a well devised supervised classifier, when the number of discriminant features increases, the performance is enhanced. However, if this number is too large with regard the size of the training set, the classifier just learns to succeed over this set but it is not able to generalize. Therefore, we should keep small the length of the feature vector, namely, the number of signature terms. For this reason, we propose to test the following choices, illustrated in figure 4 :

- F.a** Compute only the energy terms:  $c_{ii}^{kk} \forall i, \forall k$ . This is the most frequent choice in the literature.
- F.b** Calculate all correlation signatures between levels but only within the same channel :  $c_{ij}^{kk} \forall i, j, \forall k$ .
- F.c** Calculate all correlation signatures between channels but only within the same level :  $c_{ii}^{kl} \forall i, \forall k, l$ . This is the approach taken in [9].

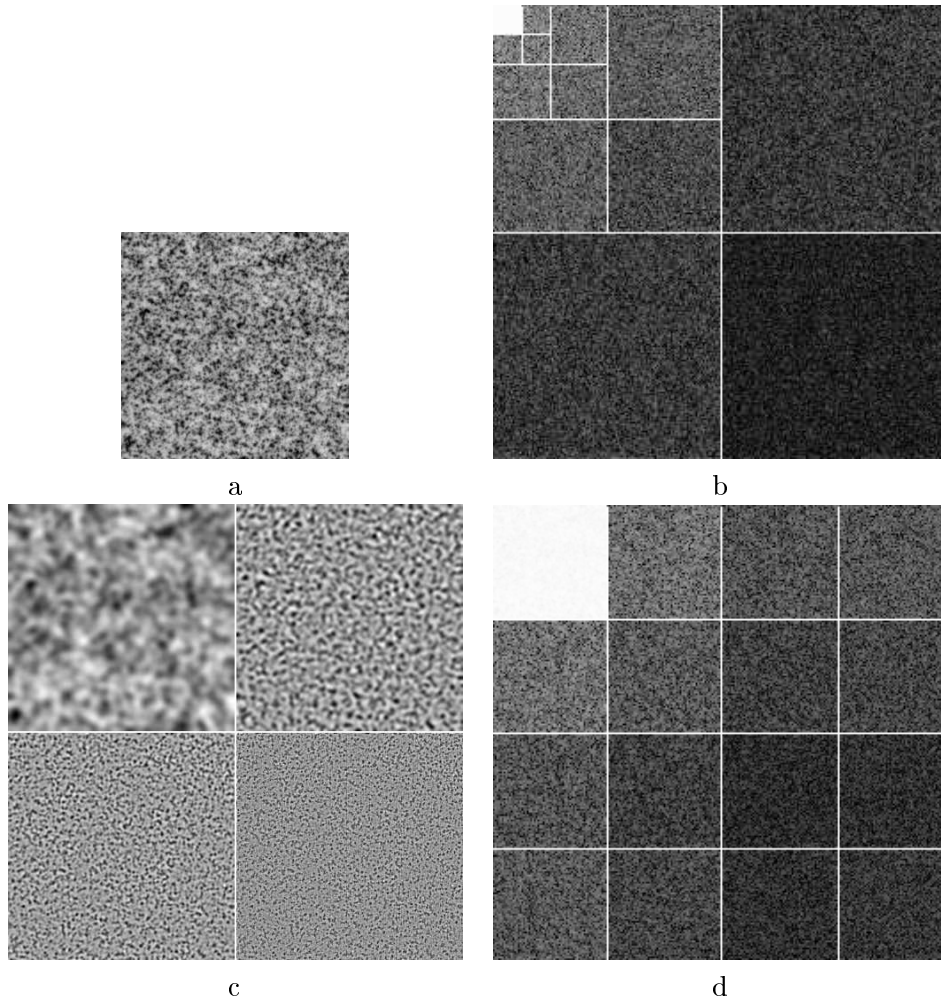


Figure 3: (a)  $256 \times 256$  portion of a tile red channel. (b) Zoom over a fourth level of Mallat's wavelet transform. A logarithm transformation has been applied for the sake of visualization. (c) Approximation and detail levels of the à trous decomposition. All of them have been contrast maximized separately. (d) Zoom of one of the possible two level wavelet packet decompositions (leafs of the tree at the second level).

**F.d** Collect all possible correlation signatures between channels and levels :  $c_{ij}^{kl} \forall i, j, \forall k, l$ .  
 In order to select relevant features, we take into account the former observation of section 2 which related some correlation signatures with the inhibition energies of the opponent color model.

Mallat's and wavelet packets transforms in which the decomposition tree has different levels can not be combined with options F.b end F.d. The reason is that coefficient images at different detail levels have different size due to decimation, thus being not possible to cross-correlate them.



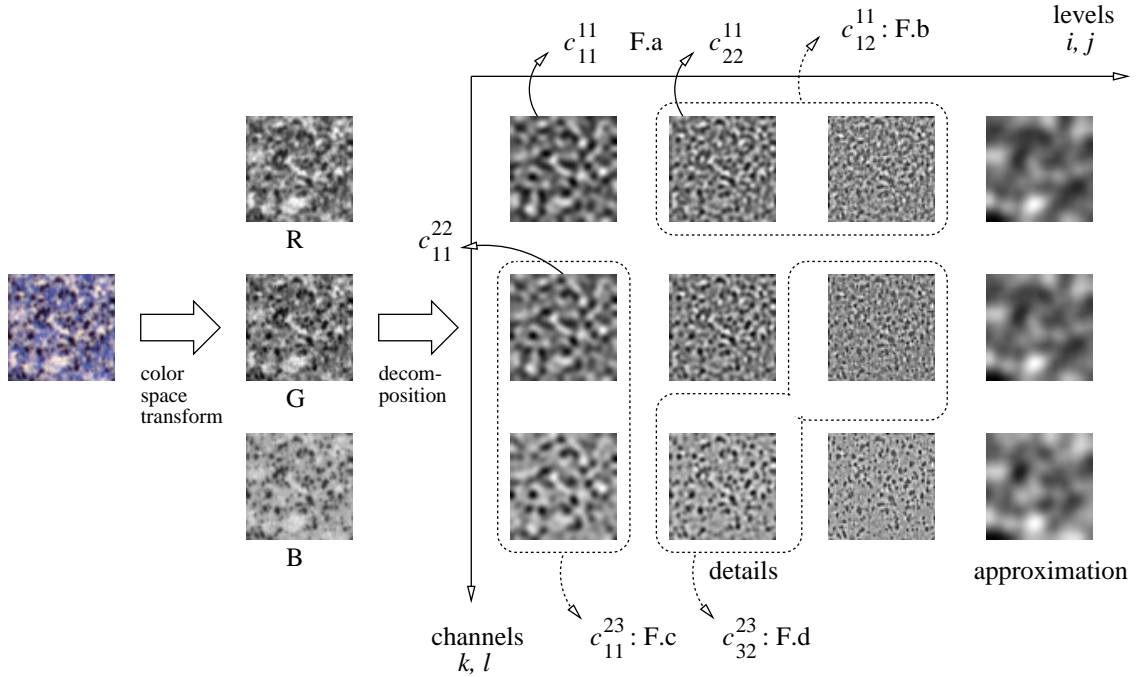


Figure 4: Features are selected among four types of correlation signatures.

### 3.4 Models

As we stated before, our approach to color texture classification is to first select a suitable space for color representation, a multiresolution decomposition scheme of the image represented in this space, and finally a set of discriminant features derived from this decomposition. However, it does not make sense to try every possible combination of choices for the three former items. Instead, we must select them according to an **image model** which explains how texture is related or mixed with color. We propose the following three models:

**M.a** : Images resulting from the addition of a grey level texture plus a uniform background color. Thus, only energy terms **F.a** from approximation and detail coefficients at different levels and **F.b** make sense. Furthermore, as this model in fact assumes a same texture for each channel, the former features must be computed just over one of the channels or the intensity image (mean of R, G and B).

**M.b** : Now, we assume that each channel contributes with a different texture to the final visual aspect of the image. But we further suppose that these textures are statically independent. Therefore, only **F.a** and **F.b**, this time over each channel, are candidates to be discriminant features with regard to a classification task. This is the model used in [15] for texture synthesis.

**M.c** : Conversely to M.b, we suppose now that textures along each channel are dependent, and in particular linearly dependent. Thus, besides **F.a** and **F.b**, correlation signatures between approximation or detail coefficients of different levels and channels, **F.c** and **F.d**, must be taken into account as potentially discriminant features.

### 3.5 Classification method

The classification method is a non-parametric discriminant analysis. In order to classify new samples we need a set of prototypes representing each possible class. Afterwards, the distance between the sample and each class can be calculated and the most similar class assigned. Given that classes are not known a priori, we need some method to learn the prototypes from a set of samples.

One of the methods that fits our constraints is that of Fisher discriminant functions, because, without any a priori knowledge of data, it is able to select the best representation maximizing the ratio between the inter-class covariance and the intra-class covariance [22]. A linear transform  $W$ , is applied to the feature vector  $\mathbf{x}$  of a particular image obtaining a new representation,  $\mathbf{y} = W^t \mathbf{x}$ , in a space where the discriminant capacity is maximized

From an image of a tile, we extract its feature vector,  $\mathbf{x}$ , and we assign it to class  $j$  if

$$|W^t \mathbf{x} - W^t \mu_j| < |W^t \mathbf{x} - W^t \mu_i| \quad \forall i \neq j ,$$

where  $\mu_i$  are the prototypes of the classes. Further details on the classifier can be found in [18]

## 4 Sorting of ceramic tiles

### 4.1 The problem

Tile manufacturing needs of pigments and clay which are mixed, melted, sprayed on to the tile substratum, and finally baked. Unavoidable variations in the pigments color, temperature, humidity and pressure conditions provoke subtle variations of the tile aspect when tiles are placed on the floor, one next to other. These visual changes are due to small differences in color and texture, and are seen as defects by customers. A system was thus needed to automatically sort tiles from a given model into perceptually homogeneous classes. At present, several trained workers at the end of the production line sort the tiles into perceptually homogeneous stacks. In each production line only a model of tiles is produced. Thus, classification must be done among classes of each model and not among models. As it is a tedious, time-consuming and subjective task, an automated system is needed.

We have built a system prototype to acquire and analyze images from tiles. Tile images are acquired with a three CCD digital line scan camera which yields 10 bits per channel. This allows us to distinguish color details invisible to the human eye, though a very stable lighting is required. We have designed a line light system which integrates several halogen sources and optical fiber light guides. In addition, we adapt the spectral content of the light to the camera CCD sensitivity by placing a set of color filters in front of the lens. Tiles move on a conveyor-belt with controlled speed under the linear camera, and this allows us to adjust the vertical resolution of the images to be the same as in the horizontal direction. The horizontal resolution is 5 pixels/mm, and it is given by the camera height above the conveyor belt and the lens fixed focal length.

Tile samples used in this study have been drawn from three different models which we will refer as A, B and C, (figure 5). Each model has a dominant color (A brown, B green and C blue) and spans into eight classes, according to the trained workers criteria. Thus,

each particular tile of a model must be labelled as belonging to a certain class. We have been provided a stack of 15 tiles per model, and for each one we have captured an image of the upper and lower half at a resolution of  $512 \times 512$  pixels.

## 4.2 Test images

We have been provided 15 tiles per class, for each of the 8 classes of the three models, summing up 360 tiles. From each class of each model, he have reserved 5 tiles for learning and used the 10 remaining tiles for testing. We have captured two  $512 \times 512$  RGB images for each tile, corresponding to the middle part of the upper and lower half. Thus, one test run means to find out which is the class of 480 images, given their models. In order to summarize the results, we will focus just on two figures : the total percentage of success and the percentage per model.

## 4.3 Simple features

Before comparing different schemes of multiresolution decomposition, space colors and types of features, we have first assessed the performance of two simple measures which do not involve any decomposition at all: the energy and the combination of mean and variance for each channel. They should provide a reference against which all subsequent more complex methods were to be compared, in order to better assess their real performance. Quite surprisingly, they reached high classification rates (table 1), despite of the fact that the color and textural content of classes are visually hard to discern, as shown in figure 5. However, they were not yet sufficient.

Feature	Model			Total
	A	B	C	
energy	73.1%	71.9%	75.0%	73.8%
mean and variance	94.4%	79.4%	87.5%	87.1%

Table 1: Percentage of right classification for simple measures. 1% of total corresponds approximately to five images and 0.6% within a model to one image.

A further test was carried out in order to make sure that our planned tests were worth the effort. We normalized the images to have all of them zero mean and variance 1.0 and next we decomposed each channel with the *à trous* algorithm, with 3 levels and first order B-Spline as base, and with the Mallat's scheme with one level and Daubechies 4 as base. Results reported on table 2 showed that these two chosen decomposition methods already got better results, though they were not conclusive.

Scheme	Model			Total
	A	B	C	
<i>à trous</i>	91.9%	82.5%	74.4%	82.9%
Mallat	91.9%	81.9%	83.8%	85.9%

Table 2: Percentage of right classification for images with zero mean and variance 1.

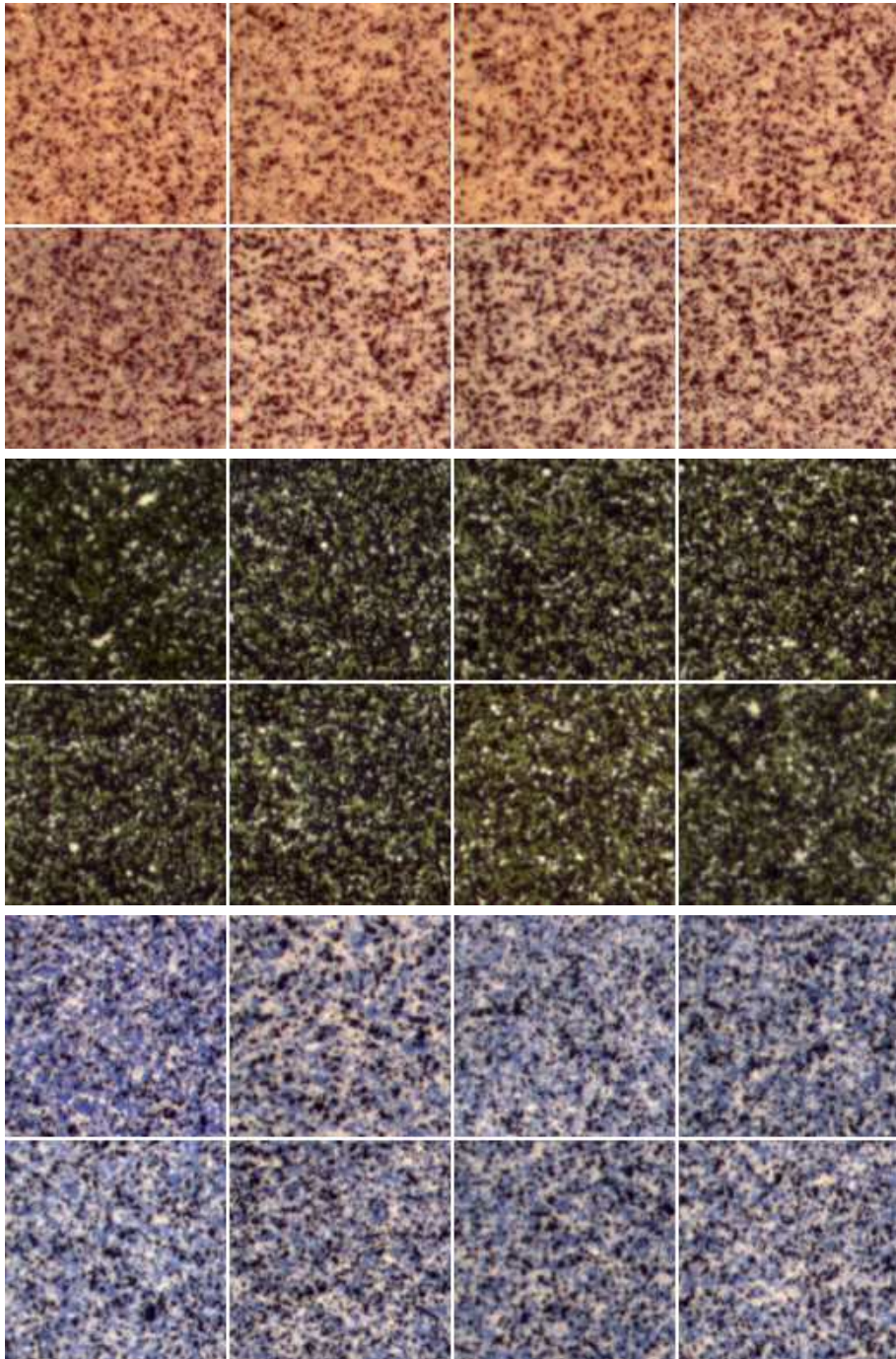


Figure 5: One sample of each class for tile models A, B and C.

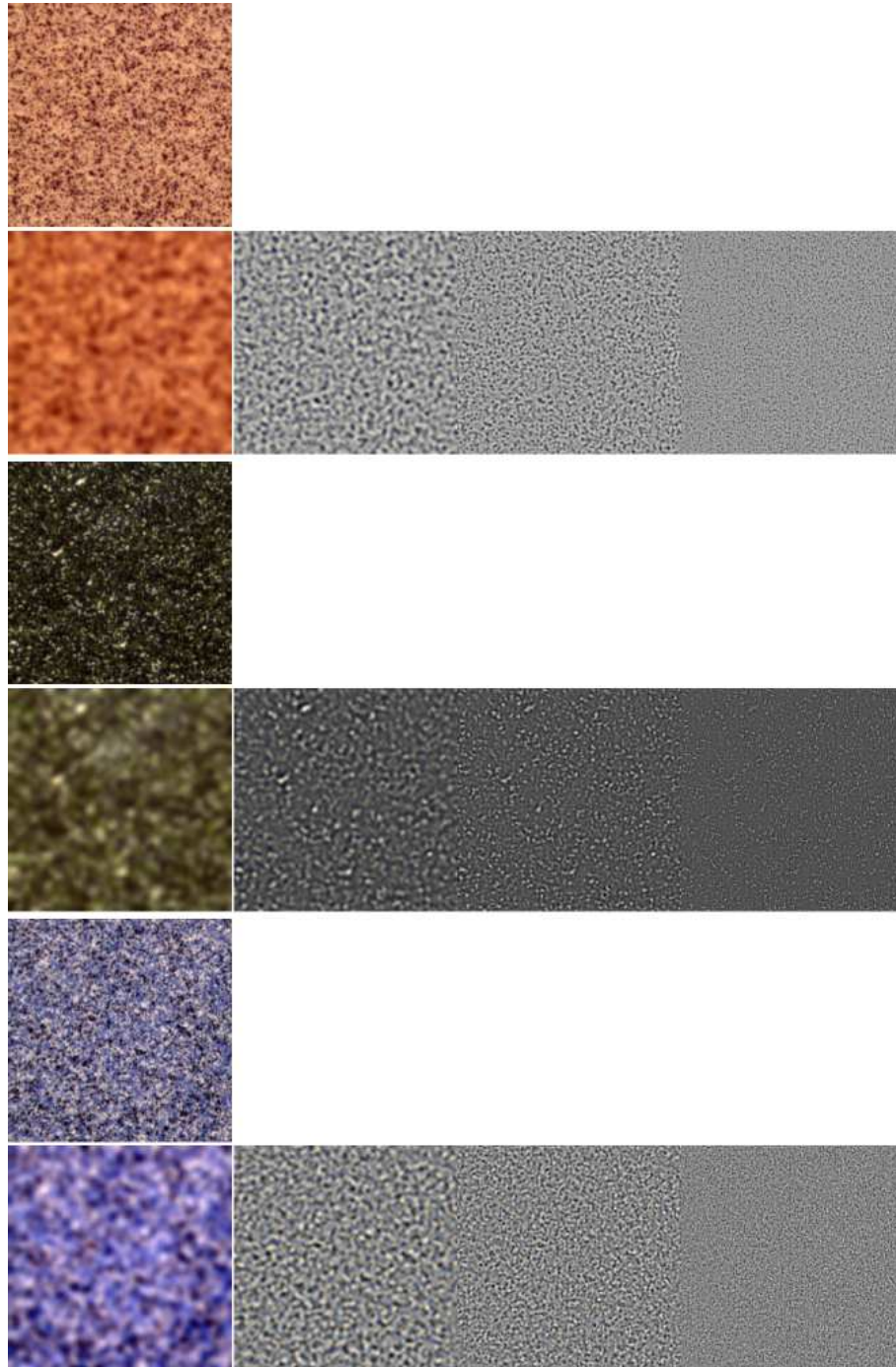


Figure 6: Three level *à trous* decomposition for one tile of each model. First row: a portion of  $256 \times 256$  pixels of the image. Second row: approximation and details at three levels.

## 4.4 Results

A huge number of tests are possible because the type of features, the decomposition scheme and the color space must be selected. Let's fix the type of feature to the simplest one, F.a, in order to explore the other factors. With regard to the second item, a number of base families are possible, and within each family, there are yet many possible selectable bases and decomposition levels. We have limited the number of tests by constraining the family of each scheme and the number of levels as listed in table 3, and then looking for the best particular base.

Actually, other families were tried like symlets and coiflets but they attained similar results to those selected. Table 4 shows the worst and best case for each scheme among all combinations of base and number of levels of table 3. Features are just the vectors of energies of the approximation and details of each channel, even for model A. Now, if we observe that the *à trous* scheme achieves the best global result (91.9 %), even though it is closely followed by other two schemes at a distance of 1.3%. Actually, this means just 7 more images wrongly classified over a total of 480. In order to prune the search for the best method, we will stick to this scheme from now on.

The next issue is to choose the classification features. Now, we are going to take advantage of the former three generic models of color textures. Images of tile model A follow quite closely M.a conditions. This can be seen in figure 6, which shows the *à trous* decomposition of three levels with first order B-Spline for a tile of each model. This redundant decomposition has the remarkable property of representing an image as the pointwise sum of the approximation and the detail images at different levels. We can appreciate how for the first tile, the approximation is a rather uniform ochre background to which mostly monochrome details must be superimposed in order to recover the original image. These detail images alone would give rise to almost all of the textural component of the original image if we had decomposed two or three more levels.

Tile models B and C are closer to M.b or M.c than to M.a, as can be deduced from their slightly more colored detail images and less uniform approximation. This can be more clearly appreciated in the image of model C. Therefore, all types of features F.a to F.d are envisageable. On one hand, however, it is not possible to compute correlation signatures between images of different size, as would be the cases of details/approximation images at different levels in the Mallat's multiresolution analysis and wavelet packets. On the other hand, F.d is the set of all possible correlation signatures between channels and levels, which, for three channels and a modest number of levels, means a huge number of features, outnumbering the images of the training set. For these two reasons, we have decided to discard features of type F.d and to take into account just F.a, F.b and F.c for models B and C. This will improve the best scores of table 4, whose features were limited to F.a for the three tile models. Best results are obtained for 3 levels and first order B-Spline, as illustrates table 5.

All previous tests were performed over the RGB representation. The last step was to check whether other spaces would further reduce the classification error of the *à trous*, 3 levels, first order B-Spline and most suitable features for each tile model. Thus, the two KL transforms of section 3.1 , generic and specific, were applied before decomposing. Results, however, did not improve significantly, see table 5.

Scheme	family of bases	bases	levels
<i>à trous</i>	B-Spline	0-th, 1-st, 2nd order	1 – 7
Mallat	Daubechies	D2 – D20	1 – 7
Wavelet packets	Daubechies	D2 – D20	1 – 2

Table 3: The three different decomposition schemes used with family bases and number of levels studied.

Scheme	Worst/Best global results						
	levels	base	number of features	A	B	C	global
<i>à trous</i>	7	B1	24	92.5%	80.6%	85.6%	86.2%
	3	B1	12	95.6%	84.4%	95.6%	91.9%
Mallat multiresolution analysis	5	D8	48	95.6%	76.9%	82.5%	85.0%
	2	D12	21	96.3%	83.1%	92.5%	90.6%
Wavelet packets	2 <sup>(1)</sup>	D2	48	95.0%	71.9%	82.5%	83.1%
	2 <sup>(2)</sup>	D6	60	93.1%	75.0%	90.6%	86.2%

Table 4: Results of the decomposition schemes tests. Features are energies of approximation and detail coefficients (F.a). B1: first order B-Spline, D: Daubechies. <sup>(1)</sup> only leaves of the wavelet packet tree are taken into account, <sup>(2)</sup> all tree nodes.

Features	A	B	C
F.a, F.b	<b>98.1%</b> <sup>(1)</sup>	84.4%	83.1%
F.a, F.b, F.c	95.6%	<b>87.5%</b> <sup>(2)</sup>	<b>95.0%</b> <sup>(2)</sup>
C. space	A	B	C
generic KL	98.1%	88.1%	93.3%
specific KL	98.1%	86.9%	95.0%

Table 5: Results of correct classification with different set of features and the same features with a color space transform applied to data. <sup>(1)</sup> only over channel R, having these only 9 features. <sup>(2)</sup> 24 features.

## 5 Paint recognition

The second problem we have addressed is the reverse engineering of metalized paints. They are a mix of one base and several effect pigments. The first one provides a background color whereas effect pigments (usually no more than three) produce changes in color and reflexion depending on the viewing angle. The goal is to find out a combination of base and effect pigments that best matches a given sample part, even though its pigments are different from those available. To our knowledge, this is still an open problem for the paint industry due to its complexity. We believe that it can be solved by combining the outcomes of two kinds of comparisons : the spectral responses of the sample under different lighting and viewing angles, and the microscopy images showing the pigments texture, both with regard the sample and the database of pigments. As this is an on going project, we will only report on the second part, which is again a texture recognition problem. However, it has an interest on its own, because there is a widespread application or color texture recognition, car refinishing.

Test images have been taken from a Ford paints card, which was readily available. A set of 14 samples (target classes) were selected, all of them appearing as different greyish colors. For each one, five images  $768 \times 576$  of non-overlapping fields were acquired with a Zeiss Olympus microscope, a 3 CCDs Sony camera and a  $\times 10$  magnification lens. Figure 5 shows an image of each class. Each one was divided into six disjoint  $256 \times 256$  images in order to increase the number of samples. Hence, we had 30 images per class, 12 of them for training and the rest for testing.

Which features should we use ? Again, this depends on the assumed image model. Some *à trous* decompositions quickly show that texture is uncorrelated to color, that is, images can be thought as the addition of a background color to a grey level texture, which is our M.a model. Therefore, only F.a features will be computed. In addition, when we examine figure 5 we realize that colors are similar and, besides texture, the main difference among classes is their contrast and brightness. Though all images were taken under constant lighting conditions, we want the classifier to be independent of it, that is, to rely only on the particles texture. For this reason, we have computed the intensity of each color image and normalized it to zero mean and unit variance.



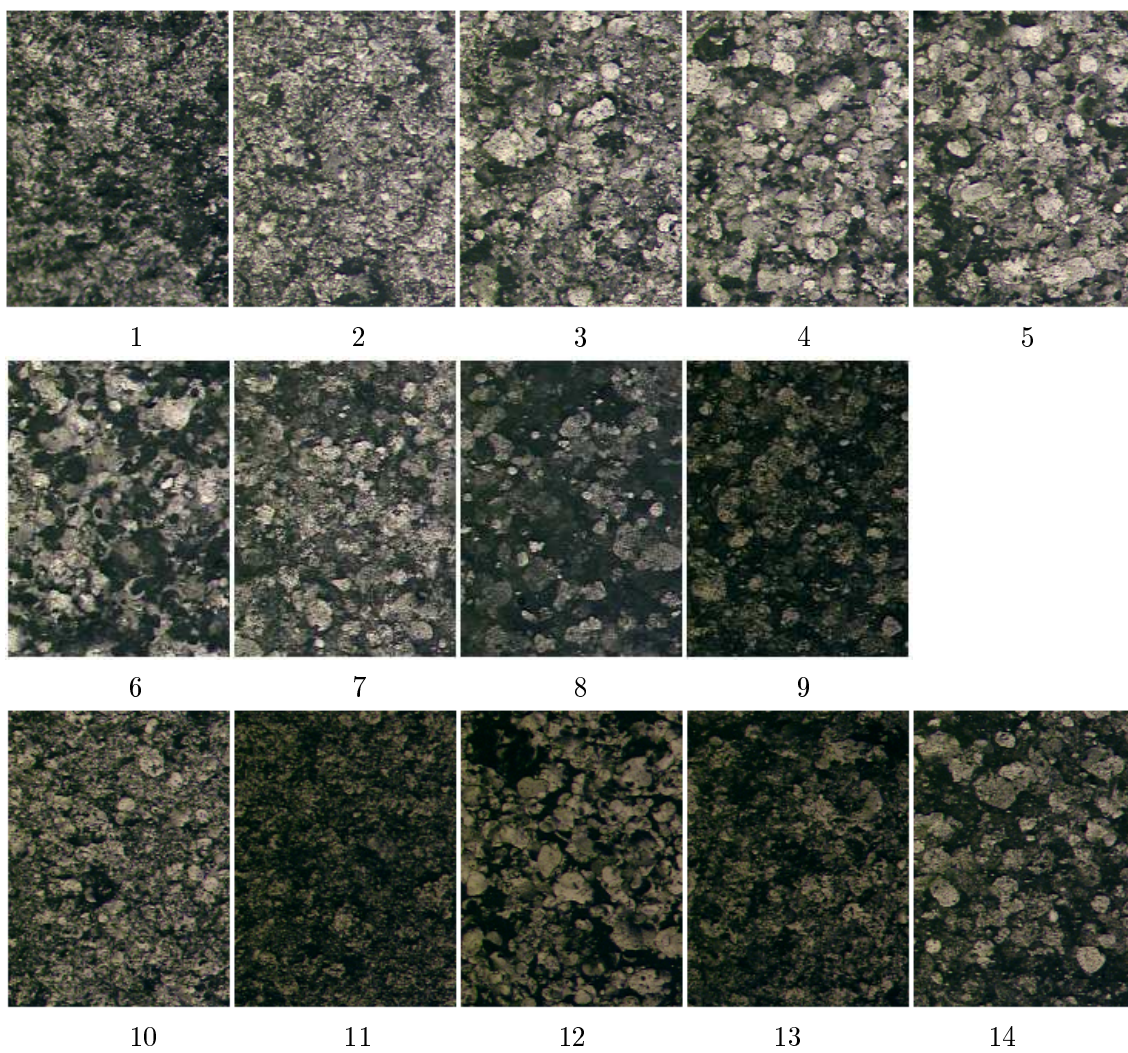


Figure 7: One sample image per class.

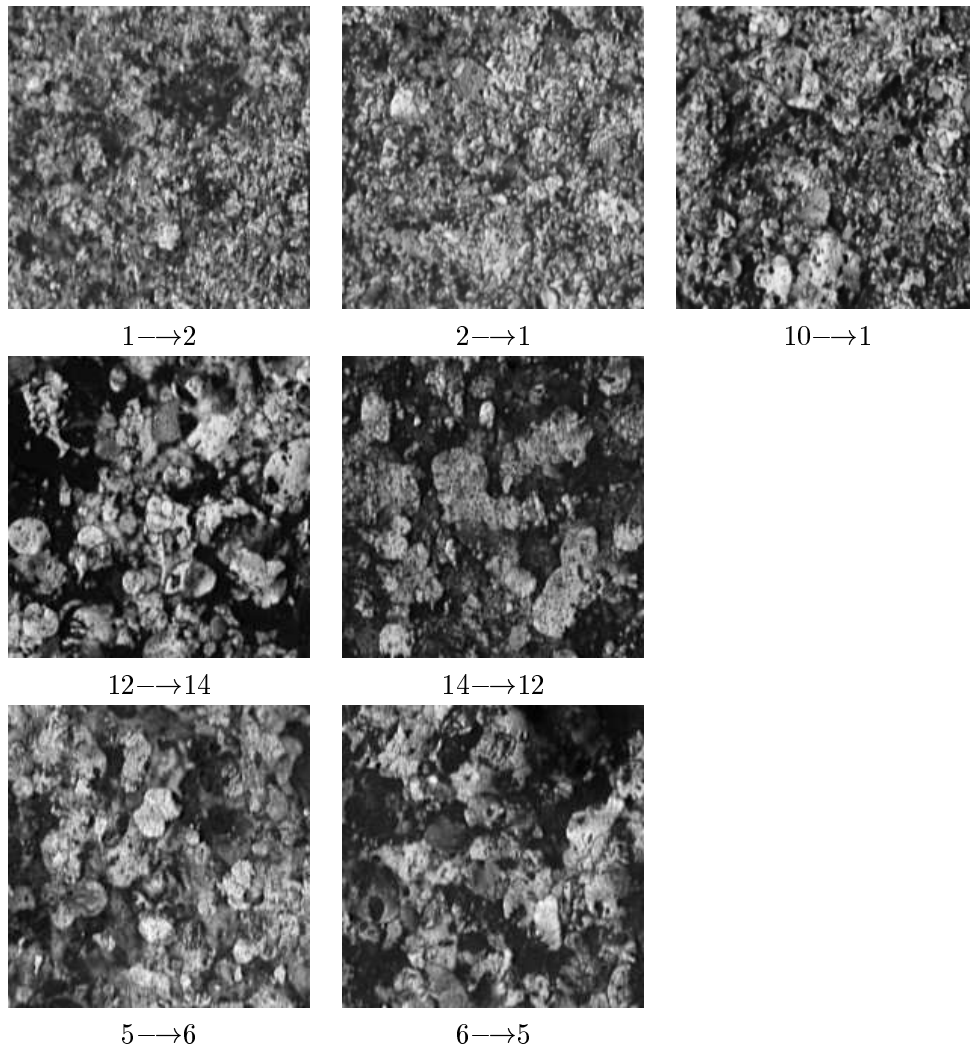


Figure 8: Examples of confusion. Numbers below are the actual and assigned class.

The decomposition scheme, number of levels and base are those most successful in the former application, that is, *à trous*, 3 and B1 respectively. Table 6 shows the classification results. Whats more remarkable is that a high recognition rate is achieved with only four parameters (energies of details at three levels and approximation), given the high visual similarity of the textures. In addition, errors happen when classes are harder to discern, even by a human observer (figure 8).

Actual class	Assigned	2nd most similar	3rd	4th
1	2 (33.3%)	1 (27.8%)	10 (22.2%)	7 (16.7%)
2	2 (83.3%)	1 (11.1%)	10 (5.6%)	
3	3 (77.8%)	4 (16.7%)	2 (5.6%)	
4	4 (72.2%)	5 (27.8%)		
5	5 (66.7%)	6 (33.3%)		
6	6 (88.9%)	4 (11.1%)		
7	7 (100%)			
8	8 (94.4%)	14 (5.6%)		
9	9 (94.4%)	13 (5.6%)		
10	10 (83.3%)	1 (11.1%)	2 (5.6%)	
11	11 (100%)			
12	12 (100%)			
13	13 (94.4%)	9 (5.6%)		
14	14 (66.7%)	12 (33.3%)		

Table 6: Paint recognition results over 18 images per class.  $n$  ( $x\%$ ) means that  $x\%$  of the 18 tested images of that class were assigned to class  $n$ . One image over 18 is 5.6%

## 6 Conclusions

We have addressed a problem of color texture classification through multiresolution decomposition techniques. Our aim was to find an optimal combination of color representation, decomposition scheme plus base and number of levels, and discriminant features. Through the search strategy described in the results section we have arrived to the conclusion that, for the specific images of our study, only the decomposition scheme substantially influences the final result. The family of bases and the specific base within it do not play a significant role, as all tests varying them get similar percentages of success. Nevertheless, we have been able to tune them in order to slightly (1% or 2%) reduce the classification error. Likewise, color spaces do not achieve a noticeable improvement.

From a more theoretical point of view, we have proposed three image models according to which several types of spatio-chromatic features are or not meaningful. These models refer to how texture is embedded into color and how texture in each channel relates to texture of the other ones. In this way, given images following one of such models, we know that only certain features computed from the multiresolution decomposition should be taken into account. This idea has been supported by actual results showing that selecting the right features achieves the smallest classification error.

Future work will address the assessment of features of type F.d as well as other measures of dependence between the images of the decomposition. In particular, the mutual information measure as an extension to entropy is being examined.

## Acknowledgments

This work has been partially funded by grant TIC2000-1123 from the CICYT, and the *Consejería de Educación y Cultura de la Comunidad Autónoma de Madrid* through the tile manufacturer Alcalagres SA.

## References

- [1] Todd R. Reed and J.M. Hans Du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372, May 1993.
- [2] Brian A. Wandell. The synthesis and analysis of color images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(1), January 1987.
- [3] T. Caelli and D. Reye. On the classification of image regions by colour, texture and shape. *Pattern Recognition*, 26(4), 1993.
- [4] H. Greenspan C. Carson, S. Belongie and J. Malik. Region-based image querying. *CVPR'97 Workshop on Content-Based Access of Image and Video Libraries*, 1997.
- [5] D.K. Panjwani and G. Healey. Markov random field models for unsupervised segmentation of textured color images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(10), 1995.
- [6] G. Healey and L. Wang. Illumination-invariant recognition of texture in color images. *Journal of the Optical Society of America*, 12(9), 1995.
- [7] A. Gagalowicz, S. De Ma, and C. Tournier-Lasserve. Efficient models for color textures. *Eighth International Conference on Pattern Recognition*, pp. 412–414, 1986.
- [8] Graham D. Finlayson, Subho S. Chatterjee, and Brian V. Funt. “Color angular indexing.” *European Conference on Computer Vision*, pp. 16–27, 1996.
- [9] G. Van de Wauwer, S. Livens, P. Scheunders, D. Van Dyck. Wavelet correlation signatures for color texture characterization. *Pattern Recognition*, 32(3), 1999.
- [10] T. Randen, J.H. Husoy. Filtering for texture classification : a comparative study. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(4), 1999.
- [11] A. Jain and G. Healey. “AMultiscale Representation Including Opponent Color Features for Texture Recognition”. *IEEE Trans. on Image Processing*, 7(1), pp. 124–128, 1998.
- [12] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7), 1989.
- [13] M. Unser. Texture Clarification and Segmentation Using Wavelet Frames”. *IEEE Trans. on Image Processing*, 4(11), 1995.
- [14] J. S. De Bonet. Multiresolution sampling procedure for analysis and synthesis of texture images. *Computer Graphics. ACM SIGGRAPH*, 1997.
- [15] D.J. Heeger and J.R. Bergen. Pyramid-Based Texture Analysis/Synthesis. *Computer Graphics. ACM SIGGRAPH*, pp. 229–238, 1995.

- [16] C.R. Boukouvalas. Colour Shade Grading and its Applications to Visual Inspection. *PhD Thesis*, University of Surrey, 1996.
- [17] J.A. Peñaranda, L. Briones, J. Florez. Colour machine vision system for process control in ceramics industry. *in Proc. of SPIE: New image Processing Techniques and Applications: Algorithms, Methods and Components*. Vol. 3101, pp. 182–192, 1997.
- [18] R. Baldrich, M. Vanrell and J.J. Villanueva. Texture-colour features for tile classification. *EuroOpto Series: European Symposium on Industrial Laser and Inspection*, Germany, June 1999.
- [19] J. L. Starck and A. Bijaoui. Filtering and deconvolution by the wavelet transform. *Signal Processing*, 35, 195–211, 1994.
- [20] M. V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. A. K. Peters, 1994.
- [21] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [22] K.V. Mardia, J.T. Kent and J.M. Bibby. *Multivariate Analysis*. Academic Press, 1997.