

**On the assessment of texture feature descriptors
in intravascular ultrasound images: A boosting
approach to a feasible plaque classification**

Oriol Pujol and Petia Radeva

Contents

Boosting in IVUS images	i
Introduction	2
0.1 Feature Spaces	7
0.1.1 Co-occurrence matrix approach	7
0.1.2 Derivatives of Gaussian	10
0.1.3 Cumulative Local Moments	11
0.1.4 Local Binary Patterns	13
0.2 Adaboost classification process	14
0.2.1 AdaBoost procedure	15
0.2.2 Behavior of the Adaboost procedure	18
0.2.3 The role of the "weak" classifier	19
0.3 Results and Conclusions	21
0.3.1 Experimental settings and database building	21
0.3.2 Discrimination Reference	23
0.3.3 The Fibrous vs Calcium problem	23
0.3.4 The Soft vs Calcium problem	24
0.3.5 The Mixed vs Calcium problem	25
0.3.6 The Soft vs Fibrous problem	26
0.3.7 The Fibrous vs Mixed problem	26
0.3.8 The Soft vs Mixed problem	27
0.3.9 Characteristic curves	29
0.4 Discussion and Conclusions	29

Introduction

Myocardial infarction, sudden cardiac death and unstable angina as a consequence of coronary thrombosis developed as a result of a ruptured vulnerable or an eroded atherosclerotic plaque. Plaque rupture or endothelial erosion with subsequent thrombosis formation are the most frequent cause of acute coronary syndromes. A study reported a high correlation between multiple plaque ruptures in acute coronary syndrome (ACS) patients. Another study shows that plaque ruptures occur not only in this case but in patients with stable angina or asymptomatic ischemia too. Moreover, there are studies showing plaque rupture in patients with non-cardiac death. Hence, it is not clear why some plaque ruptures in coronary arteries of patients lead to severe consequences meanwhile other remain asymptomatic and heal. To understand the mechanisms of plaque destabilization and guide a pharmacological treatment, it is of high interest to image the fragile part of the atheromatous plaque and to differentiate between low-risk and high-risk plaques.

The composition and structure of the vessel change with age, hypertension, diabetes mellitus and many other factors. Until this moment, it is feasible to discriminate different morphological structures of the vessel as calcium deposits, fatty, fatty fibrous and fibrous materials. Although from several decades investigators recognized that noninvasive imaging of coronary calcium might be useful to identify patients with unsuspected coronary artery disease, until the advance of high-resolution techniques little success has been achieved. Today, it is well-known that coronary calcium is a result of a complex, regulated and active process similar to bone formation that is related and at the same time different from atherosclerosis. On the other hand, it is not completely clear what the vulnerable plaque is. The common researcher opinion is that a vulnerable plaque consists of: lipid core, fibrous cap, presence of inflammatory cells and is affected by the vessel remodelling and its 3D morphology. Still a complete morphological, mechanical and chemical information is necessary in order to characterize the vulnerable plaque in a robust way.

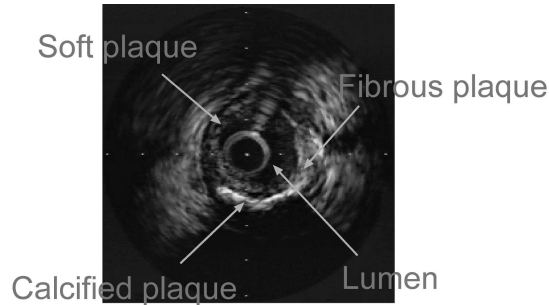


Figure 1: Typical IVUS image presenting different kind of tissues.

Coronary angiography has been so far the gold standard to assess the severity of obstructive luminal narrowing. Furthermore it serves as a decision tool to direct therapeutical procedures (as PTCA). By coronary angiography the lumen boundaries can be assessed, but no information is provided about plaque burden, plaque delineation and plaque components. The predictive power of occurrence of myocardial infarction is rather low since 70angiographically normal, and only a minority occur where there was severe stenosis. Other studies have affirmed, that the culprit lesion prior to a myocardial infarction has, in 4878a stenosis smaller than 50detected by angiography, but can be well assessed pathologically.

IVUS displays the morphology and histological properties of a cross-section of a vessel [1]. Figure 1 shows a good example of IVUS image. It is generally accepted that the different kind of plaque tissues distinguishable in IVUS images is threefold: *Calcium formation* is characterized by a very high echoreflectivity and absorption of the emitted pulse from the transducer. This behavior produces a deep shadowing effect behind calcium plaques. In the figure, calcium formation can be seen at three o'clock and from five to seven o'clock. *Fibrous plaque* has medium echoreflectivity resembling that of the adventitia. This tissue has a good transmission coefficient allowing the pulse to travel through the tissue, and therefore, providing a wider range of visualization. This kind of tissue can be observed from three o'clock to five o'clock. *Soft plaque* or *Fibro-Fatty plaque* is the less echoreflective of the three kind of tissues. It also has good transmission coefficient allowing to see what is behind this kind of plaque. Observing the figure, a soft plaque configuration is displayed from seven o'clock

to three o'clock.

There are two lines of research to describe the vessel morphology and detect plaques in IVUS images: by textural image analysis, radio-frequency analysis of IVUS data and elastograms.

Textural analysis is the most close to the physician "exercises" during IVUS analysis as a decision is taken on morphological analysis of image sequence. Visual textural analysis is a difficult, subjective and time-consuming process highly depending on the specialist. Therefore, there is an increasing interest of the medical community in developing automatic tissue characterization procedures of IVUS images. This is accentuated because the procedure for tissue classification by physicians implies the manual analysis of IVUS images frequently necessary to be done online during the therapeutical procedure.

The problem of automatic tissue characterization has been widely studied in different medical fields. The unreliability of gray level only methods to achieve good discrimination among the different kind of tissues force us to use more complex measures, usually based on texture analysis. Texture analysis has played a prominent role in computer vision to solve tissue characterization problems in medical imaging [2] [3] [4] [5] [6] [7] [8] [9].

Several researching groups have reported different approximations to characterize the tissue of intravascular ultrasound image.

Vandenberg et al, in [10], base their contribution on reducing the noise of the image, to have a clear representation of the tissues. The noise reduction is achieved by averaging sets of images when the least variance in diameter of the IVUS occurs. At the end, a fuzzy logic based expert is set to discriminate among the tissues.

Nailon et al devote several efforts to IVUS tissue characterization. In [11] they use classic Haralick texture statistics to discriminate among tissues. In [12] the author proposes the use of co-occurrence matrices texture analysis and fractal texture analysis to characterize intravascular tissue. Thirteen features plus fractal dimension derived from Brownian motion are used for this task. The conclusion shows that fractal dimension is unable to discriminate between calcium and fibrous plaque but helps in fibrous versus lipidic plaque. On the other hand, co-occurrence matrices are well suited for the overall classification. In [13], it is stated that the discriminative power of fractal dimension is poor when trying to separate fibrotic tissue, lipidic tissue and foam cells. The

method used is based on fractal dimension estimation techniques (box-counting, brownian motion and frequency domain).

Spencer et al in [14], center their work on spectral analysis. Different features are compared: mean power, maximum power, Spectral Slope and 0Hz interception. The work concludes with the 0Hz spectral slope as the most discriminative feature.

Dixon et al in [15], use co-occurrence matrices and discriminant analysis to evaluate the different kind of tissues in IVUS images.

Ahmed et al [16] uses a radial transform and correlation for pattern matching. The features used are higher order statistics such as kurtosis, skewness, and up to four order cumulants, C_4 . The results provided appear to have fairly good visual recognition rate.

The work of de Korte et al [17] opens a new proposal based in assessing the local strain of the atherosclerotic vessel wall to identify different plaque components. This line of work is based on estimating the radial strain by performing cross-correlation analysis on pairs of IVUS at a certain intra-coronary pressure. This very promising technique, is called *elastography*, and is still in development.

Probably, one of the most interesting work in this field is the one provided by Zhang, Sonka et al [18]. This work is much more complex trying to evaluate the full morphology of the vessel. Detecting the plaque and adventitia borders and characterizing the different kind of tissues, the tissue discrimination is done using a combination of well-known techniques previously reported in the literature as: co-occurrence matrices and fractal dimension from brownian motion, and adding two more strategies to the amalgam of features: run-length measures and radial profile. The experiments assess the accuracy of the method quantitatively.

Most of the literature found in the tissue characterization matters use texture features, being co-occurrence matrices the most popular of all feature extractors. Further work has been done trying to use other kind of texture feature extractors and IVUS images, and although not specifically centered on tissue characterization, the usage of different texture features in plaque border assessment is reported, that can be easily extrapolated to tissue characterization. In [19], derivative of gaussian, wavelets, co-occurrence matrices, Gabor filters and cumulative moments are evaluated and used to classify blood from

plaque. The work highlights the discriminative power of co-occurrence matrices, derivatives of gaussian and cumulative moments. Other works such as [20] provide some hints on how to achieve a fast framework based on local binary patterns and fast high-performance classifiers. This last line of investigation overcomes one of the most significant drawbacks of the texture based tissue characterization systems, the speed. Texture descriptors are inherently slow to be computed. With the proposal of the feature extractor based on local binary patterns a good discriminative power is ensured as well as a fast technique for tissue characterization.

Whatever method we use in the tissue characterization task, we follow an underlying main methodology. First, we need to extract some features describing the tissue variations. This first step is critical since the features chosen have to be able to describe each kind of tissue in a unique way so that it can not be confused with another one. In this category of feature extraction we should consider the *co-occurrence matrix measures*, *statistical descriptors*, *local binary patterns*, etc. The second step is the classification of the extracted features. Depending on the complexity of the feature data some methods will fit better than others. In most cases, high dimensional spaces are generated, so we should consider the use of dimensionality reduction methods such as *Principal component analysis* or *Fisher linear discriminant analysis*. Either a dimensionality reduction process is needed or not, this step will require a classification procedure. This procedures can be supervised, if we provide samples of each tissue to be classified so that the system "knows" a priori what the tissues are, or unsupervised, if we allow the system to try to find which are the different kind of tissues by itself. In this category we can find *clustering* methods for unsupervised classification and, for supervised classification, methods like *maximum likelihood*, *nearest neighbors*, *support vector machines*, and the center of our current analysis *adaptative boosting*. In particular, adaptative boosting techniques allow to deal with high dimensional spaces by using an intelligent feature selection process while training the classifier.

The following sections are devoted to describe: First, the texture methods we are using in our study. Second, we introduce the Adaboost classification process. And third, we describe the result of using such techniques and for plaque characterization.

0.1 Feature Spaces

The first issue when dealing with complex real problems, such as tissue characterization, is to create a representation of the data we are analyzing. The representation of the data is usually a more compact version of the original samples, for the problem to be analytically feasible. While centering in some aspects of the original samples, we restrict ourselves to that kind of features, and expect them to fully describe the problem. Though not always happens.

Plaque recognition is usually approached as a texture discrimination problem. This focuses our work, in the analysis of the variability, size, granularity of the original tissue samples. Several texture representations are available in literature. We will focus our study on two different kind of texture descriptors. The first class of texture descriptors are formally acknowledged as fully representative and highly discriminant sets. In this class we place co-occurrence matrices descriptors [22] and a bank of filters approach based on derivatives of gaussian [24]. The second class, is less recognized since the techniques are relatively new. This class comprehend descriptors characterized by its low complexity and therefore fast computationally. This gain in speed, however has a cost, the lost in accuracy of the description. In this category we are placing, cumulative moments [23] and local binary patterns [28]. Those sets include examples of the two most important lines of work when dealing with texture, the statistical approach (co-occurrence matrices measures and cumulative moments) and the kernel-based approach (bank of filters and local binary patterns). The first line of work are concerned with density estimation techniques or parameters. The second line of work centers on sampled forms of analytic functions. In this sense, the local binary patterns approach is the less conventional of the methods, but we have chosen to include it in the kernel-based approach for sake of simplicity.

0.1.1 Co-occurrence matrix approach

Co-occurrence matrices were ideated in 1963 by Julesz [29]. They were created as tools for texture segregation using second order related statistics. They have been used in image processing, as texture descriptors with very successful results[22]. In the co-occurrence method, the relative frequencies of gray level pairs of pixels at certain relative displacement are computed and sorted in a

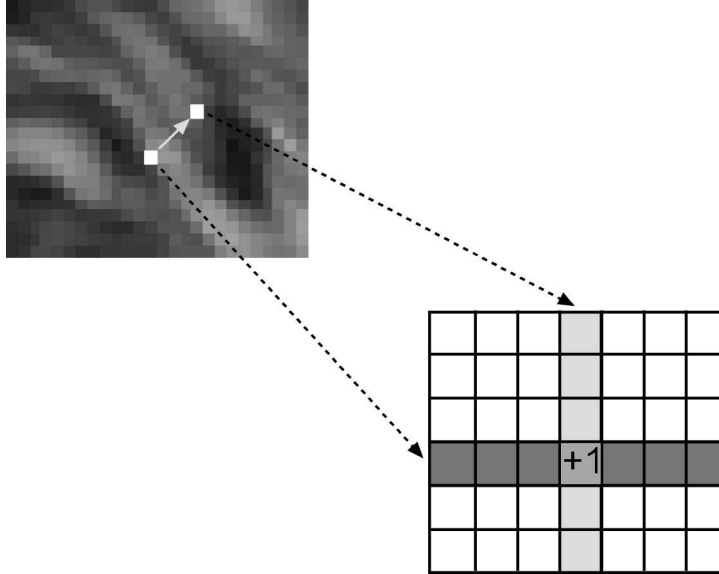


Figure 2: Co-occurrence matrix creation explanation diagram

matrix, the *co-occurrence matrix* \mathbf{P} . The co-occurrence matrix can be thought of as an estimate of the joint probability density function of gray-level pairs in an image. For G gray levels in the image, \mathbf{P} will be of size $G \times G$. If G is large, the number of pixel pairs contributing to each element, $p_{i,j}$ in \mathbf{P} is low, and the statistical significance poor. On the other hand if the number of gray levels is low, much of the texture information may be lost in the image quantization. The element values in the matrix, when normalized, are bounded by $[0, 1]$, and the sum of all element values is equal to 1.

$$P(i, j, D, \theta) = P(I(l, m) = i \text{ and } I(l + D \cos(\theta), m + D \sin(\theta)) = j) \quad (1)$$

where $I(l, m)$ is the image at pixel (l, m) , D is the distance between pixels and θ is the angle. It has been proved by other researchers [30] [21] that the nearest neighbor pairs at distance D at orientations $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ are the minimum set needed to describe the texture second-order statistic measures.

Figure 2 provides a graphical explanation of the co-occurrence method. It is based on a discrete approximation of the probability density function of the occurrences of the appearance of two pixels of certain gray levels at a certain distance with a fixed angle. In practice, this is done using a matrix with as much files and columns as the gray levels. This matrix is the co-occurrence

matrix, and is filled adding one to the cell pointed by the gray levels of the pixels pair located at a certain distance D with angle θ .

This matrix is still difficult to deal with, and the data it contains is usually sparse. Therefore, a description of the matrix shape and information is useful for reducing the complexity and dimensionality of the problem. The most important measures that characterize the co-occurrence matrices are: Energy, Entropy, Inverse Difference Moment (IDM), Shade, Inertia and Promenace.[30]

Let us introduce some notation for the definition of the features:

$P(i, j)$ is the (i, j) th element of a normalized co-occurrence matrix.

$$\begin{aligned} P_x(i) &= \sum_j P(i, j) \\ P_y(j) &= \sum_i P(i, j) \\ \mu_x &= \sum_i i \sum_j P(i, j) = \sum_i iP_x(i) = E\{i\} \\ \mu_y &= \sum_j j \sum_i P(i, j) = \sum_j jP_y(j) = E\{j\} \end{aligned}$$

With the above notation, the features can be written as follow:

$$\begin{aligned} Energy &= \sum_{i,j} P(i, j)^2 \\ Entropy &= -\sum_{i,j} P(i, j) \log P(i, j) \\ IDM &= \sum_{i,j} \frac{1}{1 + (i - j)^2} P(i, j) \\ Shade &= \sum_{i,j} (i + j - \mu_x - \mu_y)^3 P(i, j) \\ Inertia &= \sum_{i,j} (i - j)^2 P(i, j) \\ Promenace &= \sum_{i,j} (i + j - \mu_x - \mu_y)^4 P(i, j) \end{aligned}$$

As a result of the co-occurrence matrices descriptors extraction process, we have a set of values (one for each descriptor) for each pixel in a certain plaque and for a certain orientation. As we are using four orientations, the final feature vector will contain 24 features per pixel.

0.1.2 Derivatives of Gaussian

Derivatives of gaussian refers to a filter bank in which the transference function of each filter is one of the directional derivatives of the 2-D gaussian function. This alone, would only provide information of the orientation, size and complexity of the region we are filtering. Since, we aim for a more accurate description of the texture, scale can be introduced to handle image structures in a consistent manner. This texture descriptor is a particularization of the linear *scale-space representation* proposed in [24], [34]. The basic idea is to embed the original signal into a one-parameter (scale) family of gradually smoothed signals, in which fine scale details are successively suppressed. It can be shown that the Gaussian kernel and its derivatives are one of the possible smoothing kernels for such scale-space. The Gaussian kernel is well-suited for defining a space-scale because of its linearity and spatial shift invariance, and the notion that structures at coarse scales should be related to structures at finer scales in a well-behaved manner (new structures are not created by the smoothing method). Scale-space representation is a special type of multi-scale representation that comprises a continuous scale parameter and preserves the same spatial sampling at all scales. Formally, the linear-space representation of a continuous signal is constructed as follows. Let $f : \mathfrak{R}^N \rightarrow \mathfrak{R}$ represent any given signal. Then, the scale-space representation $L : \mathfrak{R}^N \times R_+ \rightarrow \mathfrak{R}$ is defined by $L(\cdot; 0) = f$ so that:

$$L(\cdot; t) = g(\cdot; t) * f \quad (2)$$

where $t \in \mathfrak{R}_+$ is the scale parameter, and $g : \mathfrak{R}^N \times R_+ \setminus \{0\} \rightarrow \mathfrak{R}$ is the Gaussian kernel. In arbitrary dimensions, it is written as:

$$g(x; t) = \frac{1}{(2\pi t)^{N/2}} e^{-x^T x / (2t)} = \frac{1}{(2\pi t)^{N/2}} e^{-\sum_{i=1}^N x_i^2 / (2t)} \quad , x \in R e^N, x_i \in \mathfrak{R} \quad (3)$$

The square root of the scale parameter, $\sigma = \sqrt{t}$, is the standard deviation of the kernel g , and is a natural measure of spatial scale in the smoothed signal at scale t . From this scale-space representation, multi-scale spatial derivatives can be defined by

$$L_{x^n}(\cdot; t) = \partial_{x^n} L(\cdot; t) = g_{x^n}(\cdot; t) * f, \quad (4)$$

where g_{x^n} denotes a derivative of some order n .

The main idea behind the construction of this scale-space representation is that the fine scale information should be suppressed with increasing values of the scale parameter. Intuitively, when convolving a signal by a Gaussian kernel with standard deviation $\sigma = \sqrt{t}$, the effect of this operation is to suppress most of the structures in the signal with a characteristic length less than σ . Different directional derivatives can be used to extract orientation related structural features at different scales. It is shown in the literature [35] that a possible complete set of directional derivatives are $\partial^n = [\partial_0, \partial_{90}, \partial_0^2, \partial_{60}^2, \partial_{120}^2, \partial_0^3, \partial_{45}^3, \partial_{90}^3, \partial_{135}^3]$. So our feature vector will consist on the directional derivatives, including the zero-derivative, for each of the n scales desired:

$$F = \{\{\partial^n, G^n\}, n \in \mathfrak{R}\} \quad (5)$$

0.1.3 Cumulative Local Moments

Cumulative Local Moments is a simple approach to moment calculation. Geometric moments have been used effectively for texture segmentation in many different application domains [23]. In addition, other kind of moments have been proposed, Zernique moments, Legendre moments, etc. By definition, any set of parameters obtained by projecting an image onto a 2D polynomial basis is called moments. Then, since different sets of polynomials up to the same order define the same subspace, any complete set of moments up to given order can be obtained from any other set of moments up to the same order. The computation of some of these sets of moments leads to very long processing times, so in this section a particular fast computed moment set has been chosen. This set of moments are known as the **accumulation local moments**. Two kind of accumulation local moments can be computed, direct accumulation and reverse accumulation. Since direct accumulation is more sensitive to round off errors and small perturbations in the input data [31], the reverse accumulation moments are used.

The reverse accumulation moment of order $(k-1, l-1)$ of matrix \mathbf{I}_{ab} is the value of $\mathbf{I}_{ab}[1, 1]$ after bottom-up accumulating its column k times (i.e., after applying k times the assignment $\mathbf{I}_{ab}[a-i, j] \leftarrow \mathbf{I}_{ab}[a-i, j] + \mathbf{I}_{ab}[a-i+1, j]$, for $i = 0$ to $a-1$, and for $j = 1$ to b), and accumulating the resulting first row from right to left l times (i.e., after applying l times the assignment $\mathbf{I}_{ab}[1, b-j] \leftarrow \mathbf{I}_{ab}[1, b-j] + \mathbf{I}_{ab}[1, b-j+1]$, for $j = 1$ to $b-1$). The reverse accumulation

moment matrix is defined so that $\mathbf{R}_{mn}[k.l]$ is the reverse accumulation moment of order $(k - 1, l - 1)$.

Consider the matrix in the following example,

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 4 & 2 & 3 \end{pmatrix}$$

According to the definition, its reverse accumulation moment of order (1,2) requires two column accumulations,

$$\begin{pmatrix} 5 & 4 & 6 \\ 5 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 14 & 9 & 13 \\ 9 & 5 & 7 \\ 4 & 2 & 3 \end{pmatrix}$$

and three right to left accumulation of the first row

$$\begin{pmatrix} 36 & 22 & 13 \\ & & \\ & & \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 71 & 35 & 13 \\ & & \\ & & \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 119 & 48 & 13 \\ & & \\ & & \end{pmatrix}$$

Then it is said that the reverse accumulation moment of order (1,2) of the former matrix is 119.

The set of moments alone are not sufficient to obtain good texture features in certain images. Some iso-second order texture pairs which are pre-attentively discriminable by humans, would have the same average energy over finite regions. However, their distribution would be different for the different textures. One solution suggested by Caelli is to introduce a nonlinear transducer that maps moments to texture features [32]. Several functions have been proposed in literature: logistic, sigmoidal, power function or absolute deviation of feature vectors from the mean [23]. The function we have chosen is the hyperbolic tangent function, which is logistic in shape. Using the accumulation moments image I_m , and a non linear operator $|\tanh(\sigma(I_m - \bar{I}_m))|$ an 'average' is performed throughout the region of interest. The parameter σ controls the shape of the logistic function. Therefore each textural feature will be the result of the application of the non-linear operator to the computed moments. If $n = k \cdot l$ moments are computed over the image, then the dimension of the feature vector will be n . Hence, a n -dimensional point is associated with each pixel of the image.

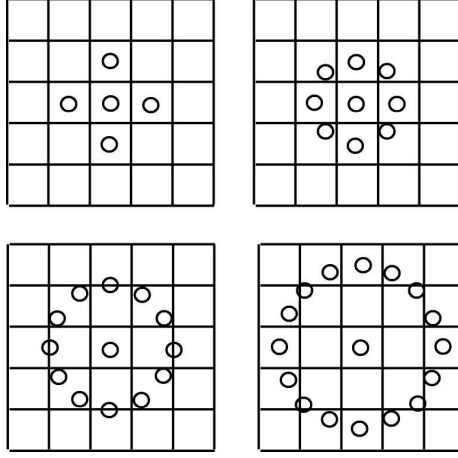


Figure 3: Typical neighbors (Top-Left) $P = 4$, $R = 1.0$ (Top-Right) $P = 8$, $R = 1.0$ (Bottom-Left) $P = 12$, $R = 1.5$ (Bottom-Right) $P = 16$, $R = 2.0$.

0.1.4 Local Binary Patterns

Local Binary Patterns [28] is a feature extraction operator used for detecting "uniform" local binary patterns at circular neighborhoods of any quantization of the angular space and at any spatial resolution. The operator is derived based on a circularly symmetric neighbor set of P members on a circle of radius R . It is denoted by $LBP_{P,R}^{riu2}$. Parameter P controls the quantization of the angular space, and R determines the spatial resolution of the operator. Figure 3 shows typical neighborhood sets. To achieve gray-scale invariance, the gray value of the center pixel (g_c) is subtracted from the gray values of the circularly symmetric neighborhood g_p ($p = 0, 1, \dots, P - 1$) and assigned a 1 value if the difference is positive and 0 if negative.

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

By assigning a binomial factor 2^p for each value obtained, we transform the neighborhood into a single value. This value is the $LBP_{R,P}$:

$$LBP_{R,P} = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p$$

To achieve rotation invariance the pattern set is rotated as many times as necessary to achieve a maximal number of the most significant bits, starting always from the same pixel. The last stage of the operator consists on keeping the information of "uniform" patterns while filtering the rest. This is achieved using a transition count function U . U is a function which counts the number of transitions 0/1, 1/0 while we move over the neighborhood.

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$$

Therefore,

$$LBP_{P,R}^{riu2} = \begin{cases} LBP_{P,R}^{ri} & \text{if } U(LBP_{P,R}) \leq 2 \\ P + 1 & \text{otherwise} \end{cases}$$

Therefore, Local binary patterns look for the longest chain of homogeneous-valued pixels given a circular sampled neighborhood. This description is gray level independent since it just binarize the neighborhood pixel value according to the central pixel value. However, we can be further improve this description by finding the local variance of the neighborhood set in gray level.

0.2 Adaboost classification process

At this point, we have a set of features that represent each instance of the problem. In our particular case, they are samples of plaque. Each sample of plaque has a feature vector associated, that contains useful information about this plaque. The feature vector describes a multi-dimensional space. Therefore, for each pixel we have an n -dimensional point in the feature space, where n is the number of features. This set of data is the input to the classification process.

The Adaboost process is a supervised learning and classification tool, since we know exactly the classes we are seeking. Adaboost is created as a method for combining simple classifiers to obtain a very accurate decision. Roughly, it is an iterative assembling process in which each classifier is devoted to find a good division of the sub-set of points formed by the samples that are more difficult classified up to that point.

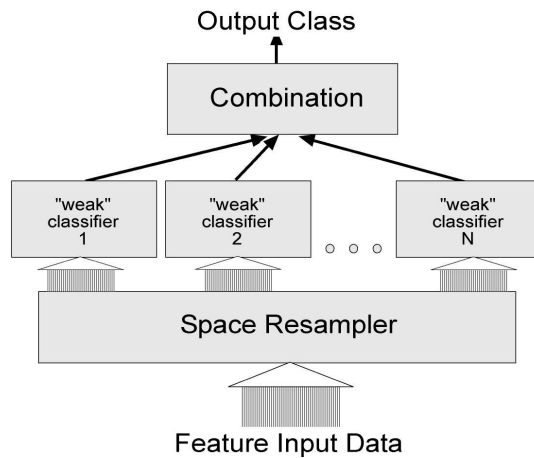


Figure 4: Block diagram of the AdaBoost procedure.

In particular Adaboost is a shortening for *Adaptative Boosting (AdaBoost)*, and is widely recognized as one of the most accurate processes for high accuracy classification.

0.2.1 AdaBoost procedure

Adaptative Boosting (AdaBoost) is an arcing method that allows the designer to keep adding "weak" classifiers until some desired low training error has been achieved [41] [42]. At each step of the proces, a weight is assigned to each of the feature points. These weights measure how accurate the feature point is being classified at that stage. If it is accurately classified, then its probability of being used in subsequent learners is reduced, or emphasized otherwise. This way, AdaBoost focuses on difficult training points at each stage.

The classification result is a linear combination of the "weak" classifiers. The weight of each classifier is proportional to the amount of data that classifies in a correct way.

Figure 4 shows a diagram of the general process of boosting. The adaboost process uses weights that modify the probability density function of the appearance of each sample data point. This fact can be troublesome, since we need to find classifiers that allow weighing the samples points. Another possibility is to resampled the data set according to the weights of each feature data. The new

set of feature points are inputs of the new classifier to be added to the process. Although, this last method is more general is unadvisable to use it, since after several iterations, the training set can be trimmed to very little data points. Therefore, it hinders the classification process.

As an additional feature AdaBoost is capable of performing a feature selection process while training. In order to perform both tasks, feature selection and classification process, a weak learning algorithm is designed to select the single features which best separate the different classes. That is, one classifier is trained for each feature, determining the optimal classification function (so that the minimum number of feature points is misclassified). And then, the most accurate classifier-feature pair is stored at that stage of the process. If feature selection is not desired, the weak classifier focuses on all the features at a time.

The general algorithm is described as follows:

- Determine a supervised set of feature points $\{x_i, c_i\}$ where $c_i = \{-1, 1\}$ is the class associated to each of the features classes.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $c_i = \{-1, 1\}$ respectively, where m and l are the number of feature points for each class.
- For $t = 1..T$:

- Normalize weights

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,i}}$$

so that w_t is a probability distribution.

- For each feature, j train a classifier, h_j which is restricted to using a single feature. The error is evaluated with respect to w_t , $\epsilon_j = \sum_i w_i |h_j(x_i) - c_i|$.
- Choose the classifier, h_t with the lowest error ϵ_t .
- Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{e_i}$$

where $e_i = 1$ for each well-classified feature and $e_i = 0$ otherwise. $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$. Calculate parameter $\alpha_t = -\log(\beta_t)$.

- The final "strong" classifier is:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, the strong classifier is the ensemble of a series of simple classifiers ("weak"). Parameter α_t is the weighting factor of each of the classifiers. The

loop ends whether the classification error of a "weak" classifier is over 0.5, the estimated error for the whole "strong" classifier is lower than a given error rate or if we achieve the desired number of "weak". The final classification is the result of the weighted classifications of the "weak". The process is designed so that if $h(x) > 0$, then pixel x belongs to one of the classes.

0.2.2 Behavior of the Adaboost procedure

Analyzing the Adaboost process, we can figure out the error rate behavior when adding new "weak" classifiers. As we have described in the former section, the probability of each sample to be used in a "weak" classifier raises if it has been misclassified up to that moment by the "strong" classifier. So if we want to add a classifier $h_{t+1}(x)$, we take the misclassified points according $h(x) = \sum_{t=1}^T \alpha_t h_t(x)$ and raise its probability in $t + 1$. As the set with higher probability is composed by the difficult data points, the "weak" classifier will easily fail in assigning the correct label to each sample. This fact, tells us that the error rate will increase the more classifiers we add. This is true for the transient time. To further understand the behavior of the stationary time, we will now describe the behavior of the "strong" classifier error rate.

One of the conditions that stop the process is the fact that the "weak" classifier must perform better than the random guess. That is, we always want the error rate of the "weak" classifier to be under 0.5. If this condition is granted at each step, it means that although the "weak" is focusing on the most difficult data set, it still manages to find a usable solution. This translates in the fact that some misclassified points will now be correctly assigned to the true label. This, of course, is decreasing the error rate of the compound of "weak". This is true up to the point that if no other stop condition is met, the error rate tends asymptotically to zero.

Resuming the "weak" classification error rate in the stationary stage. It is expected that the classifier will be able to classify correctly at least one of the samples. If this happens, the error will be better than random guess, though tending to 0.5. Otherwise, the "weak" can not be used to train the "strong" classifier and the process will end.

Figure 5 shows the evolution of the error rates for the training and the test feature points. Figure 5(a) shows the test error rate. One can observe, that the overall error has a decreasing tendency as more "weak" classifiers

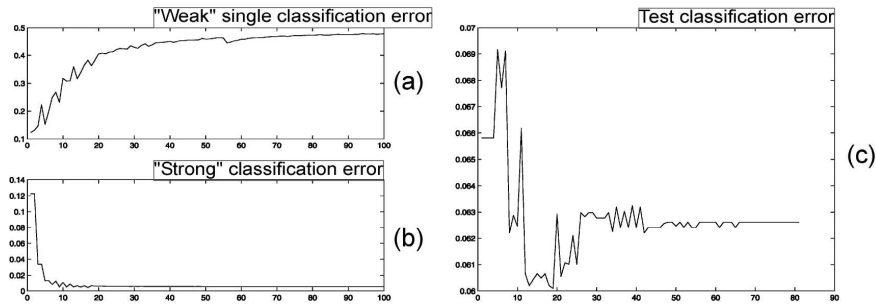


Figure 5: Error rates associated to the AdaBoost process. (a) Test error rate. (b) "Weak" single classification error (c) Strong classification error on the training data.

are added to the process. Figure 5(b) shows the error evolution of each of the "weak" classifiers. The figure illustrates how the error increases as more "weak" classifiers are added. Figure 5(c) shows the error rate of the system response on the training data. As it is expected, the error rate decreases to very low values. This, however does not ensure a test classification error of such accuracy.

One question arises at this point. What will happen with the test error rate? The answer is not simple. While we expect the test error rate to decrease in the same way as the training error rate does, one can not guarantee this behavior. However, we realize that if the training set is meaningful, in the sense that it correctly represents the problem, the test error rate should decrease according with the "strong" error rate. But we also must take into account that we have a finite amount of samples, and therefore, the "weak" classifiers could try to distinguish among not representative and conflictive points due to the sampling. This fact can lead to overtraining stages, in which, though the training is correctly classified, as it is not a good representation of the reality, the test samples are misclassified.

0.2.3 The role of the "weak" classifier

The weak classifier has a very important role in the procedure. Different approaches can be used, however it is relatively interesting to center our attention in low time-consuming classifiers.

The first and the most straight forward approach to a "weak" is the per-

ceptron. The perceptron is constituted by a weighed sum of the inputs and an adaptative threshold function. This scheme is easy to embed in the adaboost process since it relies on the weights to make the classification.

Another approach to be taken in consideration is to model the feature points as Gaussian distributions. This allows us to define a simple scheme by simply calculating the weighed mean and weighed covariance of the classes at each step of the process:

$$\mu_{i,t}^j = \sum w_{i,t} x_i \quad \Sigma_{i,t}^j = \sum w_{i,t} (x_i - \mu_{i,t}^j)^2$$

for each x_i^j point in class C_j . $W_{i,j}$ are the weights for each data point.

If feature selection is desired, this scheme is highly constrained to the N features of the N-dimensional feature space. If N is not enough large, the procedure could not improve its performance. Therefore we propose another classifier for relatively low dimensional spaces (2 magnitude orders). Because the selection of a single feature for each of the classifiers is quite a hard constraint, we can look for the most significant pair of features which discriminates better the different classes.

For each pair of features of our space we use linear discriminant analysis to find the transformation which leads to the most discriminant axis. We chose the pair of features with the lowest error. We can describe this "weak" classifier as follows.

$$h(x) = \begin{cases} 1 & \text{if } p_j W_j^t x < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$$

where p_j and θ_j are the parity and threshold parameters and W_j is defined as follows:

$$W_j = \Sigma_j^{-1} (\mu_{-1,j} - \mu_{1,j})$$

which is the canonical variate. W_j is the principal axis of the solution of the linear discriminant analysis system which maximizes

$$J(W) = \frac{W^t S_B W}{W^t S_W W}$$

, where S_B is the between-class scatter $S_B = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^t$ and S_W is the within-class scatter $S_W = \sum_{i=1}^C \sum_{x \in C_i} (x - \mu_i)(x - \mu_i)^t$, where μ is the mean value of the whole data, c is the number of classes and N_i is the number of samples in class i .

Another approach to a "weak" classifier relies on the use of the ROC curves. The ROC curves show the amount of false positives and false negatives for each possible parameter of the classifier. In particular, if we use a threshold value, it shows the curves for each threshold value. At this point, the optimal threshold value is the minimum of the sum of both curves. This is the optimal trade-off between the misclassification in both classes. This process can be done for each feature using a feature selection plus classification process. This is the approach we have used in this article.

In general, the weak classifier, $h_j(x)$ consists of a feature f_j , if the feature extraction process is desired, and the parameters of the classifier Θ_j . Those parameters are a threshold θ_j , a parity p_j and W_j , a classifier trained by linear discriminant analysis. Although the threshold separates the two classes it is not enough to identify which class is in either side of the threshold. Therefore a parameter p_j (parity) is needed to indicate the direction of the inequality sign when classifying:

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$$

Both, the feature extraction and the classification processes, are the central parts of any classification system. We will use the framework explained in the former sections in order to classify exhaustively the plaque in the IVUS image.

0.3 Results and Conclusions

The assessment of the classification process is presented in this section. First, the database building and experimental settings are introduced. Then, a discrimination classification reference for methodology comparison is shown. After the reference, we are explaining the multiple pair-wise problems and the results of their classification. We finish this section showing the characteristic curves of the behavior of the Adaboost process.

0.3.1 Experimental settings and database building

One of the main problems in the IVUS scientific community is the lack of a standard reference set for validation of the IVUS tissue classification. Regarding this matter, we have devoted a great amount of time in collaboration with

Feature Set	Computational Complexity	Discriminative Power
Filter bank	High	High
Co-occurrence set	High	High
LBP	Low	High
Cumulative Moments	Low	Average

Table 1: Computational complexity and performance for each set.

expert physicians to create a database with ten thousand samples of each of the four tissues acknowledged by experts, soft tissue, fibrous tissue, mixed tissue and calcium. Those samples have been extracted from 20 different patients, using a **nombre del aparatito y megahertz de funcionamiento y otras características tecnicas chachis :P** . Using this database, several texture descriptors have been selected.

Particularly, we have chosen

- A derivatives of gaussian filter bank, up to the third derivative. A five level multi-resolution framework is used, with scales $\{0.2, 0.5, 1, 2, 4\}$. For each scale, a set of directional derivatives is extracted. Particularly, this set is $\partial^n = [\partial_0, \partial_{90}, \partial_0^2, \partial_{60}^2, \partial_{120}^2, \partial_0^3, \partial_{45}^3, \partial_{90}^3, \partial_{135}^3]$, where the subindex points the direction of the partial derivative in degrees. To this set we also include the zero-derivative image, that is a smoothed version at the corresponding level of the original image.
- A set of descriptors of the co-occurrence matrices at angles $\{0, 45, 90, 135\}$ with neighborhoods of 11×11 pixels and distance for the co-occurrence pair of $D = 2$ and a 17×17 pixels neighborhood with a distance of $D = 3$.
- A tissue description set based on local binary patterns and local variance, using radius 1 with 8 samples, radius 2 with 16 samples and radius 3 with 24 samples.
- A feature space based on cumulative moments, with moments up to $(9, 9)$.

All these feature spaces are well-known and well-suited spaces for texture description. They usually differ from each other in terms of computational complexity and discriminative power. Table 1 shows how the performance of each feature set.

Regarding the Adaboost procedure, we use a composition of 500 classifiers in the original feature space, for each description set.

0.3.2 Discrimination Reference

To compare the performance of the boosting method we have selected a well-known classifier, Fisher Linear Discriminant Analysis. The results of this classifier is our ground-truth, to which we will refer in order to compare the results of the Adaboost technique.

Plaque pair	BOF	COOC25	COOC38	LBP	MOM
Fibrous vs Calcium	35.18%	24.55%	23.84%	48.53%	46.43%
Soft vs Calcium	25.30%	10.07%	9.94%	42.97%	44.10%
Mixed vs Calcium	29.64%	18.51%	17.83%	46.55%	46.14%
Soft vs Fibrous	34.83%	35.81%	34.57%	48.46%	47.67%
Fibrous vs Mixed	44.56%	47.68%	49.84%	49.91%	49.70%
Soft vs Mixed	42.73%	40.30%	49.36%	46.94%	47.87%

Table 2: Error rate for each plaque pair, using the different tissue descriptors.

Table 2 shows the test error rate for each pair of plaques. Having a look at the figures, the table reveals that there are some sets with a great amount of overlapping samples. This is particularly true in the discrimination between mixed tissue and soft or fibrous tissue, since mixed tissue implies an ill-defined mixture of both fibrous and soft tissues. We can also see that some feature spaces tend to perform better for our problem than others. In fact, co-occurrence matrices descriptors (COOC25, COOC38) and derivatives of gaussian filter bank (BOF), outperform clearly the low complexity descriptors, Local Binary Patterns (LBP) and Cumulative Moments (MOM). This means that both feature spaces, describe better the tissue properties than the other two. However, this fact could have been easily predicted from table 1, since there is always a trade-off between complexity and performance.

0.3.3 The Fibrous vs Calcium problem

The characterization of the calcium tissue seems to be the less difficult one since the calcium tissue has a very high echo-reflectivity and homogeneity. On the other hand, fibrous plaque is also high reflective, but have much more rugosity.

Feature Set	Initial Error	Final Error
BOF	33.13%	13.09%
COOC25	20.90%	13.74%
COOC38	20.67%	11.04%
LBP	24.76%	21.81%
MOM	43.62%	38.04%

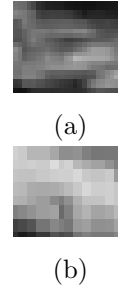


Figure 6: Examples of the fibrous tissue (a) and the calcium plaque (b).

Figure 6 shows an example of the fibrous plaque fig.6.a and the calcium plaque fig.6.b . The table from figure 6 shows the results for this test. As expected the initial error in the overall feature space of the best performing sets and the error using discriminant analysis are quite close to each other. However, the Adaboost procedure refines the classification thus increasing the recognition rates to an average of 88%. Unexpectedly, LBP has a relative good performance, close to 80%, making it an ideal candidate if we aim for fast processing.

0.3.4 The Soft vs Calcium problem

This problem is by far the most simple one since the plaques we are distinguishing are the more different kind of plaques. In particular, the soft tissue has low echo-reflectivity and high granularity, while the calcium plaque is just the opposite.

Feature Set	Initial Error	Final Error
BOF	17.75%	5.80%
COOC25	9.81%	7.27%
COOC38	8.88%	4.29%
LBP	15.31%	14.68%
MOM	45.49%	33.00%

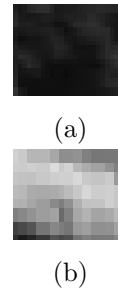


Figure 7: Examples of the soft tissue (a) and the calcium plaque (b).

Figure 7 shows an example of the fibrous plaque fig.7.a and the calcium plaque fig.7.b . The former statement is confirmed by the table shown in figure

7. That shows the figures for the error rate in this problem. Again, the recognition rate of the high complexity spaces is pretty high, and further increased by the Adaboost process, up to an average over 95%. Three important remarks can be made looking at the figures. First, there is a huge improvement in performance using derivatives of gaussian (BOF), of about 12%. Second, LBP still has pretty good results, over 85%. Third, and the most surprising, MOM still performs bad in this stage. Looking at the reference table 2 we can see that there is a huge improvement in LBP performance and BOF performance. LBP lowers its error rate by 30% and BOF lowers its error rate by 20%.

0.3.5 The Mixed vs Calcium problem

Since the behavior of the soft and fibrous plaque against calcium tissue is fairly good, we expect this problem to be an "average" of the above ones.

Feature Set	Initial Error	Final Error
BOF	26.29%	9.79%
COOC25	16.36%	12.44%
COOC38	15.91%	7.46%
LBP	20.54%	19.15%
MOM	44.16%	35.75%

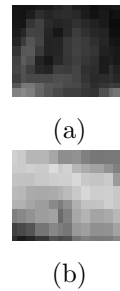


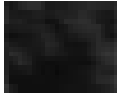
Figure 8: Examples of the mixed tissue (a) and the calcium plaque (b).

Figure 8 shows an example of the mixed plaque fig.8.a and the calcium plaque fig.8.b . Certainly, this is what happens, the results are not as good as the soft versus calcium problem, (see table in figure 8) but are better than the fibrous versus calcium one. This is logical if we recall that the mixed tissue is a combination of both fibrous tissue and lipid tissue in an interleaved way. At this stage, we have clearly a good vision of what is the performance of each feature space as well as the influence of the adaboost process in the problem. BOF and COOC38 performs the best after adaboost, granting high recognition rates. COOC25 seems to perform the worst of the trio formed by the high complexity classifiers. If we compare this results to the ones obtained using FLD, BOF lowers its error rate by 20%, and COOC38 by 10%.


0.3.6 The Soft vs Fibrous problem

One of the most interesting and less obvious problems is the soft vs fibrous problem. In this case, both plaques has a good amount of texture involved in them (high granularity). Although, there is a difference in the reflectivity of the transducer pulse.

Feature Set	Initial Error	Final Error
BOF	28.63%	26.41%
COOC25	27.58%	27.53%
COOC38	26.57%	25.98%
LBP	31.62%	30.93%
MOM	44.41%	38.43%



(a)



(b)

Figure 9: Examples of the soft tissue (a) and the fibrous plaque (b).

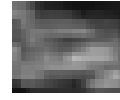
Figure 9 shows an example of the soft plaque fig.9.a and the fibrous plaque fig.9.b . Table in figure 9 shows the performance of the Adaboost procedure when applied to this problem. We can conclude from the figures, that in this case, the Adaboost process does not help much. This fact, seems to show that the way data is distributed in the feature spaces is clearly entwined. This fact, hinders the process of the combination of classifiers, since, presumingly, each weak classifier is focusing on a really low amount of bad strong classified data. In this case, the comparison of the results with the reference of Fisher, improves the recognition rate by 10%.

0.3.7 The Fibrous vs Mixed problem

This problem and the soft vs mixed problems are by far the most complex ones. The fibrous and the mixed plaques really resemble each other in terms of local distribution features. The difference between both is simply the spatial overall distribution of the tissues. Most of the methods we have tried are purely local, and therefore are destined to fail in this problem. In fact, we have seen that the mixed label is also the most disagreeable of the plaques among the experts labelling. However, we also attach the results for this two problems.

Figure 10 shows an example of the fibrous plaque fig.10.a and the mixed plaque fig.10.b . Table in figure 10 shows the error rates for this problem.

Feature Set	Initial Error	Final Error
BOF	37.74%	36.28%
COOC25	39.99%	37.33%
COOC38	39.40%	35.65%
LBP	41.31%	40.90%
MOM	43.42%	40.92%



(a)



(b)

Figure 10: Examples of the fibrous tissue (a) and the mixed plaque (b).

0.3.8 The Soft vs Mixed problem

In the same way than the former case. The soft versus mixed plaque problem is ill-posed from the local texture point of view.

Feature Set	Initial Error	Final Error
BOF	40.44%	37.36%
COOC25	37.72%	33.09%
COOC38	35.42%	29.29%
LBP	39.35%	39.01%
MOM	46.45%	41.26%



(a)



(b)

Figure 11: Examples of the soft tissue (a) and the mixed plaque (b).

Figure 10 shows an example of the soft plaque fig.10.a and the mixed plaque fig.10.b Table in figure 11 show the results for this case. It is remarkable the fact that COOC38 is able to distinguish both plaques with an average recognition rate of over 70%. This is due to the fact that COOC38 uses a 17×17 neighborhood and therefore is susceptible to pick up the spatial distribution of the entwined fibrous and soft plaques.

The fibrous vs. mixed and soft vs. mixed using linear discriminant analysis can not be made, since the results show that the decision is nearly random (recognition rates of about 55%). However, using adaboost the problem seems to have a weak solution, that is, a solution of nearly 70% of recognition.

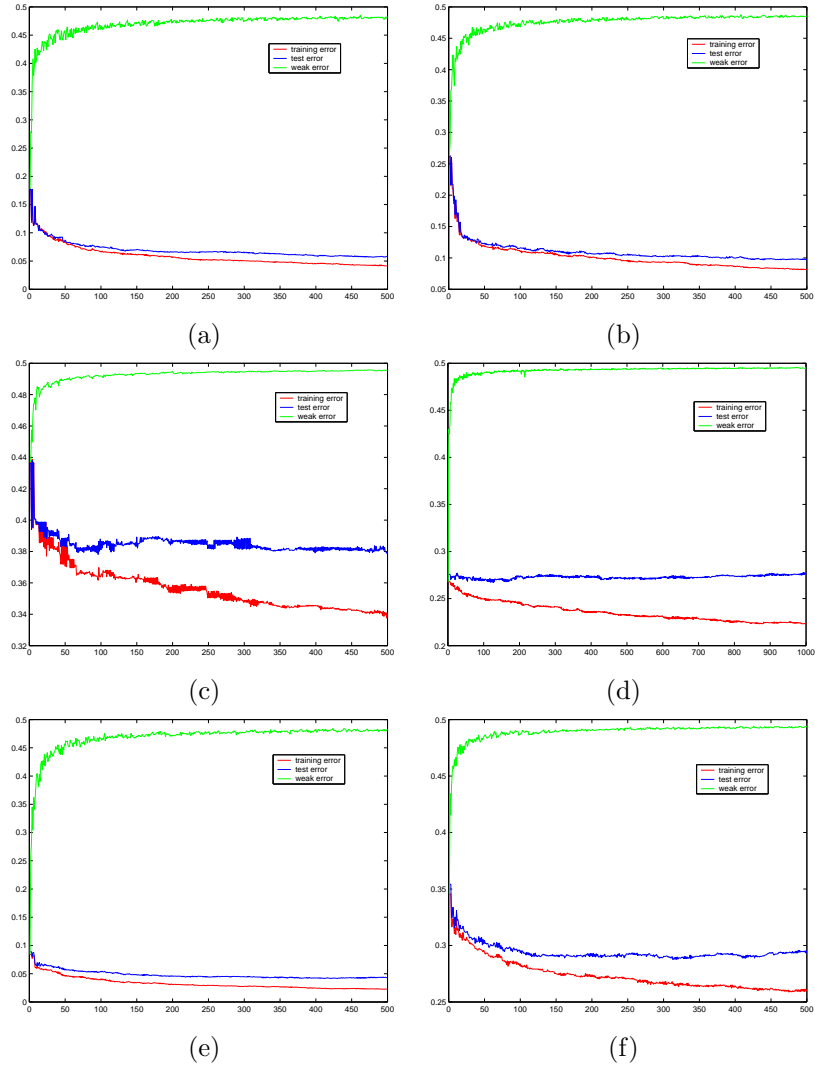


Figure 12: Examples of the characteristic curves for the adaboost process. (a) BOF in soft-calcium problem. (b) BOF in mixed-calcium problem. (c) MOM in fibrous-calcium problem. (d) COOC25 in soft-fibrous problem. (e) COOC38 in soft-calcium problem. (f) COOC38 in soft-mixed problem.

0.3.9 Characteristic curves

As stated in the adaboost description section, the error rate of the weak classifiers increases at each iteration (every time we add a classifier) due to the fact that it has to classify the previously erroneous classified data (the errors of the combination of weak classifiers up to this moment). The other characteristic of this process is that the overall error rate on the training data tends to zero as the number of weak classifiers increases. This fact does not imply that the test classification error also tends to zero. In fact, it has been shown, and we will see in the figures, that it can worsen or keep nearly constant.

Figure 12 shows some examples of the characteristic behavior of the adaboost process. Figures 12.a and 12.b refers to the behavior of the derivatives of gaussian in front of the soft-calcium and mixed-calcium problems respectively. Figure 12.c shows the error curves for the accumulation moments in front of the fibrous-calcium problem. Figure 12.d shows the aforementioned constant behavior of the test error rate although the training error rate clearly decreases. This figure refers to the co-occurrence 2,5 when characterizing soft versus fibrous plaque. Figures 12.e and 12.f show co-occurrence matrix descriptors for parameters (3,8) in front of the problems soft-calcium and soft-mixed plaque, respectively.

0.4 Discussion and Conclusions

The first conclusion that arises from the experiments is the beneficial influence that has the adaboost procedure in the classification of the plaque. Thus, rendering it to a very good classifier if we aim for plaque characterization.

However, although adaboost is a very high performance classifier, the results show that plaque characterization based only on texture can not be made accurately if we want recognition rates over 85%. Furthermore, the most different kind of tissue, calcium is easily identified even without context information, with an overall accuracy of over 95%. However, mixed plaques are really difficult to distinguish. This points out that if we want to classify mixed plaques texture descriptors alone are not suitable for the task.

Regarding the feature spaces, co-occurrence matrices descriptors performs better than the rest, closely followed by the bank of filters approach. Local binary patterns is the third in recognition rate.

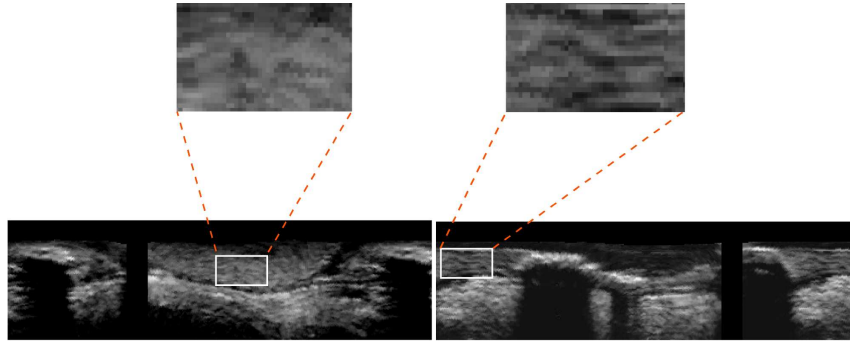


Figure 13: Illustration of how the local classification of tissue can be deceiving.

Several aspects have to be taken into account when judging the results. First of all, regarding the database. The data included in the database is widely variable and from a good amount of different patients. Analyzing the plaques, we have seen that different plaques from different patients can be quite different among each other, although they are grouped under the same label. It is surprising how, some plaques labelled as different can resemble to each other. This fact, also seems to point out that more information must be taken into account when dealing with plaque characterization. We have seen that physicians heavily rely on the context information, and the location of the plaque. Therefore, the results presented here are useful as a guideline of the performance of the classifiers but they are by no means a criterion to avoid the less accurate feature descriptors.

In particular, we can see from the study that MOM based classifiers lead to low recognition rates. However, the performance of the MOM classifier is deceiving. While it provides bad results as a feature space, we have seen that using contextual information most of the misclassified data can be seen as speckle noise caused by the decision of the classifier. That means that in the context of an IVUS image, the misclassified data points are scattered and can be re-classified taking into account neighboring plaques, since more of them are in isolated groups of two or three pixels. If this is done accurate results can be extracted from them. This reinforces the idea shown at the previous paragraph, the analysis of the contextual information can easily improve the recognition rate of the methods.

Figure 13 shows an example of how the local information can not be enough

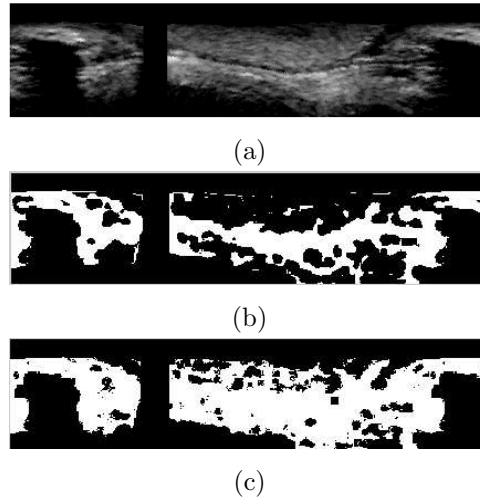


Figure 14: Adaboost overtraining and fake-plaque effect. a) Original image b) 50 classifiers c) 500 classifiers

for a good classification. In fact, if we look at the top images, we can see that both images are quite similar and we probably will classify them as some kind of plaque, in particular as fibrous plaque. However if we look at the original images we can see that one is in fact a tissue plaque, but the other it is neither a tissue, it is part of the blood pool (lumen).

The effect of the previous casuistic can lead to undesirable situations such as the one showed in figure 14. This figure is a simplification of the tissue problem. In particular, this is a blood vs tissue classification, but it will serve to easily understand some concepts. The figure is an example of two effects. On one hand, the final recognition rate is low because of the fake-plaque effect, as mentioned in the former paragraph. On the other hand, this is a classical example of an overtrained classifier. Figure 14.b shows the classifier results in a middle stage of learning. Figure 14.c shows the final result. As we can see, as the blood is relatively different from the blood expected, and because the last group of classifiers added in the adaboost process are more focused on very few samples of the training data, they can degrade the final classification result if the particularities of the training data do not coincide with the ones of the test set.

This effect lead us to create a new kind of classification process, currently under development, that takes into account the particular test set to infer

context information and therefore adapt the classification process to the particularities of the test set.

Bibliography

- [1] Wickline, S., Beyond intravascular imaging: Quantitative ultrasonic tissue characterization of vascular pathology, IEEE Ultrasonics simposium, pp. 1589-1597, 1994.
- [2] Arul, P. and Amin, V., Characterization of beef muscle tissue using texture analysis of ultrasonic images, Proceedings of the Twelfth Southern Biomedical Engineering Conference, pp. 141-143, 1993.
- [3] Mojsilovic, A. and Popovic, M., Analysis and characterization of myocardial tissue with the wavelet image extension [us images], Image Processing, 1995. Proceedings, Vol. 2, pp. 23-26, 1995.
- [4] Jin, X. and Ong, S., Fractal characterization of kidney tissue sections, Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers, Proceedings of the 16th Annual International Conference of the IEEE, Vol. 2, pp. 1136-1137, 1994.
- [5] Cohen, F. and Zhu, Q., Quantitative soft-tissue characterization in human organs using texture/attenuation models, Proceedings in Multidimensional Signal Processing Workshop, pp. 47-48, 1989.
- [6] Mavromatis, S. and Boi, J., Medical image segmentation using texture directional features, Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE, Vol. 3, pp. 2673-2676, 2001.
- [7] Mavromatis, S., Mammographic mass classification using textural features and descriptive diagnostic data, Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on, Vol. 1, pp. 461-464, 2002.
- [8] Donohue, K. and Forsberg, F., Analysis and classification of tissue with scatterer structure templates, Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on, Vol. 46, No. 2, pp. 300-310, 1999.

- [9] Ravizza, P., Myocardial tissue characterization by means of nuclear magnetic resonance imaging, *Computers in Cardiology 1991. Proceedings*, pp. 501-504, 1991.
- [10] Vandenberg, J., Arterial imaging techniques and tissue characterization using fuzzy logic, *Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems*, pp. 239-243, 1994.
- [11] Nailon, W. and McLaughlin, S., Comparative study of textural analysis techniques to characterize tissue from intravascular ultrasound, *Proc. Of the IEEE International Conference of Image Processing, Switzerland. IEEE Signal Processing Society:USA*, pp. 303-305, 1996.
- [12] Nailon, W. and McLaughlin, S., Intravascular ultrasound image interpretation, *Proc. Of the International Conference on Pattern Recognition, Austria. IEEE Computer Society Press:USA*, pp. 503-506, 1996.
- [13] Nailon, W., Fractal texture analysis: An aid to tissue characterization with intravascular ultrasound, *Proceedings 19th International Conference - IEEE/EMBS*, pp. 534-537, 1997.
- [14] Spencer, T., Characterization of atherosclerotic plaque by spectral analysis of 30mhz intravascular ultrasound radio frequency data, *IEEE Ultrasonics symposium*, pp. 1073-1076, 1996.
- [15] Dixon, K., Characterization of coronary plaque in intravascular ultrasound using histological correlation, *19th International Conference-IEEE/EMBS*, pp. 530-533, 1997.
- [16] Ahmed, M. and Leyman, A., Tissue characterization using radial transform and higher order statistics, *Nordic Signal Processing Symposium*, pp. 13-16, 2000.
- [17] de Korte, C. L. and van der Steen, A. F. W., Identification of atherosclerotic plaque components with intravascular ultrasound elastography in vivo: a yucatan pig study, *Circulation*, Vol. 105, No. 14, pp. 1627-1630, 2002.
- [18] Zhang, X. and Sonka, M., Tissue characterization in intravascular ultrasound images, *IEEE Transactions on Medical Imaging*, Vol. 17, No. 6, pp. 889-899, 1998.

- [19] Pujol, O. and Radeva, P., Automatic segmentation of lumen in intravascular ultrasound images: An evaluation of texture feature extractors, Proceedings for IBERAMIA 2002, pp. 159-168, 2002.
- [20] Pujol, O. and Radeva, P., Near real time plaque segmentation of ivus, Proceedings of Computers in Cardiology, pp. 159-168, 2003.
- [21] Randen, T. and Husoy, J. H., Filtering for texture classification: A comparative study, Pattern Recognition, Vol. 21, No. 4, pp. 291-310, 1999.
- [22] Haralick, R., Shanmugam, K. and Dinstein, I., Textural features for image classification, IEEE Trans. System, Man, Cybernetics, Vol. 3, pp. 610-621, 1973.
- [23] Tuceryan, M., Moment based texture segmentation, Pattern Recognition Letters, Vol. 15, pp. 659-668, 1994.
- [24] Lindeberg, T., Scale-Space Theory in Computer Vision, Kluwer Academic Publishers, 1994.
- [25] Jain, A. and Farrokhnia, F., Unsupervised texture segmentation using gabor filters, Systems, Man and Cybernetics, 1990. Conference Proceedings, pp. 14-19, 1990.
- [26] Mallat, S., A theory for multiresolution signal decomposition: The wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, pp. 674-694, 1989.
- [27] Mandelbrot, B., The Fractal Geometry of Nature, W H Freeman and Co. New York, 1983.
- [28] Ojala, T., Pietikainen, M. and Maenpaa, T., Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp. 971-987, 2002.
- [29] Julesz, B., Visual pattern discrimination, IRE Transactions on Information Theory, Vol. IT-8, pp. 84-92, 1962.
- [30] Ohanian, P. and Dubes, R., Performance evaluation for four classes of textural features, Pattern Recognition, Vol. 25, No. 8, pp. 819-833, 1992.

- [31] Martinez, J. and Thomas, F., Efficient computation of local geometric moments, *IEEE Trans. Image Processing*, Vol. 11, No. 9, pp. 1102-1111, 2002
- [32] Caelli, T. and Oguztoreli, M. N., Some tasks and signal dependent rules for spatial vision, *Spatial Vision*, No. 2, pp. 295-315, 1987.
- [33] Chaudhuri, B. and Sarkar, N., Texture segmentation using fractal dimension, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 1, pp. 72-77, 1995.
- [34] Lindeberg, T., Scale-space theory: A basic tool for analysing structures at different scales, *Journal of Applied Statistics*, Vol. 21, No. 2, pp. 225-270, 1994.
- [35] Rao, R. and Ballard, D., Natural basis functions and topographic memory for face recognition, *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 10-17, 1995.
- [36] Lumbreras, F., PhD Thesis: Segmentation, Classification and Modelization of Textures by means of Multiresolution Descomposition Techniques, CVC, UAB, 2001.
- [37] Jain, A. and Farrokhnia, F., A multi-channel filtering approach to texture segmentation, *Computer Vision and Pattern Recognition*, 1991. *Proceedings CVPR '91*, pp. 364-370, 1991.
- [38] Fukunaga, K., *Introduction to statistical pattern recognition*, Academic Press, 1971.
- [39] Duda, R. and Hart, P., *Pattern Classification*, Wiley-Interscience, 2001. Second Edition.
- [40] Belhumeur, P., Eigenfaces vs fisherfaces: Recognition using class specific linear projection, *IEEE Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720, 1997.
- [41] Schapire, R. E., The boosting approach to machine learning. an overview, *MSRI Workshop on Nonlinear Estimation and Classification*, 2002.

- [42] Viola, P. and Jones, M., Rapid object detection using a boosted cascade of simple features, Conference on Computer Vision and Pattern Recognition, p. 511-518, 2001.
- [43] Sonka, M. and Zhang, X., Segmentation of intravascular ultrasound images: A knowledge-based approach, IEEE Transactions on Medical Imaging, Vol. 17, No. 6, pp. 889-899, 1998.
- [44] von Birgelen, C., Computerized assessment of coronary lumen and atherosclerotic plaque dimensions in three-dimensional intravascular ultrasound correlated with histomorphometry, Amer. J. Cardiology, Vol. 78, pp. 1202-1209, 1996.
- [45] Klingensmith, J., Shekhar, R., and Vince, D., Evaluation of three-dimensional segmentation algorithms for identification of luminal and medial-adventitial borders in intravascular ultrasound images, IEEE Trans. on Medical Imaging, Vol. 19, No. 10, pp. 996-1011, 2000.
- [46] McInerney, T. and Terzopoulos, D., Deformable models in medical images analysis: a survey, Medical Image Analysis, Vol. 1, No. 2, pp. 91-108, 1996.