

SEGMENTATION OF ANATOMICAL STRUCTURES BY CONNECTED STATISTICAL MODELS

MARKO BUKOVEC, BOŠTJAN LIKAR AND FRANJO PERNUŠ

Laboratory of Imaging Technologies, University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, SI-1000 Ljubljana, Slovenia

e-mail: marko.bukovec@fe.uni-lj.si, bostjan.likar@fe.uni-lj.si, franjo.pernus@fe.uni-lj.si

(Accepted June 6, 2011)

ABSTRACT

This paper presents a framework for the segmentation of anatomical structures in medical imagery by connected statistical models. The framework is based on three types of models: first, generic models which operate directly on image intensities, second, connecting models that impose restrictions on the spatial relationship of generic models, and third, a supervising model that represents an arbitrary number of generic and connecting models. In this paper, the statistical model of appearance is used as the generic model, while the statistical model of topology, obtained by applying principal component analysis (PCA) on aligned pose and shape parameters of the generic model, is used as the connecting model. The performance of such connected statistical model is demonstrated on anterior-posterior (AP) X-ray images of the hips and pelvis and compared to the modelling by one and six unconnected generic models. The most accurate and robust results were obtained by two-level hierarchical modelling, wherein connected statistical models were used first, followed by unconnected statistical models.

Keywords: connecting anatomical structures, image segmentation, statistical appearance model, statistical shape model; statistical topology model.

INTRODUCTION

In the past two decades the role of medical imaging has expanded beyond simple visualization and inspection of anatomical structures for diagnostic purposes. Nowadays, medical imaging has become a fundamental tool for monitoring the progress of disease and effects of treatment, surgery planning and simulation, intra-operative navigation, radiotherapy planning and quantitative analysis. In many medical imaging applications image segmentation plays a crucial role, by automating or facilitating the delineation of anatomical structures and other regions of interest (ROI) (Bankman, 2000; Pham *et al.*, 2000; Sonka and Fitzpatrick, 2000). For example, ascertaining the detailed shape and organization of anatomical structures enables a surgeon to preoperatively plan an optimal approach to operate on the target structure while avoiding vital healthy structures. In radiotherapy, segmentation allows planning the delivery of necrotic dose of radiation to tumors with minimal collateral damage to healthy tissue.

Automated segmentation of anatomical structures, by which accurate, repeatable and quantitative data are efficiently extracted, remains a difficult task due

to the complexity and tremendous natural variability of shapes of anatomical structures. Moreover, the position of the patient during image acquisition, the imaging device itself, and the imaging protocol induce additional variations in shape and appearance. To address these difficulties, deformable models of anatomical structures have been extensively studied and widely used in medical image segmentation with promising results (McInerney and Terzopoulos, 2000; Xu *et al.*, 2000). Deformable models maintain the essential characteristics of a class of structures they represent but can also deform to fit a range of examples. Particular representatives of deformable models which incorporate a priori knowledge on pose, shape and appearance are the statistical models of shape (Cootes *et al.*, 1994) and appearance (Cootes *et al.*, 2001). A statistical model is a deformable model which in a compact way describes the information contained in a training dataset. These models were successfully applied to the segmentation of bony structures, *e.g.*, vertebrae (Hill *et al.*, 1996), spine (Smyth *et al.*, 1997), knee joint (Cootes *et al.*, 1998), hand (Mahmoodi *et al.*, 2000), rib cage (van Ginneken and Haar Romeny, 2000) and hip/pelvis (Bernard and Pernuš, 2001).

Organizations of anatomical structures exhibit two kinds of shape variations, *i.e.*, variations in shapes of individual structures and variations in spatial relationships of the structures. Such combined variations cannot be optimally described by a single statistical model unless variations of spatial relationships are sufficiently small and a sufficiently large training dataset is used (Smyth *et al.*, 1997). Therefore, alternative approaches are required to describe the non-linear shape variations of structures organizations. One approach is to apply piecewise linear models (Heap and Hogg, 1997) and the other is to separately model variations of shapes of individual parts and the variations of their spatial relationships (Mahmoodi *et al.*, 2000). The problem with piecewise linearization as described by Heap *et al.* (1997) is that it only approximates the non-linear shape variations because it is not using a priori knowledge on the organization of structures. However even though Mahmoodi *et al.* (2000) did use the a priori knowledge, it was used only to initialize a model and not throughout the segmentation process. Bernard *et al.* (2001) introduced the statistical model of topology to model the spatial relationships of anatomical structures. They applied the principal component analysis on pose and shape parameters of all structures to extract the most significant eigen-topologies that describe anatomically plausible topological variations of structures. Their strategy can be viewed upon as a hierarchical PCA. The statistical model of topology that describes plausible topological variations of connected anatomical structures at the upper level of hierarchy is constructed from the sets of shape and pose parameters describing plausible variations of shapes and poses of individual structures at the lower level of hierarchy. Pham *et al.* (2006) used a Gaussian model on translation invariant shape descriptions to describe spatial relationships between object parts. They also proposed to model objects appearance by learning long-distances dependencies between pixel values using Bayesian networks. De Bruijne *et al.* (2005) used the normal probability distribution to model spatial relationships between shape particle set and its expected neighbour. Additionally, multi-scale probabilistic Markov models have also been employed to characterize inter-object relationships (Lu *et al.*, 2007). Tsai *et al.* (2004) applied PCA to the collection of multiple signed distance functions, wherein zero level sets are implicit representations of shapes, to obtain a coupling between the multiple shape models. Another type of spatial modelling was used in level-set segmentation (Litvin and Karl, 2005), where relative inter-object distances were used in an energy functional for a curve evolution. The coupling constraint dependent on all contours is

introduced as an overlap penalty in energy functional (Zimmer and Olivo-Marin, 2005). The principal geodesic analysis (PGA) was applied to object pose and medial m-rep shape description (Styner *et al.*, 2006). The problem of splitting image data into sub-models for a given set of training images was addressed by Langs *et al.* (2007), where minimum description length (MDL) principle was proposed as a criterion function. Modelling of variations in spatial relationships of the anatomical structures depends on the type of the model used (linear, nonlinear, probabilistic model, etc.) and on information (pose, shape) that is being used to model spatial relationships.

In this paper, we present a general framework for simultaneous segmentation of a number of anatomical structures by connected statistical models. In contrast to Bernard *et al.* (2001), vectors of pose and shape parameters are aligned before they are submitted to PCA to generate the statistical model of topology. In this way, the spatial relationships of connected structures are better modelled. We demonstrate the feasibility of the proposed segmentation method by connected statistical models on anterior-posterior X-ray images of the hips and pelvis. We show that better segmentation results can be obtained with a two-level hierarchical segmentation procedure using statistical models of shape, appearance and topology than with unconnected models.

METHODS

GENERIC, CONNECTING AND SUPERVISING MODELS

The framework for image segmentation by connected statistical models consists of three types of models, *i.e.*, a generic, a connecting, and a supervising model. A *generic model* (G) is a statistical model, which operates directly on image intensities and models an object of interest. A *connecting model* (C) imposes restrictions on spatial relationships of generic models. A *supervising model* (S) represents an arbitrary number of generic and connecting models. An arbitrary hierarchy of models may be built by connecting supervising models.

Generic model

A generic model G , $G(\mathbf{u}, f_G, O_A)$, is defined by vector $\mathbf{u}$, containing all parameters of the generic model, criterion function f_G , which defines how well the model matches the underlying image $\mathbf{I}$ and an optimization procedure O_A , which optimizes the vector

$\mathbf{u}$ in order to optimize the criterion function f_G . Optimization O_A starts with a set $\mathbf{u}_0$ of initial parameters and when f_G reaches the optimum, ends with a set of optimized parameters $\mathbf{u}_{OPT}$. If generic models are not connected, each generic model is optimized independently of other generic models (Fig. 1).

As a generic model of an anatomical structure we use the statistical model of appearance (Cootes and Taylor, 2001) which models both the shape and texture of an anatomical structure. The statistical model of appearance consists of two sub-models – the statistical model of shape maintaining the structure's shape variation behaviour, and the statistical model of texture maintaining the structure's texture variation behaviour. The statistical model of shape of an anatomical structure is constructed by first establishing N corresponding sets $\mathbf{x}_i$, $\mathbf{x}_i = [x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{in}, y_{in}]^T$, $i=1, \dots, N$, of coordinates of n landmark points representing the shape of the structure in the i -th training image. In each of the N training images, points should be placed in the same way on the object's boundary. Next, the sets $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ are aligned with respect to translation, rotation, and scaling (Goodall, 1991). The result of the alignment are aligned point sets $\mathbf{X}_i$, $\mathbf{X}_i = [X_{i1}, Y_{i1}, X_{i2}, Y_{i2}, \dots, X_{in}, Y_{in}]^T$, $i=1, \dots, N$, and corresponding pose vectors $\mathbf{v}_i = [v_x, v_y, v_\phi, v_s]^T$. After the alignment, there is still a substantial amount of variability in the coordinates of corresponding points, which is due to differences in shape of an anatomical structure across different training images. Given N aligned shapes, described by shape vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$, the mean shape $\bar{\mathbf{X}}$, is defined to be:

$$\bar{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i. \quad (1)$$

Singular value decomposition (SVD) (Press *et al.*, 1992) is next used to compute the eigenvalues and corresponding eigenvectors of the covariance matrix:

$$\Psi_x = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T. \quad (2)$$

The eigenvectors corresponding to the largest eigenvalues of the covariance matrix represent the most significant modes of shape variation. It is sufficient to keep only some of the first t eigenvectors to sufficiently describe the variability of shapes in the set of training images. Any aligned shape $\mathbf{X}$ of a structure in the training set can be approximated by:

$$\mathbf{X} \approx \bar{\mathbf{X}} + \Theta_x \mathbf{p}_x, \quad (3)$$

where Θ_x is the matrix of the first t eigenvectors and $\mathbf{p}_x = \Theta_x^{-1}(\mathbf{X}_i - \bar{\mathbf{X}})$, is a vector of shape parameters, *i.e.*, projections of $(\mathbf{X}_i - \bar{\mathbf{X}})$ to the first t eigenvectors of Θ_x . Let the vector $\mathbf{u}$ of a generic model contain the pose and shape parameters $\mathbf{u}$, $\mathbf{u} = [\mathbf{v}^T, \mathbf{p}_x^T]^T$, corresponding to one image. To be comparable to the shape parameters, the pose parameters $\mathbf{v}$ are normalized. Usually the number of the first t eigenvectors is selected in such a way to describe some predetermined amount of variability. However, this is not the case in the connecting model, where the number of shape parameters (t) is fixed according to all generic models in order to achieve the same dimensionality of the pose and shape vector ($\mathbf{u}$).

The next step in building a statistical model of appearance is to warp, for instance with the thin-plate spline interpolation method (Bookstein, 1989), each training image so that its landmarks match those of the mean shape. In this way, "shape-free" images are obtained. The images are intensity normalized to eliminate global intensity variations between the images and raster scanned into texture vectors $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$. Next, PCA is applied to the shape-free and intensity-normalized texture vectors, yielding a linear model that characterizes the intensity variations:

$$\mathbf{Y} \approx \bar{\mathbf{Y}} + \Theta_y \mathbf{p}_y, \quad (4)$$

where $\bar{\mathbf{Y}}$ is the mean normalized texture vector, Θ_y , a matrix of significant modes of intensity variations of covariance matrix Ψ_y , and, $\mathbf{p}_y$ a vector of texture parameters.

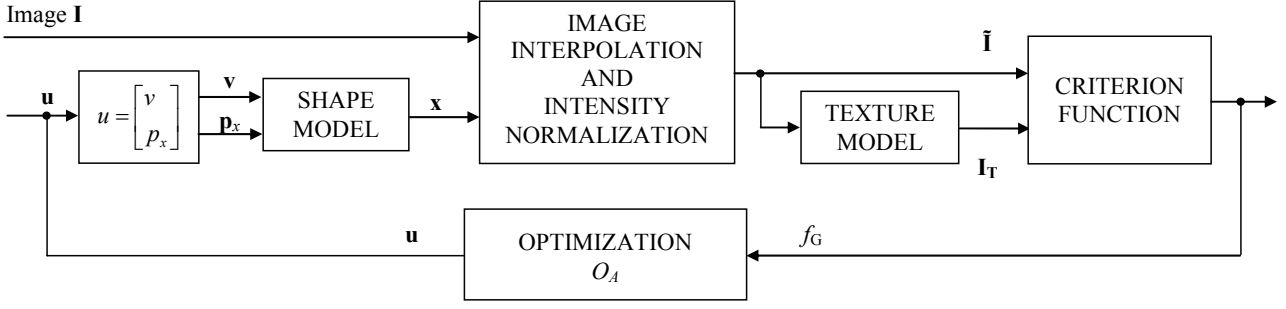


Fig. 1. Generic model and steps involved in image segmentation by unconnected statistical model.

Connecting model

A connecting model is a parametric model which models the spatial relationships of an arbitrary number, say M , of anatomical structures. A connecting model C , $C(\mathbf{z}, \mathbf{T}, f_C)$ is defined by topology vector $\mathbf{z}$, $\mathbf{z} = [\mathbf{u}_1^T, \dots, \mathbf{u}_M^T]^T$, which is composed of unaligned parameters of M generic models that C connects, geometrical transformation $\mathbf{T}$ and criterion function f_C . The geometrical transformation $\mathbf{T}$ is used during training to obtain the aligned topology $\mathbf{Z}$, which is a vector composed of aligned parameters of the M generic models. The inverse geometrical transformation $\mathbf{T}^{-1}$ maps the aligned topology $\mathbf{Z}$ back to the unaligned topology $\mathbf{z}$ and is used during segmentation to generate plausible spatial relationships among structures. The criterion function f_C represents the energy of the connecting model.

As a connecting model we use the statistical model of topology which is obtained by the same principle as the statistical model of shape. The construction of the statistical model of topology thus consists of two steps: topology alignment and principal component analysis. Let $\mathbf{z}_i$, $\mathbf{z}_i = [\mathbf{u}_{i,1}^T \mathbf{u}_{i,2}^T \dots \mathbf{u}_{i,M}^T]^T$, $i = 1, \dots, N$, be a vector containing all or just some of the pose and shape parameters of M generic models extracted from each of the N images. A vector $\mathbf{u}_{i,j}$ corresponding to image i and model j may contain all parameters of pose and shape $\mathbf{u}_{i,j} = [\mathbf{v}_{i,j}^T, \mathbf{p}_{i,j}^T]^T$, only some pose and shape parameters, or only parameters of pose $\mathbf{u}_{i,j} = \mathbf{v}_{i,j}$. The sets $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N$ are first aligned (Fig. 2) which results in aligned topo-

logies $\mathbf{Z}$, $\mathbf{Z}_i = [\mathbf{U}_{i,1}^T \mathbf{U}_{i,2}^T \dots \mathbf{U}_{i,M}^T]^T$, $i = 1, \dots, N$, corresponding transformations $\mathbf{T}_i$, $\mathbf{U}_{i,j} = \mathbf{T}_i \mathbf{u}_{i,j}$, and mean topology $\bar{\mathbf{Z}}$. Topology alignment clusters the sets of pose-shape vectors $\mathbf{u}$ before they enter PCA. Using the PCA, any aligned topology $\mathbf{Z}$ in the training set can be approximated by:

$$\mathbf{Z} \approx \bar{\mathbf{Z}} + \Theta_z \mathbf{p}_z, \quad (5)$$

where $\mathbf{p}_z$ is a vector of topology parameters.

Besides applying PCA to the aligned topologies, PCA may also be performed on the set $\mathbf{W}$ of parameters of the geometrical transformation $\mathbf{T}^{-1}$:

$$\mathbf{W} \approx \bar{\mathbf{W}} + \Theta_w \mathbf{p}_w. \quad (6)$$

This results in a reduced set of parameters $\mathbf{p}_w$ that approximates the transformation $\mathbf{T}^{-1}$, which may prove useful during segmentation, because segmentation process involves optimization of parameters.

Supervising model

A supervising model S , $S(\mathbf{p}_{sm}, f_{SM}, \mathbf{G}, \mathbf{C}, O_A)$ is a union of generic models $\mathbf{G}$ and connecting models $\mathbf{C}$ (Fig. 3). It is composed of a vector $\mathbf{p}_{sm}$, criterion function $f_{SM} = f_{SM}(\mathbf{p}_{sm})$, vector of generic models $\mathbf{G}$, vector of connecting models $\mathbf{C}$, and optimization algorithm O_A . The vector $\mathbf{p}_{sm}$, $\mathbf{p}_{sm} = [\mathbf{p}_z^T, \mathbf{p}_w^T]^T$ contains the topology parameters $\mathbf{p}_z$ and parameters $\mathbf{p}_w$ of the transformation $\mathbf{T}^{-1}$ that are used to generate plausible topologies in the image segmentation process (Fig. 2).

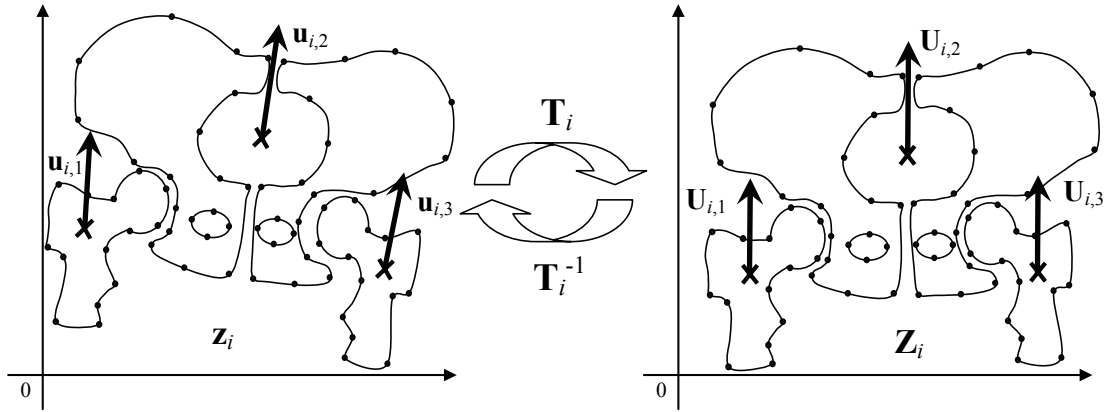


Fig. 2. Topology alignment (left to right) of three connected models in training phase and plausible topology generation (right to left) during image segmentation.

The supervising model represents geometrical (pose and shape vectors) and textural representation of the image (Fig. 3). Using the statistical connection model, whole collection of generic models is steered in the optimization process. At the level of supervising model, other segmentation strategies can be applied. For example, one could apply sequential generic model fitting, *e.g.* fit one or few models and then use connection model to predict initial poses and shapes for subsequent generic models. Another approach could be to connect generic models between two or more images of the same anatomical structures but with different modalities.

IMAGE SEGMENTATION BY UNCONNECTED AND CONNECTED STATISTICAL MODELS

The goal of segmentation is to find the landmarks of each of the M anatomical structures present in a previously unseen image $\mathbf{I}$. Segmentation of an image $\mathbf{I}$ with statistical models is in effect an optimization procedure. We have explored two different statistical model-based segmentation strategies: one strategy with unconnected statistical models, and another strategy with connected statistical models.

Image segmentation by unconnected statistical models

By unconnected statistical models, each anatomical structure is segmented independently of other structures. We have adopted the segmentation approach of Bernard *et al.* (2001), which does not require optimization of texture parameters $\mathbf{p}_y$ (Fig. 1). Given

the image $\mathbf{I}$ to be segmented and the initial pose and shape parameters $\mathbf{u}_0$, $\mathbf{u}_0 = [\mathbf{v}^T, \mathbf{p}_x^T]^T$, of a statistical model describing one anatomical structure, a shape-free and intensity normalized image $\tilde{\mathbf{I}}$ and corresponding texture vector $\hat{\mathbf{y}}^0$ are created and used to obtain the vector of texture parameters: $\mathbf{p}_y = \Theta_y^{-1} \cdot (\hat{\mathbf{y}}^0 - \bar{\mathbf{Y}})$.

Next, the texture vector $\hat{\mathbf{y}}^0$ is transformed to texture vector $\mathbf{Y}_T$ by the texture model with only t eigen-textures: $\mathbf{Y}_T \approx \bar{\mathbf{Y}} + \Theta_y \mathbf{p}_y$ and the image $\mathbf{I}_T$ corresponding to the $\mathbf{Y}_T$ is obtained. By optimizing the energy function f_G :

$$f_G = \alpha \cdot f_{\text{INT}} + (1 - \alpha) \cdot f_{\text{EXT}}, \quad (7)$$

the optimal parameters $\mathbf{u}_{\text{OPT}}$, $\mathbf{u}_{\text{OPT}} = [\mathbf{v}^T, \mathbf{p}_x^T]^T$ of pose ($\mathbf{v} = [v_x \ v_y \ v_\phi \ v_s]^T$) and shape ($\mathbf{p}_x$) are found. The internal energy f_{INT} measures the extent of deformation of the model:

$$f_{\text{INT}} = \|\mathbf{X} - \bar{\mathbf{X}}\|^2, \quad (8)$$

while the external energy f_{EXT} measures the similarity between images $\tilde{\mathbf{I}}$ and $\mathbf{I}_T$:

$$f_{\text{EXT}} = 1 / (1 + MI(\tilde{\mathbf{I}}, \mathbf{I}_T)), \quad (9)$$

where $\tilde{\mathbf{I}}$ is a shape-free and intensity normalized image, $\mathbf{I}_T$ is a best representation of an image $\tilde{\mathbf{I}}$ obtained by texture model, and MI denotes mutual information between two images. However, different similarity measures, such as normalized mutual information or correlation coefficient, may be used. The external energy function f_{EXT} (Eq. 9) is value always positive and suitable for minimization.

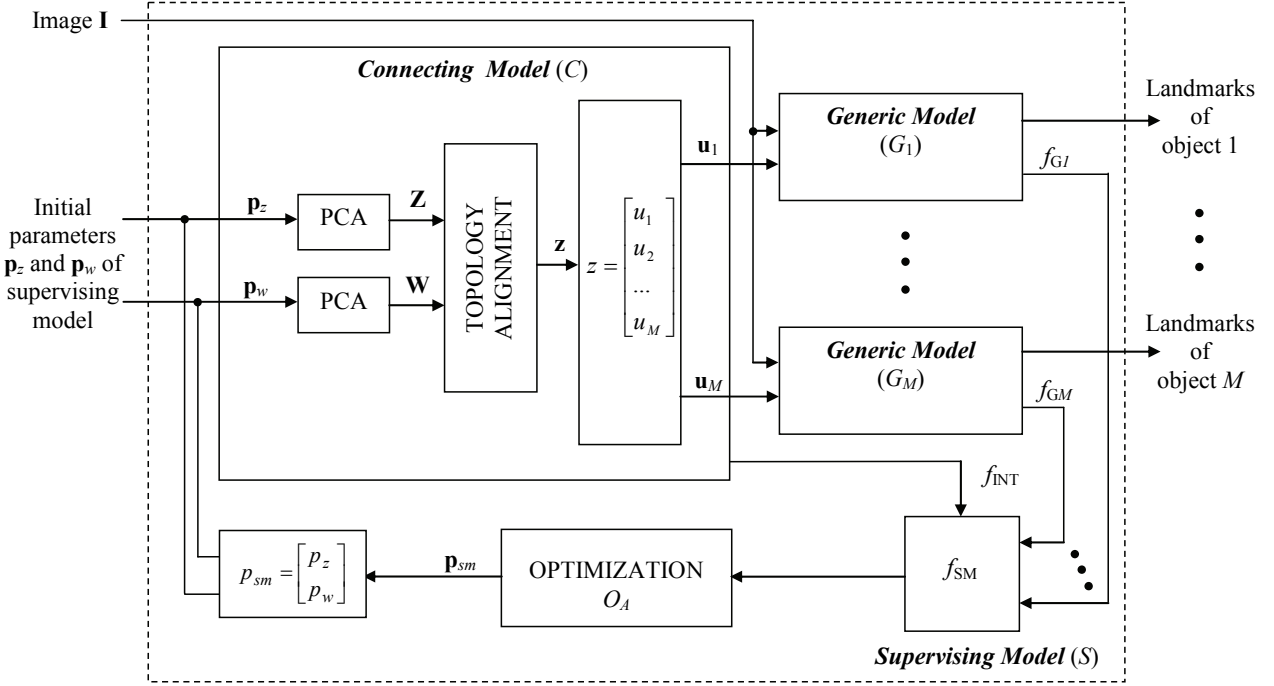


Fig. 3. Steps involved in segmentation of anatomical structures by connected statistical models. A detailed scheme of a generic model used in this strategy is same as shown in Fig. 1 without the optimization loop of the generic model.

Image segmentation by connected statistical models

Segmentation by connected statistical models is illustrated in Fig. 3. Training the statistical models of appearance and topology on a sufficiently large image database results in matrices Θ_x , Θ_y , Θ_z and Θ_w of eigenvectors and corresponding mean vectors of $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$, $\bar{\mathbf{Z}}$ and $\bar{\mathbf{W}}$ of shape, appearance, topology and geometrical transformation, respectively. The segmentation starts with some initial values of topology $\mathbf{p}_z$ and transformation $\mathbf{p}_w$ parameters. Using the statistical models of topology (Eq. 5) and geometrical transformation (Eq. 6), the aligned topology $\mathbf{Z}$ and the set $\mathbf{W}$ of parameters of the geometrical transformation are first obtained. The parameters $\mathbf{W}$ are next used to derive the unaligned topology $\mathbf{z}$ which is further decomposed into pose and shape parameters $[\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_M^T]^T$, $\mathbf{u}_j = [\mathbf{v}_j^T, \mathbf{p}_{xj}^T]^T$, $j = 1, \dots, M$, of the M generic models. The pose and shape parameters together with the image $\mathbf{I}$ are used to derive the energy functions $[f_{G1}, \dots, f_{GM}]$ and landmark positions of the M generic models according to the scheme shown in Fig. 1, where each generic model's optimization loop is omitted. The optimization procedure O_A (Fig. 3) optimizes a criterion function f_{SM} to achieve the best

fit of all generic models while keeping the deformation of the topology small. The function f_{SM} of a supervising model is defined as the weighted sum of internal f_{INT} and external f_{EXT} energies:

$$f_{SM} = \alpha \cdot f_{INT} + (1 - \alpha) \cdot f_{EXT}, \quad (10)$$

The external energy of the supervising model is a weighted sum of all M generic models' energies f_{Gi} :

$$f_{EXT} = \sum_{i=1}^M \beta_i f_{Gi}. \quad (11)$$

The internal energy f_{INT} corresponds to the deformation of the statistical model of topology:

$$f_{INT} = \|\mathbf{Z} - \bar{\mathbf{Z}}\|^2 + \|\mathbf{W} - \bar{\mathbf{W}}\|^2, \quad (12)$$

where $\mathbf{Z}$ and $\mathbf{W}$ contain the parameters of the aligned topology and geometrical transformation, respectively.

EXPERIMENTS AND RESULTS

We have applied the above segmentation strategies, *i.e.*, strategies using unconnected and connected statistical models, to segment anterior-posterior X-ray images of hips and pelvis.

IMAGE DATABASE AND LANDMARKS

The statistical models were trained on the training image database consisting of 42 anterior-posterior X-ray images of hips and pelvis. The resolution of an image was 1700x1200 pixels. To construct statistical models in total 68 landmarks were defined on each image; 16 on left hip, 16 on the right hip, and 36 on pelvis, by three experts using a purpose-specific software tool. The position of each of the 68 landmarks was selected as the position of that landmark out of three which was closest to the average of three. An example of an annotated X-ray image is shown in Fig. 4.

GENERIC AND CONNECTING MODELS OF THE HIPS AND PELVIS

Six generic models G_1 - G_6 were chosen to model the hips and pelvis and the 68 landmarks were assigned to these models (Fig. 4, left). To derive the statistical model of appearance, rectangular regions of interest, each covering all corresponding landmarks, were selected (Fig. 4, right). The statistical model of topology was used as the model that connected the 6 generic models. Fig. 5 shows unaligned and aligned topology vectors for 6 models and 41 images. The affine transformation was used for topology alignment. It is obvious that by topology alignment much more compact clusters are obtained than without alignment. Besides, if more pose and shape

parameters are used to describe the topology more compact classes are obtained. However, as the number of pose and shape parameters (R) and the number of parameters of the affine transformation ($R(R+1)$) are related, a more complex topology model will result in optimization of more parameters. To reduce the number of parameters of the affine transformations we applied PCA on the parameters $\mathbf{W}$, which resulted in vector $\mathbf{p}_w$. In special case, when $M R$ equals $R(R+1)$, the affine transformation maps all of the M pose and shape vectors of generic models to the average topology with zero variance (Fig. 5, right) and, therefore, optimization of topological parameters can be omitted. In this case ($M=R+1$), the number of unknowns ($R(R+1)$) equals the number of equations ($M R$). In any other case (Fig. 5, middle), geometrical information stored in landmarks positions is divided between topological and transformation parameters (Fig. 3), where former describes residual information after transformation/alignment and the latter the information about transformation that brings structures close to the average arrangement. Figure (Fig. 6) illustrates the effect of changing the first 4 PCA parameters of affine transformation ($\mathbf{p}_w$) on the landmark position for the case shown in figure (Fig. 5, right), where topology parameters ($\mathbf{p}_z$) have no effect. The non-linear effects of PCA parameters of the affine transformation ($\mathbf{p}_w$) on the landmarks positions can be observed, e.g. first parameter correlates with generic models rotation parameter (Fig. 7).

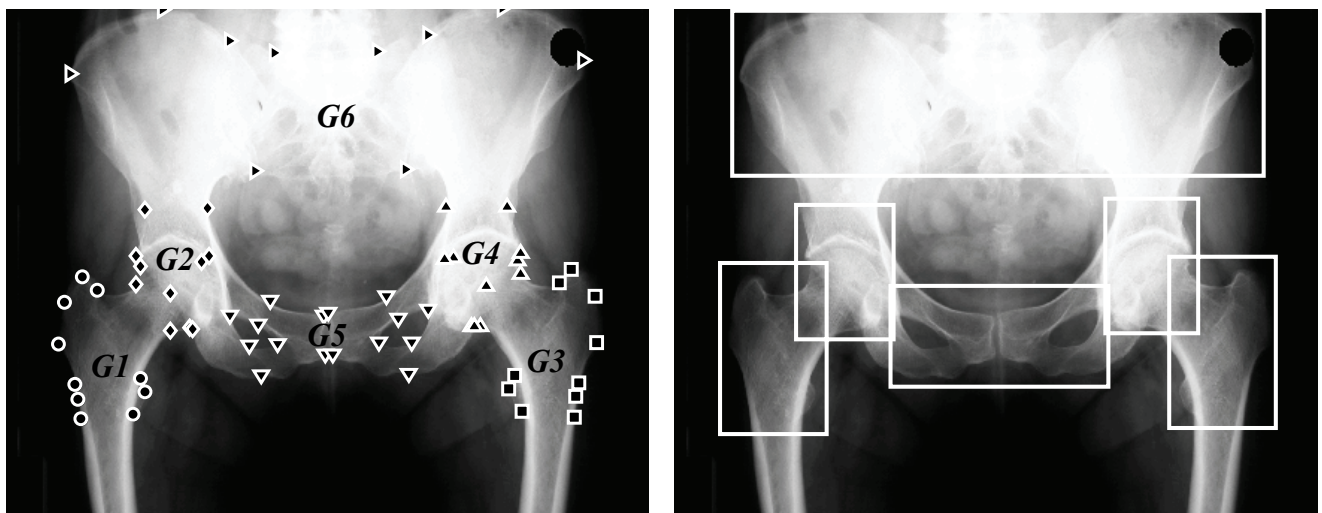


Fig. 4. AP X-ray image of the hips and pelvis with superimposed landmarks belonging to 6 generic models (G_1 - G_6) (left) and the 6 regions of interest used to obtain the statistical models of texture (right). Different symbols represent landmarks belonging to different generic models.

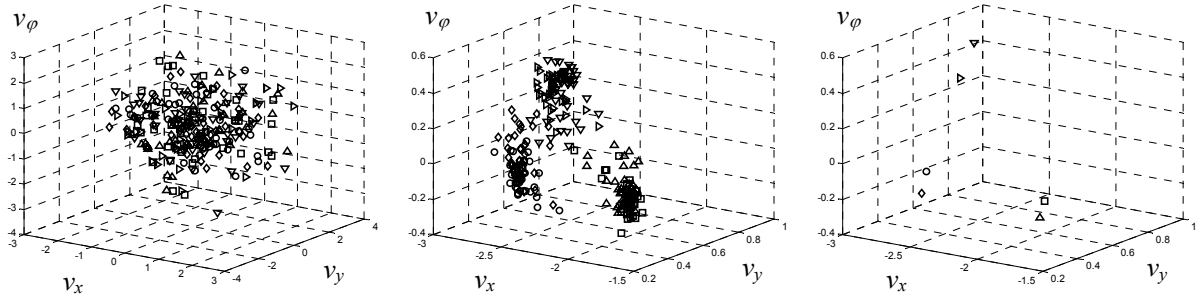


Fig. 5. Unaligned topology $\mathbf{z}$ consisting of 4 pose and one shape parameter (left), aligned topology $\mathbf{Z}$ consisting of 4 pose parameters (middle) and aligned topology $\mathbf{Z}$ consisting of 4 pose and one shape parameter (right) for 6 models and 41 images. Only three dimensions are used to visualize the vectors.

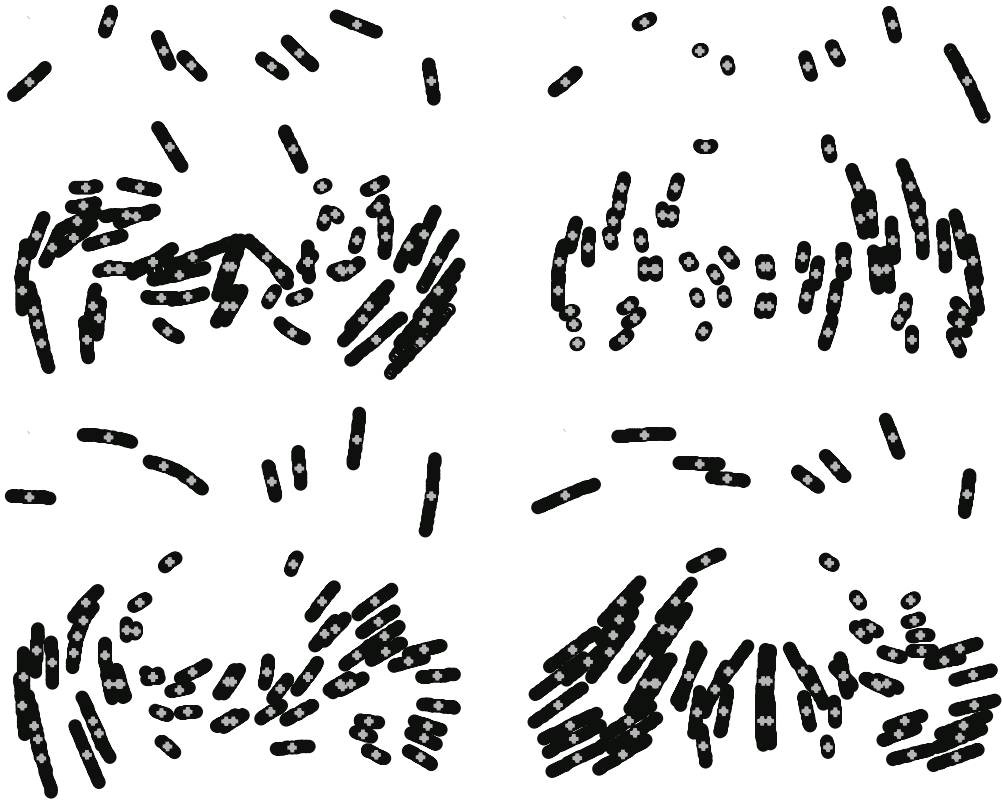


Fig. 6. Traces of landmark positions when changing the first four PCA parameters of the affine transformation (left to right) from -3 to 3 by a step of $1/3$.

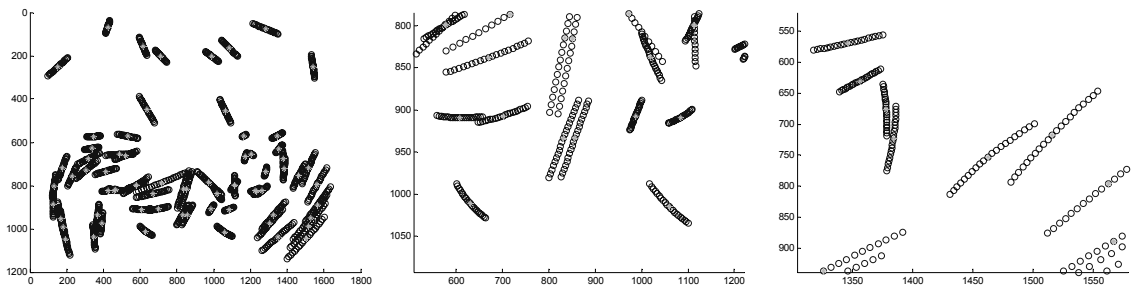


Fig. 7. Traces of landmark positions when changing the first parameter from -3 to 3 by a step of $1/3$ (left). The non-linear effects on the landmarks positions can be observed (middle and right).

EXPERIMENTS

Four statistical model-based segmentation methods have been applied to the 42 AP X-ray images of the hips and pelvis. The first method (M1) modelled the whole image by only one statistical model of appearance. The number of shape parameters ($\mathbf{p}_x$) was set to 15 to allow the single model to be flexible since it had to model 68 landmarks. The second method (M2) modelled the anatomical structures by 6 unconnected statistical models of appearance (Fig. 4). Each of them was independently optimized using 4 pose ($\mathbf{v}$) and 5 shape parameters ($\mathbf{p}_x$) and the strategy shown in Fig. 1. In the third (M3) method, the strategy, illustrated in Fig. 3, was applied. In method M3, the topology vector consisted of 4 pose parameters $\mathbf{v}$ and only the most significant shape parameter $p_{x,l}$, yielding $\mathbf{u}=[v_x, v_y, v_\phi, v_s, p_{x,l}]^T$. Because in M3 the topology parameters $\mathbf{p}_z$ had no effect on landmark positions (see Fig. 5, right), the relations between generic models' pose-shape vectors were modelled by the first twenty parameters $\mathbf{p}_w$ of the PCA affine transformation. The fourth method (M4) used a two-level hierarchical strategy. First, the method M3 was applied, which was then followed by the segmentation strategy shown in Fig. 1 using 4 pose ($\mathbf{v}$) and 5 shape parameters ($\mathbf{p}_x$). The goal of the two-level strategy was to initialize the generic models' parameters within their capturing ranges by first optimizing the parameters of the supervising model and then let the unconnected generic models to fine-tune their parameters. All experiments were conducted on 42 AP X-ray images with a leave-one-out test. In all experiments, the number of texture parameters, used to represent a shape-free and intensity normalized image $\bar{\mathbf{I}}$, was set to 3. For the supervising and generic models we used the downhill simplex optimization method (Press *et al.*, 1992) with 300 iterations. The texture vectors of generic models were sub-sampled in horizontal and vertical directions by 5 pixels for G_1 , G_3 , G_5 , by 4

pixels for G_2 , G_4 , and by 8 pixels for G_6 . In the first method M1, the texture was sub-sampled by 5 pixels. In all methods, we used the mutual information to estimate the external energy. Segmentation results obtained by each method were expressed as the average Euclidean distance (AED) in pixels between automatically and manually derived landmarks positions. Fitting of supervising model takes approximately 10 sec on Intel i7-860 processor, where main computation burden represents computation of generic models cost function (image interpolation, computation of image representation with texture model, cost function computation). The same amount of time is spend on sequential fitting of unconnected models. However, computation of generic models cost function could be parallelized in both cases (connected and unconnected) to speed up the optimization.

RESULTS

Fig. 8 illustrates the segmentation results in the form of box-whiskers diagrams, showing the minimum, maximum, median, 1st and 3rd quartile of the distances between automatically and manually derived landmarks positions. Segmentation using only one model for the hips and pelvis (method M1) performed the worst. Using individual, unconnected, models (method M2) resulted in more accurate landmark positions. Even better results were, however, obtained by the optimization of the supervising model (method M3) using the statistical model of topology. Finally, the results of the two-level optimization (method M4) in which individual generic models were independently optimized after they had been initialized at the first level, show that the landmarks have been most accurately segmented. Figure (Fig. 9) shows the manually defined landmarks and the landmark positions obtained by methods M1, M2 and M4 for 5 X-ray images.

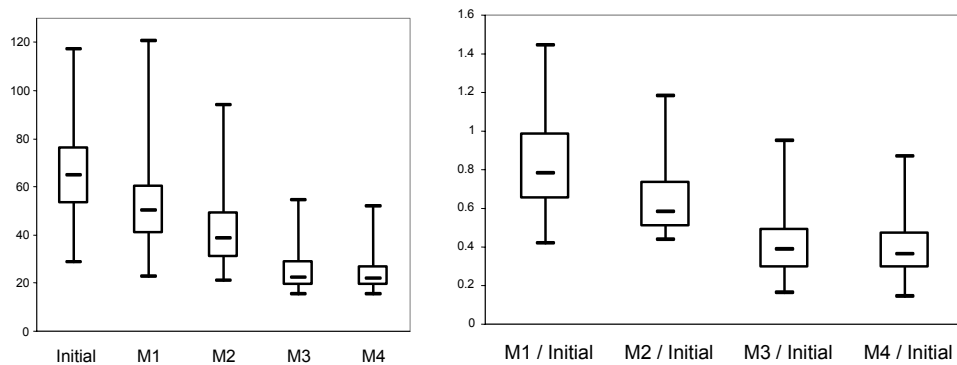


Fig. 8. Box-whiskers diagrams, showing the minimum, maximum, median, 1st and 3rd quartile of distances in pixels between automatically (initial and four methods, M1-M4) and manually derived landmarks positions (left) and box-whiskers diagrams of distances relative to the initial displacement (right).



Fig. 9. Results in the form of landmarks (o) obtained by methods M1 (left), M2 (middle), and M4 (right) and their displacements from the manually (x) defined positions shown for 5 X-ray images (rows).

Table 1 shows the segmentation performance (median of AED distances in pixels) as a function of the number of PCA components of the affine transformation used in the statistical connection model. The results

of this sensitivity analysis show that the number of PCA components in statistical connection model affects the segmentation performance, which was to be expected. Method M4 shows the best performance

Table 1. *The median of AED distances in pixels between automatically and manually derived landmarks positions for the supervising model (method M3) and the two-level optimization (method M4), given as a function of the number of PCA components of the affine transformation used in the statistical connection model.*

Method	Number of PCA components								
	3	5	8	10	13	15	20	25	30
M3	54.9	39.6	28.1	24.9	24.1	24.0	22.9	24.7	23.6
M4	41.8	30.6	26.0	23.4	22.7	22.6	22.3	23.6	23.4

and was always better than M3 method. However, the differences in the performances were larger at small number of PCA components, while the differences at larger number of PCA components were much smaller. Both methods (M3 and M4), which are based on statistical connection model, yielded much better results (Fig. 8) than the method with one model (M1) or the method with unconnected models (M2). The two-layer scheme (Fig. 3) composed of statistical connection model with statistical shape model yields more accurate positions of the landmarks and makes more compact models with fewer parameters, which consequently yields better and more robust segmentation of anatomical structures.

CONCLUSION

We have presented a general framework for segmentation of anatomical structures by connected parametrical models. The framework with the statistical model of topology enables modelling of large variations of spatial non-linearity, which are due to complexity of anatomical structures and inter-patient and inter-acquisition variations in shape, appearance, and spatial arrangement of anatomical structures. The framework, incorporating the statistical model of topology has been tested on 42 AP X-ray images of hips and pelvis. It was shown that the two-level modelling of landmark positions with the statistical model of topology improved accuracy and robustness of the whole model. The results of the presented framework indicate that the framework and statistical modelling of topology is a promising and general tool for constructing robust supervising models, by which a number of generic models can be connected.

ACKNOWLEDGMENTS

We would like to thank the Medical IT - Advanced Development team from Philips Medical Systems for their support and the anonymous reviewers for their constructive comments. The images were provided by Ludvik Travnik, MD, Department of Orthopaedic Surgery, Clinical Center, Ljubljana, Slovenia, and by

Boris Rjazancev, MD, Jesenice Hospital, Jesenice, Slovenia.

REFERENCES

- Bankman IN (2000). Handbook of medical imaging : processing and analysis, San Diego, CA, USA ; London, UK, Academic Press, 901.
- Bernard R, Likar B, Pernuš F. (2001). Segmenting articulated structures by hierarchical statistical modeling of shape, appearance, and topology. In: Nissen W, Viergever MA, eds. Medical Image Computing and Computer Assisted Intervention - MICCAI 2001, 2001; Utrecht, The Netherlands. 499–506.
- Bernard R, Pernuš F. (2001). Statistical approach to anatomical landmark extraction in AP radiographs. In: Sonka M, Hanson KM, eds. Medical Imaging 2001: Image Processing, 2001; San Diego, CA, USA. SPIE, 537–44.
- Bookstein FL (1989). Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. IEEE Trans. Pattern Anal. Mach. Intell., 11, 567–85.
- Cootes TF, Edwards GJ, Taylor CJ. (1998). Active appearance models. In: Burkhardt H, Neumann B, eds. 5th European Conference on Computer Vision, 1998. Springer-Berlin, 484–98.
- Cootes TF, Edwards GJ, Taylor CJ (2001). Active appearance models. IEEE Trans Pattern Anal Mach Intell 23:681–5.
- Cootes TF, Hill A, Taylor CJ, Haslam J (1994). Use of active shape models for locating structure in medical images. Image Vision Comput., 12, 355–65.
- Cootes TF, Taylor CJ. (2001). Statistical models of appearance for medical image analysis and computer vision. In: Sonka M, Hanson KM, eds. Medical Imaging 2001: Image Processing, 2001; San Diego, CA, USA. SPIE, 236–48.
- De Bruijne M, Nielsen M (2005). Multi-object segmentation using shape particles. In: Christensen GE, Sonka M, eds. Information processing in medical imaging, Proceedings. Berlin: Springer-Verlag Berlin, 762–73.
- Goodall C (1991). Procrustes Methods In The Statistical-Analysis Of Shape. Journal Of The Royal Statistical Society Series B- Methodological 53:285–339.
- Heap AJ, Hogg DC. (1997). Improving specificity in PDMs using a hierarchical approach. In: Clark AF, ed. BMVC'97 Proceedings of the 8th British Machine Vision Conference 1997, 80–9.

- Hill A, Cootes TF, Taylor CJ (1996). Active shape models and the shape approximation problem. *Image Vision Comput* 14:601–7.
- Langs G, Peloschek P, Donner R, Bischof H (2007). Multiple appearance models. *Pattern Recognit* 40: 2485–95.
- Litvin A, Karl WC (2005). Coupled shape distribution-based segmentation of multiple objects. *Information Processing in Medical Imaging*, 345.
- Lu CL, Pizer SM, Joshi S, Jeong JY (2007). Statistical multi-object shape models. *Int J Comput Vision* 75: 387–404.
- Mahmoodi S, Sharif BS, Chester EG, Owen JP, Lee R (2000). Skeletal growth estimation using radiographic image processing and analysis. *IEEE Trans Inf Technol Biomed* 4:292–7.
- Mcinerney T, Terzopoulos D (2000). Deformable models. In: Bankman IN, ed. *Handbook of medical imaging*. San Diego, CA, USA; London, UK: Academic Press, 127–45.
- Pham DL, Xu C, Prince JL (2000). Current methods in medical image segmentation. *Annu Rev Biomed Eng* 2:315–37.
- Pham TV, Smeulders AWM (2006). Learning spatial relations in object recognition. *Pattern Recognit Lett* 27:1673.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992). *Numerical recipes in C: The art of scientific computing*. Cambridge: University Press, 59–70, 408–12.
- Smyth PP, Taylor CJ, Adams JE (1997). Automatic measurement of vertebral shape using active shape models. *Image Vision Comput* 15:575–81.
- Sonka M, Fitzpatrick JM (2000). *Medical image processing and analysis*, Bellingham, WA, USA: SPIE Press, 1218.
- Styner M, Gorczowski K, Fletcher T, Jeong JY, Pizer SM, Gerig G (2006). Statistics of pose and shape in multi-object complexes using principal geodesic analysis. In: Yang GZ, Jiang T, Shen DG, Gu L, Yang J, eds. *Medical imaging and augmented reality*. Berlin: Springer-Verlag Berlin, 1–8.
- Tsai A, Wells W, Tempany C, Grimson E, Willsky A (2004). Mutual information in coupled multi-shape model for medical image segmentation. *Med Image Anal* 8:429.
- Van Ginneken B, Haar Romeny BM. (2000). Automatic delineation of ribs in frontal chest radiographs. In: Hanson KM, ed. *SPIE*, 825–36.
- Xu C, Pham DL, Prince JL (2000). Image segmentation using deformable models. In: Sonka M, Fitzpatrick JM, eds. *Handbook of Medical Imaging*. Bellingham, WA, USA: SPIE Press, 129–74.
- Zimmer C, Olivo-Marin JC (2005). Coupled parametric active contours. *IEEE Trans Pattern Anal Mach Intell* 27:1838–42.