Algorithms for the Recovery of the 3-D Shape of Anatomical Structures from Single X-Ray Images

Riccardo Poli
School of Computer Science
The University of Birmingham
Birmingham B15 2TT
United Kingdom
Phone: +44-121-414-3739
Fax: +44-121-414-4281
E-mail: R.Poli@cs.bham.ac.uk

Guido Valli
Dipartimento di Ingegneria Elettronica
Università degli Studi di Firenze
Via Santa Marta, 3
Florence
Italy

August 1, 2000

1 Introduction

The recovery of the three-dimensional shape of anatomical structures is one of the most important problems in the field of medical imaging as the quantitative, computer-based assessment of such a shape and its changes plays an important role in clinical and research studies on a number of diseases.

A frequently studied class of solutions to this problem consists of performing the regional segmentation of a big enough sequence of tomographic (e.g. magnetic resonance (MR), computed tomography (CT) or echographic) images and stacking the segmented slices (sometimes with interpolation) to obtain volumetric representations of the 3-D shape of the imaged structures (Robb et al., 1983; Robb and Barillot, 1989; Higgins et al., 1990; Ylä-Jääski et al., 1991; Coppini et al., 1992a; Joliot and Mazoyer, 1993). Alternatively, detailed surface representations can be obtained from the boundaries of the structures of interest detected in each slice by means of surface fitting or interpolation methods (Brinkley, 1985; Azhari et al., 1987; Brevdo et al., 1987; Xu and Lu, 1988; Lin et al., 1990; Chang et al., 1991; Cohen, 1991; Cohen et al., 1992; Poli et al., 1994). Surface fitting approaches can also be used for the recovery of the approximate shape of an anatomical structure from a small set of tomographic or radiographic images when the boundaries of such a structure or other surface landmarks (obtained, for example, by matching fiducial points in different views) are available (Dumay et al., 1988; Calamai et al., 1990; Young and Hunter, 1989; Coppini et al., 1995).

The above-mentioned methods use the output data of a segmentation or a feature-matching algorithm as geometric constraints for the recovery of 3-D shape. However, in the case of X-ray projective imaging an additional, important source of 3-D information is available: the selective absorption of X-ray photons by the different tissues being imaged.

Such a source of information is used, for example, in computed tomography for the reconstruction of a density image from a complete set of projections on the ground of Radon's theorem (Rosenfeld and Kak, 1982). Unfortunately, when the number of projections available is small, image reconstruction becomes an extremely ill-posed problem that can be (approximately) solved only with the formulation of a set of strong assumptions about the structures in the scene (Rangayyan et al., 1985; Andersen, 1989; Klifa and Lavayssière, 1990).

A typical assumption formulated when a very small number (2–4) of projections is available is that only a single structure is present in the scene.[1] The additional hypotheses that the structure has a constant density, that the rest of the imaged volume is empty and/or that the cross sections of the structure respect predefined constraints allow the reconstruction of such a structure. For example, the assumption that most of the imaged volume is empty allows for 3-D reconstruction of vessels from a very small number of X-ray images (Stiel et al., 1993). In (Kitamura et al., 1988), the assumptions of elliptical cross section and constant density of vessels allow for the estimation of the shape

[1] It should be noted that, under the hypothesis of a single structure in the scene, the reconstruction of an image representing the density of a structure is actually equivalent to the regional segmentation of a cross section of such a structure and, thus, the reconstruction of several parallel slices is equivalent to instantiating a 3-D volumetric shape representation.

of coronary arteries from two X-ray projections (biplane angiograms). The hypothesis of elliptical cross section can be removed if other (less restrictive) assumptions on the characteristics of vessels are formulated (Tran et al., 1992; Weixue and YuanMei, 1993; Pellot et al., 1994). Convex symmetric cross sections with piecewise linear boundaries are assumed in (Chang and Chow, 1973) for heart reconstruction from two X-ray projections. Fewer restrictions but much more *a priory* knowledge on the expected shape of cross sections are used in (Onnasch, 1978) for the reconstruction of ventricular shape from biplane angiocardiograms. Regularity of ventricular cross sections with respect to the two projection directions is hypothesised in (Bai et al., 1989).

In most of the above-mentioned methods the space that does not belong to the structure of interest is assumed to be empty by hypothesising that such a structure is much denser than the other structures in the scene or that a kind of background subtraction can be performed. However, the first hypothesis is generally not valid for anatomical structures and background subtraction can be performed only for cave structures injected with iodine dye (in such a case, after linearisation, a pre-injection image can be subtracted from the post-injection one). Actually, most of the methods described in the literature have been applied to cardiovascular structures.

The strong assumptions described above limit quite severely the applicability of these methods to most anatomical structures of diagnostic interest. In addition, the requirement of having at least two projections rules out completely the most interesting possibility of recovering the shape of anatomical structures starting from single, conventional radiograms: the cheapest, most widespread kind of medical images.

In this chapter we describe a strategy to solve the problem of recovering the 3-D shape of anatomical structures from single X-ray images, i.e. the problem of *Shape from Radiological Density* (SFRD). In order to overcome the non invertibility of the process of image generation, we formulate a minimal set of physical assumptions, that are used to constrain SFRD and to transform it into a well-posed problem. The method presents the following features: a) only a single X-ray image is required (even if the method is adequate for any number of projections); b) shape recovery is performed by exploiting both geometric and densitometric constraints; c) the shape of more than one structure can be recovered, d) in addition to the structure(s) of interest, background and undesired structures can be present in the image, thus allowing also for shape recovery of a large class of non-cave structures; e) shape recovery is not performed on a slice-by-slice basis but all the input data cooperate in the instantiation of a single 3-D surface model; f) noisy, incomplete and inconsistent data are acceptable inputs.

Our shape recovery strategy involves four steps: a) linearisation of the process of X-ray image generation, b) image segmentation, c) estimation of a map of the local thickness of each anatomical structure of interest, and d) recovery of the 3-D shape of each structure from its boundaries and thickness map. The last two form our SFRD method.

The theory behind the estimation of the local thickness of each anatomical structure of interest and the recovery of the 3-D shape of each structure is described in Section 3. In Section 4 we provide more details on the practical implementation of such a theory. However, before our SFRD algorithms can be applied, linearisation and segmentation have to be performed.

While linearisation is relatively simple task for a given radiographic device, due to the presence of image noise, masking structures, biological shape variability, tissue inhomogeneity, imaging-chain anisotropy and variability, etc. the segmentation of medical images is a very hard problem.

To obtain reliable segmentation algorithms almost invariably researchers have been obliged to exploit as much *a priori* information as possible. For example, the statistical properties of the gray levels of the image is a kind *a priori* information that has been extensively exploited in the case of magnetic resonance and computed tomography images (see for example (Raya, 1990; Lei and Sewchand, 1992; Gerig et al., 1992; Amartur et al., 1992; Özkan et al., 1993)). Spatial correlation is at the basis of other methods, such as those based on mathematical morphology operators (Higgins et al., 1990; Joseph W. Klingler et al., 1988; Thomas et al., 1991; Joliot and Mazoyer, 1993), on rule-based expert systems (Catros and Mischeler, 1988; Manos et al., 1993; Li et al., 93), on special purpose computer vision techniques (Raman et al., 1993; Coppini et al., 1993; Deklerck et al., 1993) or on neural nets trained with the backpropagation algorithm (Silverman and Noetzel, 1990; Toulson and Boyce, 1992; Coppini et al., 1992b).

It is interesting to note that most of the recent literature on medical image segmentation is about methods for the segmentation of tomographic images. We believe this has happened for two reasons. Firstly, the segmentation of tomographic images is easier. In fact, tomographic images are usually less noisy than radiograms and tissues tend to have the same appearance independently of the slicing plane (for any given tomographic device). Secondly, computer vision algorithms for the segmentation of natural scenes can often be used as a basis for the segmentation of tomographic images, while these are much harder to use on radiograms. This happens because in the segmentation of a tomographic image the objective is isolating regions which represent single anatomical structures. This corresponds to finding a meaningful tessellation of the image into regions representing separate tissues. In radiography different structures are semi-transparent to X-rays. Therefore the gray level of each pixel may actually represent a number of different tissues superimposed. So, the segmentation of a radiogram may not be well represented by tessellations based on non-overlapping regions.

To overcome these problems we adopted a different approach inspired by biological vision (Poli and Valli, 1997a). In Section 2 we describe the approach and a neural architecture for segmenting radiographic images derived from it.

The experimental results of our segmentation algorithm and our method for SFRD are given in Section 5 while in Section 6 we draw some final comments.

2 Image Segmentation

Vision is ruled by principles, such as perceptual grouping, selection and discrimination, which mostly depend on regularities of nature such as cohesiveness of matter or existence of bounding surfaces (Marr, 1982; Reuman and Hoffman, 1986). As these properties are valid also for the anatomical structures present in medical images, they can be exploited to build segmentation systems for such images. If no other source of information is used, the resulting segmentation algorithms are independent of the acquisition parameters, of

the imaged district, etc. and, therefore, can be used for general-purpose radiographic-image segmentation.

Regularities of nature can be exploited in a very simple way by using grouping or discrimination criteria based for example on the idea that pixels which are close to each other and have similar gray levels have a higher probability of representing the same object than pixels which are far apart and have different gray levels. Therefore, such pixels should probably be grouped together. However, even if the strategy is simple, in order to design a segmentation algorithm for radiographic images a number of requirements must be met which can make the actual implementation of the strategy quite complex. Let us analyse these requirements:

- The segmentation algorithm should be maximally sensitive to small structures or to structures with a low contrast (possible lesions or tumours in early stages).
- The algorithm should be maximally robust with respect to the noise, texture and slow intensity changes typically present in medical images.
- The algorithm should take the physics of image generation for X-ray projective images into account.
- An algorithm to be used to segment large images such as those resulting from the digitalisation of radiograms (which can easily produce images with up to 10 million pixels) should be suitable for parallel, high-speed implementation.

The first two requirements counteract each other and, therefore, any segmentation algorithm can only produce results that represent a trade-off between them. In order to achieve optimum compromises it is first necessary to define a quantitative criterion of goodness of segmentation which takes sensitivity and robustness into account, and then to optimise it for any specific image. Therefore, the problem of radiographic image segmentation can be seen as a problem of combinatorial optimisation.

Unfortunately, for any given image the space of possible solutions to this optimisation problem is huge and conventional optimisation techniques tend to fail on it. Therefore, following recent approaches in the field of natural scene segmentation (Darrell et al., 1990; Reed, 1992; Wang et al., 1992), we decided to solve it by using an architecture based on continuous Hopfield neural nets, a computational paradigm which can effectively search huge solution spaces. We recall the properties of such networks and describe the steps necessary to use them for X-ray image segmentation in the following sub-sections.

2.1 Continuous Hopfield Neural Networks

Continuous Hopfield networks (Hopfield, 1984) are recurrent neural nets ruled by the following motion equation

$$\begin{cases} \frac{du_i}{dt} = \sum_{j=1}^{N} T_{ij} v_j - \frac{u_i}{\tau_i} + i_i \\ v_j = \frac{1}{1+e^{-\frac{u}{u_0}}} \end{cases} \tag{1}$$

where v_i is the output of neuron i, i_i is its external input, u_i its net input and T_{ij} is the weight of the connection from neuron j to neuron i. In the case of symmetric connections, the equation is stable and an energy function E_{net} exists, the minima of which are the stable states of the net. In the case, considered here, of $u_0 \gg 1$, such states are approximately the vertices of the hypercube $[0,1]^N$ and E_{net} can be approximated as:

$$E_{net} \approx -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij} v_i v_j - \sum_{i=1}^{N} i_i v_i.$$

Thanks to their minimum-seeking dynamics, Hopfield networks can be used to solve optimisation problems (Hopfield and Tank, 1985; Hopfield and Tank, 1986). The basic strategy is as follows: a) to pre-process, when needed, the input data, b) to find a binary representation for the solutions of the problem so that they can be mapped into the stable states of the neurons of a Hopfield net, c) to define a quadratic (symmetric) energy function whose minimisation leads to an optimum solution of the problem and then calculate weights and external inputs, d) to initialise and let the network relax into a stable state to be then mapped back into a solution for the original problem.

In the following we describe how these steps, applied to the problem of radiographic image segmentation, lead to an architecture that not only provides the optimum sensitivity/robustness trade-off but also meets the other requirements listed above.

2.2 Preprocessing

In the hypothesis of orthographic projection, the process of formation of X-ray projective images is approximately ruled by the following equation (Heintzen, 1971; Heintzen and Bürsch, 1978; Macovski, 1983)

$$N_o(x,y) = N_i(x,y) \exp\left(-\int_0^{s(x,y)} \mu(x,y,z)dz\right)$$

where $\mu(x,y,z)$ is the linear absorption coefficient of the tissue at coordinates (x,y,z), $s(x,y)$ the thickness of the body in (x,y), $N_i(x,y)$ is the number of X photons entering the tissue in $(x,y,s(x,y))$, and $N_o(x,y)$ is the number of photons that exit such a tissue in $(x,y,0)$.

The input data of our segmentation algorithm is a 2-D X-ray image whose gray levels are approximately related to $N_o(x,y)$ by a linear equation. The objective of preprocessing radiographic images is the approximate linearisation of the image generation process. This is obtained by performing an appropriate logarithmic transformation of the gray levels of the original image after which we can express

$$I(x,y) = \int_0^{s(x,y)} \mu(x,y,z)dz.$$

We define *radiological density* the expression on the right-end-side of this equation.

It should be noted that, once the process of image formation has been linearised, we can hypothesise without loss of generality that (no more than) N anatomical structures

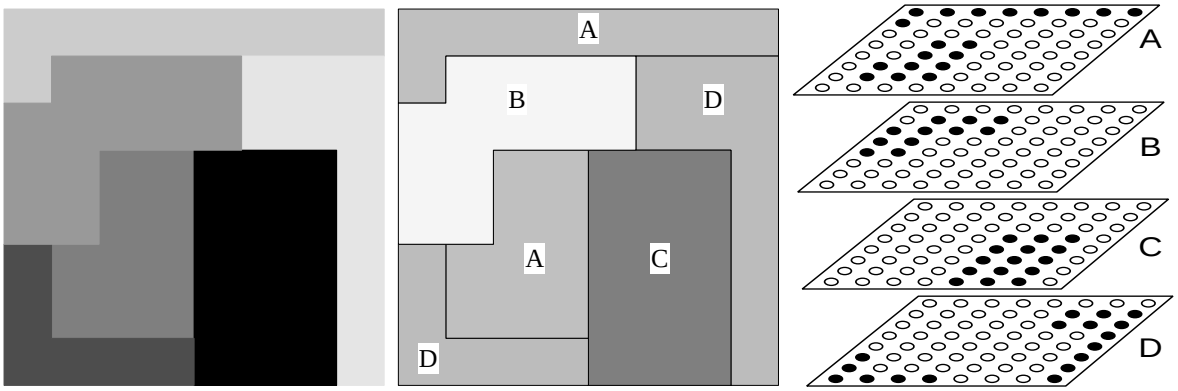


Figure 1: Synthetic 8×8 image (left), a possible labelling with 4 colours (centre), and the related binary representation with 4 layers of neurons (right). (Active neurons are represented as filled circles.)

are present in the image, each structure having thickness $s_i(x, y)$. If we also assume that the density of each structure has an approximately constant linear absorption coefficient μ_i, we can rewrite the previous equation as

$$I(x, y) = \sum_{i=1}^{N} \mu_i s_i(x, y). \tag{2}$$

As a first approximation the constant density assumption is reasonable for many anatomical structures. We will use it both in our segmentation algorithm and in our SFRD method.

The next step for solving the segmentation problem is finding a binary representation for its solutions.

2.3 Binary Representation

Our representation was suggested by the analogy of the segmentation process of a standard (non-projective) image with that of colouring geographic maps (Bilbro et al., 1987). This analogy indicates that, in order to represent the regions ("states") obtained from the segmentation of an image, only a reduced number of labels ("colours") are needed, as long as different labels are associated to connected regions ("bordering states"). Therefore, as shown in Figure 1, a segmentation can be represented with a small set of 2–D layers of neurons, each layer representing a different label. This can be seen as a 3–D array of neurons in which exactly one neuron is active in each column (only and only one colour can be associated to each state). The activations of the neurons in the array will be denoted with v_{xyc} (where the index c stands for "colour").

As in projective images anatomical structures which are overlaid or inside one another are represented by the same pixels, regions are not constrained to form a tessellation of the image but can overlap. To represent in binary form a segmentation with overlapping regions we adopted a set of 2–D layers of neurons like those mentioned above, with the important difference that each layer does not represent a different "colour" but a different

anatomical structure. This means that more than one neuron in each "column" is allowed to be active.

2.4 Energy Function

The next step is the definition of a quadratic energy function E_{net} whose minimisation gives an optimal solution to the segmentation problem. We adopted an energy function partially inspired by the one suggested in (Hopfield and Tank, 1985; Hopfield and Tank, 1986) for the solution of the travelling-salesman problem and the one proposed in (Bilbro et al., 1987) for the segmentation of signals with simulated annealing. As the fixed points of Hopfield nets tend to be the vertices of the hypercube $[0,1]^N$, we were able to design E_{net} on the hypothesis of binary neurons, i.e. $v_{xyc} \in \{0,1\}$.

E_{net} includes two parts: the *syntax energy* E_{syntax} which enforces the syntactic correctness of the solutions (i.e. prevents the network from settling into non-binary states or states which cannot be mapped back to solutions of the segmentation problem), and the *semantics energy* $E_{semantics}$ which is our criterion of goodness of segmentation. The two parts are added so that

$$E_{net} = E_{syntax} + E_{semantics}.$$

The syntactic correctness of solutions requires that, in stable states, each neuron of the network be completely excited ($v_{xyc} = 1$) or inhibited ($v_{xyc} = 0$). This constraint can be enforced by including in E_{syntax} a term of the form $v_{xyc}(1 - v_{xyc})$ for each neuron. As a result:

$$E_{syntax} = \frac{K_1}{2} \sum_x \sum_y \sum_c v_{xyc}(1 - v_{xyc})$$

where K_1 is a constant value.

The goal of the semantics energy is that of driving the network towards segmentations that represent an optimum compromise between sensitivity and robustness. Therefore the semantic energy includes two terms, the *sensitivity energy* $E_{sensitivity}$ and the *robustness energy* $E_{robustness}$, which are summed up to give

$$E_{semantics} = E_{sensitivity} + E_{robustness}.$$

The function of $E_{sensitivity}$ is maximising the consistency of segmentation with respect to the image gray levels expressed by Equation 2. Unfortunately, to obtain a quadratic $E_{sensitivity}$ we had to add the hypothesis (only approximately valid) that the thickness of the structures shown in the X-ray image is constant, i.e. $s_i(x,y) = s_i$. On this hypothesis we can define the quantity $S_i = \mu_i s_i$ (estimated on the basis of the typical density and thickness of the structures of interest) and express $I(x,y) = \sum_c v_{xyc} S_c$. To force the network to settle into solutions (approximately) consistent with this equation we defined

$$E_{sensitivity} = \frac{K_2}{2} \sum_x \sum_y (\sum_c v_{xyc} S_c - I(x,y))^2,$$

K_2 being a proper constant value.

The aim of $E_{robustness}$ is to reduce the effects of noise and texture. Since noise and texture tend to produce very small regions, $E_{robustness}$ should favour the construction of large regions which have a high probability of representing single anatomical structures. This can be obtained using the constraint: *pixels which are close to each other should have the same label.* The constraint can be implemented using terms of the form $-\sum_c v_{xyc} v_{\hat{x}\hat{y}c}$, for all the pixels $(\hat{x}, \hat{y})$ in a 4-connected neighbourhood N^{xy} of any given pixel (x, y).

Unfortunately, these terms alone can induce the diffusion of the activation of the neurons representing a given structure outside the boundaries of that structure. This happens because $E_{sensitivity}$ does not include any terms which force the neurons of a region to change their state in proximity of the boundaries of the structure represented by that region. This can be overcome by including also the constraint: *if a structure is not present in a given pixel, it is not present also nearby.* The resulting robustness energy turns out to be

$$E_{robustness} = -\frac{K_3}{2} \sum_x \sum_y \sum_{(\hat{x},\hat{y}) \in \mathcal{N}^{xy}} \sum_c v_{xyc}\, v_{\hat{x}\hat{y}c} - \frac{K_4}{2} \sum_x \sum_y \sum_{(\hat{x},\hat{y}) \in \mathcal{N}^{xy}} \sum_c (1 - v_{xyc})(1 - v_{\hat{x}\hat{y}c}).$$

where K_3 and K_4 are constant values.

2.5 Network Structure and Initialisation

Once E_{net} is defined, the weights and the external inputs of the network can be computed. If we denote with the symbols $T_{xyc\,\hat{x}\hat{y}\hat{c}}$ and i_{xyc} the weight between two generic neurons xyc and $\hat{x}\hat{y}\hat{c}$ and the external input to neuron xyc, respectively, differentiation of the previous energy expressions gives the following weights and external inputs:

$$T_{xyc\,\hat{x}\hat{y}\hat{c}} = \delta_{xy\,\hat{x}\hat{y}}(K_1 \delta_{c\hat{c}} - K_2 S_c S_{\hat{c}}) + \delta_{c\hat{c}}(K_3 + K_4) d_{\hat{x}\hat{y}}(\mathcal{N}^{\hat{x}\hat{y}})$$

$$i_{xyc} = K_2 S_c I(x, y) - \frac{K_1}{2} - 4K_4$$

where δ_{st} is the Kronecker delta, $\delta_{s_1 t_1 s_2 t_2} = \delta_{s_1 s_2} \delta_{t_1 t_2}$ and d is a membership function such that $d_{st}(\mathcal{S}) = 1$ if (s, t) belongs to the set $\mathcal{S}$, $d_{st}(\mathcal{S}) = 0$ otherwise.

Once weights and inputs are known, the network can be simulated by simply integrating numerically Equation 1 until a stable state is reached. However, before doing that the state of the net has to be initialised. As the standard random initialisation method gives, in the present case, poor segmentation results, we adopted the strategy suggested in (Chen et al., 1991) which consists of initialising the network in an area of state space where a good solution is present. In this way the network has only to improve the solution instead of looking for it in the whole state space. The initialisation algorithm used is similar to the one described in (Poli and Valli, 1997a).

3 Shape from Radiological Density: Theory

From a mathematical point of view, SFRD is an inverse, severely ill-posed problem that has an infinite number of solutions. To transform it into a problem with a single solution,

the typical characteristics of the structures to be recovered must be taken into account. To avoid an over-restriction of the solution space, we have considered only the following two assumptions: the density of each structure is approximately constant and the surface of each structure is smooth. These assumptions are valid for a large class of anatomical structures and also for many other natural or man-made objects. The first assumption was used also in our X-ray image segmentation method.

Using such assumptions and assuming that the approximate linearisation and segmentation of the radiogram have been performed, the local thickness and the 3-D shape of each imaged structure can computed as explained in the following two subsections.

3.1 Thickness Map Estimation

The first step of the method is the estimation of an image, termed *thickness map*, for each anatomical structure of interest, whose gray levels $s_i(x, y)$ are proportional to the thickness of such a structure along the ray of projection that crosses the image plane in (x, y). The thickness $s_i(x, y)$ is zero for points (x, y) outside the boundary γ_i of structure i (the curves γ_i's are provided as input along with the radiological density).

By applying the gradient operator $\nabla = [\partial/\partial x, \partial/\partial y]$ to both sides of Equation 2 we obtain

$$\nabla I(x, y) = \sum_{i=1}^{N} \mu_i \nabla s_i(x, y).$$

On the ground of the smoothness assumption we can hypothesise that the largest changes in the thickness of an anatomical structure occur near the apparent contour of such a structure, i.e.

$$\nabla s_i(x, y) \begin{cases} \neq 0 & \text{if } (x, y) \text{ is near } \gamma_i, \\ \approx 0 & \text{otherwise.} \end{cases}$$

To corroborate this hypothesis, let us consider the example of a sphere. The thickness of such a structure is given by $s(x, y) = 2\sqrt{r^2 - (x - x_c)^2 - (y - y_c)^2}$, r and (x_c, y_c) being the radius of the sphere and the projection of its center on the image plane, respectively. Simple calculations can show that $s(x, y)$ reaches 25% of its maximum value at a distance of $0.03r$ from the boundary, 50% at $0.13r$ and 75% at $0.34r$.

Now, if we consider that the contours of different structures are usually close to one another only for small tracts near the crossings (if any) of the curves γ_i, from the previous two equations we obtain

$$\nabla I(x, y) \approx \begin{cases} \mu_h \nabla s_h(x, y) & \text{if } \exists h \text{ such that } (x, y) \text{ is near } \gamma_h, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, near the boundary of an anatomical structure the image gradient is proportional to the gradient of the thickness of such a structure only.

Following (Wu and Li, 1988), if we hypothesise on the basis of the smoothness assumption that the thickness maps $s_h(x, y)$ are differentiable to the second order with continuous second derivatives, thanks to Schwartz's and Green's theorems we have

$$s_h(x_1, y_1) - s_h(x_0, y_0) = \int_{(x_0, y_0)}^{(x_1, y_1)} \left(\frac{\partial s_h(x, y)}{\partial x} dx + \frac{\partial s_h(x, y)}{\partial y} dy \right),$$

independently of the adopted integration path.

If we choose as starting point (x_0, y_0) a point belonging to γ_h (the boundary of the h-th structure), so that $s_h(x_0, y_0) = 0$, and as ending point a generic point (x, y) inside γ_h, the previous equation transforms into

$$s_h(x, y) = \int_0^l \nabla s_h(x(t), y(t)) \cdot \vec{n}(t)\, dt = \frac{1}{\mu_h} \int_0^l \nabla I(x(t), y(t)) \cdot \vec{n}(t)\, dt, \tag{3}$$

where $[x(t), y(t)]$ is an arbitrary curve such that $[x(0), y(0)] = (x_0, y_0)$ and $[x(l), y(l)] = (x, y)$, and $\vec{n}(t) = [\dot{x}(t), \dot{y}(t)]$ is the tangent vector of such a curve. This equation provides a method for estimating the local thickness of a structure from the image gradient.

The coefficient μ_h in Eq. 3 is unknown. Therefore, in theory, the local thickness of each structure of interest can be recovered up to a scaling factor only. However, in biological tissues, μ_h usually has a value in the range 0.21–1.05cm^{-1} (Johns, 1974). This provides a lower and upper limit for the aforementioned degree of freedom. Much narrower limits for μ_h can be assumed if the boundaries provided as input for SFRD are labelled, e.g. as bone, soft tissue, fat, etc., since, in this case, apart from a small inter-patient variability, the linear absorption coefficient is known. Alternatively, the linear absorption coefficient of a structure can be guessed on the basis of the dimensions of its contour or estimated from the boundaries of the same structure in images taken from different viewpoints.

3.2 3-D Shape Recovery

The data provided by the process of thickness estimation and the boundaries of the structures of interest are often noisy, incomplete and inconsistent. The phase of shape recovery is aimed at integrating all the information about each structure into a single consistent 3-D model.

Following an approach which is quite common in computer vision (see for example (Terzopoulos, 1986; Terzopoulos, 1988; Pentland and Sclaroff, 1991; Poli et al., 1994)), in order make this problem well-posed, we have used an implementation of physical inspiration for the smoothness assumption. The surface has been modeled as an elastic thin-surface S under the action of external forces generated by springs which deform the surface so as to make it best explain the data. Therefore, the problem of shape recovery is solved by mathematically modelling the process of relaxation of the elastic surface towards a state of minimum potential energy and finding the minimum energy surface.

The potential energy E_{total} of the surface-plus-springs system is the sum of three terms: the elastic energy of the springs which account for boundary points $E_{boundary}$, the elastic energy of the springs which model the action of thickness data $E_{density}$ and the internal energy of the surface $E_{surface}$. In order to express such terms mathematically, a reference system and a mathematical representation for the surface S have first to be chosen.

In this work we have adopted a cylindroidal model of S, similar to the one used in (Terzopoulos et al., 1987; Terzopoulos, 1988), which can be represented by the following parametric equation:

$$\vec{x}(u, v) = v\vec{e}_1 + f(u, v)\cos(u)\vec{e}_2 + f(u, v)\sin(u)\vec{e}_3$$

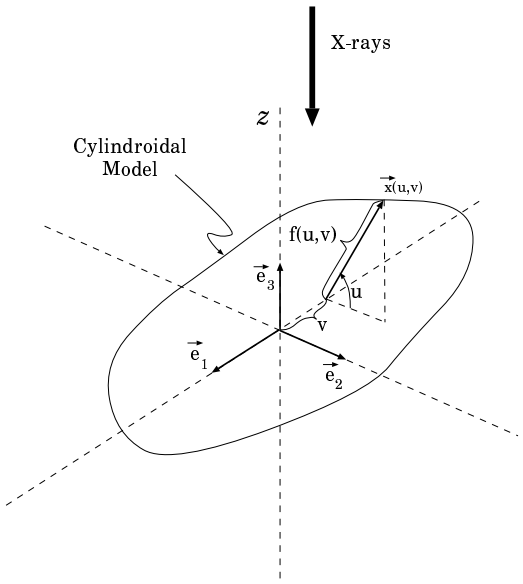


Figure 2: Cylindroidal surface model.

where $f(u, v)$ is a twice differentiable function, periodic with respect to u, $(u, v) \in D \equiv [0, 2\pi) \times [0, L]$, L being the length of the cylindroidal model, and $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ the unit vectors of an object-centred Cartesian reference (see Figure 2). The unit vector $\vec{e}_3$ of such reference is parallel to the X-ray projection direction (the z axis), $\vec{e}_1$ is parallel to the axis of inertia of the contour of the structure being recovered and the origin is aligned with the centroid of such a contour. We have chosen a cylindroidal model as it is particularly suited for representing elongated structures, such as cardiovascular structures and most of the bones of the human body, as well as many compact structures. However, with minor changes, the shape-recovery procedure described below could be used for other surface models such as the one described in (Poli et al., 1994).

3.2.1 Evaluation of $E_{boundary}$

The boundary of a given structure in the image plane is the projection of a 3-D curve belonging to the surface of the structure, termed *apparent contour*. The apparent contour does not necessarily lie on a plane parallel to the image plane; therefore this fact must to be taken into account when modelling the action of boundary points. To avoid over-constraining the surface to be recovered, we have modeled boundary points as springs attached to the apparent contour of the surface and to trolleys that can move along the z axis, as shown in Figure 3(a). (Similar constraints have been used in (Terzopoulos et al., 1987; Terzopoulos, 1988; Fua and Leclerc, 1994).)

According to this model, if we denote with (v_i, r_i) the coordinates of the i-th boundary point in the reference system of unit vectors $\vec{e}_1$ and $\vec{e}_2$, we can express the elastic energy of the springs which account for boundary points as

$$E_{boundary}(f) = \frac{1}{2} \sum_{i=1}^{P} \beta_b [f(u_i, v_i) \cos u_i - r_i]^2$$

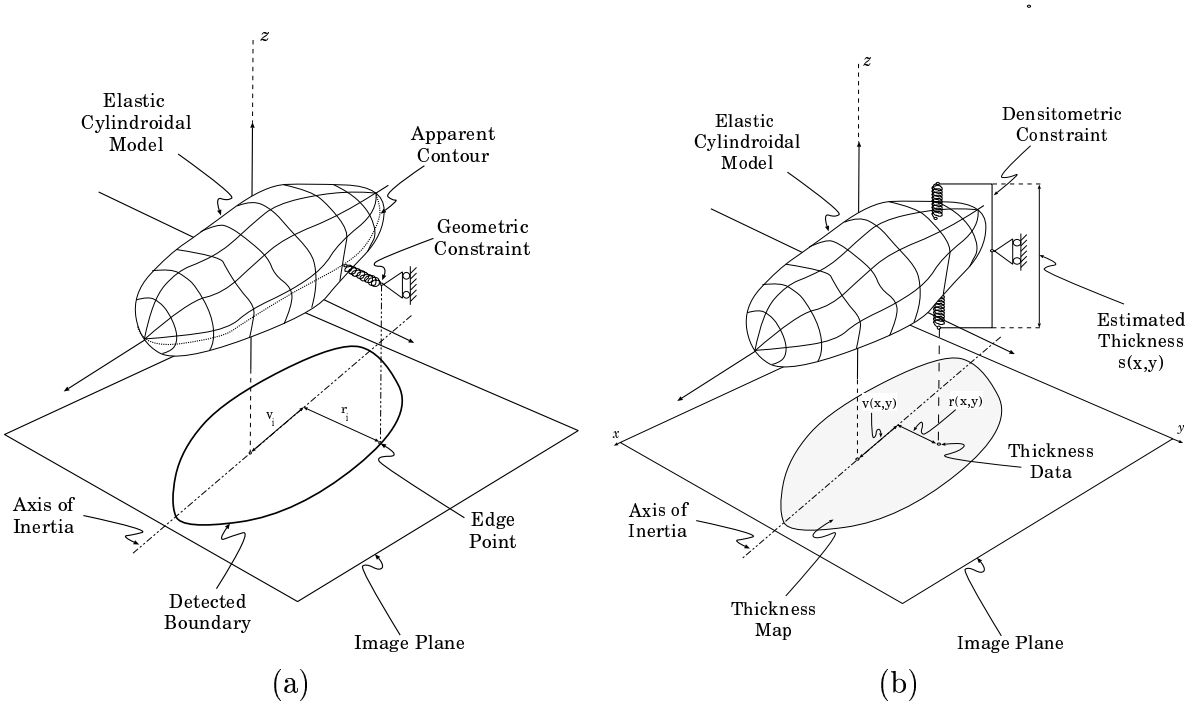


Figure 3: Action of boundary points (a) and thickness data (b).

where β_b is the stiffness of such springs and u_i is such that

$$f(u_i, v_i)\cos u_i = \begin{cases} \max_u f(u, v_i)\cos u & \text{if } r_i > 0, \\ \min_u f(u, v_i)\cos u & \text{otherwise.} \end{cases} \tag{4}$$

3.2.2 Evaluation of $E_{density}$

The estimated thickness map provides information on the local thickness of the structure whose shape is to be recovered. Therefore, thickness data should be modeled as springs, having a rest-length equal to the estimated thickness, attached to the lower and upper surfaces of the model (see Figure 3(b)).

According to this model, if we denoted with $(v(x,y), r(x,y))$ the coordinates of a point (x,y) of the thickness map $s_h(x,y)$, in the reference system of unit vectors $\vec{e}_1$ and $\vec{e}_2$, we could express the elastic energy of the springs which model the action of thickness data as

$$\begin{aligned} E_{density}(f) \;=\; & \frac{1}{2}\sum_{xy}\beta_d\Big[\sin u_u(x,y) f(u_u(x,y), v(x,y)) \\ & + \sin u_d(x,y) f(u_d(x,y), v(x,y)) - s_h(x,y)\Big]^2 \end{aligned}$$

where the summation is performed for all the points (x,y) such that $s_h(x,y) \neq 0$, β_d is the stiffness of the springs, and $u_u(x,y)$ and $u_d(x,y)$ are such that

$$f(u_u(x,y), v(x,y))\cos u_u(x,y) - r(x,y) = \min_{u\in[0,\pi)} \left[f(u, v(x,y))\cos u - r(x,y)\right] \tag{5}$$

and

$$f(u_d(x,y),v(x,y))\cos u_d(x,y) - r(x,y) = \min_{u\in[\pi,2\pi)} \left[f(u,v(x,y))\cos u - r(x,y)\right]. \qquad (6)$$

This formulation for the densitometric constraints would not overconstrain the surface to be recovered. However, when a single radiogram is used, densitometric constraints of such a form, along with the geometric constraints described above, would leave the 3-D model underconstrained. In addition, this formulation would induce more severe relative errors in the areas of S closest to the apparent contour, that are exactly the parts on which more (and more precise) information is available. Therefore, we have adopted a different formulation in which each thickness constraint (the trolley with springs shown in Figure 3(b)) is replaced with two radial springs, one attached to the upper part and the other to the lower part of S, which force the surface to be more consistent with thickness data near the boundary.

According to this formulation, we can express the elastic energy of the springs which model the action of thickness data as

$$\begin{aligned} E_{density}(f) \;=\; & \frac{1}{2}\sum_{xy}\beta_d\Big\{\Big[f(u_u(x,y),v(x,y)) - \frac{s_h(x,y)\lambda(x,y)}{\sin u_u(x,y)}\Big]^2 \\ & + \Big[f(u_d(x,y),v(x,y)) - \frac{s_h(x,y)(1-\lambda(x,y))}{\sin u_d(x,y)}\Big]^2\Big\} \end{aligned}$$

where $\lambda(x,y)$ is the ratio between the thickness of the upper part of the surface and the total thickness:

$$\lambda(x,y) = \frac{\sin u_u(x,y) f(u_u(x,y),v(x,y))}{\sin u_u(x,y) f(u_u(x,y),v(x,y)) + \sin u_d(x,y) f(u_d(x,y),v(x,y))}. \qquad (7)$$

3.2.3 Evaluation of $E_{surface}$

If we imagine the surface S to result from bending a thin plate (Courant and Hilbert, 1953), so as to obtain a sort of deformed cylinder, its potential energy is proportional to the functional

$$E_{surface}(f) = \int_S [2H^2 - K]dS \qquad (8)$$

where H and K are the mean and Gaussian curvatures of S, respectively.[2] From differential geometry (Lipschutz, 1969) it is known that

$$H = \frac{EN + GL - 2FM}{2\Delta S^2}, \quad K = \frac{LN - M^2}{\Delta S^2}, \quad dS = \Delta S du dv$$

where E, F, G and L, M, N are the first and second fundamental coefficients of S, respectively, and $\Delta S = |\vec{x}_u \times \vec{x}_v| = (EG - F^2)^{\frac{1}{2}}$. For our cylindroidal surface straightforward calculations show that

$$E = \vec{x}_u \cdot \vec{x}_u = f^2 + f_u^2, \quad F = \vec{x}_u \cdot \vec{x}_v = f_u f_v, \quad G = \vec{x}_v \cdot \vec{x}_v = 1 + f_v^2$$

[2] An alternative expression for the potential energy of S is $E_{Surface} = \frac{1}{2}\int_S (k_1^2 + k_2^2)dS$, where k_1 and k_2 are the principal curvatures of the surface.

$$L = \vec{x}_{uu} \cdot \vec{n} = -\frac{f^2 + 2f_u^2 - f f_{uu}}{\Delta S}, \quad M = \vec{x}_{uv} \cdot \vec{n} = \frac{f f_{uv} - f_u f_v}{\Delta S}$$

$$N = \vec{x}_{vv} \cdot \vec{n} = \frac{f f_{vv}}{\Delta S}, \quad \Delta S = \sqrt{f^2 + f_u^2 + f^2 f_v^2},$$

where subscripts denote the partial derivatives of $\vec{x}$ and f with respect to u and/or v, and $\vec{n} = \frac{\vec{x}_u \times \vec{x}_v}{|\vec{x}_u \times \vec{x}_v|}$ is the surface normal. Substitution of these equations into Eq. 8 leads to an expression for $E_{surface}(f)$ of the form:

$$E_{surface}(f) = \int \int_D \mathcal{I}_{surface}(f)\ dudv, \tag{9}$$

where $\mathcal{I}_{surface}(f)$ is a complicated, non-quadratic, differential operator as in (Poli et al., 1994).

4 Shape from Radiological Density: Implementation

Some of the details of our implementation of the methods for thickness estimation and shape recovery described in the previous sections are given in the following two subsections.

4.1 Thickness Map Estimation

In our implementation we have chosen as integration paths for Eq. 3 line segments. The line segments originate from contour points and have a direction orthogonal to the principal axis of inertia of the contour (the axis of minimum momentum). The tangent vector of these paths can be expressed as $\vec{n}(t) = [\cos\theta, \sin\theta]$, θ being a suitable constant, so that

$$s_h(x,y) = \frac{1}{\mu_h} \int_0^l \frac{\partial I}{\partial x}\Big|_{(x(t),y(t))} \cos\theta + \frac{\partial I}{\partial y}\Big|_{(x(t),y(t))} \sin\theta\, dt.$$

The estimation of the thickness map of a structure is performed by estimating the partial derivatives of $I(x,y)$ and numerically integrating the previous equation. Regularised estimates of the partial derivatives of $I(x,y)$ are obtained by convolving $I(x,y)$ with the two kernels that result from evaluating the partial derivatives of a two-dimensional Gaussian function with standard deviation $\sigma = 1$. Numerical integration is performed by bilinearly interpolating such estimates at the sampling points of an extended trapezoidal integration formula. Integration is stopped when the axis of inertia or the boundary of the current or another structure is encountered.

It should be noted that this algorithm attempts to estimate the thickness of a structure also in its internal parts. Of course, estimation errors are greater than for peripheral points, but this fact has been accounted for in the method for 3-D shape recovery. Integration is stopped only when the boundary of another object is encountered, since in this case gray level variations have to be attributed to the presence of such an object. In such a situation $s_h(x,y)$ of the first structure is considered to be undetermined for points within the contour of the second structure.

4.2 3-D Shape Recovery

In Section 3.2 we found the analytic expressions of the three components of the energy functional $E_{total}(f)$. In order to recover the shape of a structure from its boundary and thickness map, the energy functional has to be discretized and then minimized. Unfortunately, as all the components of $E_{total}(f)$ are non-quadratic, direct discretization leads to a non-quadratic discrete functional that is hard to calculate and minimize. Therefore, we have adopted a strategy, similar to the one described in (Poli et al., 1994), in which first a quadratic functional that approximates E_{total} is derived, and then a discrete form of it is obtained. As the approximation and discretization of the three components of E_{total} (especially $E_{surface}$) require a considerable amount of calculations, for the sake of brevity, in the following subsection we provide only an outline of the steps needed for such operations. Then, we describe strategies for the elimination of the degrees of freedom present in models recovered from a single radiogram.

4.2.1 Approximation and Discretization of E_{total}

We first hypothesise that the shape to be recovered, represented by the function f_{min} that minimizes $E_{total}(f)$, can be obtained with small deformations from a known reference configuration, represented by the function f_0.

In this case, the non-quadratic differential operator $\mathcal{I}_{surface}(f)$ in Eq. 9 can be approximated with its truncated Taylor expansion $\tilde{\mathcal{I}}_{surface}(v, f, f_0)$ about a function f_0 (Kolmogorov and Fomin, 1980). The resulting approximation for $E_{surface}(f)$ is a quadratic functional $\tilde{E}_{surface}(f)$ that is discretized with the Finite Element Method (FEM) (Bathe, 1982). Basically, FEM consists in dividing the domain D into a set of small rectangular sub-domains and hypothesising that the function f_{min} can be properly represented by quadratic functions in such sub-domains. As each of such quadratic functions depends uniquely (and linearly) on the values, termed *nodal variables*, taken by f in the vertex of the sub-domains, the substitution of such functions into $\tilde{E}_{surface}(f)$ leads directly to the required discrete quadratic approximation for $E_{surface}(f)$.

If the values u_i, $u_u(x, y)$, $u_d(x, y)$ and $\lambda(x, y)$ present in the expressions of $E_{boundary}(f)$ and $E_{density}(f)$ were known and constant, such functionals would already be quadratic and no approximation would be needed. This suggests that, in the hypothesis of small deformations with respect to the reference configuration represented by f_0, we can obtain a quadratic approximation for $E_{boundary}(f)$ and $E_{density}(f)$ by using f_0 instead of f in Eqs. 4, 5, 6 and 7. Discretization is then obtained by substituting the values $f(u_i, v_i)$, $f(u_u(x, y), v(x, y))$ and $f(u_d(x, y), v(x, y))$, present in the expressions of $E_{boundary}(f)$ and $E_{density}(f)$, with the corresponding nearest nodal variables.

As suggested in (Poli et al., 1994), the hypothesis of small deformations can be relaxed by using the following procedure. First a reference function f_0 of simple form is selected. (In this work we have used a cylinder having an axis parallel to the axis of inertia of the detected boundary and a radius equal to the average distance of the edge points from the axis of inertia.) Then, the function f_0 is used to evaluate a first approximation of the energy of the model and to find a first approximation f_1 of the function f_{min} that minimizes $E_{total}(f)$. Such an approximation is then used in place of f_0 to find a better

approximation f_2 for f_{min}, and the procedure is iterated.

As f_0 usually is a rather rough approximation of f_{min}, the springs that account for thickness data should not be applied immediately to the model as most of them would be attached to wrong nodes. This could trap the surface model in a local energy minimum in which boundary points are not properly fit. To avoid running this risk, we initially set the spring stiffness $\beta_d = 0$ and slowly increase its value as better and better approximations for f_{min} become available.

4.2.2 Elimination of Degrees of Freedom

When a single X-ray image is used, the recovery of the shape of anatomical structures is only possible up to a few degrees of freedom. In fact, the position of the models along the projection direction and the shape asymmetries with respect to a plane orthogonal to such a direction cannot be estimated. In addition, as already mentioned in Section 4.1, if the linear absorption coefficient is not known *a priori*, the local thickness of a structure can be recovered up to a scaling factor only. In our experiments (if not otherwise stated) we have resolved some of these ambiguities by arbitrarily setting the z position of the center of mass of the model and estimating the linear absorption coefficient heuristically.

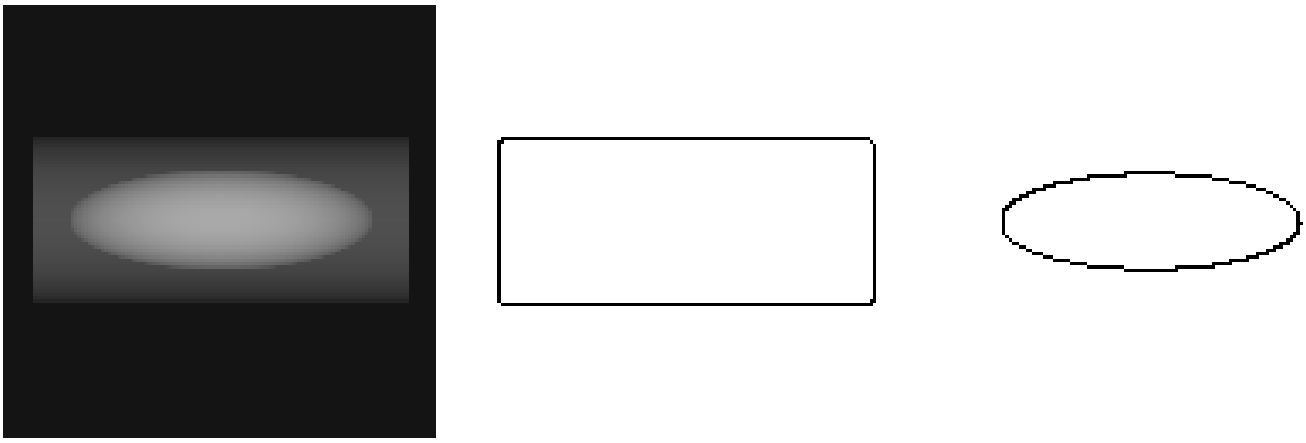
Asymmetry-related degrees of freedom do not cause any problems to the recovery procedure. The reason is that in the absence of explicit constraints the elasticity of the cylindroidal surface forces it to reach configurations approximately symmetric with respect to the image plane (i.e. $\lambda(x, y) \approx 0.5$). As a result, when a single projection is used the apparent contour maintains its initial position at the middle of the structure in the z direction during the relaxation of the model towards its minimum energy state. In turn, this implies that geometric constraints tend to remain on a plane parallel to the image plane and do not disrupt the symmetry of the model.

Obviously, if the original structure is non-symmetric, then the recovered model is only an approximation of such structure. However, some properties of the structure, such as the volume, are preserved in the recovered model. The symmetry hypothesis is acceptable for elongated structures with cylinder-like shape or other structures known to be symmetric. It is also acceptable in any applications in which a first-approximation model is sufficient.

It should be noted that, if additional views of the imaged structures are available, most of the degrees of freedom present in the single-projection case can be eliminated in other ways. For instance, the linear absorption coefficient, the position along the z axis and additional asymmetry constraints can be directly derived form images taken by projecting along an axis orthogonal to $\vec{e}_3$, such as the ones often used in cardiovascular imaging. Such constraints can easily be integrated in the shape recovery process as additional springs and allow for a nearly ambiguity-free shape recovery.

5 Experimental Results

The segmentation and SFRD methods described in the previous sections have been experimented, separately and in conjunction, on synthetic projection images, on X-ray images



(a) (b) (c)

Figure 4: Synthetic X-ray image of an ellipsoid surrounded by a cylinder (a) and the boundaries of such structures (b)–(c).

of structures of known shape and density, and on real diagnostic radiograms.

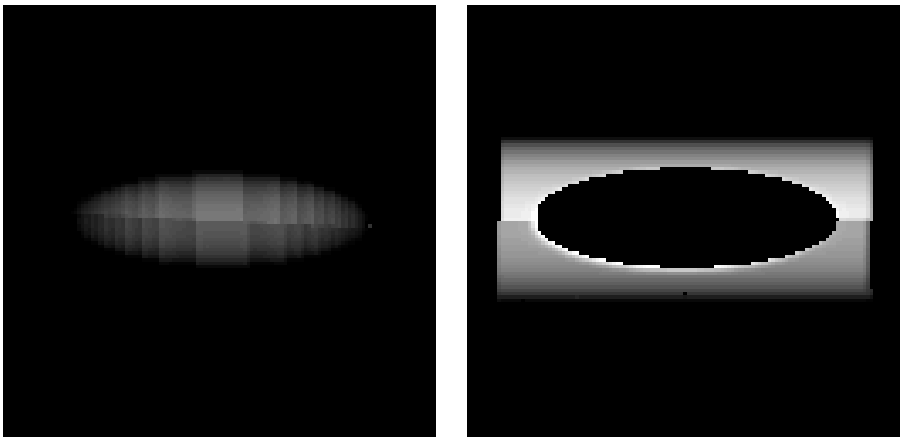
As roughly speaking $\sum_{i=1}^{P} \beta_b$ and $\sum_{xy} \beta_d$ play the role of regularisation parameters, they should be chosen on the basis of the characteristics of the noise expected to affect the data (Poggio et al., 1985). For the experiments described in this section we have adopted as default values $\sum_{i=1}^{P} \beta_b = \sum_{xy} \beta_d = 10000$; smaller values have been used only in the presence of real X-ray images of very poor quality.

Figure 4(a) shows a synthetic 128×128 X-ray image of a cylinder surrounding an ellipsoid whose shapes and linear absorption coefficients are known. The image was generated under the hypothesis that the process of image formation is linear and the projection is orthographic. The boundaries of the cylinder and the ellipsoid, detected as the zero crossings of the convolution of the original image with a Laplacian-of-Gaussian mask, are reported in Figures 4(b)–(c).

From these data, the algorithm described in Section 4.1 has estimated two thickness maps shown in Figures 5(a)–(b). Despite the inaccuracy in the computation of the axes of inertia, the approximation in the calculation of the partial derivatives of the image, the accumulation of integration errors and the inexact location of some edge points, both maps are nearly correct. Of course, the map of the cylinder includes an area (the elliptic hole) in which thickness cannot be estimated.

The models recovered from the boundary points in Figures 4(b)–(c) and the thickness maps in Figures 5(a)–(b) are shown in Figure 6. The 24×24 grid adopted for FEM discretization has been superimposed on the surfaces and the cylinder has been cut to show the internal ellipsoid. The same discretization grid has been used also for the other experiments, if not otherwise stated. On a DEC Alpha 300MHz workstation, the computation needed to reconstruct such models varies between 4 and 5 seconds of CPU time depending on the number of constraints available.

Figure 7 shows a 128×128 X-ray image of two vials containing iodine dyes of known densities (a reference metallic object is also present) and the boundaries of the two vials detected by an automatic segmentation system described in Section 2. The boundaries



(a) (b)

Figure 5: Estimated thickness maps of the ellipsoid (a) and the cylinder (b) shown in Figure 4(a).

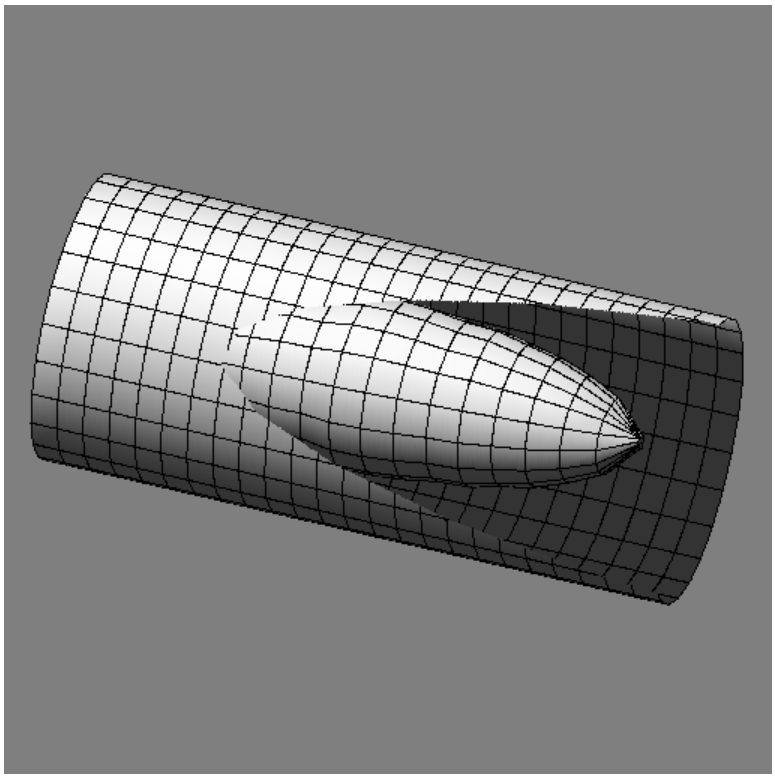
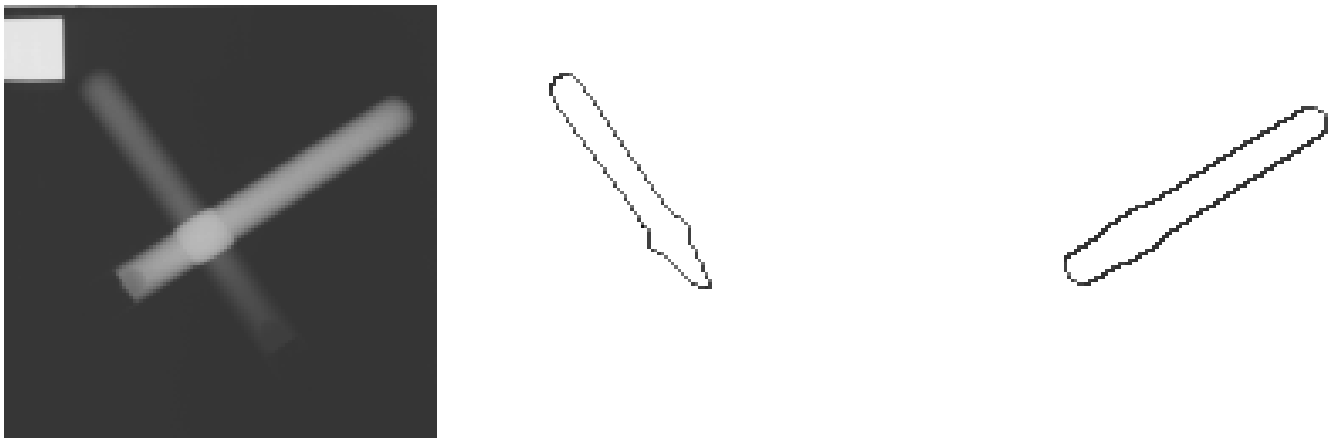


Figure 6: Recovered shape of the structures shown in Figure 4(a). The cylinder has been cut to show the internal ellipsoid.



(a) (b) (c)

Figure 7: X-ray image of two vials filled with iodine dye (a), and the boundaries of such structures (b)–(c).

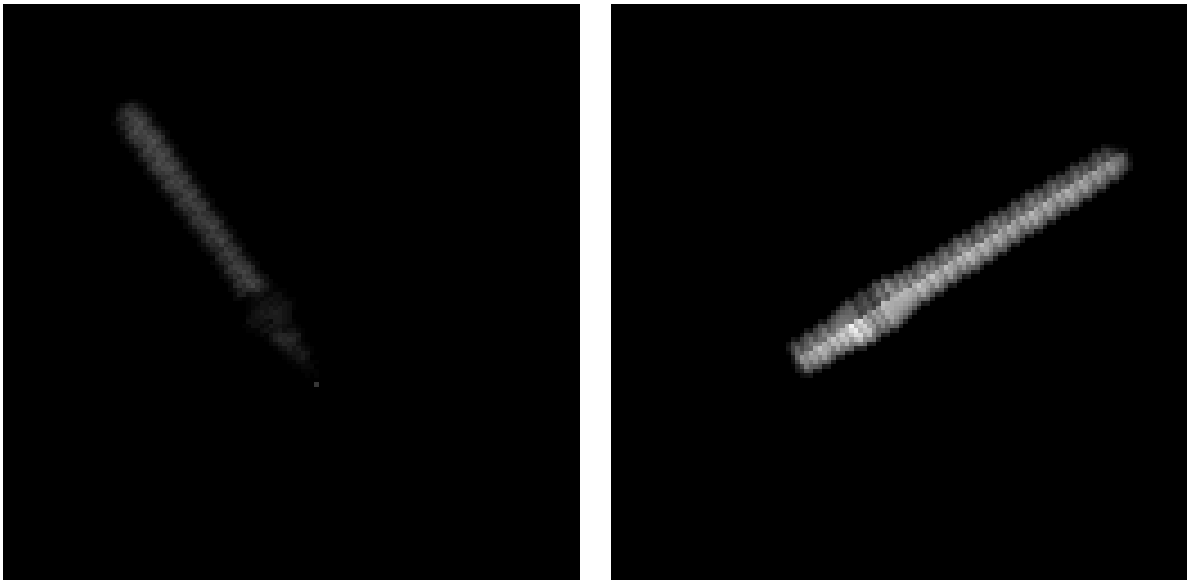
are incorrect only where the two vials overlap and near the ends.

Figures 8(a)–(b) show the thickness maps recovered from the image in Figure 7(a) before normalization by the factors $1/\mu_h$. The maps are partially affected by the errors in the position of the boundary points and by the low resolution of the original image. Low resolution is responsible for the asymmetry of the thickness estimates with respect to the axis of inertia of the denser vial.

Boundary and thickness data of the vials were integrated by the models shown in Figures 9(a)–(b). In Figure 9(a) the viewing direction is parallel to the projection axis; in Figure 9(b) the models have been rotated by 45°.

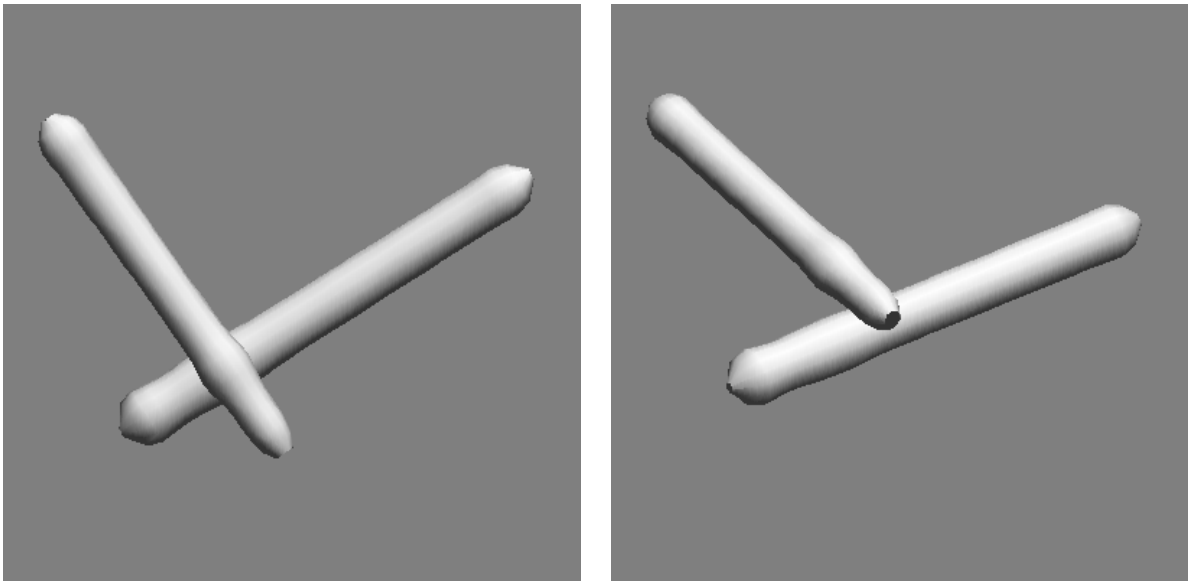
Figures 10(a) shows a 128×128 contrastographic X-ray image of the left ventricle of the heart. The largest structures inside the circular area representing the borders of the image intensifier are: the left ventricle with the descending aorta (centre), the diafragm muscle (lower left) and a metallic filter (upper right). To perform the segmentation of this kind of images we utilised three layers of neurons: one to represent the image intensifier, one for the background (soft tissues with a low density) and one for the structures just mentioned (they have approximately the same value of S_c). Figure 10(b) shows the activation of the neurons of such a layer. Diafragm muscle, left ventricle with aorta and metallic filter have been correctly represented as disjunct regions. Figure 10(c) shows the boundaries of the ventricle detected by taking the largest connected component in the image in Figure 10(b). Due to the presence of noise, ribs, disuniform distribution of the iodine dye inside the chamber and non-linearities, the detected boundaries are incorrect in several places. As a consequence, the estimated thickness map shown in Figure 11 results rather noisy and inaccurate, especially near the axis of inertia (the linear absorption coefficient of the ventricle has been guessed on the basis of the average distance of the boundary points from such an axis).

Despite such inaccuracies, as illustrated in Figure 12, the integration of the thickness map with boundary data provides a rather reasonable shape for the ventricle. In addition to the discretization grid, in the figure we have drawn (as white line segments) the springs



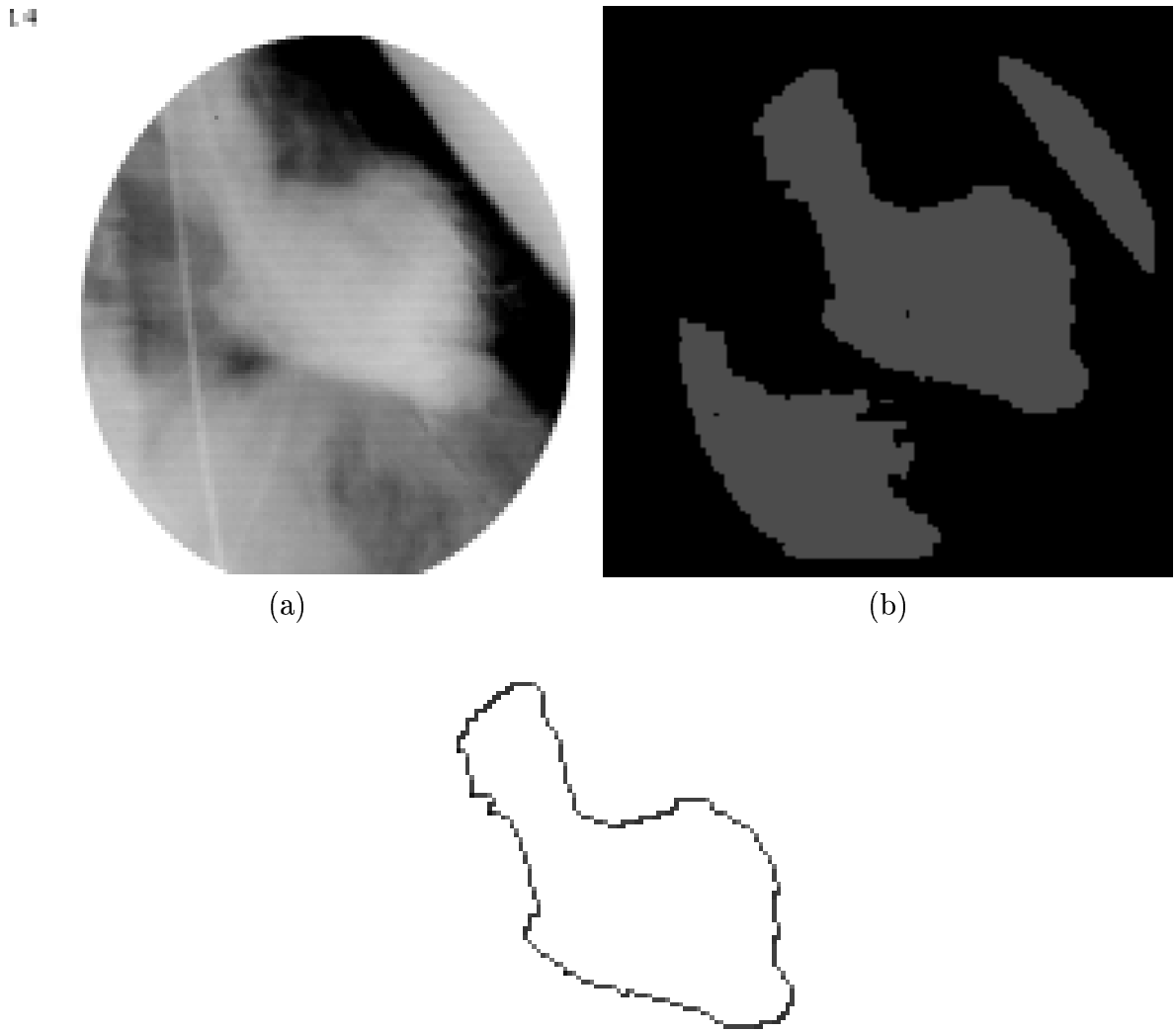
(a) (b)

Figure 8: Estimated thickness maps (before normalization) of the two vials (a)–(b) shown in Figure 7(a).



(a) (b)

Figure 9: Two views (a)–(b) of the shape of the two vials shown in Figure 7(a).



(a) (b)

(c)

Figure 10: Contrastographic X-ray image of the left ventricle of the heart filled with iodine dye (a), the connected components in the "foreground" layer of the segmentation network (b) and the boundary of left ventricle and aorta (c).

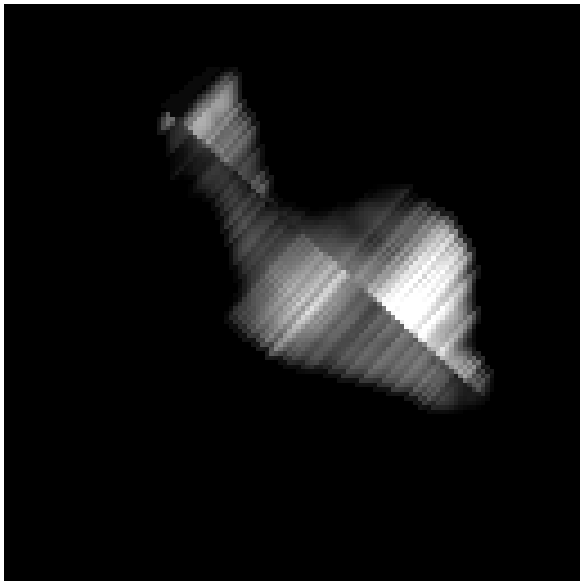


Figure 11: Estimated thickness map of the left ventricle shown in Figure 10(a).

that model the effect of boundary data.

Figure 13 shows an 853×640 radiogram of a hand that has been used for two experiments. In the first experiment we used the 128×128 sub-image shown in Figure 14(a) (the contrast has been enhanced for displaying purposes). The automatic segmentation algorithm in Section 2 produced the boundaries in Figures 14(b)–(c). Although, in this case, the network has not been capable of splitting the bone part of the finger into its anatomical components because of the very limited inter-bone space, the important discrimination between soft tissue and bone is correct, even where bone and soft tissue overlap. These data were then used as input for SFRD.

Despite the incorrect segmentation (fusion) of the three bones, the thickness maps of bone and soft tissue seem quite accurate. The recovered 16×24 3-D models are shown in Figures 16(a)–(b). Of course, being present in the input boundaries, fusion is also present in the recovered bone surface.

In the second experiment the boundaries of the bones of the image in Figure 13 have been hand-segmented and sequentially provided as input for SFRD. The resulting 3-D models have been then simply collected to form a model of the skeleton of the hand. Figure 17 shows two views of the model.

In order to assess the robustness of the two steps of our method for SFRD and to check the effects of the use of multiple views, we have performed some experiments involving the recovery of the thickness map and the 3-D shape of a synthetic ellipsoid starting from one or two orthogonal projections affected by increasing amounts of random noise. The semiaxes of the ellipsoid were oriented along the x, y and z axes of the reference system; their lengths were 80, 100 and 60 pixels, respectively. Uniform noise with zero mean and appropriate standard deviation was added to the projections of the ellipsoid so as to get the required signal-to-noise ratio (SNR). Statistics were collected by performing ten experiments for each SNR. The results of these experiments are summarised in Table 1.

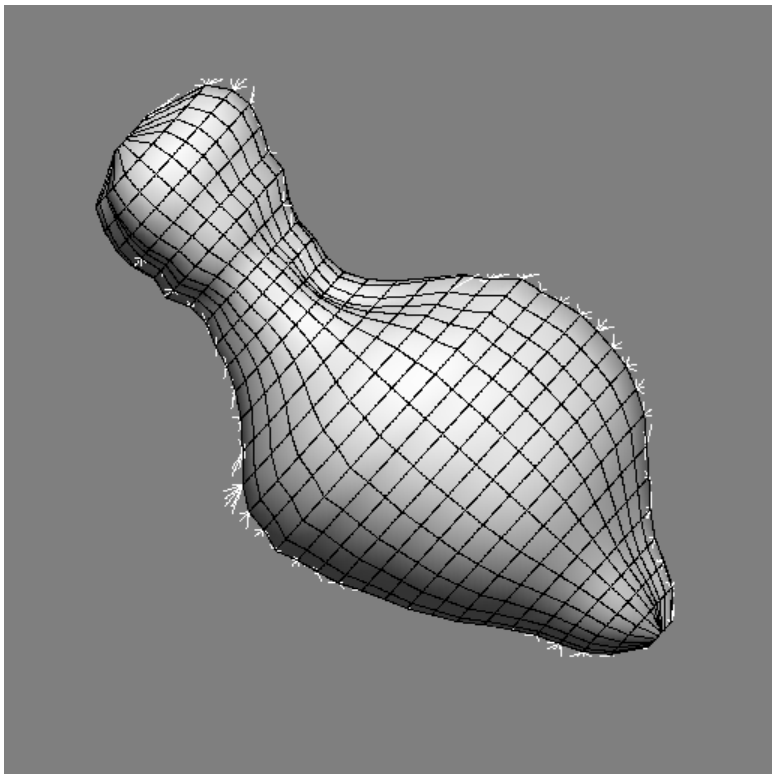
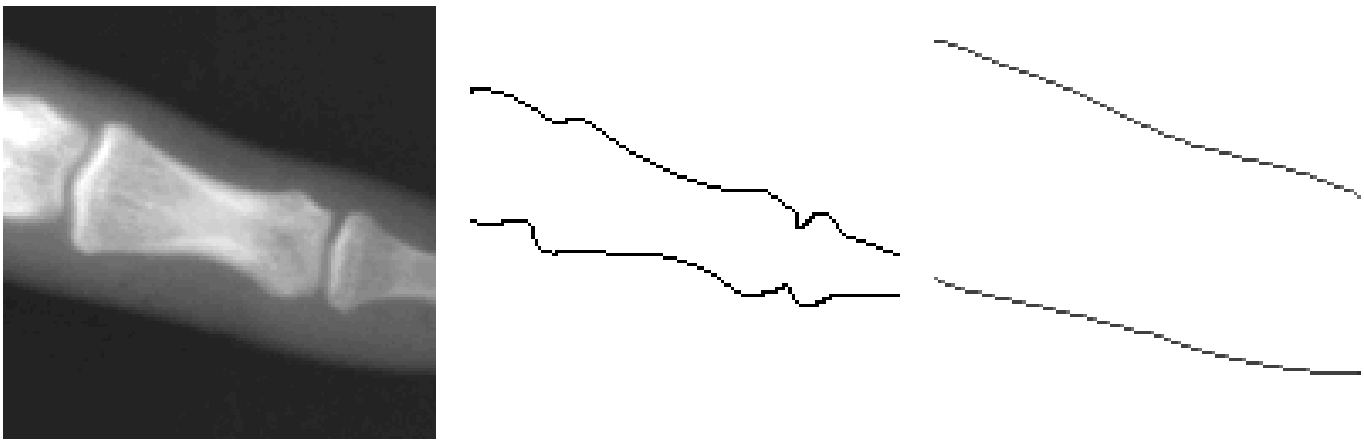


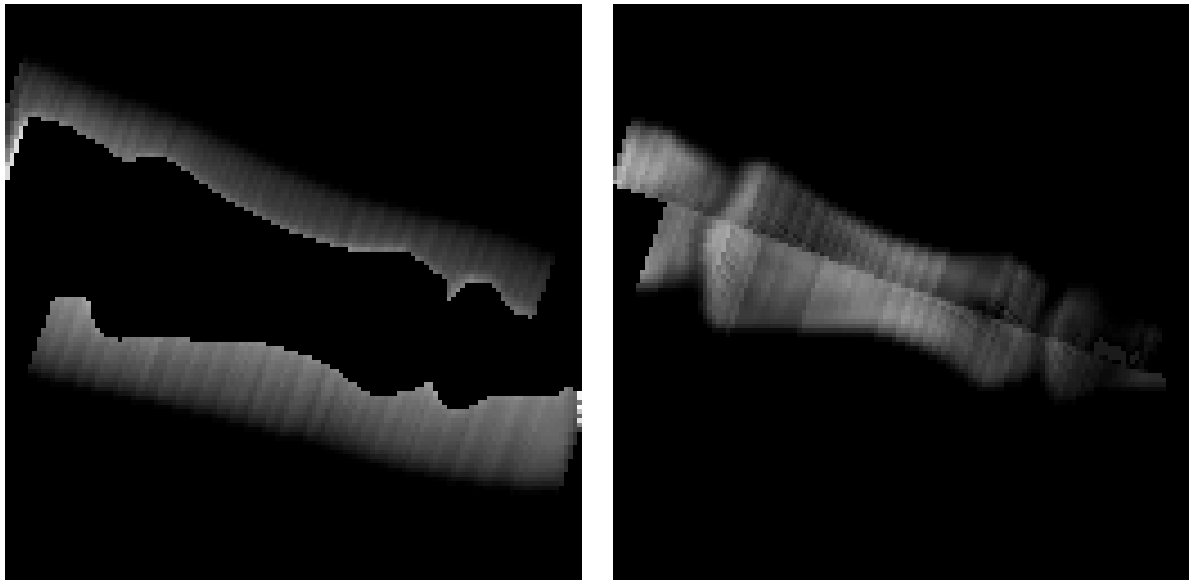
Figure 12: Recovered shape of the ventricle in Figure 10(a).

Figure 13: X-ray image of a hand.



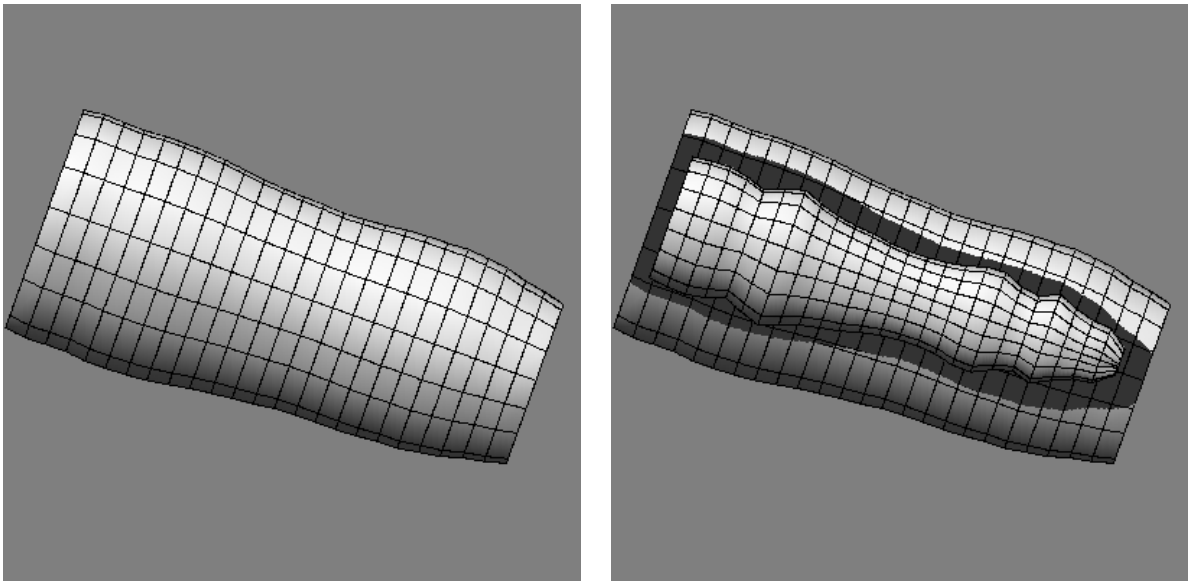
(a) (b) (c)

Figure 14: Sub-image of Figure 13 representing a tract of finger (a), along with the boundary of the bones (b) and of soft tissue (c).



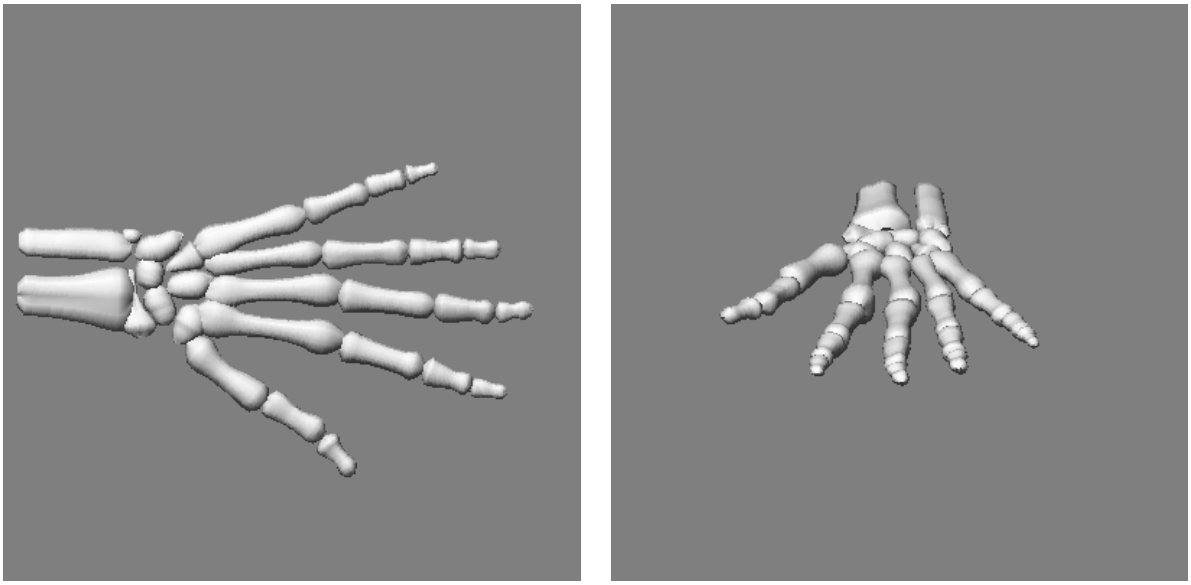
(a) (b)

Figure 15: Estimated thickness maps of the bone (a) and the soft tissue (b) shown in Figure 14(a).



(a) (b)

Figure 16: Two views (a)–(b) of the finger in Figure 7(a). In (b) the surface of soft tissue has been cut to show the bone surface.



(a) (b)

Figure 17: Two views (a)–(b) of the bones of the hand shown in Figure 13.

	Thickness Map Normalised RMS Error (%)		Recovered Model Cross-sectional difference (%)			
SNR	Projection along z	Projection along y	Single projection		Two projections	
			xz section	xy section	xz section	xy section
2	14.8	13.1	18.5	22.0	13.8	12.3
5	6.5	5.4	10.1	8.8	7.3	5.4
10	3.7	2.6	6.9	5.1	4.2	1.7
20	2.8	1.8	6.6	4.5	3.1	1.8
50	2.5	1.8	6.6	4.7	2.8	1.8
∞	2.5	1.8	6.3	4.5	2.8	1.7

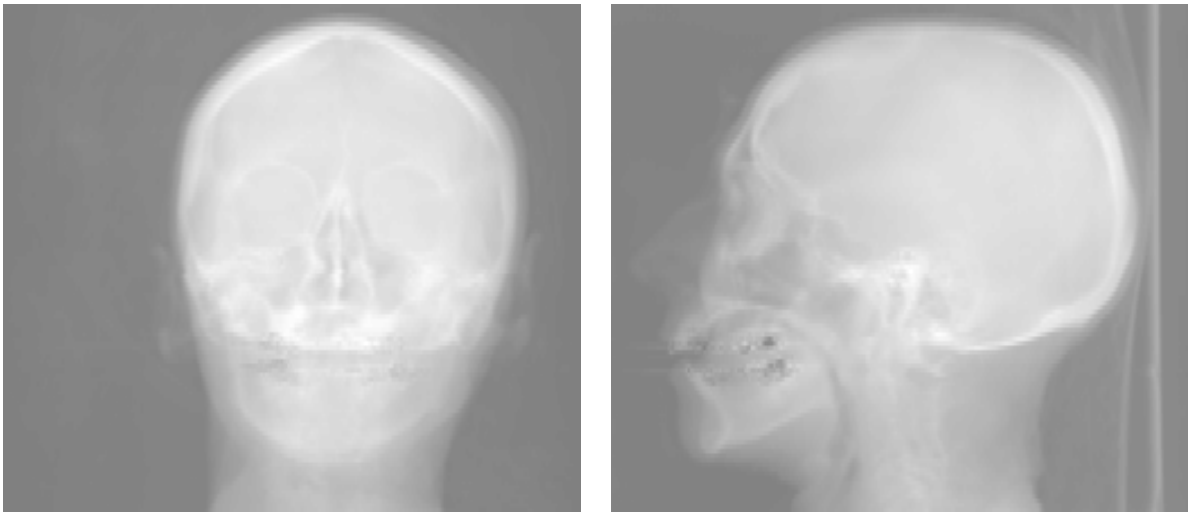
Table 1: Comparison between the exact data and the data obtained by SFRD applied to synthetic X-ray projections of an ellipsoid in the presence of uniform noise of different intensities.

The second and third columns of the table illustrate the behaviour of the method for the estimation of thickness maps. Namely, they report the Root Mean Squared (RMS) error between to the exact thickness map and the recovered one (normalised with respect to the maximum thickness) for two different projections of the ellipsoid. It should be noted that although our integration algorithm seems to be slightly biased, it is quite robust with respect to noise (the degradation of the thickness map is relatively small even at the lowest SNR).

Columns 4–7 illustrate the behaviour of the 3-D shape recovery method when only a single projection (columns 4 and 5) or two orthogonal projections (columns 6 and 7) are used. In particular the columns report the difference between the exact and the recovered cross sections of the ellipsoid.[3] Also in this case, the recovery process, although slightly biased, seems nearly insensitive to the image noise with the only exception of the cases in which SNR is extremely poor. The bias is due to the inaccuracies of the thickness map and to the tendency of the cylindroidal model to relax towards a non-elliptical (maximally flat) shape. (Increasing the elasticity of the springs modelling densitometric constraints would reduce the bias, but would also increase the sensitivity to image noise.) It should also be noted how the exploitation of two orthogonal projections leads to a dramatic reduction of recovery errors.

We have also checked the effectiveness of the method in the presence of multiple projections by using two orthogonal X-ray images of a head obtained by projecting a real 3-D CT scan (including 113 256×256 CT slices) which is part of the Chapel Hill Volume Rendering Test Dataset (SoftLab Software Systems Laboratory, University of North Carolina). Figure 18 shows the resulting X-ray images. The plots of the corresponding thickness maps estimated by our integration algorithm are shown in Figure 19. It should be noted that, due to the presence of the skull, only 21% of the thickness maps of the

[3]The section are "equatorial", minimally constrained sections in which maximum recovery errors are to be expected.



(a) (b)

Figure 18: Two orthogonal projections (a)–(b) of the 3-D CT scan of a head.

head could be estimated.

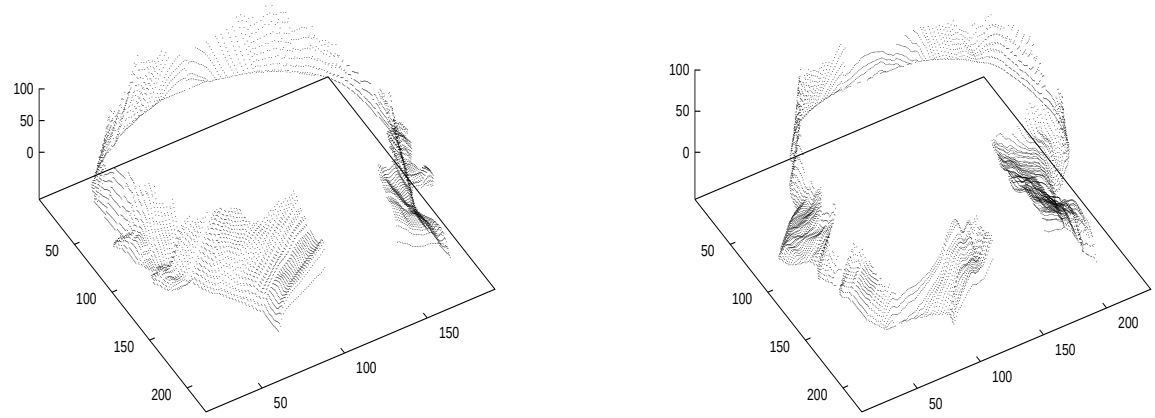
Figure 20 shows two views of the model recovered from this data. Given the relative quality of the simulated projections, in this experiment we have adopted $\sum_{xy} \beta_d = 30000$ and a 60×40 discretization grid. The good accuracy of the model can be assessed by comparing it with the images in Figure 21 obtained by interpolating, segmenting and rendering the slices included in the original 3-D CT scan.[4] The comparison shows that, although some fine shape details are lost, our SFRD method is able to recover the global shape and all the most important features of a head from the two projections despite the severe incompleteness of the related thickness maps.

6 Final Remarks

In this chapter we have presented a method for solving the problem of shape from radiological density along with its implementation and experimentation. The method is based on only two physical assumptions (constant density and smoothness) that are true for many anatomical structures. Under these assumptions a mathematical formulation has been derived which allows for the recovery of the local thickness of each structure of interest. By means of a cylindroidal model of elastic surface, such information, integrated with the geometric constraints imposed by the boundaries detected in the image, has allowed for reliable 3-D shape recovery despite the unavoidable inaccuracies in segmentation and linearisation of the input data.

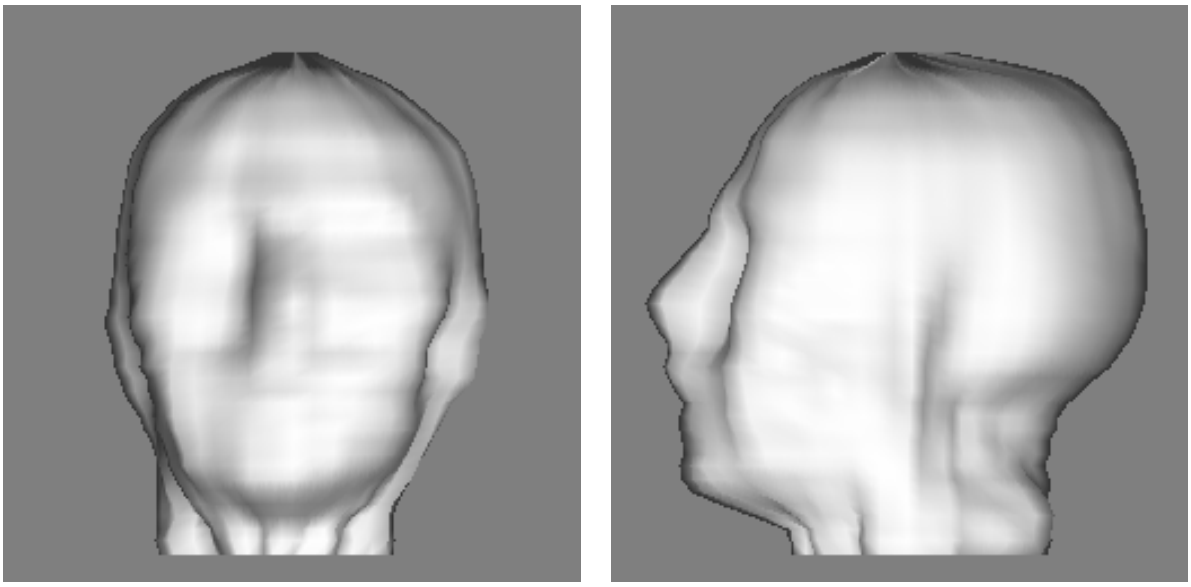
When single X-ray images were available, the recovery of the shape of the imaged anatomical structures via SFRD has been possible up to the degrees of freedom described in Section 4.2.2. However, it should be noted that shapes recovered from synthetic images

[4] The artifacts in the rendered images are due to corresponding artifacts in the original CT slices caused by lead tooth-fillings.



(a) (b)

Figure 19: Plots of the thickness maps (a)–(b) obtained from the x-ray projections in Figure 18.



(a) (b)

Figure 20: Two views (a)–(b) of the head shown in Figure 18.

(a) (b)

Figure 21: Three-dimensional rendering of the 3-D CT scan of the head in Figure 18.

and radiograms of reference objects differ only slightly from the original structures and that the shapes of anatomical structures, although approximate, have been judged to be very realistic by expert physicians.

The robustness of the method increases dramatically when more than one projection of the same structure is available. Thanks to this property the recovery of complex structures becomes possible even when incomplete thickness maps can only be obtained.

In the form presented in this chapter the method is particularly suited to recover the shape of elongated structures, such as cardiovascular structures and most of the bones of the human body. However, as shown in the experiments, it works very well also for compact anatomical structures (such as a head or a brain). Extensions to non-elongated structures are also possible.

The method requires that an approximate segmentation of the structures of interest be available. For X-ray projective images this may be a very hard task. In the chapter we have described a neural architecture for the segmentation of such images which, thanks to the exploitation of the physics of image formation and of the regularities of natural objects, can produce segmentations with a unique feature: regions are not constrained to form a tessellation of the image, but may overlap. This is very important as the objects that the regions represent may indeed be projected one on top of the other or may be one inside another.

Our segmentation architecture is based on Hopfield neural networks. Thanks to the properties of these networks, the segmentations produced by the architecture have an additional feature: they are optimum with respect to a goodness criterion which establishes the tradeoff between sensitivity and robustness.

When used together the segmentation network and our SFRD method form a reliable automatic system for the recovery of the 3-D shape of anatomical structures from single radiograms: a task that would have been considered to be totally impossible only a few years ago.

Acknowledgements

This work was partially supported by the Italian Ministry of University and Scientific and Technologic Research (MURST) and the Italian National Research Council (CNR).

References

Amartur, S., Piraino, D., and Takefuji, Y. (1992). Optimization neural networks for the segmentation of magnetic resonance images. *IEEE Transactions on Medical Imaging*, 11(2).

Andersen, A. H. (1989). Algebraic reconstruction in CT from limited views. *IEEE Transactions on Medical Imaging*, 8(1):50–55.

Azhari, H., Sideman, S., Beyar, R., Grenadier, E., and Dinnar, U. (1987). An analytical descriptor of three-dimensional geometry: application to the analysis of the left ventricle shape and contraction. *IEEE Transactions on Biomedical Engineering*, 34(5):345–355.

Bai, Z. D., Krishnaiah, P. R., Rao, C. R., Reddy, P. S., Sun, Y. N., and Zhao, L. C. (1989). Reconstruction of the left ventricle from two orthogonal projections. *Computer Vision, Graphics, and Image Processing*, 47:165–188.

Bathe, K. J. (1982). *Finite Element Procedures in Engineering Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey.

Bilbro, G., White, M., and Snyder, W. (1987). Image segmentation with neurocomputers. In Eckmiller, R. and v.d. Malsburg, C., editors, *Neural Computers*, Berlin. Springer-Verlag.

Brevdo, L., Sideman, S., and Beyar, R. (1987). A simple approach to the problem of 3D reconstruction. *Computer Vision, Graphics, and Image Processing*, 37:420–427.

Brinkley, J. F. (1985). Knowledge-driven ultrasonic three-dimensional organ modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(4):431–441.

Calamai, R., Coppini, G., Demi, M., Poli, R., and Valli, G. (1990). A computational approach to medical imaging. *Journal of Nuclear Medicine and Allied Sciences*, 34(1):42–50.

Catros, J. and Mischeler, D. (1988). An artificial intelligence approach for medical picture analysis. *Pattern Recognition Letters*, 8:123–130.

Chang, L.-W., Chen, H.-W., and Ho, J.-R. (1991). Reconstruction of 3D medical images: A nonlinear interpolation technique for reconstruction of 3D medical images. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 53(4):382–391.

Chang, S.-K. and Chow, C. K. (1973). The reconstruction of three-dimensional objects from two orthogonal projections ant its application to cardiac cineangiography. *IEEE Transactions on Computers*, 22(1):18–28.

Chen, C. T., Tsao, E. C. K., and Lin, W. C. (1991). Medical image segmentation by a constraint satisfaction neural network. *IEEE Transaction on Nuclear Science*, 38(2):678–686.

Cohen, I., Cohen, L. D., and Ayache, N. (1992). Using deformable surface to segment 3-D images and infer differential structures. *CVGIP: Image Understanding*, 56(2):242–263.

Cohen, L. D. (1991). On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218.

Coppini, G., Demi, M., Poli, R., and Valli, G. (1993). An artificial vision system for X-ray images of human coronary trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(2):156–162.

Coppini, G., Poli, R., Rucci, M., and Valli, G. (1992a). A neural network architecture for understanding 3D scenes in medical imaging. *Computer and Biomedical Research*, 25:569–585.

Coppini, G., Poli, R., and Valli, G. (1992b). Methods for medical image understanding. In *Topics on Biomedical Physics, Atti VI Congresso Nazionale dell'Associazione Italiana di Fisica Biomedica*, pages 49–56, Singapore. World Scientific.

Coppini, G., Poli, R., and Valli, G. (1995). Recovery of the 3-D shape of the left ventricle from echocardiographic images. *IEEE Transactions on Medical Imaging*, 14(2):301–317.

Courant, R. and Hilbert, H. (1953). *Methods of mathematical physics*. Interscience, New York.

Darrell, T., Sclaroff, S., and Pentland, A. (1990). Segmentation by minimal description. In *IEEE International Conference on Computer Vision III*, Osaka, Japan.

Deklerck, R., Cornelis, J., and Bister, M. (1993). Segmentation of medical images. *Image and Vision Computing*, 11(8):486–503.

Dumay, A. C., Gerbrands, J. J., Zijlstra, F., Minderhoud, H., Essed, C. E., Levenbach, W. A., Serruys, P. W., and Reiber, J. H. C. (1988). Three-dimensional reconstruction of myocardial contrast perfusion from biplane cineangiograms. In Gelsema, E. S. and Kanal, L. N., editors, *Pattern Recognition and Artificial Intelligence*, pages 155–168. Elsevier Science.

Fua, P. and Leclerc, Y. G. (1994). Using 3–dimensional meshes to combine image-based and geometry-based constraints. In *Proceedings of the 3rd European Conference on Computer Vision*, number 801 pt 2 in Lecture Notes in Computer Science, page 281, Stockholm, Sweden.

Gerig, G., Martin, J., Kikinis, R., Kubler, O., Shenton, M., and Jolesz, F. A. (1992). Unsupervised tissue type segmentation of 3D dual-echo MR head data. *Image and Vision Computing*, 10(6):349–360.

Heintzen, P. H., editor (1971). *Roentgen-, Cine- and Videodensitometry.* Georg Thieme Verlag, Stuttgart.

Heintzen, P. H. and Bürsch, J. H., editors (1978). *Roentgen- Video- Techniques*, Stuttgart. Georg Thieme Verlag.

Higgins, W. E., Chung, N., and Ritman, E. L. (1990). Extraction of left-ventricular chamber from 3-D CT images of the hart. *IEEE Transactions on Medical Imaging*, 9(4):384–395.

Hopfield, J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81:3088–3092.

Hopfield, J. and Tank, D. (1985). "neural" computation of decisions in optimization problems. *Biological Cybernetics*, 52:141–152.

Hopfield, J. and Tank, D. (1986). Computing with neural circuits: a model. *Science*, 233:625–633.

Johns, H. E. (1974). *The Physics of Radiology.* C. C. Thomas, Springfield, Illinois, USA.

Joliot, M. and Mazoyer, B. M. (1993). Three-dimensional segmentation and interpolation of magnetic resonance brain image. *IEEE Transactions on Medical Imaging*, 12(2):269–277.

Joseph W. Klingler, J., Vaughan, C. L., Theodore D. Franker, J., and Andrews, L. T. (1988). Segmentation of echocardiographic images using mathematical morphology. *IEEE Transactions on Biomedical Engineering*, 35(11):925–935.

Kitamura, K., Tobis, J., and Sklansky, J. (1988). Estimating the 3-D skeletons and transverse areas of coronary arteries from biplane angiograms. *IEEE Transactions on Medical Imaging*, 7:173–187.

Klifa, C. and Lavayssière, B. (1990). 3D reconstruction using a limited number of projections. In *SPIE Vol. 1360 Visual Communications and Image Processing '90*, pages 443–454.

Kolmogorov, A. N. and Fomin, S. V. (1980). *Elementy teorii funktsij i funktsional'nogo analiza.* Nauka, MIR, Moscov.

Lei, T. and Sewchand, W. (1992). Statistical approach to X-ray CT imaging and its applications in image analysis—part ii: A new stochastic model-based image segmentation technique for X-ray CT image. *IEEE Transactions on Medical Imaging*, 11(1):62–69.

Li, C., Goldgof, D. B., and Hall, L. O. (93). Knowledge-based classification and tissue labeling of MR images of human brain. *IEEE Transactions on Medical Imaging*, 12(4):740–750.

Lin, W., Liang, C., and Chen, T. (1990). Improvement on dynamic elastic interpolation technique for reconstructing 3-D objects from serial cross sections. *IEEE Transactions on Medical Imaging*, 9:71–83.

Lipschutz, M. M. (1969). *Differential Geometry*. McGraw-Hill, New York.

Macovski, A. (1983). *Medical Imaging Systems*. Prentice Hall, Englewood Cliffs.

Manos, G., Cairns, A. Y., Ricketts, I. W., and Sinclair, D. (1993). Automatic segmentation of hand-wirst radiographs. *Image and Vision Computing*, 11(2):100–111.

Marr, D. (1982). *Vision*. W.H. Freeman & Co., New York.

Onnasch, D. (1978). A concept for the approximative reconstruction of the form of the right or left ventricle from biplane angiocardiograms. In Heintzen, P. H. and Bürsch, J. H., editors, *Roentgen-Video-Techniques for the Dynamic Studies of Structure and Function of the Heart and Circulation*, pages 235–242, Stuttgart. Georg Thieme Verlag.

Özkan, M., Dawant, B. M., and Maciunas, R. J. (1993). Neural-network-based segmentation of multi-modal medical images: A comparative and prospective study. *IEEE Transactions on Medical Imaging*, 12(3):534–544.

Pellot, C., Herment, A., Sigelle, M., Horain, P., Maitre, H., and Peronneau, P. (1994). A 3D reconstruction of vascular structures from two X-ray angiograms using an adaped simulated annealing algorithm. *IEEE Transaction on Medical Imaging*, 13(1):48–60.

Pentland, A. and Sclaroff, S. (1991). Closed-form solution for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 715–729.

Poggio, T., Torre, V., and Koch, C. (1985). Computational vision and regularization theory. *Nature*, 317:314–319.

Poli, R., Coppini, G., and Valli., G. (1994). Recovery of 3-D closed surfaces from sparse data. *Computer Vision Graphics and Image Processing: Image Understanding*, 60(1):1–25.

Poli, R. and Valli, G. (1997a). Hopfield neural nets for the optimum segmentation of medical images. In Fiesler, E. and Beale, R., editors, *Handbook of Neural Computation*, chapter G.5.5. Oxford University Press.

Poli, R. and Valli, G. (1997b). Shape from radiological density. *Computer Vision and Image Understanding*, 65(3):361–381.

Raman, S. V., Sakar, S., and Boyer, K. L. (1993). Hypothesizing structures in edge-focused cerebral magnetic resonance images using graph-theoretic cycle enumeration. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 57(1):81–98.

Rangayyan, R., Dhawan, A. P., and Gordon, R. (1985). Algorithms for limited-view computed tomography: an annotated bibliography and a challenge. *Applied Optics*, 24(23):4000–4012.

Raya, S. P. (1990). Low-level segmentation of 3-D magnetic resonance brain images — a rule-based system. *IEEE Transactions on Medical Imaging*, 9(3):327–337.

Reed, T. R. (1992). Region growing using neural networks. In Wechsler, H., editor, *Neural Networks for Perception, Vol. 1*, pages 386–397. Academic Press, San Diego, CA.

Reuman, S. R. and Hoffman, D. D. (1986). Regularities of nature: the interpretation of visual motion. In Pentland, A. P., editor, *From Pixels to Predicates*, pages 201–226. Ablex, Norwood, New Jersey.

Robb, R. and Barillot, C. (1989). Interactive display and analysis of 3-D medical images. *IEEE Transactions on Medical Imaging*, 8:217–226.

Robb, R., Hoffman, E., Sinak, L., Harris, L., and Ritman, E. (1983). High-speed three-dimensional X-ray computed tomography: the dynamic spatial reconstructor. *Proceedings of the IEEE*, 71:308–319.

Rosenfeld, A. and Kak, A. (1982). *Digital Picture Processing, Vol. 1 and 2.* Academic Press, New York.

Silverman, R. H. and Noetzel, A. S. (1990). Image processing and pattern recognition in ultrasonograms by backpropagation. *Neural Networks*, 3:593–603.

Stiel, G. M., Stiel, L. S. G., Koltz, E., and Nienaber, C. A. (1993). Digital flashing tomosynthesis: A promising technique for angiocardiographic screening. *IEEE Transactions on Medical Imaging*, 12(2):314–321.

Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424.

Terzopoulos, D. (1988). The computation of visible-surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):417–438.

Terzopoulos, D., Witkin, A., and Kass, M. (1987). Symmetry-seeking models and 3D object reconstruction. *International Journal of Computer Vision*, 1(3):211–221.

Thomas, J. G., II, R. A. P., and Jeanty, P. (1991). Automatic segmentation of ultrasound images using morphological operators. *IEEE Transactions on Medical Imaging*, 10(2):180–186.

Toulson, D. L. and Boyce, J. F. (1992). Segmentation of MR images using neural nets. *Image and Vision Computing*, 10(5):324–328.

Tran, L. V., Bahn, R. C., and Sklansky, J. (1992). Recontructing the cross sections of coronary arteries from biplane angiograms. *IEEE Transactions on Medical Imaging*, 11(4):517–529.

Wang, T., Zhuang, X., and Xing, X. (1992). Robust segmentation of noisy images using a neural network model. *Image and Vision Computing*, 10(4):233–240.

Weixue, L. and YuanMei, W. (1993). Multicriteria regularizing neural network approach to implicit image information extraction from two projections. In Lemke, H. U., Inamura, K., Jaffe, C. C., and Felix, R., editors, *Computer Assisted Radiology, CAR'93*, pages 309–314, Berlin. Springer-Verlag.

Wu, Z. and Li, L. (1988). A line-integration based method for depth recovery from surface normals. *Computer Vision, Graphics, and Image Processing*, 43:53–66.

Xu, S. B. and Lu, W. X. (1988). Surface reconstruction of 3D objects in computerized tomography. *Computer Vision, Graphics, and Image Processing*, 44:270–278.

Ylä-Jääski, J., Klein, F., and Kübler, O. (1991). Fast direct display of volume data for medical diagnosis. *CVGIP: Graphical Models and Image Processing*, 53(1):7–18.

Young, A. A. and Hunter, P. J. (1989). Epicardial surface estimation from coronary angiograms. *Computer Vision, Graphics and Image Processing*, 47:111–127.