

Medical Image Retrieval Based on Plaque Appearance

*Jaume Amores , Petia Radeva*¹

0.1 Introduction

Images provide a powerful means to represent data, and many applications have as fundamental components the acquisition, processing and storing of huge amount of images. A typical example of such application is medical imaging. In hospitals working with medical images, large amounts of image data are received daily for processing, analysis and archiving. This arises the necessity of constructing database management systems able to organize, analyze and retrieve this type of data. The first image retrieval systems were based on information retrieval methodologies applied to textual annotations. As it has been seen in the last decade, describing images by textual annotations is not satisfactory for three main reasons: i) there is too much information in an image, so that small sets of words cannot describe it well; ii) labelling an image by words is a subjective task, as different individuals can apply different labels (words) depending on what they consider more relevant in the current image; iii) a complete manual labelling in a large collection of images is a very tedious task. These reasons have led to the use of image content for processing and organizing the database, constituting the so-called content-based image retrieval systems.

Such content-based image retrieval (CBIR) systems are necessary for a wide range of applications dealing with vast volumes of images. In the medical image field, retrieval by content is necessary for performing guided diagnosis and therapy. Having medical images archived along with their descriptions, retrieval by content allows the physician to present a query image representing

¹Computer Vision Center, Dept. Informàtica, UAB, Bellaterra, Spain

the current clinical case, and obtain the most similar stored images and their associated descriptions, so that the diagnosis of the image at hand is more easily taken by comparison. Indeed, whenever the type of the medical image is of great complexity it is common for the physician to base its diagnostics on manuals containing for each pathology its representative images. Other applications are the construction of electronic atlases with a high number of examples for each case, and educational uses. In this chapter, we overview the work on medical image retrieval, underline the peculiarities of medical image retrieval as opposed to image retrieval of general domains, and present a general framework of medical image retrieval based on plaque appearance.

A fundamental aspect of any CBIR system is the use of an appropriate feature space. This feature space should be able to represent all aspects of the image relevant to its description. For doing so, not only global information, but also local and contextual information is necessary. Global information is necessary for describing the statistics of relevant features inside the image. Local information is fundamental for dealing with images in which important parts are localized in small regions such as the pathology bearing regions of medical images. In addition to local and global information, the feature space should also be able to describe the spatial relations of the different structures conforming the overall image. In retrieval of general scope, this contextual information is good for describing semantically complex scenes as the relation between the objects conforming them. In the medical image retrieval field, the relative disposition of pathological structures respect to other structures play key roles in diagnosis. Finally, in complex domains such as retrieval of medical images the design of the feature space must be flexible enough to incorporate descriptors specific to the concrete problem, as general descriptors perform poorly in describing them.

0.1.1 State of the art in general content-based image retrieval and medical image retrieval

Most of the general CBIR systems up to day try to characterize the whole image by using only global information, i.e. a set of global signatures such as histograms of color, texture or shape [30],[13]. Pentland et al ([24]) use different global descriptors depending on the type of the image. Their global description is based on an orthogonal feature space representing the content,

using eigenimages for objects such as faces, eigenmodes for retrieving shapes and the wold-based features for textured images. In retrieval of general scope, some authors [28], [7], [31], [9] have included recently local information able to differentiate between images in which an important part of the discriminant information is localized in small regions. Schmid et al ([28]) extract characteristic points localized in discriminant parts, and compute a set of invariant local features around each characteristic point. Carson et al [7] segment the images using expectation-maximization based on texture and color. They achieve weak segmentation: the image is segmented into blobs which may represent objects as a whole, parts of objects, or unions of different objects. This weak segmentation, although does not represent isolated objects, approaches the semantics by which users organize images, and permits better retrieval results when the user queries for discriminant objects in the image. Wang et al [31] computes a weak segmentation based on k-means clustering of color and texture features. Their algorithm is faster than the used by Carson in [7], but achieves more inaccurate segmentations. They propose a similarity measure among images that compensates inaccurate segmentations by allowing multiple matchings from regions of the query to regions of the target image. Chen et al [9] following the same idea compensates inaccurate segmentations by treating the segmented regions as fuzzy sets and using a unified fuzzy matching between regions. Contextual information or (relative) spatial layout have also been explored by some authors. A typical way to deal with spatial organization is generalizing the concept of string to a plane, conforming the so-called 2D string descriptor [8] and 2D C string [17]. In this descriptor, objects are represented by nodes and their spatial relationship by strings. 2D strings require very accurate segmentations of the objects in the image, what leads in practice to manual segmentations. Another descriptor for describing context is the correlogram. A correlogram is an histogram measuring the distribution of features such as color in the image as well as their spatial relationship (or their co-occurrence). Huang et al. use these correlograms in retrieval of images by color, and Belongie et al. use another type of correlogram, called shape context, for taking into account the statistics of the distributions of points in shape matching and retrieval. These descriptors are good for taking into account the spatial distribution of the colors [15] or the spatial distribution of binary shapes [4], but do not permit to take into account context of different structures. We will

show how generalizing them we can have this contextual information without needing manual segmentations.

The extension of image retrieval to medical image applications is a challenging and emerging field where few works have been published. Y. Liu et al [19] characterizes retrieval as a classification problem. This is correct when the images can be categorized as belonging to known classes (e.g. representing different diseases in medical diagnosis) and the user wants to retrieve images that belong to the same class as the example presented. Following this idea the authors begin with a big set of descriptors and find the set of weights that achieves the lowest classification error. They basically use global descriptors specifically chosen for CT brain images, exploiting the symmetry in the brain as a basic property of normal (not diseased) brains. Local information is also extracted for asymmetrical regions.

Other authors such as Korn et al [16] explore the use of fast spatial access methods such as R-trees and fast nearest neighbor search. Their work is based on artificial sets of data that simulate tumor-like shapes.

Focusing on the design of appropriate feature spaces, Kak and Brodley [29], [11] take local and specific information for each of the pathology bearing regions (PBR) previously delineated by the user (the physician). Their work is specifically applied to high-resolution computer tomographies (HRCT) of the lungs. For each image, low-level features are extracted locally at every region and globally for the whole image. They use an exhaustive set of descriptors that includes the usual features based on texture, gray-level and shape, but also includes very specific descriptors designed to describe the different lung diseases and their appearance in high-resolution computed tomographies (HRCT). Finally, they take only the most relevant information by sequential forward selection search, which reduces the dimension of the feature set. Very specific contextual information is also extracted by recording the relative position of the manually segmented region respect to manually extracted fissures of the lung ([29]). The high cost of the manual delineation makes it impractical for regularly actualized large image collections.

Liu and Sclaroff [18] perform segmentation of blood cell micrographs based on deformable shape models. The image is over-segmented and candidate regions are then merged into whole objects by using a deformable shape model and the regions color compatibility. Finally, they perform population-based retrieval of

cells. For every image, an histogram of the shapes of the cells in the image is computed, and retrieval is achieved by shape histogram similarity. Paredes, Lehmann et al [23] represent each image by several square windows which may overlap, and use them as a set of local appearances for the image. A condition for robust retrieval using local appearance is the preservation of shape.

Regarding contextual information P. Liu et al [14] use the centers of mass of manually segmented objects and their geometric relations to deal with the contextual information in retrieval of magnetic resonance images (MRI) of the chest. Petrakis et al [25] use graphs, for retrieval of MRI images of the brain. Both 2D strings used by Chang et al [8] and graphs used by Petrakis [25] permit a flexible form of describing any type of image in terms of its structures (including their local attributes) and their spatial relations. The authors achieve in these works highly discriminative descriptors that can be indexed efficiently when the images have structures segmented manually, but not robust for automatic (non exact) segmentations.

In addition to having an appropriate feature space, it is highly desirable to have invariance to a particular family of spatial transformations determined by the application domain. For achieving this we can employ either invariant feature spaces or similarity measures that are not affected by these spatial transformations. This invariance is very important in medical images, as there is a high degree of shape and appearance variability inter and intra subject. Despite this fact, there are few works in retrieval that deal with this type of invariance. In medical image retrieval, Lehmann et al [10] use a distance measure invariant to small global transformations, and a distortion model that compensates more local deformations. This distortion model allows each pixel to be matched to any pixel of the destination image around a local neighborhood, without guaranteeing any regularity or topology preservation in the mapped object. Tagare et al [27] compute the similarity between pairs of shapes by first registering them. They forbid changes in topology by setting a set of constraints in the matchings, and solve the combinatorial problem by dynamic programming. Their method, however, is only valid for matchings between contours, and does not attempt to achieve any smoothness in the mapping. Y. Liu et al [20] make retrieval of 3-D CT brain volumes by first registering them, using a specific method based on the symmetric properties of the brain and an affine transformation, non invariant to local elastic deformations.

In this chapter we present a general framework of medical image retrieval based on plaque appearance, and consider a specific image modality: intravascular ultrasound images.

0.1.2 Medical image retrieval based on plaque appearance, application to IVUS

An IntraVascular UltraSound (IVUS) image is obtained by inserting a catheter into the artery with a transducer on its tip. The transducer emits ultrasonic waves as it rotates, and these waves are propagated according to the physical properties of the media the catheter is in. Depending on the echogenic impedance of some structure, and its proximity to the tip, the wave will be reflected with some amplitude and some delay, which allows to form a cross-sectional image of the artery: the delay accounts for the distance of the structure from the tip, and the amplitude accounts for the intensity of the gray level used for representing the structure in the image. Fig. 1 shows an IVUS image with two calcium plaque structures (regions of high level with a shadow behind), the catheter, and the usual adventitia tissue. The shadow behind the calcium plaques is due to the high impedance of the calcium, which does not allow the ultrasonic wave to pass through. We refer to [12] for an introduction on the topic.

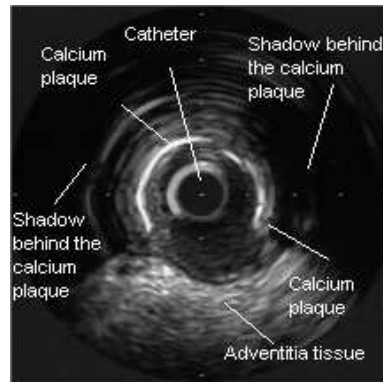


Figure 1: IVUS image with different structures

IVUS images of the coronaries are a novel and, at the same time, key tool in the correct diagnosis of coronary diseases such as atherosclerosis, which can provoke heart attacks. The great difficulty of these images makes it particularly

interesting for the physician to perform analysis based on a history of similar cases, which motivates the construction of CBIR systems of IVUS. This CBIR should take into account the spatial relative distribution of structures such as atherosclerotic plaques, as studies have shown ([21]) that spatial characteristics such as the size of these atherosclerotic plaques, their eccentricity, and degree of embracement around the vessels; play important roles in the study of heart diseases. Furthermore, the intrinsic elastic properties of the arteries make these objects have an extraordinary high variation intra-subject and inter-subject, which demands invariance to elastic transformations in retrieval similarity between IVUS images.

0.1.3 Context Based Plaque Retrieval

In contrast to the discussed approaches, the Context Based Plaque Retrieval method uses a feature space that integrates all types of information important in describing the content of the image: local, contextual and global information. This makes it appropriate to deal with medical images that hold complex information. For doing so we make a generalization of the correlograms to deal with images presenting different types of structures. Correlograms are histograms that take into account not only the statistics of the features in the image, but also the relative spatial distribution of these features. Using this generalization of the correlograms to conform our feature space has various advantages: First, it incorporates all the types of information mentioned above. Second, it does not need accurate segmentation/classification of the structures inside the image. This fact permits to make an automatic segmentation of the structures instead of making manual segmentation of the image, as opposite to the majority of descriptors used for dealing with this contextual information. Finally, the generalized correlograms permit easily to incorporate local information specific to the image application domain, which is mandatory in complex domains such as medical images [8],[25],[14].

For dealing with invariance to elastic transformations, an efficient elastic matching method is proposed, aligning each pair of images before their comparison. This is achieved by using a sparse set of landmarks placed around salient regions of the image, and computing a fast transformation such as the thin-plate spline based on matching between these landmarks. The thin-plate spline produces elastic alignments modelling both global and local deformations, and

restrains the object from undergoing unnatural deformations (e.g changes in topology) by performing regularization. In the feature space component of the registration, we also make use of the proposed generalized correlograms. This allows us to match structures attending not only to their local properties (the type of the structure), but also to the global attributes (such as their size) and their context. Furthermore, these correlograms introduce spatial coherence and remove ambiguities in the computation of the correspondences, which makes the algorithm achieve good solutions with few iterations.

0.2 Feature Space

As mentioned before, it is important that the feature space takes into account all types of information relevant to retrieve the image. Thus, we include local, global and contextual information, and we do so by using generalized correlograms.

0.2.1 Local information

By using local information we aim at describing the different types of structures inside the image. In the IVUS case, the discriminating structures are placed around the wall of the vessel [12]. A snake is placed at the center of the image after applying an anisotropic diffusion, and it is attracted to the wall. The set of landmarks is then obtained by sampling the snake (see fig. 2). Associated to each landmark, a local feature vector is computed that describes the type of structure where the landmark lies. Here is where we include specific information about our domain, as these local feature vectors must be specifically chosen for characterizing the structures of our particular medical domain [11]. In our IVUS case, we take the gray level profile along the normal to the wall at the landmark, in the direction from the landmark to the outward part of the vessel (fig. 2). This descriptor is specially appropriate for characterizing IVUS structures, see [1] for further details. Finally, these local feature vectors are classified and labels are assigned to each landmark, giving more compact local information. In our case non-parametric discriminant analysis [5] is performed followed by K-NN classification [1].

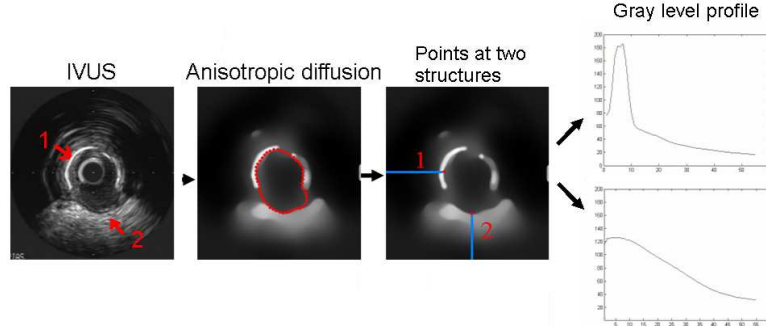


Figure 2: Landmark and local feature vector extraction

0.2.2 Global and contextual information

We incorporate this local information into a generalization of correlograms that allows to provide this information along with contextual and global information about the image. Correlograms are histograms which not only measure statistics about the features of the image, but also take into account the spatial distribution of these features. For doing so a spatial quantization of the points inside the image must be done. Regarding this spatial quantization we provide two different definitions of the correlogram: a bidimensional definition and a periodic unidimensional correlogram. The bidimensional definition takes the same spatial quantization of the shape context descriptor of Belongie et al [4]. The proposed generalized correlogram is then defined by adding a dimension to the correlogram. This dimension takes into account the types of structures of the landmarks around the one being described. Let C be a set of n landmarks, and $p_i \in C$ the current one being described. Let l_j be the label of the type of structure where the landmark p_j lies. Let n_c be the maximum number of classes. For p_i the correlogram h_i is defined as:

$$h_i(r, \theta, c) = \#\{p_j \in C, p_j \neq p_i : \|(p_j - p_i)\| \in D_r, \widehat{(p_j - p_i)} \in A_\theta, l_j = c\}, \quad (1)$$

where D_r is the r -th interval of radius: $r = 1 \dots n_r$, A_θ is the θ -th interval of angles: $\theta = 1 \dots n_\theta$, and c represents the class: $c = 1, \dots, n_c$. The sets of intervals $\{D_r\}_{r=1}^{n_r}$ and $\{A_\theta\}_{\theta=1}^{n_\theta}$ constitute a spatial quantization of the possible angles and possible distances of the relative positions around the current landmark. Thus, the first two dimensions of the generalized correlogram take into account the *relative* spatial position of the points (or landmarks), and the third

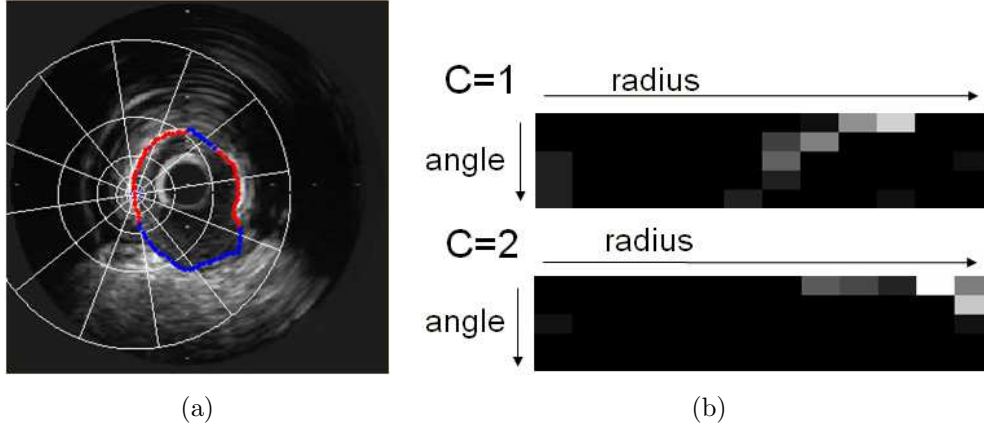


Figure 3: Bidimensional correlogram

dimension takes into account their type of structure. The correlogram can be interpreted then as an histogram measuring the density over relative positions of the different types of structures. The landmarks whose distance and angle relative to p_i lie at D_r and A_θ constitute a cell in the plane (see fig. 3-(a)). The size of these cells is incremented exponentially as the positions move away from p_i , so that more importance is given to local context. Fig. 3-(a) shows a bidimensional correlogram applied to a landmark of an IVUS image. The landmarks in this image have been classified into two classes: those belonging to calcium plaque (red points), and those belonging to adventitia (blue points). This correlogram has 12 intervals of angles ($n_\theta = 12$) and 5 intervals of radius ($n_r = 5$). Fig. 3-(b) shows a log-polar representation of the correlogram for each of the values of the third dimension: type of structure $c = 1$ and $c = 2$. In this plot, cells with a high density of points from a particular type of structure are represented by a high gray level.

This correlogram has scale and orientation invariance by normalizing the distances $\|p_i - p_j\|$ by the size of our object and orientating the correlogram along the tangent of the shape (see [4]). As said above, this bidimensional definition of the correlogram is a generalization of the one used by Belongie, by extending to take into account landmarks from different types of structures. At the same time, this spatial quantization defined by Belongie can be regarded as a generalization of the one used by Huang et al [15], as the latter takes only into account distances in the spatial organization.

The main disadvantage of this bidimensional quantization is that the resulting

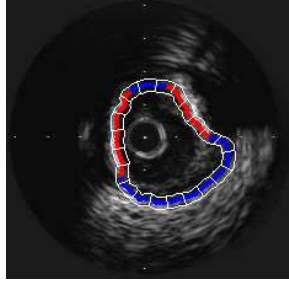


Figure 4: Unidimensional correlogram

correlograms are not robust against large shape changes of the object. This low robustness is accused before registering the images, however it can be avoided by using an appropriate feedback scheme (see below).

Still, we make another definition of a correlogram by taking another type of spatial organization, robust to shape changes. This second definition represents a change of manifold from all the plane to the closed curve where the landmarks are placed, as in our case all the landmarks lie along the wall of the vessel. Now the landmarks p_i are represented as the arc-length position inside the curve. Let $\Gamma : [0, 1) \rightarrow \mathbb{R}^2$ be the arc-length parameterized curve. The representation of the i -th characteristic point p_i is now taken as its arc-length parameter u_i inside this curve: $\Gamma(u_i) = p_i$, $u_i \in [0, 1)$. Let $C_\Gamma = \{u_i\}_{i=1}^n$ be the set of arc-length parameters corresponding to the set of landmarks C . The correlogram for the i -th characteristic point is defined as:

$$h_i(s, c) = \#\{u_j \in C_\Gamma, u_j \neq u_i : (u_j - u_i) \in I_s, l_j = c\}, \quad (2)$$

where I_s is the interval of arc-length positions of the s -th cell, $s = 1 \dots n_s$. By this definition, we are quantizing the possible values of arc-length differences between points of the curve. This difference must be computed with arithmetic modulus 1, in order to take into account the closed nature of the curve. Fig. 4 shows a diagram of the spatial quantization of the 1D correlogram. As can be seen the cells form a partition of the closed curve where the landmarks lie. The picture shows cells of equal size for clarity purposes, although the size of the cells is incremented exponentially from p_i to outwards, as occurred with the bidimensional correlogram.

The advantage of this new spatial definition is that it is robust against changes of shape. The contextual information is now taken along the curve, without

being affected by its changes in curvature.

The main advantage of using correlograms is that they are robust against miss-classifications, so that if some landmarks are missclassified the spatial distribution of the structures around p_i is still represented. They also allow to include specific descriptors representing the types of structures in an easy way. We only have to choose a good descriptor for the structures of a particular medical domain, and classify the landmarks according to this descriptor. Then the labels are included in the correlograms as explained before.

We also use generalized correlograms as a descriptor of the landmarks in the registration step. These contextual descriptors produce matching between structures attending not only to their type (local attributes) but also to their global attributes (such as their size) and their context. This solves ambiguities in the possible correspondences for each landmark. It also provides spatial information leading to regularity and coherence in the set of correspondences, accelerating the registration process.

In this work we provide results when using bidimensional and unidimensional correlograms. Both of them are valid for registering the images, although as we will see unidimensional correlograms are more efficient than bidimensional correlograms. The disadvantage of the former is that it is only valid when the landmarks form a closed curve.

All along we have defined the correlograms to measure the density over relative positions of the different types of structures around p_i . We can also define an auto-correlogram, as called by Huang et al. [15] which takes into account only how the points from the *same* type of structure are organized around p_i :

$$h_i(r, \theta) = \#\{p_j \in C, p_j \neq p_i : \|(p_j - p_i)\| \in D_r, (\widehat{p_j - p_i}) \in A_\theta, l_j = l_i\}$$

An analogous definition is done using an unidimensional definition as Eq.2. Auto-correlograms are used just for registration purposes. Given two sets of landmarks from two structures of the same type, using auto-correlograms we produce matchings between points in the same relative position inside the structure, e.g. matching extremum points together, middle points together, and so on. This auto-correlogram is used in a refinement step, once we have two structures of the same type and with the same context coarsely aligned. The aim is to make correspondences between homologous structures more exact and regular, discarding the information of the rest of structures.

Finally, we can use fuzzy definitions of the correlograms defined before. Until

now each point p_j lying in some cell adds 1 to the count of this cell, in the dimension $l_j = c$. We can add to the count in the dimension c the probability $Pr[class(p_j) = c]$. In this work, however, we do not use fuzzy definitions. In fig. 5 we can see a scheme illustrating the whole feature space construction.

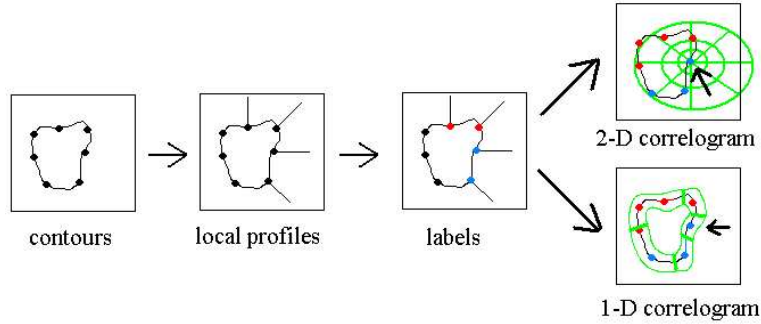


Figure 5: Scheme of the overall feature space design

Finally, regarding the similarity measure between correlograms, we use the χ^2 statistics to compare them:

$$\chi^2(h_i, h_j) = \frac{1}{2} \sum_{k=1}^d \frac{(h_i(k) - h_j(k))^2}{h_i(k) + h_j(k)}$$

where d is the dimension of the correlogram, if we treat it as a unidimensional vector (e.g. in the case of the definition of Eq. 1, we arrange the matrix of dimension $n_r \times n_\theta \times n_c$ into a vector of one dimension $d \times 1$, $d = n_r n_\theta n_c$). This distance acts like the Euclidean metric normalized by the number of points falling in the k bin of both histograms h_i and h_j . It has been used among others by Huang et al. [15], and Belongie et al. [4].

0.3 Registration: obtaining invariance against elastic transformations

We obtain invariance against elastic deformations through registering the images before their comparison. The scheme followed in the registration is the so-called point-mapping (fig. 6). First a set of landmarks is extracted from each image. The landmarks are described in some feature space, and a set

of correspondences is computed which globally minimize the distance between landmarks in this feature space. Finally, a transformation is obtained based on these correspondences. The proposed registration method also includes a search strategy of the final transformation. The feature space is the same as the explained in the previous section.

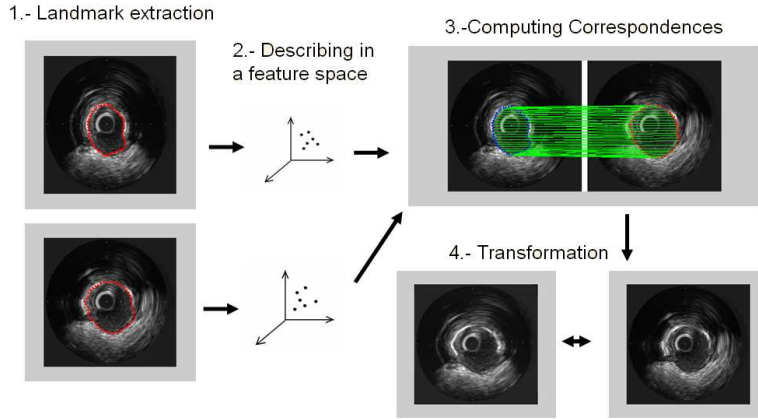


Figure 6: Scheme of the point-mapping paradigm

0.3.1 Computing correspondences

Once described the characteristic points in the feature space we compute the distance between any characteristic point in the image I_1 and any characteristic point of the image I_2 and based on this distance obtain our set of correspondences. We let for section 0.3.3 the expression of the exact distance. Let $d(i, j)$ be the distance between landmarks $p_i \in I_1$ and $q_j \in I_2$. We want to obtain a correspondence function $\phi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ that minimizes $\sum_{i=1}^n d(i, \phi(i))$. Such a function can be obtained by an assignment optimization algorithm such as the Hungarian's method [22].

0.3.2 Finding a transformation

Once we have our set of correspondences we want to find a transformation function $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, mapping the coordinates of I_1 onto I_2 . This function must map the characteristic points of I_1 close to their correspondent positions of I_2 , and produce a smooth mapping for the rest of points of I_1 . Any regis-

tration algorithm is characterized by the family of transformations to which it provides invariance. Biological bodies have a great deal of variability, making necessary the use of elastic transformations. Elastic matching methods allow to model global changes or transformations and local (elastic) deformations. Furthermore, they lead to smooth (regular) transformations avoiding changes in topology of the object. For doing so, the transformation must incorporate a regularization component.

In the present work we study the use of Thin-Plate Splines (TPS) as elastic matching method. TPS is an efficient method when based on a small set of landmarks, as the computation of the transformation is done by a closed-solution formulae which involves inverting one matrix of $n \times n$ elements, n being the number of landmarks.

A Thin-Plate Spline based transformation for an image can be expressed as $T_\lambda(x, y) = (f^x(x, y), f^y(x, y))$, where $f^x(x, y)$ and $f^y(x, y)$ are two independent surfaces obtained from the set of correspondences, and λ is the regularization parameter of the transformation. Let $\{x_i, y_i\}_{i=1}^n$ be the n landmarks of the origin, and $\{x'_i, y'_i\}_{i=1}^n$ the corresponding landmarks at the destination: $f^x(x_i, y_i) = x'_i, i = 1, \dots, n$ and $f^y(x_i, y_i) = y'_i, i = 1, \dots, n$. We explain how to obtain one of the surfaces, let $f(x, y)$ be such a surface. We call $\{t_i\}_{i=1}^n$ the known data for the function we want to interpolate, so that we restrain $f(x_i, y_i) = t_i, i = 1, \dots, n$. Hence, if the surface f to estimate is f^x , then $\{t_i\}_{i=1}^n = \{x'_i\}_{i=1}^n$. If the surface f to estimate is f^y , then $\{t_i\}_{i=1}^n = \{y'_i\}_{i=1}^n$. Then $f(x, y)$ can be expressed as:

$$f(x, y) = d_1 + d_2x + d_3y + \sum_{i=1}^n c_i k(\|(x, y) - (x_i, y_i)\|) \quad (3)$$

where k is the basis function or kernel: $k(r) = r^2 \log(r)$, $k(0) = 0$. The coefficients

$d_1, d_2, d_3, c_1, \dots, c_n$ are obtained solving a system of linear equations derived from applying the following restrictions: i) interpolation conditions (n equations): $f(x_i, y_i) = t_i, i = 1, \dots, n$; ii) in order for $f(x, y)$ to have square integrable second derivatives we require:

$$\sum_{i=1}^n c_i = 0, \quad \sum_{i=1}^n c_i x_i = 0, \quad \sum_{i=1}^n c_i y_i = 0$$

which produces a total of $n + 3$ equations expressed by the following matrix

notation:

$$\begin{pmatrix} K & P \\ P^T & 0_{3 \times 3} \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} t \\ 0_{3 \times 1} \end{pmatrix} \quad (4)$$

where $K_{ij} = k(\|(x_i, y_i) - (x_j, y_j)\|)$, the i th row of P is $(1, x_i, y_i)$, $c = (c_1, \dots, c_n)^T$, $d = (d_1, d_2, d_3)^T$, $t = (t_1, \dots, t_n)^T$ and $0_{3 \times 1}$ denotes a 3×1 matrix of zero elements. Let L be the $(n+3) \times (n+3)$ matrix of the last equation, a the vector of coefficients $a = (c, d)^T$ and $v = (t, 0_{3 \times 1})^T$. We need only to invert the matrix L for obtaining the coefficients of Eq 3, $a = L^{-1}v$, and as discussed in [26] this matrix is non singular.

The steps described above obtain an interpolating surface f , for $\lambda = 0$. If we want to obtain a regularized (i.e. approximating) surface, we just have to substitute in Eq. (4) the matrix K by $K + \lambda I_{n \times n}$, where $I_{n \times n}$ is the $n \times n$ identity matrix [2].

After applying TPS we can also obtain a measure of the bending energy this transformation has. Let a_x be the vector of coefficients a , defined above, for component x of the transformation. Let K be the kernel matrix defined also above, then the bending energy due to the component x of the transformation is:

$$E_x = a_x^T K a_x \quad (5)$$

Let E_y be the energy employed in the y mapping f^y , obtained in an analogous manner. We have as energy E of the whole transformation the sum of both components: $E = E_x + E_y$

Finally, the regularity of the TPS for different values of λ depends on the square of the scale of our objects. So if we have s as the size of our objects, we should compute $\lambda = \lambda_{norm} s^2$, and experimentally set values for λ_{norm} on normalized objects.

The λ parameter represents the tradeoff between approximation to the data (the computed correspondences) and smoothness. A high value of λ smoothes out irregularities in the correspondences, necessary in the first steps of the algorithm. At the same time, a high λ leads to a more global (coarse) transformations, whereas a low λ makes the transformation more elastic. The transformation can then accommodate more local deformations, refining the alignment. The regularity of the set of the correspondences depends on the stage of the algorithm. Initially they are quite irregular due to the complexity of the problem, and they become more and more regular through the feedback scheme we

will explain below.

Finally, the computed transformation also permits to produce a warping of image I_1 , making it more similar to image I_2 . This is necessary for taking into account the appearance similarity in the final comparison (see section 0.4).

0.3.3 Search strategy

Many registration algorithms [6] include some search strategy of the final transformation. We include this component through a simple yet effective feedback scheme performed in an iterative manner. The complexity of the images leads to irregular non-reliable initial sets of correspondences. However, the correspondences do obtain the correct global information to align coarsely the images. Thus, we introduce feedback of coarse transformations to obtain finer ones. The feedback is made by biasing the next set of correspondences towards the positions obtained by the last transformation. This is done by recomputing the distances between landmarks $p_i \in I_1$ and $q_j \in I_2$ as: $d_{ij} = d_{ij}^F + \alpha \|T_\lambda(p_i) - q_j\|$, where d_{ij}^F is the distance between landmarks in the feature space, and α is the degree of influence of the last transformation in the next set of correspondences. This feedback achieves a simultaneous maximization of the regularity of correspondences and the similarity of corresponding points throughout the iterations. The regularity is enforced by the term $\alpha \|T_\lambda(p_i) - q_j\|$, that represents the difference of the vector formed by the potential correspondence $u = (q_j - p_j)$ from a regular correspondence $v = (T_\lambda(p_i) - p_i)$: $\|v - u\| = \|T_\lambda(p_i) - q_j\|$. The similarity is enforced by the term d_{ij}^F . It can also be seen as a form of cooperation, as defined by Brown in [6]. In a cooperative scheme the correspondences computed for neighbor landmarks give information about the correspondence computed for the current landmark. Here the TPS mapping for the current point, $T_\lambda(p_i)$, depends on the obtained correspondences in its neighborhood, and we restrict a potential matching point q_i not to lie far away from this mapping.

The α and λ parameters representing the influence of the last transformation and the degree of regularization are updated throughout the iterations following an annealing scheme. The α parameter represents the confidence in the last transformation, whereas a low λ parameter represents confidence on the last set of correspondences (i.e. correspondences without many irregularities). As both the correspondences and the transformation are improved throughout

the iterations, we must increment α and decrement λ . Both are changed with an exponential ratio, see [1] for further details.

Finally, we can consider part of the search strategy the hierarchical scheme followed in the computation of global to local transformations. Furthermore, a hierarchical scheme is followed in the use of global to local information, as we use bigger correlograms for computing the first transformations (global alignments require more global information) and smaller correlograms for making more fine alignments.

0.4 Similarity measure in the final comparison between images

The registration produces as output a function T which is regular and maps the characteristic points p_i from I_1 close to their corresponding ones in I_2 . However, the mapped points are not exactly the characteristic points q_i of I_2 . For obtaining a regular final set of correspondences ϕ from $\{p_i\}_{i=1}^n$ to $\{q_i\}_{i=1}^n$, we simply take the Euclidean distances of mapped points and destination points: $d_{ij} = \|T(p_i) - q_j\|$ and compute the correspondences using the hungarian's algorithm over this matrix of distances.

Our similarity measure is based on the sum of three factors: the distance in the feature space described earlier, the amount of deformation necessary to align both objects, and a local appearance difference between the aligned image and the destination image.

The amount of deformation is measured through the energy E of the Thin-Plate Spline computed over the regular correspondences obtained before. A high energy means that the amount of deformation is too high for considering *natural* this transformation.

The distance in the feature space is in our case the distance of the correlograms. We can use either the 2-D correlograms or the 1-D correlograms. We have seen small difference between them in the results because the 2-D correlograms perform as well as the 1-D correlograms when the shapes are already aligned. As 2-D correlograms are more general, we used it. We recompute these correlograms orienting them now along the x axis of the image. This is done to reduce the little robustness that these correlograms have on the computation of the tangents, now that the alignment permits us to avoid the invariance to

orientation. The size of the correlograms is the maximum size in the hierarchical approach, so that all the image is included in every correlogram. Let d_{ij}^F be the χ^2 distance between correlograms of $p_i \in I_1$ and $q_j \in I_2$. We compute the global distance between both images, $d^F(I_1, I_2)$ as the symmetric sum over distances of best *matching* correlograms:

$$d^F(I_1, I_2) = \frac{1}{n} \sum_{i=1}^n \arg \min_j d_{ij}^F + \frac{1}{m} \sum_{j=1}^m \arg \min_i d_{ij}^F$$

For computing the local appearance difference between both images, let I_W be the image I_1 warped according to the obtained transformation from the registration. We take local windows around the mapped points $T(p_i)$ in I_W and matching points $q_{\phi(i)}$ in I_2 . The local appearance difference is expressed as:

$$d^A(I_1, I_2) = \sum_{i=1}^n \sum_{x=-w}^w \sum_{y=-w}^w G(\|(x, y)\|) [I_W(T(p_i) + (x, y)) - I_2(q_{\phi(i)} + (x, y))]^2$$

where $G(r)$ is a gaussian-like function of the radius r , more sensitive to close positions. The warped image I_W does not respect the original pattern of the textures, so it is better to remove them in the comparison. Therefore, we take as images I_1 and I_2 the anisotropic diffusion of the original images.

Finally, the total distance between both images is computed as a combination of the distance components defined above:

$$d(I_1, I_2) = \alpha^F d^F(I_1, I_2) + \alpha^A d^A(I_1, I_2) + \alpha^E E$$

This distance is used by Belongie et al. in [4], for binary objects. We apply it generalized for objects with several structures. The weights $\alpha^F, \alpha^A, \alpha^E$ are computed as the ones minimizing the classification error on the IVUS database, following a leave-one-out procedure.

0.5 Results

In this section we evaluate the system in three aspects: i) the rate of success of the local classification of tissues plaque/versus non-plaque, ii) the results of applying each of the components of the registration algorithm: the proposed feature space, the feedback scheme, and the overall registration results; and finally, iii) the overall retrieval results. All the experiments have been conducted

on a database of 100 IVUS images, all of them presenting plaque structures. Studying the registration of images presenting plaques is of high interest for the following reasons: first, there is a great difficulty in the differentiation between plaques and adventitia tissue, second, there is a high variability in the shapes of both the entire vessel and the plaque structures, and third it has been clinically seen that the relative spatial distribution of the plaques is important in diagnosis of heart diseases. This makes it interesting to study the use of the proposed contextual descriptors on these images in order to retrieve them. As explained in the introduction, retrieval of medical images is very interesting to perform computer-aided diagnosis. In particular, retrieval of IVUS images is useful to assist coronary diagnosis and intervention.

0.5.1 Local classification results

As explained in section 0.2, when building the feature space, we first obtain a set of local descriptors based on gray level profiles along the normal to the contour at this point, and then perform classification of these feature vectors, assigning labels to each characteristic point. We have designed a classification algorithm based on reducing the dimensionality by Non-Parametric Discriminant Analysis [5], and nearest neighbor classification. The scope is to achieve moderately good classification results. As the classification efficiency is different depending on the particularities of each image, we evaluate the efficiency of classification for all the landmarks in each image, and compute the mean efficiency and standard deviation of classifications across all the images. In fig. 7 we show the histogram of the efficiency of classification, the horizontal axis represents the efficiency of classification and the vertical axis the ratio of images that whose landmarks are classified with this efficiency of classification. The red line indicates the mean efficiency, which is 90.2%, and the green lines indicates the mean minus and plus the standard deviation. The standard deviation is of 7.8%.

We must note that the scope of this work is neither to design a very accurate local specific feature, nor to employ a sophisticated classification scheme over this local feature space. The results achieved in classification are good enough to work with the proposed correlograms for performing registration and retrieval. In fig. 8 we show some classification examples. Red points represent landmarks automatically classified as belonging to plaque structure, whereas blue points represent landmarks belonging to non-plaque tissue. With cyan

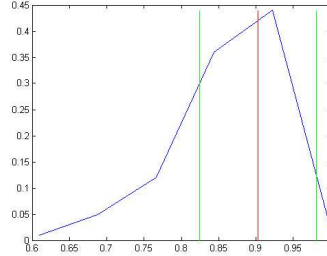


Figure 7: Histogram of the efficiency of classification

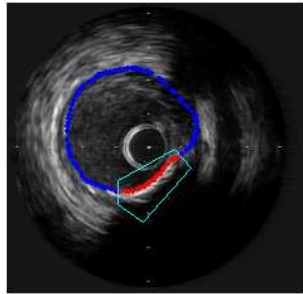
boxes we have indicated the correct regions where landmarks belong to plaque structure, in order to show the goodness of the automatical classification. The majority of images has good classification results, as in the case of figs. 8(a)-(b). The worst examples are shown in figs. 8(e)-(f). The main problem is that in some images compact regions are wrongly classified as belonging to plaque structure, and therefore, these regions are interpreted by the correlogram as structures (false structures). This is difficult to avoid because there are some regions of adventitia resembling much to plaque. These are regions of high gray level and with a black region behind it, which is what characterizes mostly calcium plaque. Figs. 8(e)-(f) show such regions.

0.5.2 Performance of the correlograms

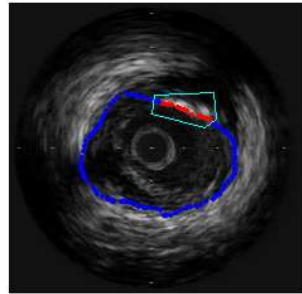
In the registration algorithm, the use of correlograms has several advantages: first, it allows to obtain more regular sets of correspondences, as it introduces spatial coherence information in the feature space. Second, it incorporates not only contextual information, but also global characteristics of the different types of structures. This produces matching of homologous structures *having the same global characteristics* such as size. We first show how correlograms obtain regular correspondences.

Fig. 9 shows a couple to be registered. Both images have only one plaque structure, the homogeneous white structure.

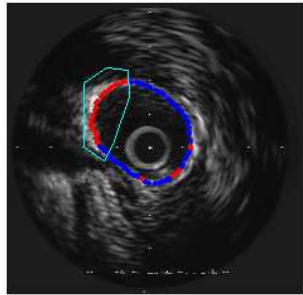
Fig. 10 shows correspondences obtained by different descriptors for the couple of fig. 9. The original landmarks are placed in a small region of adventitia in the left image of each pair. Fig. 10(a) shows the correspondences when using local feature vectors. The set of correspondences does not hold spatial regular-



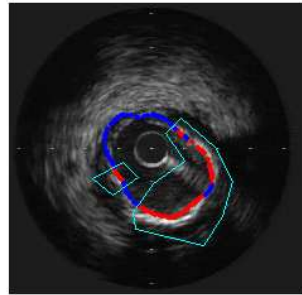
(a)



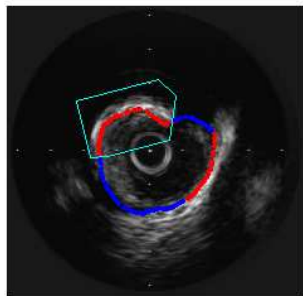
(b)



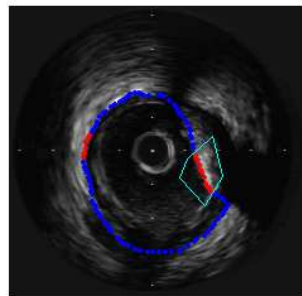
(c)



(d)

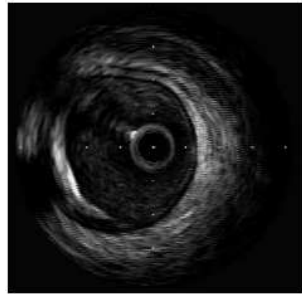


(e)

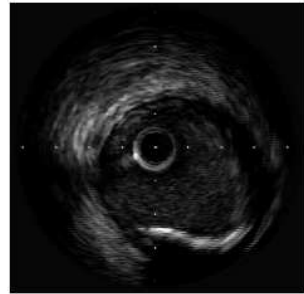


(f)

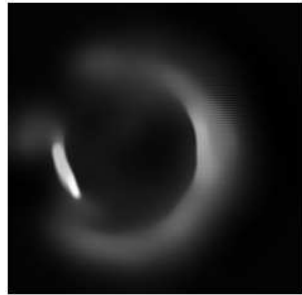
Figure 8: Some classification examples. Red points indicate classification as calcium plaque, and blue points as adventitia (the rest of the tissue). Correct calcium structures are indicated by green boxes surrounding the structure. (a-b) good results, (c-d) regular results, (e-f) bad results



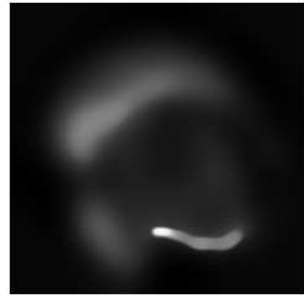
(a)



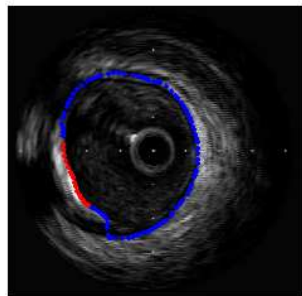
(b)



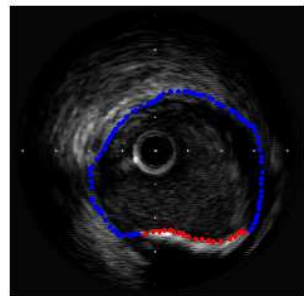
(c)



(d)



(e)



(f)

Figure 9: First couple example to be registered.(a) I_1 image to be aligned, (b) I_2 destination image; (c)-(d) anisotropic diffusions of (a)-(b) respectively;(e)-(f) Classification results for (a),(b): in red the landmarks classified as belonging to calcium plaque, in blue landmarks classified as belonging to adventitia tissue

ity as spatial information is not included in the descriptor. Fig. 10(b) shows the correspondences when using local gray level windows of size 19×19 . Taking local windows is not robust against shape changes, and gray-level windows are poor in IVUS. As can be seen, the set of correspondences is not regular. Fig. 10(c) shows correspondences using 2D correlograms. The regularity is not complete, but is much higher than with the other descriptors. Finally, correspondences using 1D correlograms are shown in fig. 10(d). The set has very few irregularities, and it is more regular than using 2D correlograms.

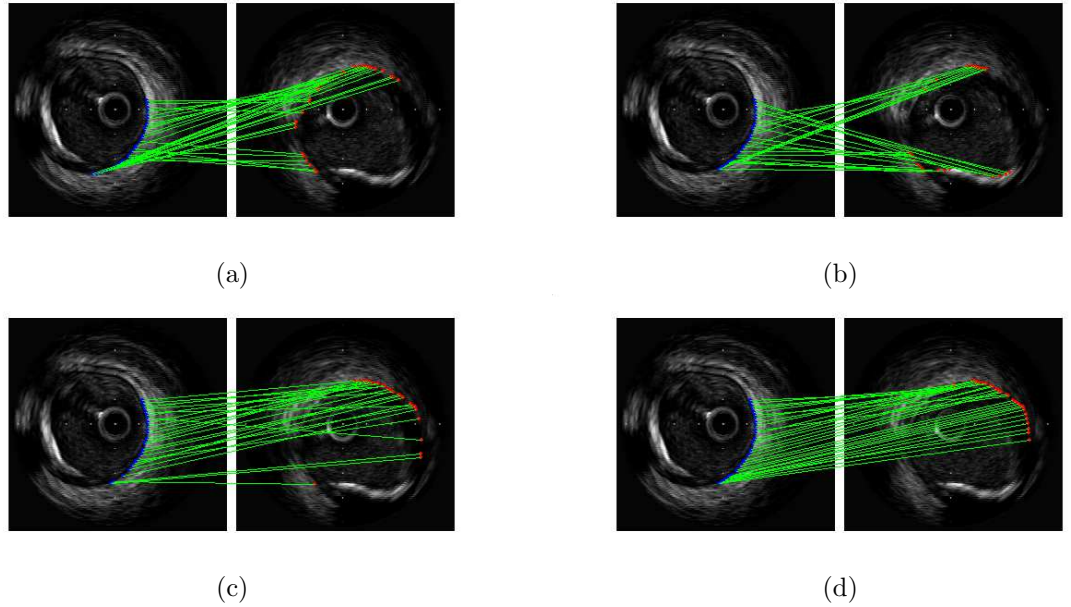


Figure 10: Correspondences using different descriptors (see text)

Now we see the effect of using correlograms for taking into account global characteristics of the structures. Figs. 11(a)-(b) show a new couple of IVUS images to be registered, (a) displays the image I_1 to be aligned and (b) the destination image I_2 . Figs. 11(c)-(d) show the anisotropic diffusion of (a)-(b) respectively. In red we superpose the contour of the vessel from which the characteristic points are extracted in each image. The image I_1 has two calcium plaques on both sides (indicated in the fig. 11(a)), and the image I_2 has three calcium plaques: two on both sides and one at the bottom (indicated in the fig. 11(b)). Taking into account global characteristics the plaques on both sides should be matched in both images, leaving alone the small plaque at the bottom of the

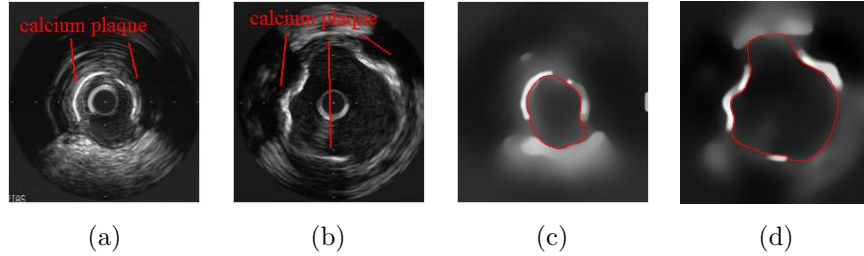


Figure 11: Couple of images to register (a)-(b). Their anisotropic diffusions (c)-(d)

image I_2 . We show how the global description is included with correlograms by comparing the result of a first coarse transformation using contextual information (2D correlograms) and then using only local information (the local feature vectors). We show transformation results on the anisotropic diffusion of the images because it is visually more clear. Fig. 12-(a) shows the warped I_1 when using correlograms. Fig. 12-(b) shows the image I_2 with the edges of the warped I_1 superposed in red. We can see how both calcium plaques of I_1 are mapped close to the big plaques of I_2 , as well as the adventitia tissue. Figs. 12(c)-(d) show the warping result when using only local feature vectors. One of the calcium plaques (indicated in fig. 12(e)) has not been mapped close to any of the plaques of I_2 , as the mapped plaque (indicated by a green arrow in fig. 12(d)) lies at an intermediate position between a big plaque and a small one (blue arrows). Using correlograms this matching is avoided as the size characteristic is included.

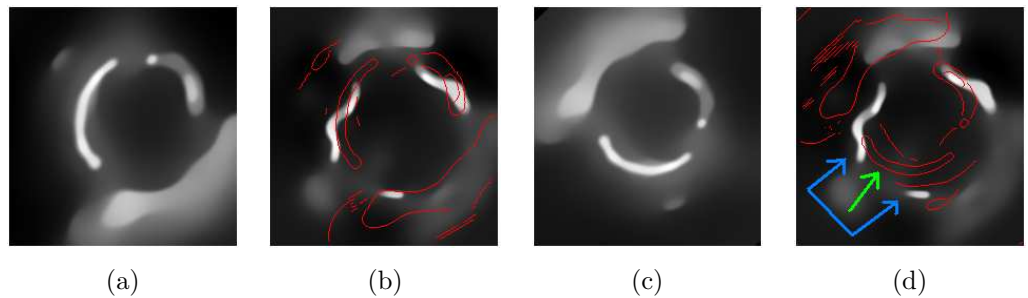


Figure 12: Coarse alignment (first step of the algorithm) using first contextual information (a)-(b), and then only local information (c)-(d).

0.5.3 Evaluation of the feedback scheme

The result of applying the feedback scheme is that the transformation becomes more and more approximate and at the same time the set of correspondences becomes more and more regular. This is illustrated first qualitatively for the pair of images shown in fig. 11, and then quantitatively by showing the evolution of the error through the iterations of the algorithm applied to 100 random registrations.

Fig. 13-(a) shows the initial set of correspondences in the registration of the pair of images of fig. 11, and fig. 13-(b) shows the final set of correspondences. The initial set is very irregular, but holds information about the correct global matching. Fig. 13-(c) shows the transformation based on this set, where the plaques are placed with the correct orientation and position (see green arrows in right image of (c)). Fig. 13-(d) and fig. 13-(f) show respectively an intermediate and final stage in the computed transformation. The intermediate transformation is coarse (rigid) but more approximated than the initial transformation, whereas the final transformation is accurate and elastic. Now we see a quantitative evolution of the approximation error and degree of irregularity of computed correspondences throughout the iterations of the algorithm. We have made 100 random registrations, and computed the median of the approximation error and irregularity at the successive iterations of the algorithm.

We measure the approximation error for an obtained transformation T as:

$$E_a(I_1, I_2) = \frac{1}{n_c} \sum_{i=1}^{n_c} \max \left(\frac{1}{n_i^1} \sum_{p \in C_1, class(p)=i} E_a^i(p, C_2), \frac{1}{n_i^2} \sum_{q \in C_2, class(q)=i} E_a^i(q, C_1) \right) \quad (6)$$

$$E_a^i(p, C) = \min_{q \in C, class(q)=i} (\|T(p) - q\|),$$

where I_1, I_2 are the registered images, n_c is the number of classes (types of structures) we deal with, n_i^1 is the number of landmarks from I_1 belonging to class i , and n_i^2 is the same for image I_2 , T is the obtained mapping, and $E_a^i(p, C)$ measures the distance of the mapped point p to points of C from the same class.

We measure the irregularity of each set of correspondences performing a TPS over this set and measuring its bending energy (see the 0.3.2). This measure is high whenever there are irregularities.

Fig. 14 shows a graphic of the median evolution of the approximation error

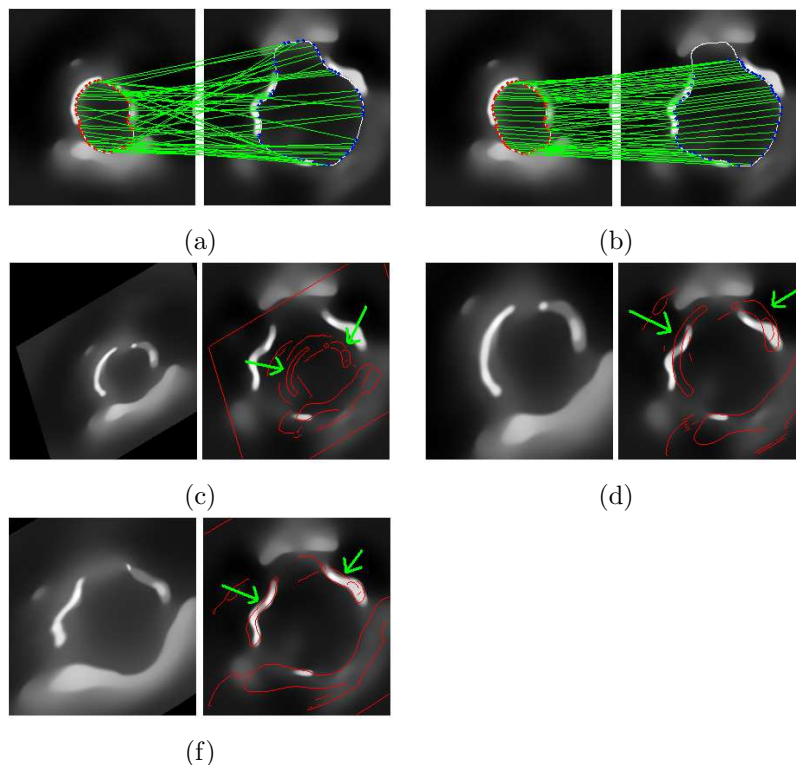


Figure 13: Evolution through the different stages of the algorithm

(fig. 14-(a)) and irregularity (fig. 14-(b)) of the set of correspondences throughout the iterations of the algorithm. We compute this median over 100 randomly chosen pairs of images to be registered. Each pair consists of two images from the same category. In blue dashed line we have this evolution when using bidimensional correlograms, and in red solid line with squares when using unidimensional correlograms. Both using bidimensional and unidimensional correlograms, the approximation error and the irregularity are decreased throughout the iterations. This is due to the simultaneous maximization, in the proposed feedback scheme, of the approximation and regularity (see section 0.3.3). Regarding the efficiency of using both types of correlograms, the unidimensional correlogram achieved better approximation errors and more regularity from the first steps, whereas the bidimensional correlogram needed around eight iterations to achieve accurate results. With unidimensional correlograms we needed just two iterations in order to achieve almost the same error than with

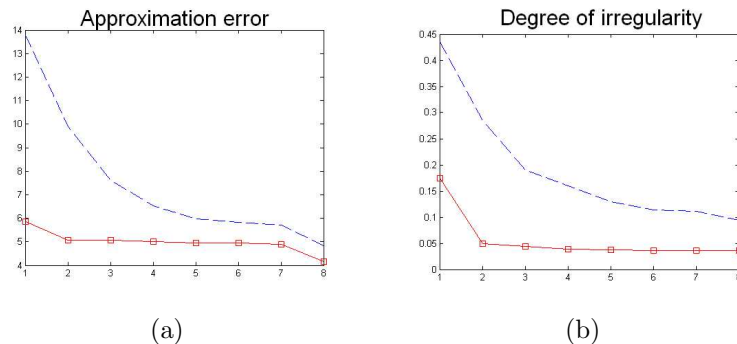


Figure 14: Evolution through the iterations of the approximation error (a) and degree of irregularity (b) using 1D and 2D correlograms. The solid line with squares display the 1D correlogram case, and the dashed line the 2D correlogram case

bidimensional correlograms. This is due to the robustness the unidimensional definition has against shape changes. However, both correlograms are valid for obtaining accurate correspondences with not many iterations.

Finally, we see how a registration algorithm such as the one employed by Belongie et al [4] without any cooperative feedback scheme can lead to poor results when dealing with the types of images we have. We take the same couple displayed in fig. 11, and only use the landmarks from one type of structure (calcium plaque in this case) for registration. We do so because the descriptor of Belongie is not suitable for different types of structures. Figs. 15(a)-(b) show the disposition of the mentioned landmarks in both images. We use Belongie’s registration algorithm with his shape-context descriptor rotated along the tangents of the points in order to achieve orientation invariance. Fig. 15-(c) shows the final set of correspondences. The points structures are not mapped close to their destination, and the final correspondences are quite irregular. Fig. 15(d) shows the transformed image I_1 according to these correspondences, and fig. 15(e) shows in red the edges of this transformed image superposed onto the target image. The warping is very irregular as it is based on irregular correspondences. The irregularity is due to the high shape difference in both objects, which arises the necessity to enforce step by step some regularity in the correspondences. If this regularity is not enforced, the resulting transformation will map the points without preserving the spatial coherence. Thus, correlograms are good for modelling contextual and global information (as shown

with the result in fig. 13), but only if we strengthen the spatial coherence of the mapping by some feedback algorithm such as the explained in section 0.3.3. Finally, fig. 16 shows the evolution of the approximation error (fig. 16(a)), and the evolution of the degree of irregularity of the correspondences (fig. 16(b)). They use 6 iterations in the algorithm. Both the approximation error and irregularity do not decrease throughout the iterations, as the correspondences become more and more irregular.

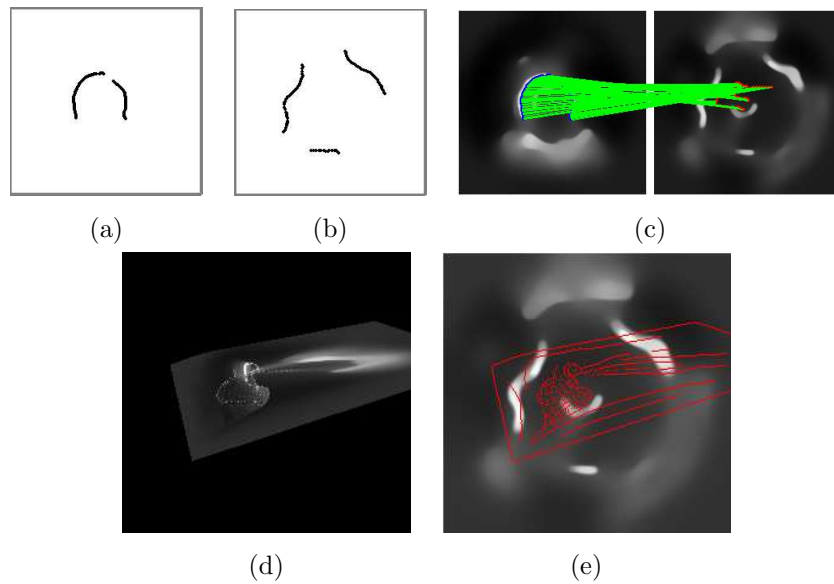


Figure 15: (a)-(b) Points from the plaque structures of fig. 11 (a)-(b). Final set of correspondences (c) and final transformation: warped query (d), and edges of warped query superposed on the complementary (e) obtained with Belongies's registration algorithm

0.5.4 Registration results

We see registration results now, using 1646 pairs of homologous images (i.e. from the same category) for computing the statistics. We obtain a mean approximation error of 4.6 pixels, a median error of 2.04 pixels and a standard deviation of 7.6 pixels. The mean distance between two neighbor characteristic points is of 3.1 pixels. Thus, the mapping is 1.5 times the distance between landmarks that are neighbors. Experimentally we have seen that errors below

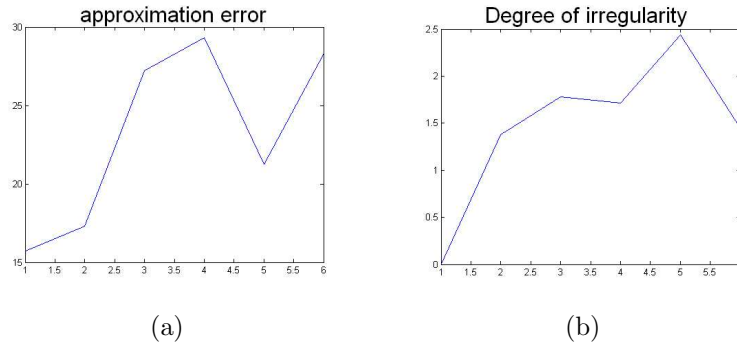


Figure 16: Approximation error (a) and degree of irregularity (b) through the iterations using the shape context matching from [4]

6 pixels are fairly good. 75% of the alignments have an error below 4.23 pixels. In fig. 17 we show the alignment with an error of 5 pixels, higher than the mean. Image (a) displays the couple, the left IVUS being the one to be aligned with the right IVUS. Red points are landmarks automatically classified as belonging to plaque structure, whereas blue points are classified as adventitia structure. In cyan we display polygons containing the correct manual classification of plaque structure. Note the high difference in the shapes of the two contours. Images (b)-(c) show the correspondences for the plaque structure, and image (d) shows the global alignment of the two contours: in blue the contour of I_2 and in red the mapped contour of I_1 .

Most of the error in alignment is due to the classification error. Although this error is not high (90.1% of success), the problem is that the points classified incorrectly can spatially concentrate. This makes false structures appear, as shown in fig. 18. Fig. 18(a) shows the automatic classification of landmarks for the couple of images (I_1 on the left and I_2 on the right), with cyan boxes indicating the correct classification of plaque structure. The classification of the landmarks from I_1 is quite correct, but in I_2 a false plaque structure is detected. Fig. 18(b) shows the computed correspondences. As can be seen, the plaque structure from I_1 is matched with the false plaque structure appeared at I_2 .

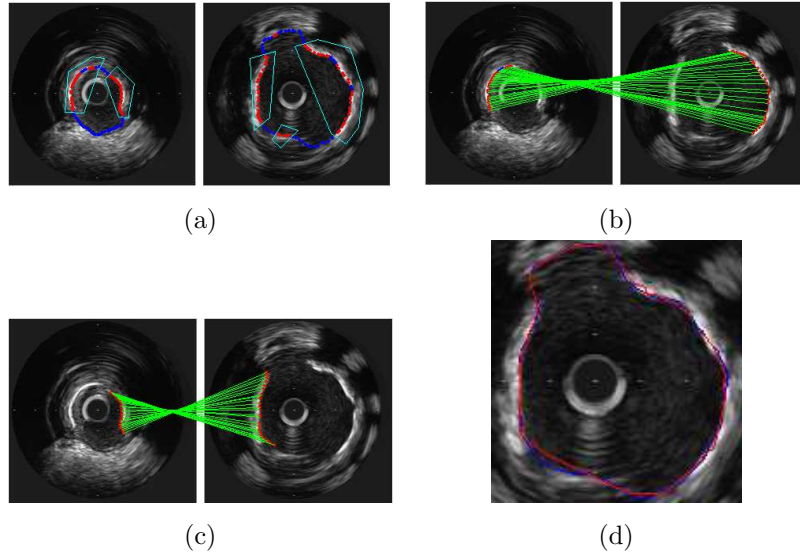


Figure 17: Registration with mean approximation error

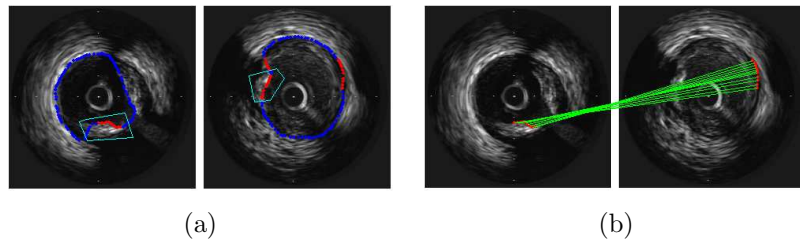


Figure 18: Registration with high alignment error due to the detection of false structures.

0.5.5 Retrieval results

For assessing the retrieval efficiency, a database of 100 IVUS images has been used. Two examples of these categories are illustrated in figs. 19 and 20. Fig. 19 shows a subset of IVUS from a first category. We can see that in all of them the plaque structure has a high degree of embracement around the vessel. Fig. 20 shows a subset of IVUS from another category. These IVUS present several plaque structures along the vessel.

In this manner, the IVUS images in the database have been classified by a group of physicians into several categories. The categories are chosen attending to clinical properties (see [1]). We extract each image from the database,

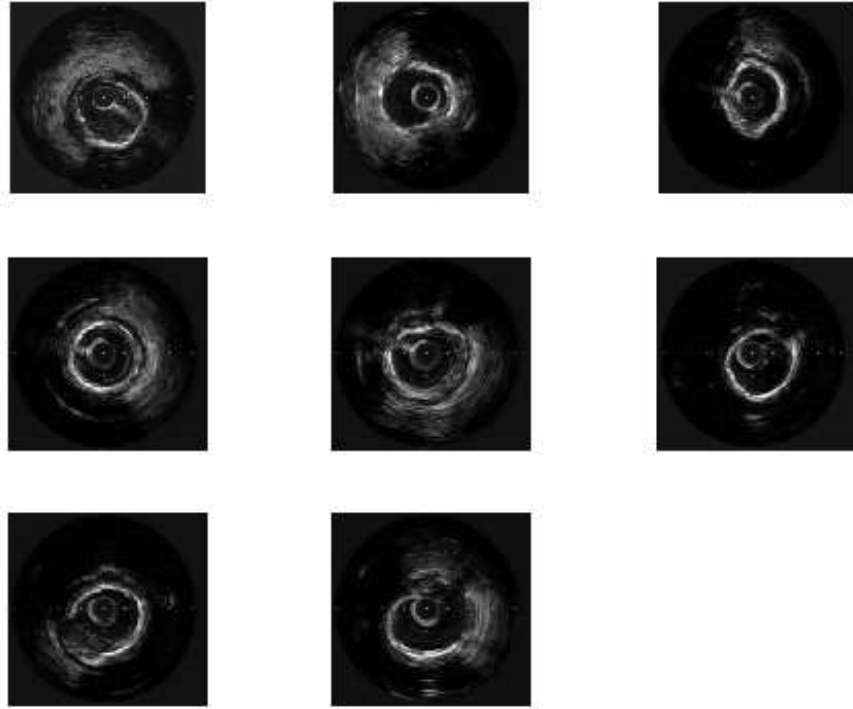


Figure 19: First example of a subset of images falling into the same category

present it as query, and the system orders the rest of the images from the database in order of similarity to the query. From this ordered list, we take only the first K images (i.e. the K most similar images to our query). Two measures of retrieval efficiency are taken. The first one is the estimated number of images we need to retrieve in order to include an image from the same category as the query. We obtained an average of 2.33 images necessary for including one of the same category. For $K = 2$ retrieved images the mean number of times in which an image from the same category is included is 89.7%. The second measure is the recall vs scope [15]: if query Q has N images from the same category, we compute for Q :

$$E^Q(K) = \frac{\#|I : rank(I) \leq K, category(I) = category(Q)|}{N}$$

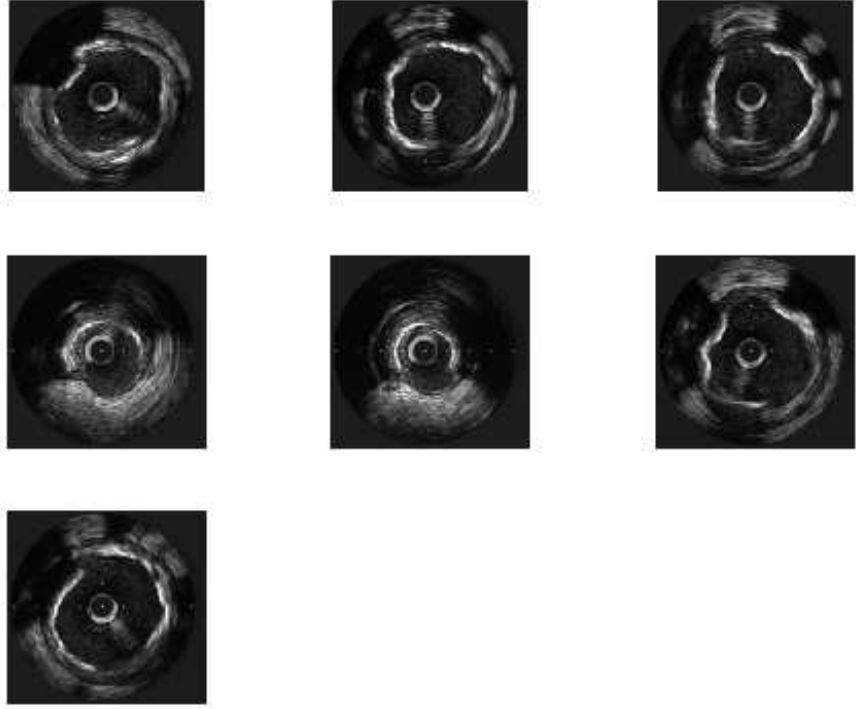


Figure 20: Second example of a subset of images falling into the same category

and average over all the queries presented:

$$E(K) = \frac{1}{N_Q} \sum_Q E^Q(K),$$

where N_Q is the number of query images presented to the database. We compare the efficiency of the Context Based Plaque Retrieval with the efficiency of the retrieval reported in Huang et al. [15]. They report results using two descriptors: color coherent vectors and auto-correlograms, both of them representing types of histograms that include spatial information. In table 1 we present the $E(K)$ measure for different values of K when using the proposed method versus the CCV, and the auto-correlogram. A color Coherent Vector consists of two histograms, one of non-coherent color pixels of the image, and

the other one of color-coherent pixels. A color coherent pixel set is a set of pixels that have a similar color and are connected spatially. Auto-correlograms are histograms measuring for some color and some distance the proportion of couples of pixels having this color and lying at this distance of each other. The results with CCV and auto-correlograms are obtained with the database reported in Huang et al [15]. Huang uses a general database, as CCV and auto-correlograms are not intended to be used in medical domains such as the presented. We only make the comparison for illustration purposes of the goodness of these numbers. As the mean number of images from the same category is 16.56, we show the numbers obtained with CCV and auto-correlogram for experiment on query 4 (see [15]). The efficiency of the proposed approach can be considered to be between the efficiency of CCV and auto-correlograms. In

K	CCV	color auto-correlograms	proposed method
10	0.19	0.38	0.22
30	0.38	0.63	0.50
50	0.38	0.75	0.73

Table 1: Recall vs scope measure

fig. 21 we see a result of three queries with high performance. In the three cases the first $K = 3$ retrieved images are all of the same category. In fig. 22 we see a result of three queries with low performance. Although in this example non of the retrieved images strictly belong to the same category as the query, the similarity of the retrieved images and the queries is quite high. For six of the nine retrieved images, the degree of embracement of the plaque in both the query and the retrieved image is the same. Most times the failure of the retrieval system is due to the interpretation of false structures (in the local landmark classification), as can be seen in fig. 18. Although the correlograms presented are quite robust against non perfect classifications, they can not succeed if an entire spatially coherent region is missclassified.

0.6 Conclusions

In this chapter, we presented a content-based image retrieval method that is able to deal with medical images of bodies with large shape and appearance

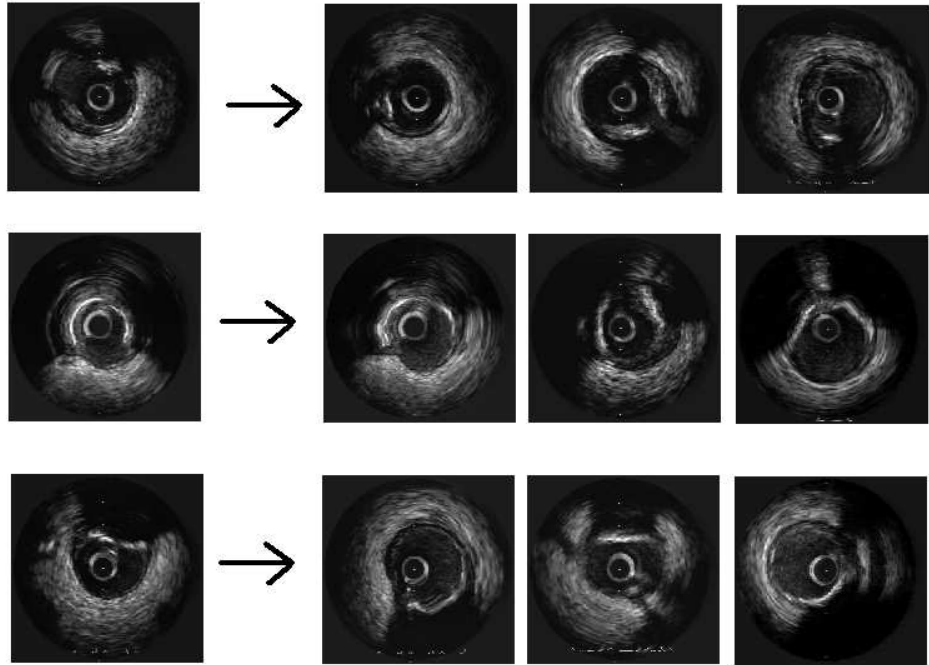


Figure 21: Example of optimal retrieval result for $K = 3$

variability. The proposed method uses a highly discriminant feature space that includes all types of information relevant to retrieve images: local, global and contextual information. This is achieved by generalizing the correlogram descriptor for dealing with different types of structures. The main advantage of the proposed generalization of correlograms is that they are robust against non-exact classifications/segmentations of the different structures inside the image. This allows automatic classifications of the structures. On the other hand, specific information can be easily introduced in this correlogram generalization. This makes the feature space designment suitable for different image domains. Specific information on the other hand is fundamental for dealing with complex domains such as medical image, as general features perform poorly in these domains. Finally, the comparison between images is made invariant to elastic transformations by means of registering the images before their comparison. The registration scheme is based on the point-mapping paradigm, the use of the same generalized correlograms, a thin-plate spline transformation based on few landmarks around discriminant regions and a feedback scheme which

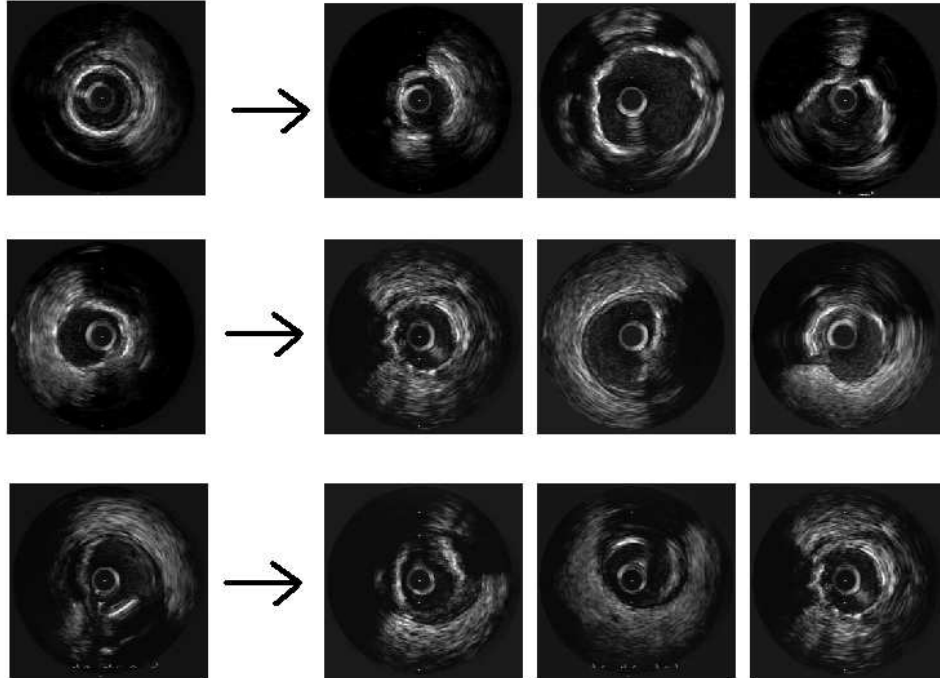


Figure 22: Example of worst retrieval result for $K = 3$

strengthens the regularity of the final transformation. Registration with TPS and few landmarks is very efficient as compared to other registration methods such as the use of the Navier-Stokes PDE [3].

There are two main lines of work for improving the presented retrieval system. In the particular IVUS domain, there is still work to do in the characterization of the different types of structures. We have seen that the main cause of error in the system is the wrong classification of entire regions in the image. In this sense, we could explore improved local descriptors. The inclusion of information along the longitudinal axis of the arteries is very important, as the physicians use this information for identifying the different types of structures. In general, considering the problem of medical image retrieval, we want to explore fast indexing techniques of the generalized correlograms by use of fast spatial access methods.

Bibliography

- [1] J. Amores and P. Radeva, *Elastic matching retrieval in medical images using contextual information*, Tech. report, CVC, September 2003.
- [2] Nur Arad., *Image warp design based on variational principles*, Ph.D. thesis, Tel-Aviv University, Tel Aviv, 1995.
- [3] R. Bajcsy and S. Kovacic., *Multiresolution elastic matching*, Computer Vision, Graphics and Image Processing (1989), no. 46, 1–21.
- [4] S. Belongie, J. Malik, and J.Puzicha., *Shape matching and object recognition using shape contexts.*, Tech. Report UCB//CSD-00-1128, UC, Berkeley, 2001.
- [5] Marco Bressan and J. Vitria, *Nonparametric discriminant analysis and nearest neighbor classification*, Pattern Recognition Letters **24** (2003), no. 15, 2743–2749.
- [6] L. Brown., *A survey of image registration techniques*, ACM Computing Surveys **24** (1992), no. 4, 325–376.
- [7] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, *Blob-world: A system for region-based image indexing and retrieval*, Third Int. Conf. on Visual Information Systems (Springer-Verlag, ed.), LNCS, 1614, 1999, pp. 509–516.
- [8] S. K. Chang, C. W. Yan, D. C. Dimitroff, and T. Arndt, *An intelligent image database system*, IEEE Transactions on Software Engineering **14** (1988), no. 5.
- [9] Y. Chen and J. Wang, *A region-based fuzzy feature matching approach to content-based image retrieval*, IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002), no. 9, 1252–1267.

- [10] J. Dahmen, T. Theiner, D. Keysers, H. Ney, T. Lehmann, and B. Wein, *Classification of radiographs in the 'image retrieval in medical applications system (irma)*, Procs 6 th Int. RIAO Canf on Content-Based Multimedia Information Access (Paris), 2000, pp. 551–566.
- [11] J. G. Dy, C. E. Brodley, A. Kak, L. S. Broderick, and A. M. Aisen, *Unsupervised feature selection applied to content-based retrieval of lung images*, IEEE TPAMI **25** (2003), no. 3.
- [12] Boston Scientific Europe. (ed.), *Beyond angiography. intravascular ultrasound: State of the art*, vol. 1, XX Congress of the ESC, August 1998.
- [13] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, *Query by image and video content: The qbic system*, IEEE Computer (1995), 23–32.
- [14] T.-Y. Hou, P. Liu, A. Hsu, and M.-Y. Chiu., *Medical image retrieval by spatial features*, IEEE Int. Conf. on Systems, Man and Cybernetics., vol. 2, 1992, pp. 1364–1369.
- [15] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih, *Image indexing using color correlograms*, Proc. CVPR., 1997, pp. 762–768.
- [16] F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Protopapas, *Fast nearest neighbor search in medical image databases*, Pr. of the 22nd VLDB Conference, 1996, pp. 215–226.
- [17] S. Y. Lee and F. J. Hsu, *2d c-string: A new spatial knowledge representation for image database systems*, Pattern Recognition **23** (1990), no. 10, 1077–1087.
- [18] Lifeng Liu and Stan Sclaroff, *Medical image segmentation and retrieval via deformable model*, International Conference on Image Processing (Thessaloniki, Greece), vol. 3, October 2001, pp. 1071–1074.
- [19] Yanxi Liu and Frank Dellaert, *Classification driven medical image retrieval*, Proc. of the Image Understanding Workshop, 1998.

- [20] Yanxi Liu, Frank Dellaert, William E. Rothfus, Andrew Moore, Jeff Schneider, and Takeo Kanade, *Classification-driven pathological neuroimage retrieval using statistical asymmetry measures*, Proceedings of the 2001 Medical Imaging Computing and Computer Assisted Intervention Conference (MICCAI '01) (Utrecht, The Netherlands), October 2001.
- [21] G. S. Mintz, J. J. Popma, C. J. Ditrano, J. Mackenzie, and L. F. Satler, *Intravascular ultrasound vs. quantitative coronary angiography: A statistical comparison of 538 consecutive target lesions (abstract)*, *Circ.* **88** (1993), 1–411.
- [22] C. Papadimitriou and K. Stieglitz, *Combinatorial optimization: Algorithms and complexity*, 1982.
- [23] Paredes, D. Keysers, T. M. Lehmann, B. B. Wein, H. Ney, and E. Vidal., *Classification of medical images using local representations*, *Bildverarbeitung für die Medizin*, 2002, pp. 171–174.
- [24] A. Pentland, R. W. Picard, and S. Sclaroff., *Photobook: Tools for content-based manipulation of image databases*, *Storage and Retrieval for Image and Video Databases (SPIE, ed.)*, 1994, pp. 34–47.
- [25] E. G. M. Petrakis and C. Faloutsos, *Similarity searching in medical image databases*, *IEEE Transactions on Knowledge and Data Engineering* **9** (1997), no. 3.
- [26] M. J. D. Powell, *A thin plate spline method for mapping curves into curves in two dimensions*, *Computational Techniques and Applications (CTAC95)* (Melbourne, Australia), 1995.
- [27] G. P. Robinson, H. D. Tagare, J. S. Duncan, and C. C. Jaffe, *Medical image collection indexing: Shape-based retrieval using kd-trees*, *Computerized Medical Imaging and Graphics* **20** (1996), no. 4, 209–217.
- [28] C. Schmid and R. Mohr, *Local grayvalue invariants for image retrieval*, *IEEE TPAMI* **19** (1997), no. 5.
- [29] C. R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, *Assert - a physician-in-the-loop content-based retrieval system for hrcr image databases*, *Computer Vision Image Understanding* (1999), no. 75, 111–132.

- [30] M.J. Swain and D.H. Ballard, *Colour indexing*, International Journal of Computer Vision **7** (1991), no. 1, 11–32.
- [31] James Ze Wang, Jia Li, and Gio Wiederhold, *SIMPLIcity: Semantics-sensitive integrated matching for picture Libraries*, IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001), no. 9, 947–963.