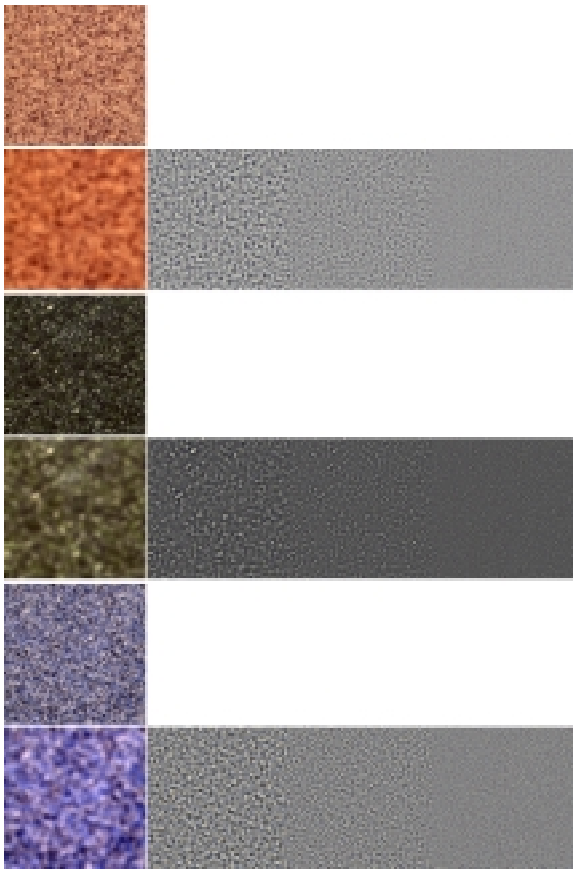

Figure 4.7: Three level *à trous* decomposition for one tile of each model. First row: a portion of 256×256 pixels of the image. Second row: approximation and details at three levels.

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 Fig. 4.7, 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. 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** can be envisaged. 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 4.4, 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 4.5.

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 K-L transforms of Section 4.3.1 , generic and specific, were applied before decomposing. Results, however, did not improve significantly, see Table 4.5.

Table 4.3: The three different decomposition schemes used with family bases and number of levels studied.

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

Table 4.4: Results of the decomposition schemes tests. The first row for each scheme correspond to the worst case and the second to the best. Features are energies of approximation and detail coefficients (F.a). B1: first order B-Spline, D: Daubechies. $^{(1)}$ only leaves of the wavelet packet tree are taken into account, $^{(2)}$ all tree nodes.

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

Table 4.5: Results of correct classification with different set of features and the same features with a color space transform applied to data. (^{R} only de red channel)

Features	Color space	#features	A	B	C
F.a	**C.a**	4	97.5%	77.5%	85.0%
F.b	**C.a**^{R}	7	**98.1%**	84.4%	83.1%
F.c	**C.b**	24	95.6%	**87.5%**	**95.0%**

Color space	Features	#features	A	B	C
Ohta (generic K-L) (**C.c**)	**F.c**	24	98.1%	88.1%	93.3%
Specific K-L (**C.d**)	**F.c**	24	98.1%	86.9%	95.0%

4.5 Paint recognition

4.5.1 The problem

The second problem we have addressed is the reverse engineering of metalized paints. In the paint manufacturing industry the specification of certain paint is done starting from a piece sample with the required paint. In most cases paint components are unknown and must be guessed from the study of the sample trough the microscopy and the colorimeter and also with the help of the experience. In these paints there are a mix of one or more base paints and several effect pigments. The first one provides a background color whereas effect pigments (usually no more than three) produce changes in color and reflection 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 to 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.

Firstly, we broach this problem as a classification problem like in the previous case. We have a set of images of different classes and we want that if we show a new image to the system it relates the sample to the closest class in the set. This is a first attempt to solve the problem. The global problem consists in finding what are the effect pigments, base pigments and its concentrations of an unknown sample with the help of the microscopy and a colorimeter. Nowadays, the problem is solved by skilled professionals that achieve a good matching between the problem sample and the final solution by successive approximation with the microscope, colorimeter and experience as only tools. If this previous classification gives good results will permit trying to solve the problem performing a huge database with the most common mixtures of effect and base pigments. Really, collecting all this possible mixtures is a colossal and nearly impossible task and it means that the problem must be started from a new point of view, trying to extrapolate information of the mixture composition from the single elements properties.

Which features should we use? Again, this depends on the assumed image model. In these preliminary proofs we start with gray paints Fig. 4.8 with no important color information. The study of some *à trous* decompositions with these samples quickly show that texture is uncorrelated to color, that is, images can be thought as the addition of a background color to a gray level texture, which is our **M.a** model. Therefore, only **F.a** features will be computed. In addition, when we examine Fig. 4.9 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. This decision is very restrictive because samples used have not an important color contents. Working with the whole problem, where color is a very important property, need the use of color information, and it means that models **M.b** and **M.c** must be explored.

4.5.2 Test images

Two different acquisition conditions have been used. Images were acquired with a Zeiss Olympus microscope, and a 3 CCD Sony camera. Firstly, we use a high magnification ($\times 500$) to observe the individual particles and secondly a low magnification ($\times 100$) to observe the texture globally. The usefulness of two resolutions is noted also in [23]. A first step in analyzing an effect pigment would be to observe it in bright field illumination at low magnification ($\times 200$). Here, a rough estimation about the effect can be made: pigment load, mica/aluminum ratio, and kind of mica. Then, a study in bright field and dark field at high magnification ($\times 400$ or $\times 500$) gives more detailed information about the pigments.

Test images have been taken from a Ford paints card, which was readily available (see Fig. 4.8). A set of 14 samples (target classes) were selected, all of them appearing as slightly different grayish colors. We try to avoid color and intensity information in these samples to relegate success on classification to the texture information.

For each one of these 14 paint samples, five images 768×576 of non-overlapping fields were acquired at $\times 100$ magnification. Figure 4.9 shows an image of each class. Subdividing each one of these images we obtain the sets used in classification. We planned three strategies form the case of few and big images to the cases of many and little images:

- A. Whole images, it means, five images per class: two for training and three for test.

- B. Each image was divided into six disjoint 256×256 subimages in order to increase the number of samples. Hence, we had 30 images per class, 12 of them for training and 18 for testing.

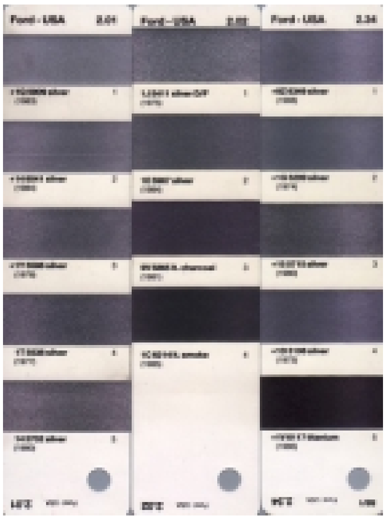

Figure 4.8: Fourteen paint classes used in the classification.

- C. Each image was divided into 24 disjoint 128×128 subimages in order to increase the number of samples. Hence, we had 120 images per class, 48 of them for training and 72 for testing.

After some preliminary results presented in Table 4.6 we discard set A because the good results obtained maybe reflect that 70 samples are few elements to start a classification, nevertheless, for a real application it is interesting to consider big images because they summarize better the texture. Set C was also discarded because they do not represent well the global texture image. The reduction of the image size reduce the capability to describe all the image and therefore the classification performance decrease, To show this behavior we perform a simple test limiting the number of samples to the minimum case (70 images: 28 for training and 42 for test). The classification results using mean and variance as features were: A (90.5%), B(78.6%), C(66.7%). We have to choose between a better performance and a sufficient class representation and, for this reason, we use the intermediate set B for an in-depth study.

For the high magnification setup (×500) we also acquire five images but, in this case, for eight of the fourteen classes. The problem that arises in this kind of images is that the whole field of view is wrongly focused. Some of the particles are in focus and another ones are badly focused. To solve this, we have acquired three images at different focusing depth. We focus the whole image using a method based on a wavelet decomposition that chooses those largest coefficients in the decomposition of each image and after that we rebuild the image. This procedure is also used in other field as in data fusion [48] explained in Section 2.2.8. Figure 4.10 shows how the images acquired at different depth have been merged to obtain a focused image. The set of focused images was further subdivided in order to achieve a sufficient number of samples for classification purposes.

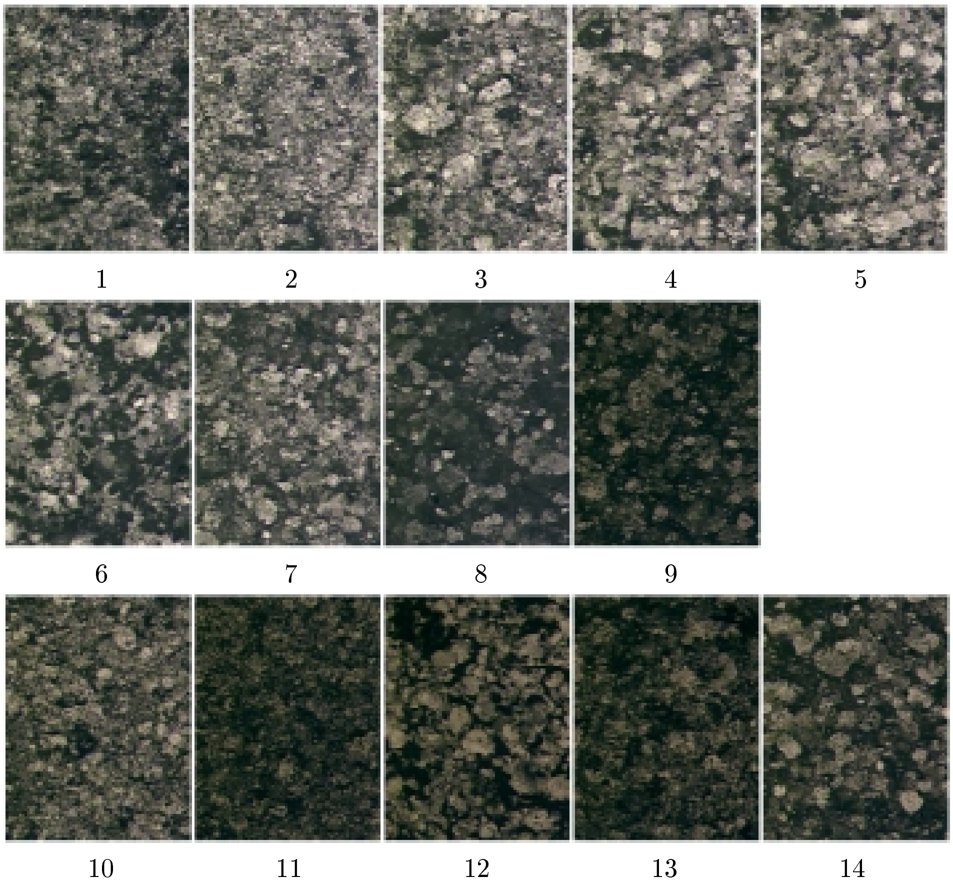

Figure 4.9: One sample image per class.

Table 4.6: Preliminary results for the three set. Now, classification tests were done with all the available samples for each set. The dash means that this proof has not been done.

Method	A	B	C
Mean and variance (m.v.)	90.5%	81.3%	69.1%
RGB, *à trous* 4 level, B1	90.5%	92.9%	—
RGB without m.v., *à trous* 4 level, B1	88.1%	87.3%	—
Grey, *à trous* 4 level, B1	95.2%	90.1%	74.6%
Grey without m.v., *à trous* 4 level, B1	85.7%	82.5%	65.0%

- D. Each image was divided into four non-overlapping subimages, that is, 160 images of 256×256 pixels.

4.5.3 Features

As in the previous case the methodology used to classify the samples starts with some simple features as mean and variance. With these preliminary results we realize that they are not intended for classification, although the good results obtained, due to the high dependency to the illumination. Next we explore simple features as energy on multiresolution decompositions to evaluate the best number of levels and decomposition scheme. These results are obtained for images without mean and variance, therefore reducing illumination problems, and we compare them with the previous ones. In this case the features express only the texture behavior and with this methodology we try to tune the best decomposition where the discrimination among classes is better. Finally we improve features adding cross terms in order to refine the classification. In this step better results are obtained due to two reason, in one hand the inclusion of cross terms increase the dimensionality and therefore in a large space classes a easier to discern reducing in this the generalization power [84]; on the other hand these new features add extra information useful for the classification.

4.5.4 Results

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 4.7 shows the classification results. What is 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 (Fig. 4.11).

For the ×500 magnification case, we do not achieve a similar classification ratio as the obtained with a lower magnification (×100). A reason to this behavior could be the fact that texture, as a global feature, is lost in this high magnification images, different images of the same class have very different aspect. This result is agree with the protocol we have presented and also pointed in [23]: firstly, we use a high magnification

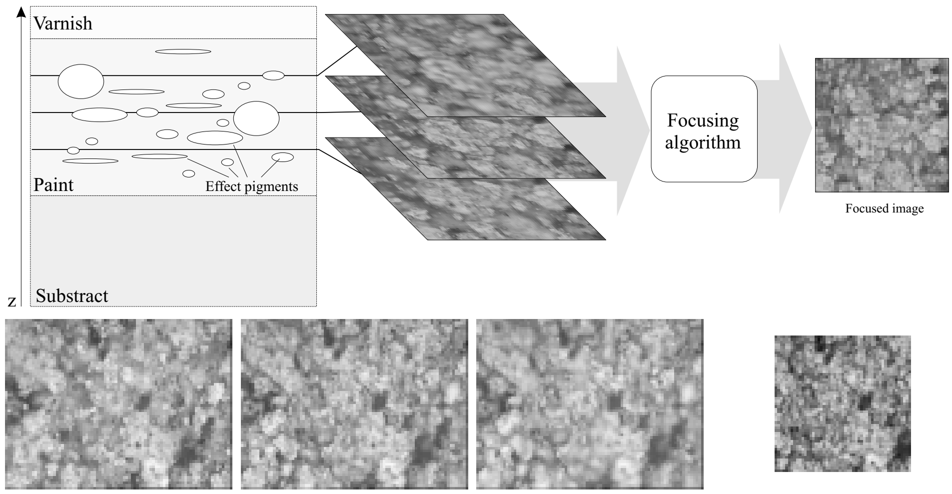

Figure 4.10: The focusing procedure selects the focused areas in each one of the input images taken at different depth. Bellow, a detail of a set of three images and its focused result image. To obtain this result we use a multiresolution analysis with a Daubechies6 as base.

($\times$500) to observe the individual particles and secondly a low magnification ($\times$100) to observe the texture globally. For the test we have done, the results for the high magnification case are with similar conditions to previous cases are 32.3% (cross-levels of *à trous* decomposition, 3 level, without color information and without mean and variance). An extended table of results is showed in Appendix A (Table A.7). This table shows how we can achieve until a 68.8% of correct classification if we relax some conditions as, (i) inclusion of color information, (ii) more levels in the decomposition. The increase of performance with the number of levels can be a clue to realize that magnification too high.

4.6 Marble recognition

Once, we have developed a protocol to classify texture, we also prove the classification of marble images used in Chapter 3. There are some handicaps that point to a failure in this results, on the one hand the few set of images (in the best case we have a couple of images for each quarry), on the second hand grains are large compared to the size of the images (a similar problem that previous $\times$500 magnification paint case). These problems arise due to fact that the problem was focused firstly as a segmentation problem and samples were taken with this goal. Nevertheless, the test were done and a slightly classification were obtained far from the previous results. We obtain results for 9 classes, 8 images per class (2 sample images that were cut each one in 4 disjoint subimages). The worst classification that we can obtain based on bad features (random) should be near of an 11.1% due to we have 8 classes. Simple features (mean and variance) give a classification ratio of 15.25% that is a very bad result. Adding

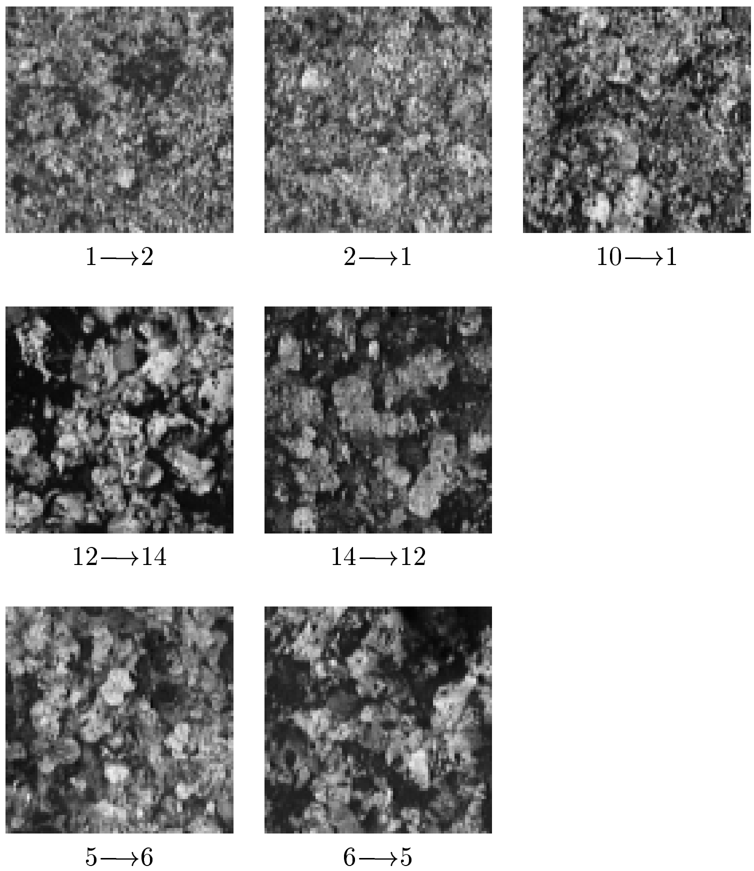

Figure 4.11: Examples of confusion. Numbers below are the actual and assigned class.

Table 4.7: 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%.

Class	Assigned	2nd	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%)		

a decomposition scheme in order to make texture as feature for classification give us a value of 36.1% that is nearly the double of the classification with simple features and threefold the worst case. The scheme used for this result was a four level *à trous* decomposition with out mean and variance, and B1 base. The classifier was the same but in this case 50% of samples were used to the learning stage and 50% to test. More results are presented in Appendix A (Table A.8). In this results we can observe how results improve as number of decomposition levels rise, it is due to the fact point before that grains in images are very big, and grains are the most important contribution to the texture features used to distinguish among classes. When the number of decomposition levels increase we can access to low frequencies and therefore to the information related with big structures in the image. Also if we do not perform a mean and variance normalization we can improve results but in a tricky way because we need in this case to control accurately the illumination and the acquisition in our system. Also we obtain better results if we use a little base in the decomposition in this way for fixed conditions if we reduce the filter we obtain an increase in the results. Taking into account all the things we explained before we can arrive to a 47.2% of classification.

4.7 Brodatz

Finally, in order to compare this scheme of texture analysis we use a standard image database of texture. Brodatz [13] is a compilation of different textures. This book is really a photo album with material useful for artist but that has been adopted by texture analysis community as a test-bed for the analysis of algorithms performance.

These images have been widely used in classification problems. Also, they have been assembled in a mosaic in order to perform and evaluate segmentation processes. A common problem in most part of the bibliography devoted to this topic is that the sets of images used in each experiment are different form one paper to another.

Another problem related to the Brodatz album that do not match in our problems is that most textures in Brodatz are very different form one image to another but there are several images that has similar appearance. Some of the album images do not present homogeneity that is a good property because we can measure the texture parameters in any point with similar results. Our problems are restricted to images that differs slightly form one image to another and also, images are homogeneous, it is to say, properties related to textures are nearly constant over the entire image. In this case as in the previous one the problem is not well adapted to our scheme but we want to quantify the classification rate.

We obtain over the net a Brodatz set with 111 images (see Fig. A.2). Each images is 640×640 pixels and we break each image in a 16 subimages of 160×160. Eight images were used as test and eight for learning. The total amount of images is 1776 subimages. The first attempt was to classify all the images with our scheme and without caution. In this case we obtain a classification rate of 23.5% for the mean and variance parameters and the best result for the studied cases was 66.05% with 7 parameters of a cross-level à trous decomposition with 3 levels with mean and variance, 47.65% for the same case with out mean and variance (see table A.9 for more results). If we try to explain these moderate results we must to analyze the images and find how there are some images that are very similar and some other images has inhomogeneities, it means that subimages from these images are not similar. In order obtain a fear results trying to distinguish among different textures we select 55 images (a half of the total, see Fig. A.1) where these problems has been avoided. Results obtained for Cross-levels, à trous decomposition with 3 levels and without mean and variance (47.65% before) was in this case 93.55% nearly the double percentage. These results are very high if we take into account that the total number of images to classify is 55. In the literature some times there are very good results closer to the 100% but for a limited number of images no greater than ten or twenty. If we reduce our set of images to a selected and reduced set of images the results we can obtain will be higher, near the 100% of correct classification.

4.8 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, in the problem we have addressed 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 spatial and 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 will be examined.