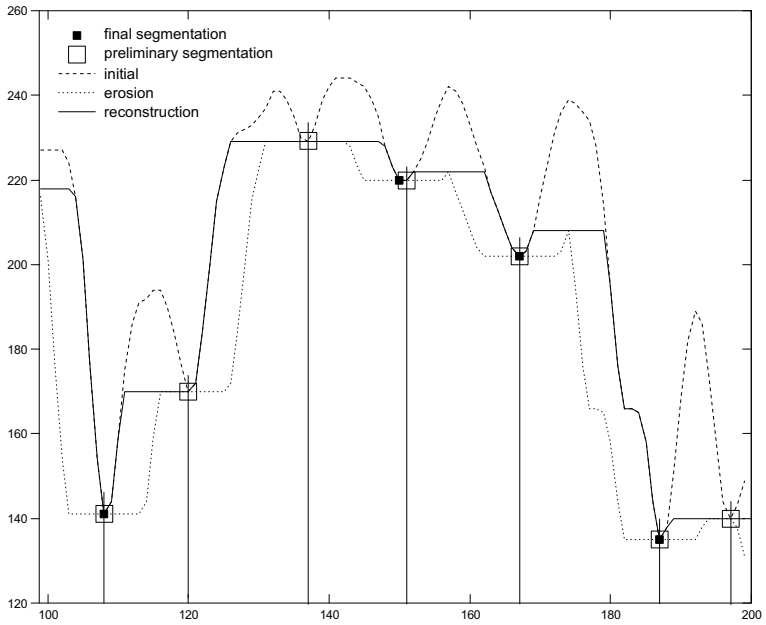

Figure 3.12: Erosion-reconstruction filtering example for a 1D signal. Erosion step removes the meaningless minima and reconstruction step restores previous result to its maximum value in a defined neighborhood. The removed minima do not contribute to the watercourses and nor the segmentation.

valleys and the filtering of the interior of grains. To this end, we design a new wavelet filtering scheme to emphasize ridges or valleys.

As we see in Sec. 2.2.3, the continuous wavelet transform [67, 20, 1] decomposes a function $f(x)$ at several scales. This decomposition is performed by convolving the function with the dilations and translations of a special function ψ named the mother wavelet:

$$w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x)\, \psi^*\left(\frac{x-b}{a}\right) dx \quad a \in \mathbb{R}^+, b \in \mathbb{R} \ .$$

The function $f(x)$ can then be recovered as

$$f(x) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} w(a,b)\, \psi\left(\frac{x-b}{a}\right) \frac{da\, db}{a^2} \ ,$$

where C_ψ is a constant depending only on ψ.

The discrete version of this transform results from the sampling of parameter space (a,b), see also Sec. 2.2.3. One of the most well-known discrete wavelet transform algorithms is the multiresolution analysis [62] (see Sec. 2.2.4). In it, a sequence of embedded function subspaces $\ldots \subset V_i \subset V_{i-1} \subset \ldots \subset V_0$ is spanned by the dilations $\phi(2^{-i}x)$ of a function $\phi(x)$ named the scaling function, such that if $f(x) \in V_{i-1}$ then $f(x/2) \in V_i$. Therefore, the sampled parameters are $a = 2^{-i}, b = k2^{-i}$. A continuous

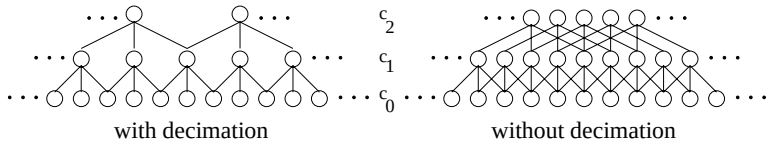

Figure 3.13: Calculation of the coefficients with and without decimation.

signal $f(x)$ is projected on each subspace by means of the scalar product with the scaling function dilated and translated to integer positions. We can see the initial discrete signal $c_0(k)$ as the projection of $f(x)$ on V_0,

$$c_0(k) =< f(x), \phi(x-k) >= \int_{-\infty}^{\infty} f(x)\,\phi(x-k)\,dx \ . \tag{3.5}$$

The projection on a subspace V_i,

$$c_i(k) =< f(x), \frac{1}{2^i}\,\phi\Big(\frac{x}{2^i} - k\Big) > \ , \tag{3.6}$$

is then an approximation of c_0 at scale or resolution i. The greater is i, the coarser the approximation will be. For scaling functions fulfilling certain conditions, the difference between two successive approximations c_{i-1} and c_i is a discrete signal belonging to a new space $W_i \subset V_{i-1}$ which is the orthogonal complement of V_i in V_{i-1}, that is, $V_{i-1} = V_i \oplus W_i$. These subspaces W_i are similarly spanned by the dilations and translations of a mother wavelet function ψ. Note that Eq. (3.6) includes a decimation: the number of coefficients at level i is half the number of coefficients at level $i-1$. Thus, the full decomposition of a signal c_0 of n samples is set of $n-1$ wavelet coefficients plus one approximation coefficient. Because of the orthogonality property, there is no redundancy among them. Thus, it is very suitable for compression purposes. However, this decomposition scheme lacks an important property for image analysis, namely spatial invariance (translation invariance).

Therefore, we have applied a decomposition algorithm, named *à trous*, which does not perform decimation and consequently produces a redundant representation [89], see Sec. 2.2.5. Figure 3.13 shows how a general algorithm uses the previous levels in the decomposition to calculate the new coefficients. At this moment one can decide among decimation with the advantage of reducing data or a non-decimation scheme. As we see in the figure the decimation step as well as any pyramid scheme do not assure the conservation of the initial volume of data, it can be achieved by a special orthogonal scheme [63, 90]. Also in the figure, we see that in the scheme without the decimation process a new level in the decomposition implies that the total amount of data increases in a volume equal to the initial data, but having the property of translation invariance.

In pattern analysis and recognition most of the time it is necessary to construct signal representations that are translation invariant. This property guarantee that any measure over the decomposition coefficients are not influenced by their positions. If we have a particular element in a specific position that we want to analyze, in our case a ridge or valley, and the same element appears a distance apart, their coefficients

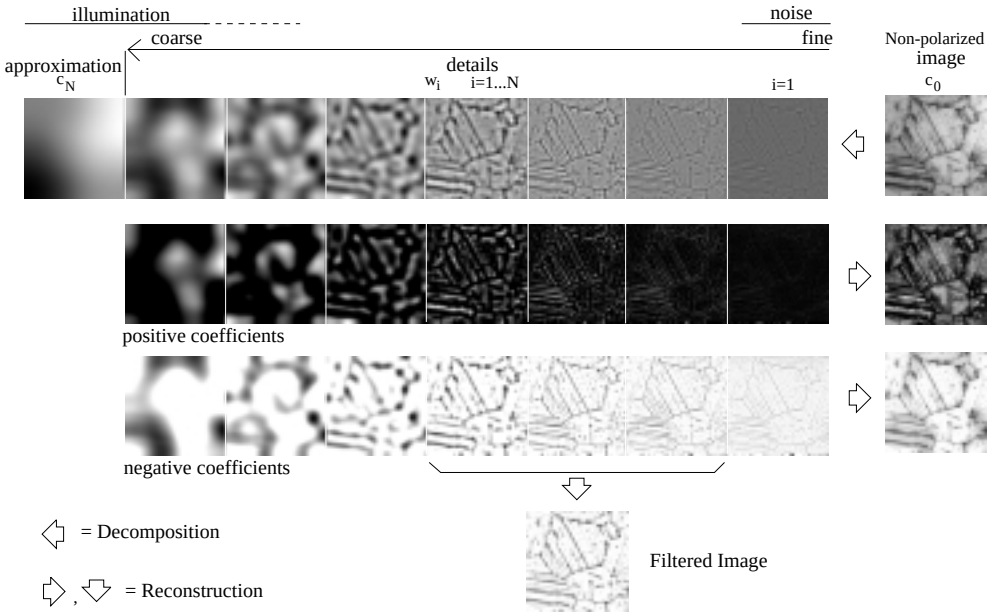

Figure 3.14: Detail of a marble image filtered with the process that we propose based on the selection of some coefficients in a wavelet decomposition.

at their relative positions are the same. CWT and STFT are translation-invariant representation and there are also some other approaches to achieve this translation invariance with wavelets [17, 50]. First one is the so-called *spin cycle*, that needs to create shifted versions of the original signals and used them for training, obtaining and average solution. The second one is a redundant invariant wavelet transforms. We have chosen the *à trous* algorithm in our problem to overcome the non-invariance drawback due to its simplicity.

In this work, we use a two-dimensional version of the *à trous* algorithm. The matrix of initial values $c_0(k,l)$ is the non-polarized image and the number of levels is $N = 4$. The scaling function in the two-dimensional case has been chosen as the separable function $\phi_2(x,y) = \phi(x)\phi(y)$, being

$$\phi(x) = \begin{cases} 1-|x| & x \in [-1,1] \\ 0 & \text{otherwise} \end{cases},$$

whose associated filter $h(n)$ is $h(-1) = h(1) = \frac{1}{4}, h(0) = \frac{1}{2}$, and zero elsewhere.

The wavelet filtering process we have devised is performed by the reconstruction of only a selection of details at some scales of the decomposition. We select the scales that express the most important image information aiming at a correct segmentation, that is, the grains. Consequently, we remove the noise and the inhomogeneous image illumination by discarding in the reconstruction the finest detail $w_1(k)$ and the coarser approximation $c_N(k)$, respectively. After that, the coefficients of the remaining scales are divided into positive and negative. Negative coefficients are roughly related to contours and the positive ones with grains. This is due to the shape of the filter

h, that fits better into grain-shaped structures and thus, gives positive coefficients. Conversely, $-h$ fits into the contours, therefore mostly represented by negative coefficients. We use a filter kernel h with a narrow region of support. Other bases, like B-splines of higher degree, mix contour and noise information in the detail coefficients of the first scales because their region of support is wider. The first row of Fig. 3.14 shows a decomposition and selective reconstruction of a marble non-polarized image for $N = 7$. Figure 3.15 compares the watershed of an inverted non-polarized image with and without wavelet filtering. Comparing results for the entire samples, this wavelet scheme gives slight better segmentation with a minor drawback of interior grains that can be easily removed with a labeling process and a neighborhood computation. Compare the two example (Figs. 3.11 and 3.15) almost taken over the same image area.

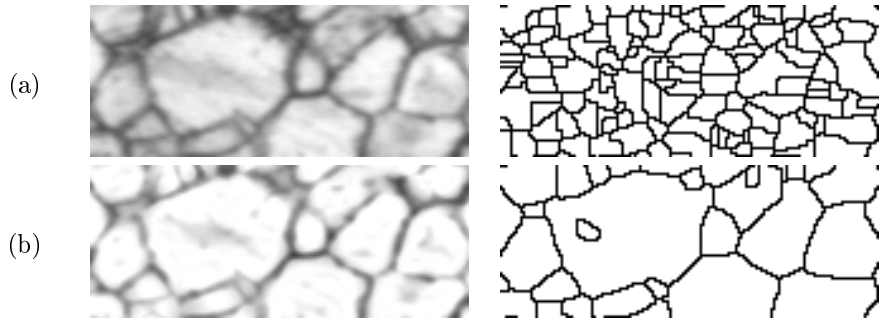

Figure 3.15: Oversegmentation and segmentation: (a) initial image and watershed of its inverse, (b) filtering by partial wavelet reconstruction and watershed of its inverse.

3.7 Region merging

In this part of the process, we perform the merging of the regions that belong to the same grain. Two adjacent regions are merged if one of the following criteria is fulfilled:

a) Their common boundary has high average grey level in the transmittance image.

b) The amplitudes and phases of both regions are similar enough in someone of the sequences.

The reason why the second criterion does not require similarity in both sequences is that the computed phase is not reliable when amplitude is small. Therefore, we shall choose the best case between both sequences, according to a similarity measure we have devised. The merging process is implemented by means of the following procedure of five steps:

Labeling. We label the regions and boundaries of the binary image resulting from the previous segmentation. Thus, we obtain an image where the grey level at each pixel is equal to the number of the region it belongs, according to the image scanning order. Labeling of regions is a direct process. However, to perform a boundary labeling we divided lines in the segmentation image into isolate boundaries removing all the juntions in this image and applying a strandard labeling process.

Graph representation. This labeled image is represented as an undirected, weighted graph where nodes are regions and arcs are boundaries between adjacent regions. Each one of the nodes and the arcs stores a label of its region in order to keep the relationship between graph and image. Also, depending on the process we want to apply to the image, each arc or node takes an appropiate value.

Removal of weak arcs. In this case we assign to each arc the averaged grey level of the boundary that represent. We apply the first criterion removing each arc with a value less than a threshold $\boldsymbol{t_1}$. These changes are propagated to the initial binary image by setting to 1 the pixels of the removed boundaries, in this way, joining regions in the image domain. Next, we label again and transform to a graph the new binary image.

Distance criterion. In addition to a region number, we now associate to each node two phases d_0, d_{45} and amplitudes m_0, m_{45} which are representative of the $0°$ and $45°$ amplitude and phase of its region. Then, we weight each arc with a dissimilarity measure d between the two nodes it connects, which is calculated from their representative values. This dissimilarity measure has been designed according to the second criterion, as following:

$$\begin{aligned} d &= \max(d_0, d_{45}) \\ d_g^2 &= m_g^2 + {m'_g}^2 - 2m_g m'_g \cos(f_g - f'_g) \ , \qquad g = 0, 45 \ . \end{aligned}$$

The distance d_g, as Fig. 3.16 illustrates, is the distance between (m_g, f_g) and (m'_g, f'_g) taken as vectors in polar form, that is, taking the amplitude as a radius and phase as an angle. The goal of d_g is to return a large dissimilarity when phases are opposite but also when phases are similar and amplitudes are very different.

The second criterion is applied by removing all those arcs with a distance criterion (d) lower than a threshold $\boldsymbol{t_2}$. These changes are propagated again to the initial binary image, as in the *Removal of weak arcs* step. Thus, we finally obtain the image of merged regions.

Until now, we have not specified how to compute representative values for each region of $0°$ and $45°$ amplitudes and phases. We need them because the phase and amplitude within a region are often not uniform. They can change due to macles, borders between twin crystals appearing in the image as straight scratches, impurities and material stress. In addition to this, a region can contain part of the real border of a grain, where the phase is uncertain. The average of their values obtains a biased estimation due to the influence of extrema. Therefore, we calculate the representative of each parameter as the mode, that is, the most frequent value inside a region. This

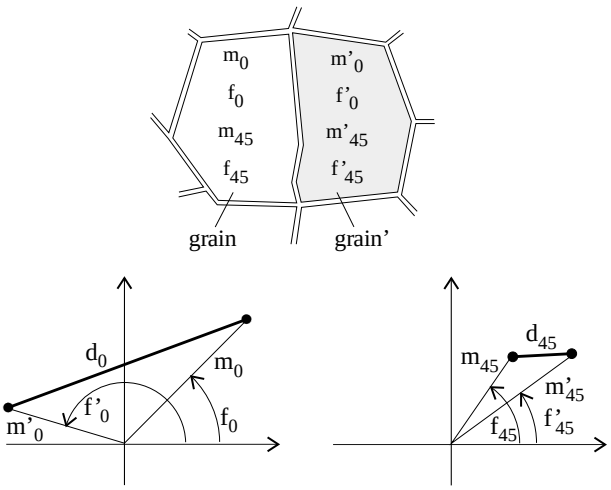

Figure 3.16: Dissimilarity measure d between two nodes calculated as the maximum of d_0 and d_{45}.

is a more robust value than mean, and it is less influenced by the undesirable effects that we mentioned before.

Removal of interior grains. Finally, after region merging, we must remove regions that are totally or mostly placed inside another region, because there are not interior grains in marble samples. These regions owe its existence to the fact that phase at boundaries, and even inside a grain, differs significantly from the rest of the grain. For the interior region removal, the ratio between the longer boundary length shared with another region and the total perimeter is calculated for each region. The more a region is included within another one, the closer this ratio will be to 1. Thus, a third threshold $\boldsymbol{t_3}$ is applied to the perimeter ratio to decide the removal of these grains.

3.8 Segmentation results

Figures 3.17 and 3.18 show the segmentation results for six sample images, each one belonging to a different type of marble (quarry). The original images show at once the two most important problems: ill-defined contours and the presence of strong and false contours due to macles. For the most part of grains and marble types, our method succeeds in finding out the right contours of grains thanks to the use of the two amplitudes and phases at each point. Nevertheless, we have to admit that no perfect segmentation is achieved in all samples. If we look at the results accurately, we can see two types of errors. On the one hand, we have different grains that appear as an only one (e.g. Carrara and Pentelikon) and, on the other hand, grains in which the merging step does not achieve to join all their regions (e.g. Paros). These errors

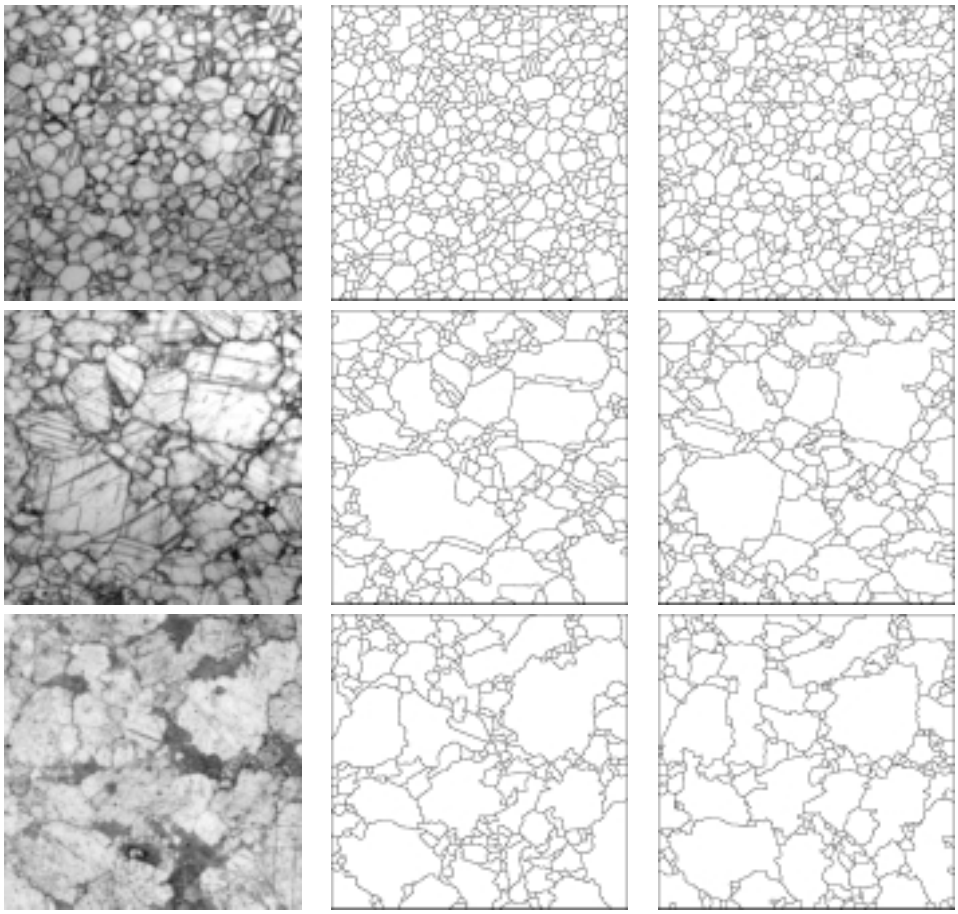

Figure 3.17: Segmentation results: Carrara, Paros, Naxos. Second column with erosion-reconstruction filtering and third column with wavelet filtering.

are mainly due to the parameters t_2 and t_3, whose values have been tuned in order to obtain a good global performance, that is, for all types of marble and for different samples within each quarry. These values are $t_2 = 3$ and $t_3 = 0.6$. The range of t_3 for which the segmentation results are not drastically affected is wide, $t_3 \in [0.5, 1]$, only t_2 has got to be accurately selected. Better results can be obtained if these parameters are tuned for each case, but this is not a good solution for our purposes because in the application we do not have a previous information about samples. The input of our program are just the different images taken from the sample, without any previous knowledge of its features. Despite of this, we believe that the internal geological structure of marbles is statistically well preserved in the segmented images.

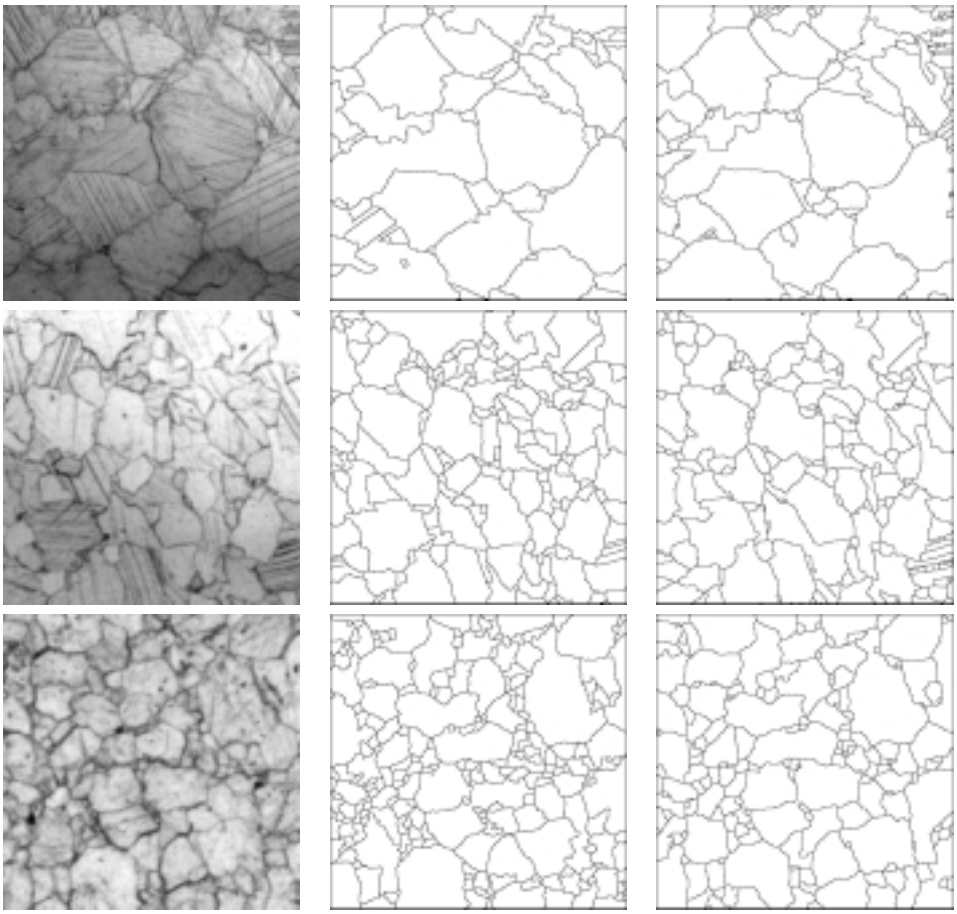

Figure 3.18: Segmentation results: Aphrodisias, Afyon and Pentelikon. Second column with erosion-reconstruction filtering and third column with wavelet filtering.

3.9 Classification

The previous segmentation gives us a set of adjacent grains that can be studied first in isolation and later in relation to their neighbors. The study of the individual grains implies to measure values related to their size and shape. All this information, individual values and topological relations, are used by an expert to determine the origin quarry of the analyzed sample. The knowledge base for rock classification using grain parameters was established long time ago [68], but the acquisition of all this kind of data manually by means of a microscope was so time-consuming that this technique was discarded.

A preliminar visual classification of our samples can be seen in Table 3.1 where shape and size features has been used. It shows how some of the quarries are clearly

Table 3.1: Classification with shape and size features.

Criterion	CA	PA	PR	AF	AY	NA	AL	PE	ES	SA	SC
Size											
Large grains				×					×		
Average grains						×					×
Small grains	×				×			×			
Non-uniform		×	×				×			×	
Shape											
Regular grain	×			×					×		×
Irregular grain		×	×		×	×	×	×		×	
Isotropy											
Oriented											×
Non-oriented	×	×	×	×	×	×	×	×	×	×	

characterized according to just four shape parameters: Carrara (CA), Naxos (NA) and the Saint-Béat (SC); some other can be sometimes confused, however: Aphrodisias (AF) ↔ Estremoz (ES), Afyon (AY) ↔ Pentelikon (PE); the remaining four classes have the same parameters and therefore they have a high confusion: Paros (PA), Proconnesos (PR), Almadén (AL), Saint-Béat (SA). These results reflect that visual inspection based on simple parameters is not enough to a good classification and that it is necessary to select and quantify accurately the classification features to surpass this problem.

According to the human expert, the fundamental criterion in the visual sample classification are: mean size, size distribution and shape of grains. Some of the parameters that contribute to characterize grain shape are:

Circularity: computed as the ratio $\frac{4\pi s}{p^2}$, where s is the area and p is the perimeter. It represents the relation between the object to be measured and a circle of equal area. For a circle it is equal to 1.

Quadrature: defined as the ratio $\frac{p}{4\sqrt{s}}$ and represent the relation between the object and a square. For the square it is equal to 1.

Irregularity: ratio between actual perimeter and the perimeter of the convex hull. For a regular object the two perimeters are the same and the irregularity is equal to 1. But also for convex objects. Thus, irregularity measures divergence from convexity or degree of concavity.

Elongation: it is the relation between the major axis and the perpendicular one to this major axis.

These four parameters plus mean size of the grains have been used for classification. Among them, size has the highest discriminating power [5]. Therefore, it is possible to establish a preliminary classification between sample marbles of small, medium

and large grains. Previous works in this field [4] calculate a discriminat analysis over five samples for each one of the ten quarries with classification parameters: size, circularity, quadrature, elongation. The result obtained in this work was an 86% of correct classification.

The fact that our system is able to deliver the result with an extra information, that is, a graph structure, adds new criteria to extend the classification process. Easily, the set of neighbours of each grain can be computed and its shape properties evaluated and new statistics of these distributions can be also supplied to the classifier. In addition, the rate of little and big grains and its relative positions can be evaluated. All this new tracks allow refining and improving the classification results, but this work must be done by skilled professionals able to extrapolate new evidences from the facts that provoke these behaviors. It is necessary to know the petrogenesis or to deduce it from the images themselves to take profit of these results. Therefore, this issue remains as a future work.

3.10 Further extensions

The previous work was oriented to solve a concrete problem, the segmentation of petrographical marble images. But this segmentation strategy can be easily adapted to other completely different problems than marble recognition. Accordingly, a similar solution has been proposed in [86] extending it to colour textures.

The segmentation in the marble case is achieved by detecting the valley lines between grains but other images may not have these kind of border structures. Valleys must be generated from edges by something like a gradient transformation or a contour detection algorithm. Furthermore, the properties we use to merge the oversegmented grains were amplitude and phase images that are specific for petrographical images and concretely form uniaxic materials. Thus, in a generic problem if segmentation is performed the results can be also improved merging regions with similar properties. These new properties could be related for example with the colour of the region or with other features as texture. These properties fit well into our proposed methodology that permits improving a preliminary segmentation taking into account the appropiate properties of the specific problem. As a demonstrative example we can see in Fig. 3.19 how the method applied with a little tuning of the parameters, achieves the segmentation on other scenes. Figure 3.19a comes from the problem explained in [76] where the aim is to extract the regions corresponding to clothes and skin of people in front of a desk, in order to be able to find out later in a different scenario. In the real application a background subtraction removes any possibility of confusion with the static objects behind people, then segmentation is restricted to those areas that change along the sequence (basically subjects). Figure 3.19b shows the segmentation of a face where a fast and easy result is achieved with no information about skin, hair, cloths and background colours. In both cases valley lines has been obtained with a morphological gradient transform, and the mean colour has been used as merging criterion.

(a)

(b)

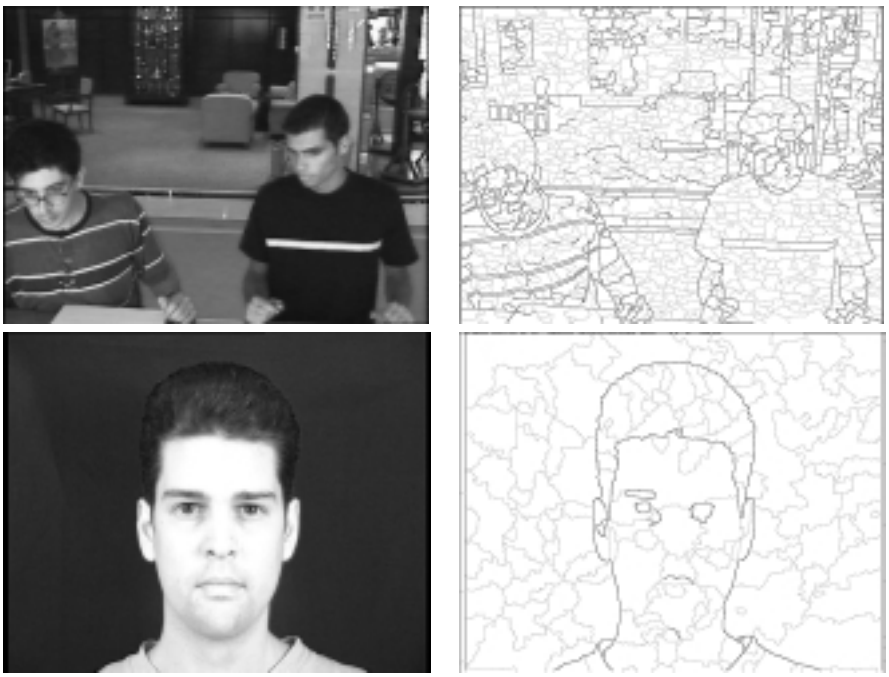

Figure 3.19: Segmentation applied to other problems: (a) people location, (b) face region detection. The left part are the original image and the right part is the segmentation with grey lines the preliminary segmentation and black lines the final segmentation.

3.11 Conclusions

Given the difficulties that presents working only with non-polarized images, our approach has been focused in emulating the procedure followed by human experts when a visual classification of samples is performed. Thus, our method extracts additional information through the illumination of samples with polarized light. In particular, we take advantage of the image formation model offered by the Johannsen's law, which relates the incident and transmitted light intensities through uniaxial crystals. This model allows calculating two parameters at each pixel (amplitude and phase) which are intrinsic of each grain. These parameters are the input of a region merging procedure that improves an initial oversegmentation to achieve a better correspondence between regions and grains.

The oversegmentation is obtained by means of the watershed transform applied to the non-polarized image, filtered in order to reduce the number of local minima. We propose two alternatives to this filter process. A morphological erosion-reconstruction that removes those valleys non-important (the filtering element do not fit inside the ridges associated to these valleys). This approach is an ad hoc solution with good results but strongly related to the filtering element. This strategy has been also used

later in some other works for other authors with similar purposes and also good results. The second filter strategy proposed is based on a partial reconstruction of a wavelet decomposition, behind all this there is a model that detaches valleys, ridges, noise and illumination inhomogeneities. We select for reconstruction purposes the scales and coefficients where the boundaries of grains (valleys) are well represented and discard the scales and coefficients which mainly contain ridges, noise and inhomogeneous illumination. Additionally, we take advantage of the relation between the three dimensional shape of the grain boundaries and the analyzing wavelet function. These boundaries mostly contribute to negative coefficients, which are the only ones that take part in the reconstruction.

Results show that our method achieves a correct segmentation for most grains in a variety of marble types, without any initial knowledge about their features. Once the segmentation is done the expert study several parameters related to the morphology of each grain and the relation among the bulk of grains. With all this information and based on his knowledge and expertise he gives a source quarry for the sample.

This work allows creating a data base coming form different marble samples of known origin. The collection of this kind of data permits to identify the geographical origin of some non catalogued samples as archaeological or ornamental stones, contributing to the knowledge of the trading routes of these materials. Provenance identification is important in discovering exchanges between populations, dating artifacts, and identifying forgeries and copies.