# Constraining Contour Deformations Using Statistical A Priori Knowledge of Shape Without Requiring Point-to-Point Correspondence

Ghassan Hamarneh and Tomas Gustavsson‡
Image Analysis Group, Department of Signals and Systems, Chalmers University of Technology

Abstract

*In this paper we present a method for constraining the deformations of active contour models during image segmentation in a way that is consistent with a previously delineated training set of example images. The training examples are carefully traced with complete or semi- manual supervision without the need for point-to-point correspondence. Then frequency domain shape descriptors (in particular, Discrete Cosine Transform coefficients) are used to establish model parameter correspondence. The main modes of shape parameter variation are then captured (using Principal Component Analysis) and used to restrict the active contour model deformations. The method was applied to segment the human left ventricle in echocardiographic images.*

# 1. Introduction

Automatically locating objects of interest in images has long been desirable in many applications. Although segmenting objects in high contrast, noise-free images can be done with simple techniques, problems do arise when the images are corrupted with noise and the object itself is not clearly or completely visible in the image. This may result in detecting erroneous object boundaries or failing to detect true ones. Snakes or Active Contour Models[5] gained large acceptance as a segmentation tool due to the way snakes consider the boundary as a single, inherently connected, and smooth structure. Snakes also support intuitive interactive mechanisms for guiding the segmentation deformations. Numerous variations to the original snakes formulation were proposed to improve their performance. For example, the addition of an inflation force[2], snakes that change topology[6], and using dynamic programming[1] or simulated annealing[4] for snake energy minimization, and many others.

In image segmentation scenarios where high levels of noise and occlusions are present, the traditional snake segmentation techniques lead to unacceptable results. A remedy is to present the snake with a priori information about the shape of the object to be segmented. This is done here by capturing the main modes of the statistical variation of an object's shape using Principal Component Analysis. Previously, PCA was performed on landmark coordinates labeled on many example images of the object of interest[3]. This is problematic, however, since it requires point-to-point correspondence between the landmarks. In medical applications, for example, it is cumbersome to obtain a training data set segmented with point-to-point correspondence. In this paper, the shapes are represented by shape descriptors that eliminate the need for such correspondence, namely the coefficients of the Discrete Cosine Transform (DCT). Demonstration and results on segmenting the left ventricle in echocardiography data are presented.

# 2. Methods

## 2.1 Active Contour Model Formulation

A snake in the continuous spatial domain is represented as a 2D parametric contour curve $\mathbf{v}(s)=(x(s),y(s))$ where $s\in[0,1]$. In order to fit the snake model to the image data we associate energy terms with the snake and aim at deforming the snake in a way that minimizes its total energy. The energy of the snake, $\xi$, depends on both the shape of the contour and the image data $I(x,y)$ reflected via the internal and external energy terms, $\alpha(\mathbf{v})$ and $\beta(\mathbf{v})$, respectively. The total snake energy is written as

$$\xi(\mathbf{v})=\alpha(\mathbf{v})+\beta(\mathbf{v}) \tag{1}$$

The internal energy term is given as:

$$\alpha(\mathbf{v})=\int_0^1 w_1(s)\left|\frac{\partial \mathbf{v}}{\partial s}\right|^2+w_2(s)\left|\frac{\partial^2 \mathbf{v}}{\partial s^2}\right|^2 ds \tag{2}$$

The weighting functions $w_1$ and $w_2$ control the tension and flexibility of the contour. The external energy term is given as:

‡ The authors would like to acknowledge the funding for this work by the Swedish VISIT (Visual Information Technology) program. Corresponding author: Ghassan Hamarneh, ghassan@s2.chalmers.se.

$$\beta(\mathbf{v}) = \int_0^1 P(\mathbf{v}(s))\,ds \tag{3}$$

For the contour to be attracted to image features, the function $P(x,y)$ is designed such that it has minima where the features have maxima. For example, for the contour to be attracted to high intensity changes (high gradients) we can choose:

$$P(x,y) = -c|\nabla[G_\sigma * I(x,y)]| \tag{4}$$

where $G_\sigma * I$ denotes the image convolved with a smoothing (e.g. Gaussian) filter with a parameter $\sigma$ controlling the extent of the smoothing (e.g. variance of Gaussian). The contour $\mathbf{v}(s)$ that minimizes the energy $\xi$ must, according to the calculus of variations, satisfy the vector-valued partial differential (Euler-Lagrange) equation:

$$-\frac{\partial}{\partial s}\left(w_1\frac{\partial \mathbf{v}}{\partial s}\right) + \frac{\partial^2}{\partial s^2}\left(w_2\frac{\partial^2 \mathbf{v}}{\partial s^2}\right) + \nabla P(\mathbf{v}(s)) = \mathbf{0} \tag{5}$$

In a dynamic setting we have $\mathbf{v}(s,t) = (x(s,t), y(s,t))$ and the corresponding constraint equation becomes:

$$\mu(s)\frac{\partial^2 \mathbf{v}}{\partial t^2} + \gamma(s)\frac{\partial \mathbf{v}}{\partial t} - \frac{\partial}{\partial s}\left(w_1\frac{\partial \mathbf{v}}{\partial s}\right) + \frac{\partial^2}{\partial s^2}\left(w_2\frac{\partial^2 \mathbf{v}}{\partial s^2}\right) + \nabla P(\mathbf{v}(s,t)) = \mathbf{0} \tag{6}$$

where $\mu(s)$ and $\gamma(s)$ are mass and damping densities, respectively. In a discrete setting the snake is defined as a set of $N$ nodes: $\mathbf{v}_i(t) = (x_i(t), y_i(t))$ where $i = 1,2,\ldots,N$, and equation (6) (with the assumption of zero mass density, constant damping, and the addition of an inflation force) can be seen as a set of discrete forces acting on each of the snake nodes at specific times according to:

$$\mathbf{v}_i(t) = \mathbf{v}_i(t-1) + \frac{\Delta t}{\gamma}\left(w_1\mathbf{F}_i^{tensile}(t) - w_2\mathbf{F}_i^{flexural}(t) - \mathbf{F}_i^{external}(t) - \mathbf{F}_i^{inflation}(t)\right) \tag{7}$$

where $\mathbf{F}_i^{tensile}(t)$ is the tensile force (resisting stretching) acting on node $i$ at time $t$ and is given in discrete form as (finite differences approximate derivatives):

$$\mathbf{F}_i^{tensile}(t) = 2\mathbf{v}_i(t) - \mathbf{v}_{i-1}(t) - \mathbf{v}_{i+1}(t) \tag{8}$$

$\mathbf{F}_i^{flexural}(t)$ is the flexural force (resisting bending) acting on node $i$ at time $t$ and is given as:

$$\mathbf{F}_i^{flexural}(t) = 2\mathbf{F}_i^{tensile}(t) - \mathbf{F}_{i-1}^{tensile}(t) - \mathbf{F}_{i+1}^{tensile}(t) \tag{9}$$

$\mathbf{F}_i^{inflation}(t)$ is the inflation force acting on node $i$ at time $t$. It is directed towards the normal of the contour and in a direction that will either inflate or deflate the snake towards the target boundary.

$F_i^{external}(t)$ is the external forces acting on node $i$ at time $t$. It is derived from the image data and is defined to be:

$$\mathbf{F}_i^{external}(t) = \nabla P(x_i(t), y_i(t)) \tag{10}$$

where $P$ is defined as in (4).

## 2.2 Snake Re-parameterization: The Discrete Cosine Transform

The one-dimensional discrete cosine transform (DCT) of the sequence of $x_i(t)$ contour coordinates (and similarly for the $y_i(t)$ coordinates) is defined as follows:

$$X(k,t) = w(k)\sum_{i=1}^{N} x_i(t)\cos\frac{\pi(2i-1)(k-1)}{2N} \tag{11}$$

where

$$w(k) = \begin{cases} \dfrac{1}{\sqrt{N}}, & k = 1 \\ \sqrt{\dfrac{2}{N}}, & 2 \le k \le N \end{cases} \tag{12}$$

The DCT was favored as the new frequency domain shape parameterization because it produces real coefficients, has excellent energy compaction properties, and correspondence of its coefficients (when transforming contours with no point-to-point correspondence) is readily available.

## 2.3 Principal Component Analysis

Principal Component Analysis is used to capture the main modes of variation found in the shape parameters (DCT coefficients) describing the objects in the training set of delineated images. The same number, say $M$ (by coefficient truncation), of DCT coefficients is obtained for the set of $x$ and $y$ coordinates that represent each shape in the training set. Performing PCA on the set of coefficients yields the principal components (PCs) or main variation modes, $\mathbf{a}_j$, and the associated variance explained by each PC, $\lambda_j$. Only the first, say $t$, PCs are used being sufficient to explain a desired variance, i.e. $j = 1, 2, \ldots, t$, and are stacked column-wise in the matrix $\mathbf{A}$. The mean of the coefficient vector $\overline{X}$ is also calculated.

## 2.4 Constraining Contour Deformations

After providing a set of images containing the object of interest, the training set of tracings is obtained (contours represented by coordinate-vectors of varying length with no point-to-point correspondence). Then DCT coefficients ( $X$ ) are obtained followed by PCA. Now, presented with a new image containing a similar object, a snake contour is initialized and left to deform (by applying the different forces as in equation (7)) to fit the object. In order to constrain the deformations and force only acceptable shapes (similar to those in the training set), we first obtain the $M$ DCT coefficients for the current active contour. Then the coefficients are forced to produce acceptable shapes by projecting them (similar to [3]) onto the subspace of principal components (allowable shape space) according to:

$$X_{proj} = \overline{X} + \mathbf{A}\mathbf{b} \qquad (13)$$

where $\mathbf{b}$ is a vector of scalars weighing the main variation modes and is calculated as:

$$\mathbf{b} = \mathbf{A}^T\,( X - \overline{X} ) \qquad (14)$$

Before performing the IDCT, we restrict the projected coefficients ( $X_{proj}$ ) to lie within $\pm 3\sqrt{\lambda_j}$ (since most of the population typically lies within three standard deviations of the mean).

# 3. Experiments and Results

The discussed methodology was tested on echocardiographic images of the human left ventricle. Five ultrasonic sequences were available. Each sequence contained 21 frames (total of 105 frames). The frames were noisy, with deterioration typically present in ultrasonography (noise, weak echo, echo dropouts). The left ventricular boundaries were traced manually by medical experts. There was no point-to-point correspondence in the manual tracings between frames (the number of points used in the manual tracings varied from 28 to 312). The DCT of the manual tracings was then obtained. Figure 1 shows an example of the manual tracings and the resulting contours after Inverse-DCT (IDCT) of the truncated DCT coefficients. The ratio ‘energy of truncated contour’/ ‘energy of the original contour’ for increasing numbers of DCT coefficients was examined in order to determine how many DCT coefficients to use.

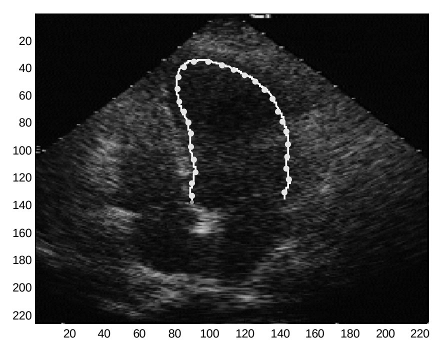


**Figure 1. Ultrasound image with the manual tracing (continuous) and the contour after IDCT of truncated DCT coefficients (dots).**

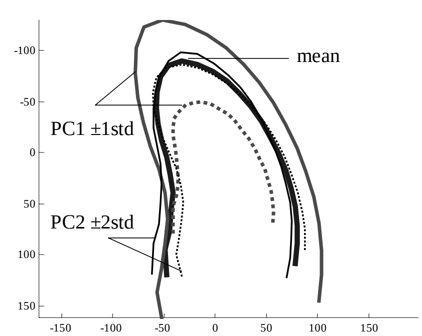


**Figure 2. Mean contour and the first and second variation modes (weighted by ±1 std).**

In order to obtain the ‘allowable shape space’ or a ‘DCT-coefficients-space’, PCA was performed on the truncated DCT coefficients for all the data. Only the main PCs were considered (11 main modes of variation or PCs were used here). The effect of the two strongest variation modes is illustrated in Figure 2. Prior to PCA, the data was represented by the truncated DCT coefficients that are in correspondence (since corresponding coefficients capture specific contour frequency) and also equal in number. Remember that the original data points of the manual tracings had no point-point correspondence and were not equal in number.

In the example shown in Figure 3, we started with one of the manual tracings, added noise to it, obtained the DCT, truncated certain DCT coefficients, projected the remaining coefficients on the allowable shape space, and then performed the IDCT. It is obvious how the constrained contour resembles a much more plausible boundary of the left ventricle.

Figure 4 shows an example of a simulated possible outcome of snake segmentation that, due to noise and drop outs in the image, gives unreasonable and unacceptable shape of the left ventricle. Similar to the previous example an acceptable left ventricle boundary was obtained after DCT-truncation-projection-IDCT.

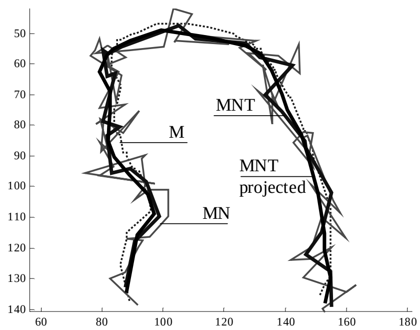


**Figure 3. Manual tracing (M), M with noise (MN), IDCT of truncated DCT coefficients of MN (MNT), and the MNT projected on allowable shape space.**

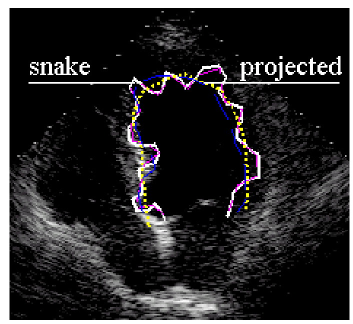


**Figure 4. A synthesized snake overlaid on an ultrasound image of the left ventricle and the result of the DCT-truncation-IDCT-projection onto the allowable shape space.**

# 4. Conclusion

We have developed a method for constraining the deformations of a snake (active contour model) in such a way that the segmentation will always give an acceptable result, i.e. similar to training examples. To capture the typical variations of the training set, a frequency domain shape representation is used to avoid the requirement of point-to-point correspondence. The method was successful when applied to echocardiographic images of the human left ventricle, however, we plan for further investigation and testing.

## References

[1] Amini A, Weymouth T, Jain R. Using dynamic programming for solving variational problems in vision, IEEE Transactions on PAMI 1990, 12(9), 855 -867.

[2] Cohen L D, On active contour models and balloons. CVGIP: Image understanding, 1991, 53(2), 211-218.

[3] Cootes T, Taylor C, Cooper D, Graham J. Active Shape Models - Their Training and Application. Computer Vision and Image Understanding 1995, 61(1), pp.38-59.

[4] Grzeszczuk R, Levin D, "Brownian strings": Segmenting images with stochastically deformable contours. IEEE Transactions on PAMI 1997, 19(10), pp.1100 -1114.

[5] Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. Int'l J Computer Vision 1988, 1(4), pp.321-331.

[6] McInerney T, Terzopoulos D. Topologically adaptable snakes. Proceedings of the Fifth International Conference on Computer Vision 1995, pp.840 -845.