3D Object Tracking Using Directional Procrustes Snake

Mehdi kamandar
Electrical Engineering Dept.
Ferdowsi university of Mashhad, Iran.
mehdi_kamandar@yahoo.com

Seyed Alireza Seyedin
Electrical Engineering Dept.
Ferdowsi university of Mashhad, Iran.
seyedin@um.ac.ir

Hossein Khoshbin
Electrical Engineering Dept.
Ferdowsi university of Mashhad, Iran.
khoshbin@um.ac.ir

Abstract— A novel method of parametric active contours with geometric shape prior is presented in this paper. The main idea of the method consists in minimizing an energy function that includes additional information on a shape reference called a prototype. Shape prior introduced through a similarity measurement between evolving contour and Procrustes mean shape of desired object. This similarity measurement is full Procrustes distance between these two contours that is invariant with respect to similarity transformations (translation, scaling, and rotation). This extra shape knowledge enhances the model robustness to noise, occlusion, complex background, similarity transformations, occlusion, and changing viewpoint of 3D object.

Keywords- parametric active contour, shape prior, similarity transformations, full Procrustes distance, Procrustes mean shape, directional Procrustes snake, 3D object tracking.

I. INTRODUCTION

Image segmentation and 3D object tracking are important research topics for many areas, such as military, medical, and digital video applications. Recently researchers have developed various algorithms for image and video segmentation. Among these algorithms, active contour models, more widely known as snakes, have been extensively used as an edge-based segmentation and tracking method [1].

Snake is a deformable contour on image plane that deforms to seek minimum value of its energy function. This energy function is defined so cleverly that takes its minimum value when fits to a closed boundary of a region in image plane. Hence, snake converts the segmentation problem to minimizing an energy function. By now, there are two kinds of active contour models: parametric and geometric active models. Parametric models explicitly parameterize active contours or surfaces in some certain ways. A geometric model based on curve evolution theory and the level set method can automatically handle topology adaptation with more complicated computing involved.

A critical problem is how to add shape prior information to energy function of snake, driving it toward boundary of desired object in image. By now, several methods have been proposed for adding shape prior information to energy function of geometric active models and region based active contours [2, 3, 4]. Parametric models have lower computation complexity. Hence, we decided to add shape prior information to energy function of them. Charmi et al. [5] have proposed a shape energy term according to distance between Fourier descriptors of evolving contour and a reference template. In our method [6], shape knowledge was introduced through a similarity measurement between evolving contour and Procrustes mean shape of desired object. This similarity is measured by full Procrustes distance between these two contours that is invariant with respect to similarity transformations (translation, scaling, and rotation) and is added to traditional energy function of snakes as an extra shape energy term. Procrustes mean shape of desired object is extracted from a training set of its sample shapes [5]. This extra shape knowledge enhances the model robustness to noise, occlusion, similarity transformations, and complex background.

In general edge information is used as an image energy term which usually represented by the gradient magnitude of an image. However, when an image has complex background or heavy noise, the snake gets confused, and finding the correct object boundary from the gradient magnitude only is not easy. Hence, we improve robustness of the method against complex background and noise, by including gradient direction information in the image energy term. Because of using full Procrustes distance and gradient directional information in energy function, we call this new snake “Directional Procrustes snake”.

Finally we perform 3D object tracking by minimizing the energy function of directional Procrustes snake. The final contour in the current frame or a prediction by Kalman filtering will be chosen as initial contour for the next frame and will be fitted the boundary of desired object by minimizing the energy function of directional Procrustes snake. The minimizing process is done by greedy algorithm [6]. Greedy is a suboptimum and fast algorithm for
minimizing the energy function of snakes. We obtained promising results of 3D object tracking by new snake, representing the robustness of our tracking method against noise, complex background, occlusion, similarity transformations, and changing viewpoint of a 3D object.

This paper organized as follows: section 2 describes Procrustes shape analysis. In section 3, we extract the energy function of directional Procrustes snake. Experimental results for 3D object tracking are presented and discussed in section 4. Conclusion is presented in section 5.

II. PROCRUSTES SHAPE ANALYSIS

Procrustes shape analysis is a particularly popular method in direction statistics and is intended to cope with 2D shapes. A discretized shape in 2D space can be described by a vector of \( n \) complex numbers:

\[
\begin{align*}
\mathbf{z} &= [z_1, z_2, \ldots, z_n]^T \\
z_i &= x_i + jy_i, \quad i = 1, 2, \ldots, n
\end{align*}
\]

\((x_i, y_i)\) are the Cartesian coordinates of the \( i \)th landmark of shape, vector \( \mathbf{z} \) is called a configuration (We will represent vectors by using bold letters). Fig. 1 represents a continuous shape and its discretization by \( n \) equidistant landmarks. For two shapes \( z_1 \) and \( z_2 \), if their configurations are equal through a combination of translation, scaling and rotation (similarity transformations), i.e.:

\[
\begin{align*}
z_{1\alpha} &= \alpha \mathbf{a}_n + \beta z_2, \quad \alpha, \beta \in C \\
\alpha &= \alpha_R + j\alpha_I \\
\beta &= \beta e^{j\angle(\beta)} 
\end{align*}
\]

We may consider \( z_1, z_2 \) represent the same shape. In (2), \( 1_n \) is a \( n \times 1 \) vector with entries 1, \( \alpha_R \times 1_n \) translates \( z_2 \) by \( \alpha_R \) units in the horizontal axis direction and \( j \times \alpha_I \times 1_n \) translates \( z_2 \) by \( \alpha_I \) units in the vertical axis direction. \( |\beta| \) scales and \( \angle(\beta) \) rotates \( z_2 \). It is very convenient to center shapes by defining the centered configuration as

\[
\mathbf{u} = [u_1, u_2, \ldots, u_n]^T, \quad u_i = z_i - \bar{z}, \quad \bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i.
\]

The full Procrustes distance between two configurations \( \mathbf{u}_1, \mathbf{u}_2 \) can be defined as [7] (we suppose that corresponding points on two contours have similar indices in two configurations):

\[
d_F^2(\mathbf{u}_1, \mathbf{u}_2) = \min_{\alpha, \beta} \left\| \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} - \alpha 1_n - \beta \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \right\|^2 (3)
\]

Hence Procrustes distance is Euclidean distance between aligned configurations with respect to similarity transformations, having similar position, scale, and orientation in image plane. Minimizing the above objective function with respect to \( \alpha \) and \( \beta \), we have: \( \alpha = 0 \),

\[
\beta = u_2^* u_1^\dagger \left( \|u_1\| \times \|u_2\| \right) \quad \text{where superscript } * \text{ represents the complex conjugation transpose. Substituting } \alpha \text{ and } \beta \text{ in (3), we have:}
\]

\[
d_F^2(\mathbf{u}_1, \mathbf{u}_2) = 1 - \frac{u_1^* u_2^* u_1^\dagger}{u_1^\dagger u_1 u_2^\dagger u_2} = 1 - \frac{|u_1^* u_2^*|^2}{\|u_1\|^2 \|u_2\|^2}
\]

Based on Cauchy-Schwarz inequality (\( |u_1^* u_2^*|^2 \leq \|u_1\|^2 \|u_2\|^2 \)) we can show that \( 0 \leq d_F(\mathbf{u}_1, \mathbf{u}_2) \leq 1 \). When \( d_F(\mathbf{u}_1, \mathbf{u}_2) = 0 \), two configurations \( \mathbf{u}_1, \mathbf{u}_2 \) represent the same shape and when \( d_F(\mathbf{u}_1, \mathbf{u}_2) = 1 \), \( \mathbf{u}_1, \mathbf{u}_2 \) represent two shapes that have no resemblance to each other. We conclude that smaller \( d_F(\mathbf{u}_1, \mathbf{u}_2) \) means that \( \mathbf{u}_1, \mathbf{u}_2 \) represent two shapes that have more resemblance to each other. Consequently, full Procrustes distance measures degree of resemblance of two shapes independent of their position, scale and rotation (similarity transformations) in image plane. For using full Procrustes distance in snake energy, we need a mean shape (prototype) for desired object, to measure resemblance of evolving contour with it. Given a training set composed of \( m \) sample shapes of desired object \( (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m) \), we can find their mean by finding \( \mathbf{u} \) that minimizes the objective function in (5) [7]:

\[
\mathbf{u} = \arg \inf_{\mathbf{u}} \sum_{i=1}^{m} d_F^2(\mathbf{u}, \mathbf{u}_i) \quad (5)
\]

Hence, \( \mathbf{u} \) is the shape that has the smallest possible Procrustes distance (maximum resemblance) from each of the sample shapes. From (4) and (5), we have:

\[
\mathbf{u} = \arg \inf_{\mathbf{u}} \sum_{i=1}^{m} \frac{1}{2} - \frac{\mathbf{u}_i^* \mathbf{u}_i^\dagger \mathbf{u}_i}{\mathbf{u}_i^\dagger \mathbf{u}_i \mathbf{u}_i^\dagger \mathbf{u}_i} = \arg \inf_{\mathbf{u}} (m - \mathbf{u}_i^* \mathbf{u}_i^\dagger \mathbf{u}_i) \quad (6)
\]

Where \( m \) is a constant representing number of sample shapes and:
\[ S = \sum_{i=1}^{m} (u_i^* u_i^*)/(u_i^* u_i) \]  

(7)

From (6) we conclude that:

\[ \hat{u} = \arg \sup_{u \in \mathcal{H}} u^* S u \]  

(8)

If we suppose that \( \lambda_i \) and \( v_i \) (\(|v_i|=1\)) are corresponding eigenvalues and eigenvectors of matrix \( S \), respectively, we will have:

\[ S v_i = \lambda_i v_i \Rightarrow v_i^* S v_i = \lambda_i \quad i = 1, 2, ..., n \]

\[ (v_i^* S v_i)_{\text{max}} = (\lambda_i)_{\text{max}} \]  

(9)

Comparing (8) and (9), we conclude that the Procrustes mean shape \( \hat{u} \) is the dominant eigenvector of \( S \), i.e., the eigenvector that corresponds to the greatest eigenvalue of \( S \), provided that distribution of sample shapes in shape space follows Gaussian distribution.

III. ENERGY FUNCTION OF DIRECTIONAL PROCRUSTES SNAKE

A traditional snake is a controlled continuity spline that moves and localizes onto a specified contour under the influence of its energy function minimization [1]. Let a snake be a parametric contour, \( v(s) = (x(s), y(s)) \), where parameter \( s \in [0, 1] \). It moves around the image spatial domain to minimize the traditional discretized energy function as defined by: [10, 11]

\[ E_{\text{shape}}(\nu) = w_i \times \sum_{i=1}^{n} (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + ... \]

\[ w_2 \times \sum_{i=1}^{n} \left( d - \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \right)^2 + ... \]

\[ w_3 \times \sum_{i=1}^{n} (x_i + 2x_i + x_{i-1})^2 + (y_i + 2y_i + y_{i-1})^2 + ... \]

\[ + w_4 \times \sum_{i=1}^{n} \left| \nabla I(x_i, y_i) \right|^2 \]  

(10)

In the above equation, \( v \) is the evolving contour with \( n \) snaxels, \( (x_i, y_i) \) are the Cartesian coordinates of \( i \)th snaxel and \( (x_0, y_0) = (x_n, y_n) \). \( w_i \) \( s \) are constant weights used to tune the impact of each energy terms. \( d \) is the average distance between adjacent snaxels on contour.

The first term in above energy function is called first order continuity. Minimization of this term reduces the distance between adjacent snaxels. Hence it prevents of gaps in contour that are due to noise and pseudoedges. Existence of this term is essential, because when snake lies on homogeneous regions of image, the image energy is negligible and only minimization of this term can move the snake toward boundary of desired object. One important problem with this term is a tendency for snaxels to bunch up on a strong portion of an edge [11]. For solving this problem, we use the second term in energy function. This term encourages even spacing of snaxels. It tends to keep the distance between each pair of adjacent snaxels equal and prevents this tendency.

The third term in above energy function is called second order continuity. If the \( i \)th snaxel is pushed toward the midpoint of two adjacent snaxels, this term will be minimized. Hence it prevents of sharp corners in contour that are usually due to noise and pseudoedges. The first three terms in above energy function that are independent of image information \((I(x, y))\), are called internal energy terms.

The fourth term (image energy term) considers the gradient magnitude. Minimizing this term, the snaxels will be attracted to locations in image with large gradient magnitude, i.e., strong edges.

Using the traditional energy function, snake suffers from many problems such as noise, strong edges in complex background, occlusion, and the limited capture range. Consequently snake cannot find correct boundary of desired object. For overcoming these problems, we add two extra terms to traditional energy function, one for shape information and one for considering gradient direction information. Shape knowledge is introduced through a similarity measurement between evolving contour and Procrustes mean shape of desired object. This similarity measurement is full Procrustes distance between these two contours that is invariant with respect to similarity transformations and is added to traditional energy function of snakes as an extra shape energy term \( E_{\text{shape}} \). The smaller \( E_{\text{shape}} \), the more similar the contour is to the sample shapes.

For using of Procrustes distance between two contours, correspond saxels on them must have similar indices in two configurations. For this purpose, we reparametrize \( v \) by changing the starting snaxel to \( i \)th snaxel (applying circular shift to \( v \)), representing it by \( v' \), and calculate \( d^2(v', \hat{u}) \) for each parametrization. The shape energy will be replaced with minimum of \( d^2(v', \hat{u}) \) over \( n \) possible parametrization. In this case the energy function is invariant with respect to the starting snaxel.

Procrustes mean shape of desired object is extracted from a set of \( m \) sample shapes according to section 2. In many situations, the distribution of sample shapes in shape space does not follow Gaussian distribution e.g., different views of a single 3D object that form distinct clusters in shape space. Hence, firstly we cluster them to \( k \) clusters by \( k \)-mean clustering algorithm (usually \( 8 \leq k < m \)). In this case the distribution of the sample shapes for each cluster, with large probability, follows Gaussian distribution. So we can extract a mean shape for each cluster (the dominant eigenvector of covariance matrix of sample shapes in each cluster) and model the 3D object by these \( k \) mean shapes \( \{\hat{u}_1, \hat{u}_2, ..., \hat{u}_k\} \).
We compute full Procrustes distances between evolving contour \( v \) and these \( k \) mean shapes of desired object and replace the shape energy term with minimum of them. In this case the final energy function is invariant with respect to changing viewpoint of 3D object.

In general edge information is used as an image energy term which usually represented by the gradient magnitude of an image (fourth term). However, when an image has complex background or heavy noise, the snake gets confused, and finding the correct object boundary from the gradient magnitude only is not easy. For more robust tracking in complex background and heavy noise, we also use an extra energy term that minimizes difference between gradient direction and contour normal direction because they are the same for boundary points.

The final energy function of directional Procrustes snake is a linear combination of six energy terms (11). Gradient magnitude term has nonlinear variations versus spatial variations, i.e., a small spatial variation can change the gradient magnitude severely, which is not the case for the variations of the boundary of desired object. Because of the discrete nature of the energy function of Directional Procrustes snake (11), we use greedy or random search algorithms for minimizing it.

IV. EXPERIMENTAL RESULTS

As an important application of directional Procrustes snake, we present experimental results of 3D object tracking. We perform 3D object tracking by minimizing the energy function of directional Procrustes snake. The final contour in the current frame or a prediction by Kalman filtering will be chosen as initial contour for the next frame and will be fitted the boundary of desired object by minimizing the energy function of Directional Procrustes snake in a predefined search region around each snaxel. The size of the search region depends on maximum expected displacement of the object in two consecutive frames. The gradient map is only computed over these search regions, not all of image plane. Hence, we have \( n \) (the number of snaxels) sub gradient maps. The minimizing process is done by greedy algorithm [9]. Greedy is a suboptimum and fast algorithm for minimizing the energy function of snakes.
shapes. Hence we cluster sample shapes to 8 clusters by \( k \)-mean clustering algorithm. Fig. 4 represents sample shapes for two clusters of 8 clusters and the Procrustes mean shape extracted for each cluster.

Fig. 5 represents 3D object tracking for different values of free parameters \((w_f)\). Everywhere the values of some \( w_i \)s are not mentioned, they are appropriate values opposite zero. The first row represents results for conventional snake \((w_5 = w_6 = 0)\). Because of complex background and occlusion, snake cannot follow the object. The second row represents results after adding shape prior to energy function \((w_5 = 1, w_6 = 0)\). Although the snake can preserve the shape of desired object, it traps in complex background \( i.e., \) snaxels are attracted to points with high gradient magnitude in background. For solving this problem, we add gradient direction information to energy function \((w_6 = 1)\). The third row represents the results for this case. We observe that many of snaxels cannot still follow the correct boundary of the object and are attracted to strong edges in complex background. This is due to nonlinear variations of gradient magnitude. Hence, we apply smoothing Gaussian filters with standard deviation equals 3 to each of sub gradient maps computed over search region around each snaxel. The fourth row represents the tracking results obtained by minimizing the complete energy function of Directional Procrustes snake \((11)\). This results show robustness of the method against complex background, occlusion, noise, and changing viewpoint of 3D object.

The computation complexity of the method depends on the number of snaxels, the size of predefined search region, and the number of iterations of minimizing process. In the above example, the number of snaxels is 12, the size of predefined search region is \(4 \times 4\) (15 pixels around each snaxel), and 2 iterations. Hence we should compute the energy function \((11)\), \(12 \times 16 \times 2 = 384\) times for each frame. For the above example the average processing time, on a serial computer (Intel Celeron 3GHz, 512 MB RAM), for each frame is only 0.1 second.

V. CONCLUSION

We have developed a parametric active contour model, named the Directional Procrustes snake, which uses information about shape prior of desired object and gradient direction in addition to its magnitude. By incorporating this information provides better segmentation and tracking results in the scenes with complex background, heavy noise, occlusion, and changing viewpoint of 3D object. Shape information is introduced through similarity measurement between evolving contour and Procrustes mean shape of desired object. This similarity measurement is full Procrustes distance between these two contours that is invariant with respect to similarity transformations. Mean shape of desired object is extracted from a set of sample shapes according to Procrustes mean shape algorithm. We also apply \( k \)-mean clustering algorithm to sample shapes and extract a mean shape for each cluster, to solve the viewpoint changing of 3D object during tracking process.
Figure 5. The first row represents the tracking results for conventional snake ($w_5=0$, $w_6=0$), the second row represents the tracking results after adding shape prior to energy function ($w_5=1$, $w_6=0$), the third row represents the tracking results after adding gradient direct information to energy function ($w_5=1$, $w_6=1$), the fourth row represents the tracking results for minimizing complete energy function of directional procrustes snake including smoothing of sub gradient maps by applying smoothing Gaussian filters with standard deviation equals 3 to them.

REFERENCES


