Chapter 5

Coronary model

This chapter deals with the building of a generic dynamic adaptive model of the coronary tree. A semantic network (graph) is proposed and developed to obtain a static structural model. Then, it is extended incorporating dynamic parameter of the coronary arteries using a statistical snake approach. A new minimizing scheme called *eigensnakes* is developed together with a contrast learning approach. The vessel shape, deformation and $3D$ reconstruction are detected in an energy-minimization framework applying the same active contour model (snake) technique.

5.1 Introduction

One of the key ideas of this thesis is to extract and use the information present in the coronary motion to provide a complete model including anatomical as well as dynamic information.

Some of the expected benefits and applications of the model are:

- The model can help to achieve a high degree of automation in the computer analysis/assessment of the coronary heart disease through angiography sequences. Parameters for assisting the image processing to enable a reasonable degree of accuracy in locating and to identify the arteries can be extracted.
- The dynamic assessment of the coronary tree enables the cardiologist to research on new parameters of the arteries and the myocardium. This is a new field of clinical research in the coronary vessel movement. The purpose is to use this model in order to achieve spatio-temporal understanding of the cardiovascular dynamics.
- In a second step, the model shall be incorporated into the workstation Care Cardiology to assist in coronary vessel analysis. It is important to say that the workstation has been evaluated in clinical daily work as part of the EU funded project TeleRegions mainly for image review and remote consultation. Therefore, there is a good perspective to maintain the workstation in use after some

changes suggested by the cardiologists, and so, the results achieved regarding the coronary model could be easily accepted by the users as a real added value.

In order to build the model, based on coronary angiography image sequences it is necessary to define a clear methodology to achieve the desired result. The work is presented following a bottom-up style, from low-level tasks upward to the final model.

- Image pre-processing: Whenever necessary the images are pre-processed to ease the following steps. The pre-processing is strongly dependent on the imaging devices. While modern digital equipment offers better quality images and hence needs a small amount of pre-processing, older equipment needs a previous analogue to digital conversion and the whole quality is often low. In such a case, some preprocessig is required, usually contrast enhancement, light correction, etc.. The sequences used in this thesis are coming from modern image equipment, so the preprocessing done is just a low pass filtering to obtain a smooth image where the snake deforms without discontinuities.

- $2D$ Vessel Detection: A correct vessel detection and segmentation is necessary to obtain a vessel description. A global vessel segmentation is achieved using snakes and probabilistic learning.

Then, a tracking technique is proposed followed by a $3D$ reconstruction.

- Movement Learning: Using the data from the vessel detection module for each pair of frames in a set of biplane sequences, a schema is proposed to obtain $3D$ vessel trajectories. A tracking method based in statistical snakes allows for the data collection of $2D$ vessel shape and trajectories.

- $3D$ reconstruction: Using biplane image sequences and the tracking method, a three dimensional reconstruction is carried out to obtain $3D$ data. Again, snakes are used in the reconstruction process to solve many known ambiguities.

Finally, the data structure to store the data is developed.

- Model building: After developing all the modules commented above, a data structure is necessary to hold the model data. A generic graph structure is developed and customized to keep the knowledge associated to the coronary tree model.

5.2 2D Vessel Detection

This section proposes the use of deformable models for vessel segmentation. The deformable models will be enhanced to cope with the drawbacks enumerated in section 3.2.7. When the objective is to extract boundary elements belonging to the same structure and integrate these elements into a coherent and consistent model of the structure, using deformable models is a natural way to regularize the ill-posed problem of segmentation and interpretation of images.

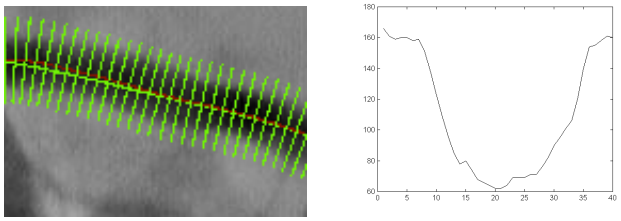
Conventional image feature detectors, either in a scanning or tracking strategy, commented in chapter 3, are too general for the purpose of vessel detection in angiography bringing too many false responses that makes difficult the image interpretation. Tracking strategy relies on simple densitometric features that are not enough to discriminate the vessel appearance. Moreover, the tracking strategy needs a continuous set of image features that is a too strong constraint for angiographies.

By its own nature coronary vessels angiography shows a flexible structure elongated and curve. The special features of the vessel structure lead, naturally, to consider the use of flexible models (snakes). Models have been used to guide pattern recognition systems, [50, 82] but all of them are dependent, sooner or later, on some kind of conventional edge detection and, so, subject to its properties. Snakes, [50] are of great interest since they are not rigid and have a physical explanation, where the energy principles are used to deform a curve under external forces coming from the image features (see chapter 3). The weak point is precisely those external forces, usually computed by an edge or ridge detector and in a posterior step translated to a distance map simulating potential energy for curve deformation. Level set based ridge/valley detectors [59] offer a very good response but they could be strongly affected by the cleaning steps because in noisy images it is not guaranteed the continuity of the detected ridge/valley, i.e. the detected feature can suffer discontinuities along its path. When many holes are present in the path and each consecutive segment length is shorter than the proposed threshold, one can loose a very long part of the desired feature. The immediate consequence of large gaps in image features of interest is the distraction of the snake usually to near, false segments coming from image noise. Figures 5.2 illustrate the problem. Summarizing, the method inherits the properties of the detector.

(a) Original image (b) Ridges (c) Ridges after cleaning

Figure 5.1: Image and crease detection.

To avoid the problems commented above a learning process allows obtaining an accurate statistical vessel description to assist the detection process. Such description has to be flexible enough to cope with vessel variability. With the knowledge accumulated in the learning process the external forces can be described through a probabilistic potential map [105].



(a) (b)

Figure 5.2: Learning of profiles perpendicular to the vessel direction (a) and a profile from the set of training samples (b).

5.2.1 A Probabilistic Framework for Structure Learning

The statistical approach is currently used in computer vision to add knowledge in different tasks from feature extraction to tracking [36, 17]. The first step is, usually, learning a feature of interest. For the vessel detection purpose, we propose two methods:

1. learning gray level profiles of the elongated structure.
2. learning contrast coherence of the image feature and of a set of derivatives.

Learning gray level profiles

A descriptor representation by local grey-level image profiles perpendicular to the vessel elongation and longer than the maximal expected linear structure width is defined (fig. 5.2).

To obtain the training samples, a "classic" snake [50] that deforms on a potential generated by a crease based distance map and corrected by the user (if necessary) is utilized. Each profile is defined by an image coordinate point (i, j) at the center of the structure and a direction $\mathbf{v}$ that indicates its orientation. The learning process is done by choosing the central points (i, j) and orientation of profiles $\mathbf{v}$ from the snake position.

Let U, V be vectorial spaces referred to the image coordinates (i, j) and to the orientations $\mathbf{v}$ respectively, where $(i, j) \in U$ and $\mathbf{v} \in V$. Then, we define a mapping:

$$\mathcal{T} : U \times V \rightarrow \mathbb{R}^d,$$

relating each extended coordinate pair and orientation $(i, j; \mathbf{v}) \in U \times V$, to a grey-level profile:

$$\mathbf{t}_n \in \mathbb{R}^d : \ \mathbf{t}_n = \mathcal{T}(i, j; \mathbf{v})$$
$$\mathbf{t}_n = \{\mathbf{t}_{n,k}\}_{k=1,\ldots,d,\ n=1,\ldots N}$$

By sampling the training set of images with a profile length d enough to cover the widest structure, we obtain a learning data set $D = \{\mathbf{t}_1, \ldots, \mathbf{t}_N\}$. All profiles are aligned in order to obtain as far as possible axial symmetry with respect to the

middle structure point. An interesting consequence of such an alignment is that the maximal probability shall be assigned to the real crease of the elongated object.

Actually, vessel profiles are characterized by a high degree of statistic regularity due to their morphological consistency.

Learning contrast coherence

The intrinsic image structure in a coronary angiography leads us to consider the contrast coherence as a parameter to reinforce an accurate vessel description.

Direction of grey-level variance (Structure detection)

The structure tensor field, applied to an integration region ρ, of the regularized gradient image (∇I_σ), under a suitable scale σ, measures the similarity between the regions and the searched structure [114].

Let K_ρ be a gaussian function as follows:

$$K_\rho(x) = \frac{1}{2\pi\rho} \exp\left\{-\frac{|x|^2 + |y|^2}{2\rho^2}\right\} \tag{5.1}$$

and I an image under analysis. The so called structural tensor is defined as follows:

$$J_\rho = K_\rho * (\nabla I_\sigma \nabla I_\sigma^T) \quad (\rho \geq 0) \tag{5.2}$$

$$J_\rho = \begin{pmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{pmatrix} \tag{5.3}$$

where the $*$ in (5.2) is the convolution operator, ∇ the gradient operator and I_σ an image previoulsy smoothed with a gaussian with standard deviation σ as follows:

$$I_\sigma(x,t) = (K_\sigma * I(.,t))(x) \tag{5.4}$$

The eigenvalues $\mu_{1,2}$ of the tensor J_ρ ($\mu_1 \geq \mu_2$) describe the average contrast variation in the eigenvectors $\mathbf{w}_{1,2}(\mathbf{w}_1 \perp \mathbf{w}_2)$. Equations (5.5) to (5.9) show the formulae to compute the eigenvalues and the eigenvectors:

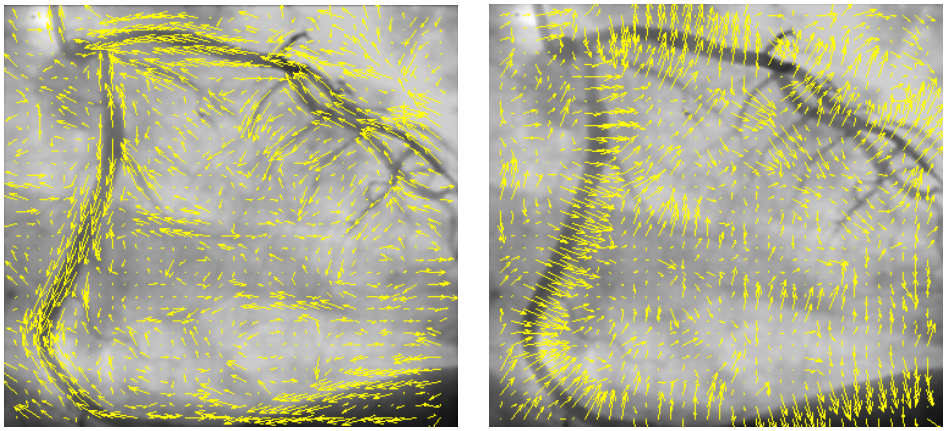
$$\mu_{1,2} = \frac{1}{2}(tr(J_\rho) \pm \sqrt{tr^2(J_\rho) - 4det(J_\rho)}) \tag{5.5}$$

$$tr(J_\rho) = j_{11} + j_{22} \tag{5.6}$$

$$det(J_{rho}) = j_{11} * j_{22} - j_{12} * j_{21} \tag{5.7}$$

$$\mathbf{w}_1 = \{\cos\varphi, \sin\varphi\} \tag{5.8}$$

$$\tan 2\varphi = \frac{2j_{12}}{j_{11} - j_{22}} \tag{5.9}$$



(a) (b)

Figure 5.3: Eigenvectors associated with the lowest (a) and the largest (b) eigenvalue.

The eigenvector $\mathbf{w}_2$ associated to the lower eigenvalue μ_2, is the orientation of the lowest fluctuation, detecting the vessel flow 5.3(a). The first eigenvector describes the direction of maximal grey-level variance, e.g. the direction coincident with the one used to learn the profiles, fig. 5.3(b). Figure 5.4 is another example: (a) is a left coronary angiography, (b) and (c) illustrate the first and second eigenvectors respectively. Not only the eigenvectors provide useful information. Areas with straight edges give $\mu_1 \gg \mu_2 = 0$, and corners give $\mu_1 \geq \mu_2 \gg 0$.

A normalized coherence measure of local image structure is obtained as follows [114]:

$$k = \frac{(\mu_1 - \mu_2)^2}{(\mu_1 + \mu_2)^2} = \frac{(j_{11} - j_{22})^2 + 4j_{12}^2}{(j_{11} + j_{22})^2} \tag{5.10}$$

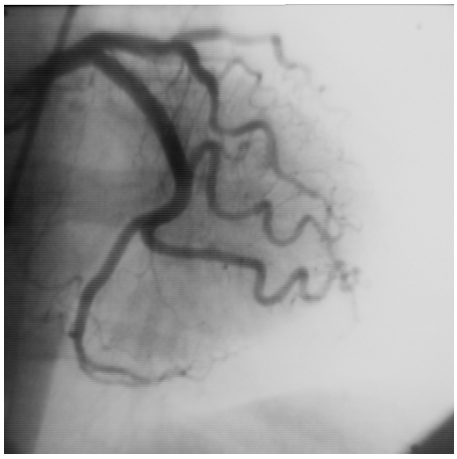
The potential discontinuity corresponding to flat image neighbours ($\mu_{1,2} \to 0$) is not a problem in our case because the equality is true only in areas where the image is flat (constant). The possible limit case $k \to 1$ because $\mu_2 = 0$ and $\mu_1 \to 0$ is eliminated by a next step through the likelihood.

The normalized coherence measure allows to focus on regions of interest with significant value of the coherence measure κ reducing, in this way, the computational cost of generation of a likelihood map for snake minimization.

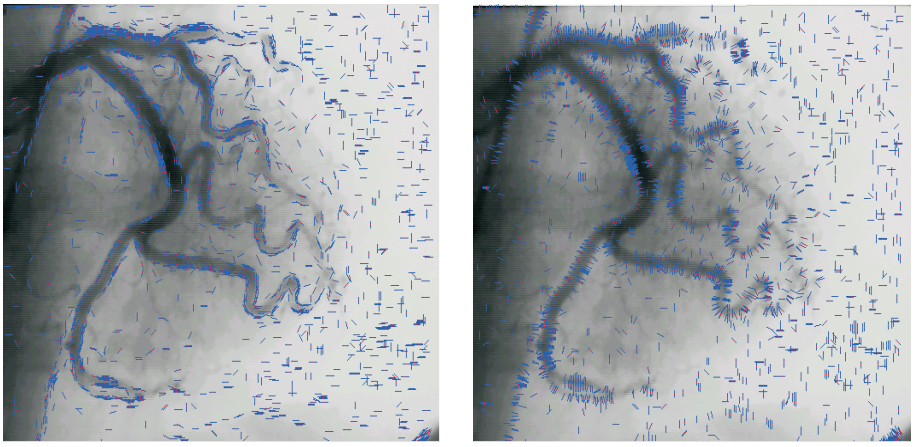
The coherence measure shall be used for:

1. The projection of gaussian filters along the direction of contrast coherence and sampling the filter responses at the features of interest (vessels).

2. To constrain the degrees of freedom of orientation profile vectors $\mathbf{v}$, since they can be searched from the vector field $\mathbf{w}_1$ with the biggest structural tensor eigenvalue (fig. 5.3(b) and 5.4 (c)).



(a)



(b) (c)

Figure 5.4: (a) Image under analysis. (b) First eigenvector. (c) Second eigenvector.

Derivative Projections

A bank of Gaussian derivative filters $\frac{d^k K_\rho}{d_x^{k1} d_x^{k2}}$ at different scales ρ is used to obtain the statistic vessel description (K_ρ is defined in (5.1)). Note that different scales are necessary to cope with the vessel diameter variability while using different derivatives allows us to generalize edge- crest- and valley detectors. Defining a mapping of the image pixels to the space of filter responses:

$$I \to \mathbb{R}^n(x,y) \to \mathbf{F}(f_1, \ldots, f_n) \tag{5.11}$$

Each sample f_i is a filter output μ_{kr} in a vessel pixel oriented by the eigenvector $\mathbf{w}_1$:

$$\begin{array}{lll} \mu_{k\rho}(x,y) & = \frac{d^k K_\rho}{d_x^{k1} d_x^{k2}} * \mu(x,y) & k = k_1 + k_2 \\ f_i & = \mu_{k\rho}(x,y) \cdot \mathbf{w}_1 & i = 1 \ldots n \end{array} \tag{5.12}$$

A matrix $\mathbf{D}_{m \times n}$ is built where each row is a sample along the vessel. Given a set of training points (fig. 5.5(a)) we get their filter responses $f_i, i = 1 \ldots n$ and construct the training data matrix. In fig. 5.5 (b) first derivative projection with r = 9 is showed. Fig. 5.5(c) depicts a training data matrix for the points in figure 5.5(a). The images were normalized to a gray level range $[0, 1]$. The vertical axis in figure 5.5(c) corresponds to the filter outputs, axis $\mathbf{x}$ represents the learned points (rows of $\mathbf{D}$)and axis $\mathbf{y}$ represent the filters applied to each vessel point (columns of $\mathbf{D}$).

5.2.2 Vessel detection embedding a Mahalanobis distance into the minimizing scheme: *eigensnakes*

The eigensnakes [107] comprises the learning, dimesional reduction and minimizing eschema.

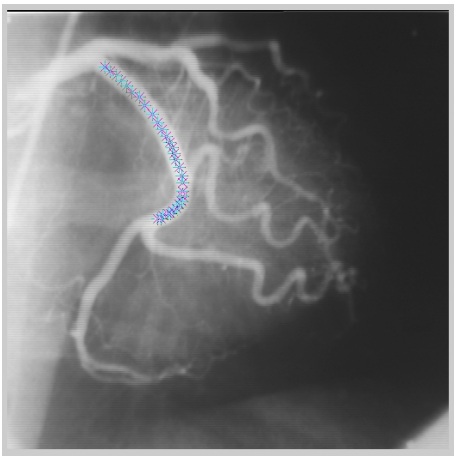
Dimensional reduction

The use of PCA in this work is mainly for searching for a few optimal linear combinations of filters representing the dimensionality of the feature space. The optimization in our case is in the sense of keeping the set of variables explaining almost all the variance of the data and discarding the less important ones. Using PCA, a dimensional reduction is carried out as follows:

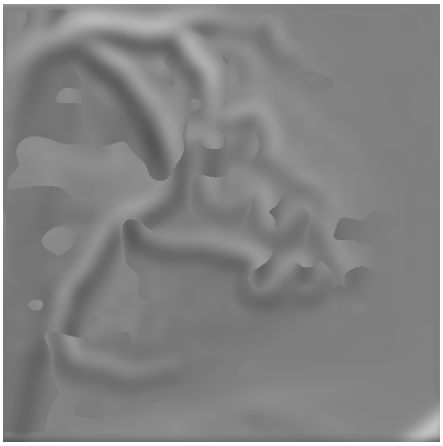
$$\Lambda : \mathbb{R}^n \to \mathbb{R}^l, \quad l < n, \quad \mathbf{f} \to y$$

To obtain the principal components, the eigenvectors of the covariance matrix $\mathbf{S}$ of $\mathbf{D}_{m \times n}$ are computed. The eigenvectors are sorted according to their associated eigenvalues (variances) and form the columns of a matrix $\mathbf{W}$.

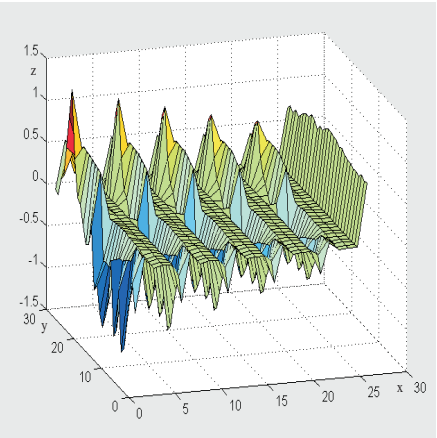
Such reduced space is used to measure the distance of an image feature to the learned ones. The measure can be regarded as a likelihood function giving the probability of each pixel belonging to a vessel category. Fig. 5.6(a) shows the training data projected onto the first two eigenvectors. Fig. 5.6 (b) shows all image features projected onto the first two eigenvectors, the center of the cluster contains the learned



(a)



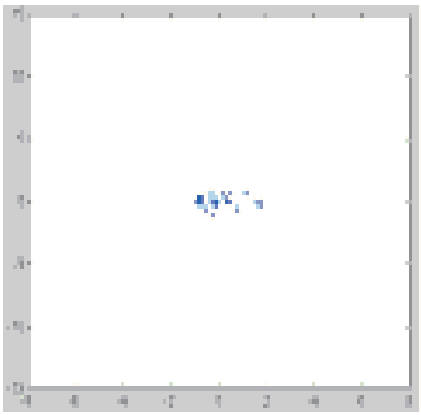
(b)



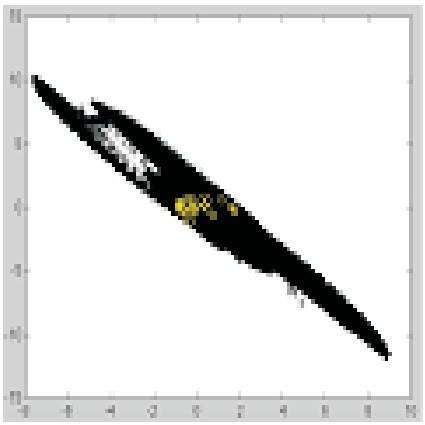
(c)

Figure 5.5: Training vessel samples (a). First Derivative directed by the first eigenvector of the image structure tensor(b). Learned data matrix (c).

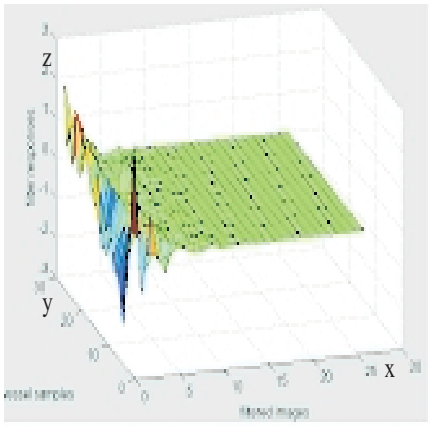
data of fig 5.6(a). Fig. 5.6(c) shows the learned data projected onto the principal component coordinate system after dimensional reduction.



(a)



(b)



(c)

Figure 5.6: Learned cluster (a). Image projection (b). Representation of training image features in the feature space determined by principal axes (c).

5.2.3 A probabilistic energy-minimizing scheme

The aim is to make the snake to be attracted by image features corresponding to the statistical description of the object. To this purpose, the external energy of the snake is defined as a function of the Mahalanobis distance of the projected image features x to the centre μ of the learned cluster. The Mahalanobis distance [61] is computed

in the reduced feature space, defined as follows:

$$E_{ext}(\mathbf{v}) = D_I^2(\mathbf{v}, \mu) = (WFI(\mathbf{v}) - \mu)^T \Lambda^{-1} (WFI(\mathbf{v}) - \mu) \tag{5.13}$$

where Λ is a diagonal matrix containing the eigenvalues of the training covariance matrix and $I(\mathbf{v})$ is a vector representation of the image neighbourhood around the snake pixels. Using (3.8) and the gradient of the probabilistic external energy in (5.13), we get a new energy minimising scheme for the snake:

$$-\frac{d}{ds}(\alpha \mathbf{v}_s) + \frac{d^2}{ds^2}(\beta \mathbf{v}_{ss}) + (cos\varphi, sin\varphi)\frac{dD}{d\mathbf{e}_1} = 0 \tag{5.14}$$

where $\frac{dD}{d\mathbf{e}_1} = (WFI(\mathbf{v}) - \mu)^T \Lambda^{-1} (W \frac{dF}{d\mathbf{e}_1} I(\mathbf{v}))$ and $\mathbf{e}_1 = (cos\varphi, sin\varphi)$ is the first eigenvector.

5.2.4 Vessel detection using feature profiles and Probabilistic Principal Component Analysis (PPCA)

In the PPCA framework [14, 15] a small number of causes are considered, that acting in combination generate the complexity of the observed data set. This leads to define a joint distribution $p(\mathbf{x}, \mathbf{f})$ over visible $\{\mathbf{x}\}$ and hidden variables $\{\mathbf{f}\}$, the corresponding distribution $p(\mathbf{x})$ for the observed data is obtained by marginalization:

$$p(\mathbf{x}) = \int p(\mathbf{x} \mid \mathbf{f}) p(\mathbf{f}) d\mathbf{f}$$

The main goal is to find the parameters that maximize the joint observed data distribution i.e. the best description under a specific generative model.

One of the basic tools is the standard factor analysis [8], which seeks to relate d-dimensional observed data vectors $\{\mathbf{x}_n\}$ corresponding to a set of q-dimensional latent variables $\{\mathbf{f}_n\}$ by a linear mapping (3.40) where latent variables are distributed into an isotropic Gaussian distribution, $\sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The noise model $\mathbf{n}$, or error, is considered also Gaussian such that $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Psi)$. The $(d \times q)$ parameter matrix Λ contains the factors loading, and $\boldsymbol{\mu}$ is a constant which, maximized the likelihood, corresponds to the mean $\boldsymbol{\mu}$ of the data. Given this formulation, the model for $\mathbf{x}$ is also normal $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$, with mean $\boldsymbol{\mu}$ and covariance matrix $\Sigma = \Lambda\Lambda^T + \Psi$. The corresponding probability density function is as follows:

$$p(\mathbf{x} \mid \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})} \tag{5.15}$$

Assuming uniformly distributed noise over the whole image and linearity assumption in (3.40) lead to the developing of a PPCA [14]. In this case Ψ endows with equal variance the principal axes (i.e. $\Psi = \sigma^2 I$). Hence, PPCA is a permissible technique when illuminant variations problem is not analyzed from variance structure. Considering this key assumption leads to consider the *conditional independence* of observed data. The underlying idea is that the dependencies between data variables $\mathbf{x}$ are explained by a small number of latent variables $\mathbf{f}$, while $\mathbf{n}$ represents the

unique variance of each observation variable. Instead, conventional PCA treats both variance and covariance identically. The corresponding distribution of observed data ($D = \{\mathbf{x}_1, ..., \mathbf{x}_N\} / \quad \mathbf{x}_n \in \mathbb{R}^d$) defines a normalized measure of distance, in a natural way, in terms of log-likelihood:

$$\mathcal{L}(\boldsymbol{\mu}, \Sigma) = -\log p(D \mid \boldsymbol{\mu}, \Sigma) \tag{5.16}$$

This measure called *normalized Mahalanobis* distance reduces the penalty on pattern differences along cluster major directions of data distribution, instead of the Euclidean distance [102]. At this point, the problem is centered on parameter estimation, which, in practice, will be given by data observations. This leads to consider the problem of *incomplete data*. For this purpose, Dempster et al. in [30] use the EM algorithm, where each observation $\mathbf{x}_n$ is associated to an unobserved state $\mathbf{f}_n$, and the main goal is to determine which component generates the observation. In this sense, the unobserved states can be seen as *missing data* and therefore the union of observations $\mathbf{x}_n$ and $\mathbf{f}_n$ is said to be complete data, $\mathbf{y}_n = (\mathbf{x}_n, \mathbf{f}_n)$. In this way the likelihood measure to be maximized is the *Complete-log-Likelihood*, i.e. $\mathcal{L} = \sum_{n=1}^{N} \log p(\mathbf{x_n}, \mathbf{f_n})$. Maximum-likelihood formulation for PPCA also allows a closed solution for the mapping matrix Λ and the noise variance σ [14]:

$$\Lambda = U_q \sqrt{(\Delta_q - \sigma^2 I)} \;\; R; \quad \sigma^2 = \frac{1}{d-q} \sum_{j=q+1}^{d} \delta_j \tag{5.17}$$

where U_q are the first q eigenvectors of the data set covariance matrix, Δ_q is a diagonal matrix with the corresponding first q eigenvalues ($\delta_i, \quad \forall 1 \leq i \leq N$) and R is an arbitrary rotation matrix.

Ought to the high number of pixels (samples) in any image, a dimensional space reduction by means of PPCA is used to statistically describe the feature.

Each sample is a grey-level profile and the covariance matrix S of the observed data is constructed from the learning data set:

$$S = \frac{1}{N} \sum_{n=1}^{N} (\mathbf{t}_n - \boldsymbol{\mu})(\mathbf{t}_n - \boldsymbol{\mu})^T, \tag{5.18}$$

where N corresponds to the profile population, $\boldsymbol{\mu}$ is the sample mean profile and $\mathbf{t}_n$ is the n-th grey-level profile.

The diagonalization of the covariance matrix S allows to reach the closed solution for the maximum likelihood estimation in (5.17). Hence, it provides the transformation between latent variables and observed data as a linear mapping $\mathbb{R}^q \rightarrow \mathbb{R}^d$, being q the latent space dimension and d the profile length, defined by the projection matrix Λ and the sample mean $\boldsymbol{\mu}$.

5.2.5 Feature detection

At this stage, the main goal is to build a local likelihood map for the snake given the image neighbourhoods projected onto the optimized space $\mathbb{R}^d$. The mapping is a probability measure of belonging to the learned structure category $\Omega = (\Sigma, \boldsymbol{\mu})$ for each

coordinate pixel (i, j) and any direction $\mathbf{v}$. Since the learned model is done through the gray-level profiles, the following functional composition to build the likelihood map $\mathcal{P}(\mathcal{T}(i, j; \mathbf{v}) \mid \Omega)$ is defined:

$$\begin{array}{lcccl} U \times V & \rightarrow & \mathbb{R}^d & \rightarrow & \mathbb{R} \\ (i, j; \mathbf{v}) & \rightarrow & \mathbf{t} & \rightarrow & \mathcal{P}(\mathcal{T}(i, j; \mathbf{v}) \mid \Omega) \end{array}$$

Computing this probability requires a factorization of the model covariance matrix: $\Sigma^{-1} = (WW^T + \sigma^2 I)^{-1}$. This can be done with low computational cost ($\mathcal{O}(q^3)$ instead of $\mathcal{O}(d^3)$) using the Woodbury's identity:

$$(WW^T + \sigma^2 I)^{-1} = \{I - W(W^T W + \sigma^2 I)^{-1} W^T\}/\sigma^2$$

In order to detect the learned feature, the image under analysis ought to be scanned searching profiles like the learned ones. Hence, the main problem arises when it comes to assign for each pixel (i, j) the profile orientation $\mathbf{v}$. However, due to the elongated structure of the objects it is not necessary to learn any curve shape or to scan the image considering all directions. Instead, the profiles are oriented with the direction of maximal grey-level variance of image neighbourhood by a structure tensor [114]. Since profile orientation $\mathbf{v}$ is fixed for each pixel (i, j), the profile extraction is done through the mapping:

$$(i, j; \mathbf{v}) \rightarrow \mathbf{t} = \mathcal{T}(i, j; \mathbf{v}).$$

Then the distance map from every profile to the learned model is obtained as a result of assigning its corresponding negative log-likelihood defined in (5.16) to each pixel (fig. 5.7(b)). The minimal probability computed in the regions of interest is assigned to all areas discarded by the threshold on the coherence measure κ (5.10). In fig. 5.8 bright regions correspond to the most probable crease points, i.e. corresponding to the minimum distance to the learned model. Dark areas refer to regions with a minimal probability of being part of a vessel (linear object).

5.2.6 Hybrid Snake Potential

The obtained likelihood map offers high responses (i.e. negative likelihood tends to -1 in image features of interest) at centrelines of elongated objects and nearby form strong slope towards the centreline (fig. 5.8). As a result, good convergence behaviour is observed once the snake falls in a neighbourhood of the elongated structure. Another advantage of the refined map is the small amount of false responses (potential points with high probability to represent learned feature and at the same time not belonging to an elongated structure). Therefore, far from the objects of interest the potential map consists of large flat areas with constant 0 likelihood. In these regions, a snake (initialised far from the objects of interest) does not suffer any external forces to deform it. It is more effective having a smooth slope towards the feature. To cope with these plain areas, the negative likelihood map is enhanced: $\hat{p}(\mathbf{x}|\Omega) = -p(\mathbf{x}|\Omega)$ obtaining a hybrid potential field as follows:

$$P_{hybrid}(x|\Omega) = \begin{cases} \hat{p}(\mathbf{x}), & \hat{p}(\mathbf{x}) \in [-1, 0) \\ d(\mathbf{x}), & \text{otherwise.} \end{cases}$$

where $d(\vec{x})$ is the distance to the closest potential point with non-zero probability. As a result, the snake moves with a constant small step towards the potential zones and converges only to the zones corresponding to learned image features. Fig. 5.7 illustrates an original image, the likelihood map according to (4), the hybrid map and a topographic map corresponding to the black square in fig. 5.7(c).

5.3 3D vessel Reconstruction using Snakes

In this section, the physics-based model to segment and reconstruct coronary vessels from biplane angiograms is explained. Using the snake technique to model the vessel, the snake deforms in space to adjust its projections to the image data. In this way, segmentation and $3D$ reconstruction can be unified into the same procedure assuring that only plausible vessel shapes will be detected. The method is general allowing to reconstruct the coronary tree from any angles and distances. The results are encouraging.

Traditionally, $3D$ vessel reconstruction techniques follow a bottom-up approach based on image feature extraction and reconstruction by interactively indicating corresponding point projections [10, 35, 113, 112] or by exhaustive feature matching [20, 93, 95, 115]. However, this approach can not cope with ambiguities in the image interpretation (e.g. in branch or intersection points) nor the lack of precision in determining corresponding image points. This approach proposes a top-down work based on a $3D$ deformable model (B-snake) of the vessels that deforms in an active way, adjusting its projections towards the image features. As a result, the spatial reconstruction and image segmentation are simultaneously obtained by detecting only plausible shapes for the vessels. These shapes are constrained by the physics-based model that implicitly incorporates spatial, continuity and structural constraints. In $3D$ vessel reconstruction techniques, spatial position of a vessel point can be recovered by calculating the intersection of projection rays to the corresponding points in biplane images [20]. In practice, due to systematic errors and difficulties in precisely determining corresponding points, a true intersection point does not exist. In these cases, the minimum distance reconstruction method is applied [35]. The approach followed in this thesis consists on optimizing the B-snake model, so that the distance between projection rays for each $3D$ point is minimized. In addition, in the B-snake approach all points of the vessels contribute to the $3D$ reconstruction instead of using a sparse set of points (usually, the bifurcation points). Furthermore, taking into account that the isocenter is not a sharp and stable point, its position is included in the optimization process to simultaneously refine global (related to the isocenter) and local (related to the vessel co-ordinates) parameters. The $3D$ angiography reconstruction approach by snakes has been proposed in [66] to recover $3D$ catheter paths in angiographic images. Here, this approach is extended refining the external force derived from image data, generalizing the method for any angles and number of views and including global (related to the isocenter) parameters into the optimization scheme.

Biplane Imaging Geometry

Standard biplane angiographic devices consist of two x-ray systems that can be rotated independently from each other. Two rotation axes are allowed: a horizontal axis in longitudinal direction with the table and a horizontal axis perpendicular to the table. The left-right movement of a system with respect to the patient defines the rotation angle. The angles by which a movement can be defined towards the head of the patient (cranial direction) or the feet (caudal direction) represent the angulation angles [35]. The projection axes of both systems intersect in the isocenter. The distances from the x-ray sources to the image intensifiers are predetermined before the image acquisition process. $3D$ reconstruction techniques require a global reference system. In our case, the isocenter coincides with the origin of the co-ordinate system (see 5.9).

The x axis is horizontally directed in longitudinal direction of the table, the y axis is horizontal directed towards the left arm of the patient and the z axis is vertically directed and upwards. The local reference systems are chosen as left-hand oriented, triple orthogonal systems determined as follows [35]:

$$\mathbf{k} = (0, \cos\alpha, \sin\alpha)^T \tag{5.19}$$

$$\mathbf{l} = (-\cos\beta, \sin\alpha, \cos\alpha \sin\beta)^T \tag{5.20}$$

$$\mathbf{c} = (\sin\beta, \sin\alpha\cos\beta, \cos\alpha\cos\beta)T \tag{5.21}$$

where α and β give the angulation and rotation angles. The vector $\mathbf{c}$ indicates the direction of the central beam of the x-ray cone while the vectors $\mathbf{k}$ and $\mathbf{l}$ are in the opposite direction of the image axes x and y. There are, also, a need to calibrate the measures to cope with the conversion between pixels and mm, and to obtain the origin of coordinates on the images; the process is quite simple and well described in [34]. In order to reconstruct a point in a $3D$ space from several projections, its image points should be identified. The intersection point of the projection rays from the x-ray sources to the image points defines the $3D$ point. Given the acquisition parameters (x-ray source position, isocenter position, angulation and rotation angles and calibration factor m defined by the ratio of the true size of an object in mm to its projected size in pixels), the exact $3D$ position can be found. Let us consider a point D and its projections $D^{'}$ and $D^{''}$ in two image planes. The intersection point of lines $F^{'}D^{'}$ and $F^{''}D^{''}$ determines the $3D$ position of the point D. However, in practice, both lines do not usually intersect (see fig. 5.10) due to the limited accuracy of the projection geometry and the calibration factor or the presence of geometric image distortion. Then, the spatial location of the point D is computed as the point of minimal distance to both lines by the back-projection expression for the point D as follows [35]:

$$\vec{OD} = \vec{OF'} + \tau\vec{F'D'} + \frac{1}{2}\vec{S'S''} \tag{5.22}$$

5.3.1 3D Reconstruction by Snakes

The equation (5.22) provides an estimation of the $3D$ position of vessel points given their corresponding image points. However, manually indicating each pair for all

vessel points is very tedious and time-consuming. To facilitate the $3D$ reconstruction, we utilize a snake that is represented as a $3D$ curve. This elastic curve is deformed in space to adjust its projections to the image data. In this way, the segmentation and the reconstruction processes are integrated and the $3D$ vessel position is directly obtained. At the same time, only plausible shapes are recovered that agree with the image data and the structural and smoothness constraint of the snake.

A snake is an elastic curve with associated energy:

$$\mathbf{E} = \int E_{int}(\mathbf{Q}(u)) + E_{ext}(\mathbf{Q}(u))du \tag{5.23}$$

where $\mathbf{Q}(u) = (x(u), y(u), z(u))$ is the spatial representation of the snake curve (see chapter 3). From its initial position, the snake is iteratively updated to adjust to the image data while maintaining a smooth shape. The convergence towards the image data is related to the definition of an external energy while the smoothness corresponds to the definition of an internal energy. The solution of the reconstruction process corresponds to a $3D$ snake that has a minimum energy, i.e. has converged to the closest set of image data that fulfills the smoothness constraint. The internal snake energy was defined in chapter two as a sum of the membrane energy that operates like springs attracting successive points of the snake and the thin-plate energy that prevents the snake from excessive twisting. The external energy of the biplane snake was defined so that it reaches a minimum when it is projected into the vessel projections in the images. To this aim, there are two approaches:

- create a potential field P corresponding to each image where the distance to the closest ridge point is assigned to each point of the potential field.
- use statistical techniques to build the potential field as explained before.

In this way, the external energy is minimal when the distance of each projected snake to the closest ridge is negligible.

Generalized $3D$ External Force

In its original formulation [66], the snake is a planar curve $\mathbf{Q}(x(u), y(u))$ that deforms in the image plane to match the object's contour. The snake energy is minimized by applying the Euler-Lagrange equation that leads to the following equation [66]:

$$Q^{t+1} = (A + \gamma I)^{-1}(\gamma Q^t + F_{ext}) \tag{5.24}$$

where $F_{ext} = -\nabla P(Q) = -(\frac{dP}{dx}, \frac{dP}{dy})(Q), \gamma$ is a damping parameter, A is a stiffness matrix, P is the potential. In our case, the snake is a $3D$ curve, in order to apply (5.24) its 3D external force should be defined based on the image external forces. We use the relations (5.19) between a local (k, l, c) and global (x, y, z) co-ordinate systems:

$$
\begin{aligned}
x_i &= -k_i = \cos\alpha_i y - \sin\alpha_i z \\
y_i &= -l_i = \cos\beta_i x - \sin\alpha_i \sin\beta_i y - \cos\alpha_i \sin\beta_i z \\
\nabla P_i &= \frac{\partial P_i}{\partial x_i} x_i + \frac{\partial P_i}{\partial y_i} y_i \\
&= \cos\beta_i \frac{\partial P_i}{\partial y_i} x + (\cos\alpha_i \frac{\partial P_i}{\partial x_i} - \sin\alpha_i \sin\beta_i \frac{\partial P_i}{\partial y_i}) y + \\
&\quad (\sin\alpha_i \frac{\partial P_i}{\partial x_i} + \cos\alpha_i \sin\beta_i \frac{\partial P_i}{\partial y_i}) z
\end{aligned}
$$

where i denotes the image view.

Taking into account the magnification due to the perspective projection, the magnitude of the image external force is re-escaled. Considering N projection planes, after a term reduction the following expression results for the generalized external force of the $3D$ snake from the N image potential fields:

$$
\begin{aligned}
F_x &= \sum_{i=1}^{N} \frac{\lambda_i m_i}{F_i D_i} \cos\beta_i \frac{\partial P_i}{\partial y_i} [|F_i O| + |O D^t|] \\
F_y &= \sum_{i=1}^{N} \frac{\lambda_i m_i}{F_i D_i} (\cos\alpha_i \frac{\partial P_i}{\partial x_i} - \sin\alpha_i \sin\beta_i \frac{\partial P_i}{\partial y_i}) [|F_i O| + |O D^t|] \\
F_z &= -\sum_{i=1}^{N} \frac{\lambda_i m_i}{F_i D_i} (\sin\alpha_i \frac{\partial P_i}{\partial x_i} + \cos\alpha_i \sin\beta_i \frac{\partial P_i}{\partial y_i}) [|F_i O| + |O D^t|]
\end{aligned}
$$

where λ_i is the potential weight of image i and D^t is the estimation of point D at iteration t.

Given an initial snake located near the vessel of interest, at each iteration the snake is projected into the image planes, the external forces from each potential field are estimated, $3D$ external force is generalized and the iterative equation (5.24) is applied. An example of vessel reconstruction of a left anterior descendent coronary artery is shown in fig. 5.11. Figure 5.12 shows how a modification of a control point affects the $3D$ curve.

5.4 Final Model

5.4.1 Graph

To represent the knowledge associated to the model a generic graph able to hold structural and dynamic data is designed.

Graph matching techniques [44, 45] have been successfully used in computer image analysis to label an extracted coronary tree after a segmentation process in angiographic images [34].

These techniques involve an anatomical representation of the coronary tree, a cost function and a segmented image to label. The labeling consist of minimizing the cost function matching the model against the image. Basically there are two ways for representing the coronary anatomy through a graph:

1. The nodes are related to the branching points and the arcs are related to the coronary segments

2. The nodes are related to the arterial segments and the arcs are related to the relationships between the segments

The attributes generally used to characterize the coronary segments have been distances, branching level, orientation, mean lumen diameter, position, and length [34]. Applying an object-oriented paradigm and using the template mechanism for the graph implementation is generic and efficient in many ways: the structure is a template class where the data objects and the internal graph implementation details are template parameters. Nodes and arcs are classes, both inherit from the same, common, base class, so the user of the graph can, by inheritance put any data in the graph. The user can decide where to keep the data objets (either in the nodes or in the arcs) at compile time. Being an object oriented approach, each node and / or edge can store not only data but behavior associated to it (objects). So the graph can hold both the anatomical and trajectories knowledge. The graph has been developed in ANSI C++ using the Standard Template Library as a base [38].

5.4.2 Data kept in the graph

The semantic network (graph), is able to hold both, anatomical and motion knowledge about a generic, average, coronary tree. The anatomical shape representation is achieved by interpolated cubic B-splines. The motion is represented as a sequence of $3D$ points keeping the average trajectory for each segment at any time. In our graph we add to the attributes described in [34], a sequence of $3D$ positions for each segment. Moreover, the points are control points to build the cubic B-splines. While the data in [34] helps to the necessary image segmentation process, the B-splines keep information about the movement and shape. The key idea is not to use point information but a curve to avoid many uncertainty problems otherwise present at later use of the model as a computer image analysis helper tool. The sequences stored in the graph represent the $3D$ trajectory of the coronary tree. At any time t, a set of $3D$ segment points modeling the tree can be obtained traversing the graph [38].

Figure 5.14 shows the coordinate system used for the $3D$ points. The origin is located at the ostium, which is the beginning of the Left Main artery, where r is the distance to the ostium, θ is the rotation angle and Φ is the angulation angle.

The graph can be edited to support any of the anatomies. Moreover, there can be differences in branches. Some of the secondary arteries can be absent in a patient. The graph editor cope with these differences. Figure 5.15 shows the user interface of the graph editor for the model. A user can add/delete any branch and choose a dominance. The application is able to cature the user actions and map them into the corresponding graph operations.

5.5 Summary and conclusions

In this chapter, the coronary model was explained. New techniques in the field of snakes were developed and used to make a $3D$ reconstruction, detect and track the vessels from coronary angiography image sequences. The data structure needed to hold the model is also explained.

5.5.1 Conclusions about vessel detection and tracking

The value of the statistic basis for linear structure detection and tracking has been established by demonstrating two methods:

1. the mechanism of PCA and Mahalanobis distances embeded into the minimization eschema of the snakes.

2. the mechanism of the PPCA embedded into the snake framework.

Snake-based tracking of image structures is based on the minimization of a functional energy term, which is usually created from the outcome of a standard feature detector. This fact constrains the applicability of the method and reduces the classes of image structures that can be analyzed. In order to manage complex objects and the variability of appearance of image structures, our techniques are supported by a learning approach to extract and detect only these "crease-like" features determined by the training set. Learned models are used in a probabilistic framework in order to build significant energy potential, resulting in less false responses of the image feature detector and more robust snake-based object tracking.

A new approach to potential computation using a likelihood map is formulated and applied to the tracking of specific structure on angiographies: coronary vessels as a difficult example of automatic analysis. As a result, the snake is less dependent on its initialisation and once placed on the hybrid potential map it converges to image features with high probability to represent learned object profiles. The obtained results and the self-training capability of the snake encourage utilizing it in different applications.

Building the likelihood maps for angioraphy image sequences, the vessels are tracked, obtaining data about the trajectory. The data is collected as a sequence of bspline control points. Finally the experiments carried out showed that there are no significative differences between the use of PCA or PPCA regarding the results (vessel detection). The PCA are faster to compute, meanwhile PPCA offers a probabilistic framework open to incorporate higher levels of reasoning.

Remarks about *eigensnakes*

The new formulation of the energy-minimising scheme allows statistically learning and detecting image features characterising different appearances of non-rigid elongated objects. Incorporating the statistical framework, the approach can be extended to the labelling task and to obtain the whole coronary tree using a likelihood matching.

Remarks about the tracker using the PPCA approach

The initialization process demands only one point to the user achieved by exploiting the coherence vector field and the likelihood function. Finally, the likelihood map is modified to avoid the uncertainty regions far from the detected feature using a distance function (5.2.6) to the high likelihood regions.

In fig. 5.16 the snake tracking is illustrated for a set of consecutive frames. Using the initial condition showed in fig. 5.7(a), the refined likelihood map guides the snake along the sequence.

Conclusions about 3D reconstruction

A snake model was applied to segment and reconstruct the coronary vessels in a semiautomatic way. The advantages of this approach are that no exact user-provided point correspondence is necessary. The snake evolves in the space to adjust to the image data; as a result the model provides such a correspondence between the image points. The results showed that the technique is optimal from the point of view of the minimal reconstruction error defined as the distance between the projection rays. Furthermore, the reconstruction is improved when isocenter co-ordinates are iteratively updated.