AN AFFINE SHAPE CONSTRAINT FOR GEOMETRIC ACTIVE CONTOURS

Mohamed Amine Mezghich, Mawaheb Saidani, Slim M'Hiri and Faouzi Ghorbel

GRIFT Research Group, CRISTAL Laboratory
École Nationale des Sciences de l'Informatique, ENSI
Campus Universitaire de la Manouba, 2010 Manouba, Tunisia
ma.mezghich@cristal.rnu.tn, {mawaheb.saidani,slim.mhiri,faouzi.ghorbel}@ensi.rnu.tn

ABSTRACT
We intend to present in this research a new method to incorporate affine invariant geometric shape prior into a level set based active contours for robust segmentation of partially occluded object. The proposed shape constraint is defined after an affine shape alignment of the active contour and a reference shape. In order to generalize our work to the multi-references case, we present a set of complete and stable invariant shape descriptors computed using Fourier transform of the contours to choose the most suitable reference according to the evolving contour. Experiments on synthetic and real data show the ability of the proposed approach to constrain an evolving curve towards a target shapes that may be occluded and cluttered under geometric transformations.

Index Terms— Affine shape alignment, affine invariant shape descriptors, shape constraint, active contours.

1. INTRODUCTION
Precise object of interest detection is of high need in many applications of image processing such as medical imaging, tracking of moving object and 3D object reconstruction. There have been several works in this field of image segmentation and object detection. Among them we can distinguish the classical active contours which are intensity-based models [1, 2, 3, 4, 5] and the constrained models by a prior knowledge that are driven both by global geometrical [6, 7, 8, 9] or statistical [10, 11, 12, 13, 14] shape information and a local image information like gradient or curvature. There have been promising results given that the classical snake models can detect object with smooth contours whereas the level set based models have the additional ability to detect simultaneously many objects in the image. However, several work of the prior knowledge driven models reached good segmentation results by detecting partially occluded shapes presenting missing part in low contrast or very noisy image. To define prior knowledge, the authors use either shape alignment [7] or registration [9] between the contour of the target shape and the shape of reference or a distance between invariant shape descriptors like [8]. All these work manage the case of Euclidean transformations (translation, rotation and scale factor). But in real situations, the object of interest can submit large transformations (affine or projective cases) and distortions. Hence the class of rigid transformations are not appropriate to model this type of movement between the shape of reference and the target one. To our knowledge, there has been only the work of Foulonneau et al. [15] that treats the case of affine transformations. Although this approach, which is based on the distance between an invariant set of Legendre descriptors of the target and reference shapes, presented good results, this method suffers from the instability of the Legendre moments and the order that has to be fixed empirically. Besides, the method requires a heavy execution time to achieve satisfactory results. In order to extend the work of shape priors presented in [16, 17]. We propose in this research a geometric approach to manage the class of affine transformations that can occur between the shape of interest and the reference one. We use an explicit affine shape alignment method which is based on the Fourier transform of the curve to define prior knowledge. At the beginning, we consider that the shape of reference is given, then we were based on a set of an affine invariant shape descriptors which are complete and stable to select the best template in case of many available references. The remainder of this paper is organized as follows : In Section 2, we will recall the used shape alignment method based on Fourier transform. Then, the proposed shape priors will be presented in Section 3. Experiments will be presented and commented in Section 4. Finally, we conclude the work and highlight some perspectives in Section 5.

2. AFFINE SHAPE AND MOTION DESCRIPTION
In order to define affine invariant shape prior from a reference shape, we have to perform shape matching between the target and the template shapes $F_1$ and $F_2$. To correctly estimate the affine motion parameters, we have to deal with the same number of points which is not relevant if we consider two different points of view. For this purpose in [18], the author proposes to use the affine reparametrization of curve which is invariant...
under affine transformation. We will start by presenting the affine reparametrization procedure of a given curve. Then we will describe our method of affine motion’s parameters estimation. By the last paragraph of this section, we present the set of invariant shape descriptors used to choose the best reference.

2.1. Reparametrization of closed curves

If we take the same object in two different camera sides, we found a different number of contour points in each image. Consequently, we proceed by normalization of curves under affine transformation. Given the shape front, its edge pixels are extracted and traversed to yield a discrete closed curve which is a parametric equation \( \gamma(t) = (u(t), v(t)) \) where \( t \in [0, ..., N - 1] \) and \( \gamma(N) = \gamma(0) \). We use the affine arc length reparametrization to normalize closed curve under affine transformation.

\[
 s_a(t) = \frac{1}{L_a} \int_{[a,b]} \| \gamma'(t) \times \gamma''(t) \| dt, \quad t \in [a, b] \tag{1}
\]

\[
 L_a = \int [\gamma'(t) \times \gamma''(t)] dt \tag{2}
\]

Where \( L_a \) is the total affine arc length of the considered curve, \( \times \) represents the cross product between two vectors and \( \| . \| \) denotes the Euclidean norm.

2.2. Contours alignment using geometrical affine parameters estimation

We consider two closed curves \( O_1 \) and \( O_2 \) which define the same shape. \( O_1 \) and \( O_2 \) are said to be related by an affine transformation if and only if

\[
 h(l) = \alpha A f (l + l_0) + B \tag{3}
\]

where \( B \) is a translation vector, \( A \) is a linear transformation, \( l_0 \) is the shift value, \( \alpha \) is the scale factor, \( f \) and \( h \) are the affine reparametrization of two contours having the same affine shape. The Fourier transform in this case corresponds to the Fourier coefficients of an affine arc length parametrization of a given curve. So, In Fourier space we get:

\[
 U_k(h) = \alpha e^{2i\pi kl/\alpha} AU_k(f) + b\delta_k \tag{4}
\]

where \( U_k(h) \) and \( U_k(f) \) are respectively the Fourier coefficients of \( f \) and \( h \). So estimation of affine motion can be resumed to the estimation of it’s three parameters : the affine matrix \( A \), the shift value \( l_0 \) and finally the scale factor \( \alpha \) if we consider a normalization under translation.

2.2.1. Estimation of the scale factor \( \alpha \)

The scale factor can be estimated using the following formula

\[
 \alpha^2 = \frac{\text{det}(U_h(k), U_h^*(k))}{\text{det}(U_f(k), U_f^*(k))} \tag{5}
\]

where \((U_h(k),U_h^*(k))\) and \((U_f(k),U_f^*(k))\) are \(2 \times 2\) matrix formed by Fourier coefficients on some fixed index \( k \) and \( U^* \) is the \( U \) complex conjugate.

2.2.2. Computation of the Shift value \( l_0 \)

Let’s consider \( M_1 = (U_h(k_1), U_h(k_2)) \) and \( M_2 = (U_f(k_1), U_f(k_2)) \)

\[
 l_0 = \frac{\text{arg}(\text{det}(M_2)) - \text{arg}(\text{det}(M_1))}{(k_1 + k_2)} \tag{6}
\]

where \( k_1 \) and \( k_2 \) are two fixed index and \( \text{arg}(U) \) is the complex argument of \( U \).

2.2.3. Computation of the Matrix \( A \)’s parameters

Since scale factor \( \alpha \) and shift value \( l_0 \) have been previously estimated (eq. 5) and (eq. 6), we are going to estimate affine matrix \( A \) parameters. Let \( O_1, O_2 \) be two curves and \( U_f, U_h \) respectively their Fourier coefficients. In Fourier space we have (eq. 7)

\[
 U_h(k) = \alpha e^{jkl_0} AU_f(k) \tag{7}
\]

Using matrix, we have the following formula (eq. 8)

\[
 \begin{pmatrix}
 u_2(k) \\
 v_2(k)
 \end{pmatrix} = \begin{pmatrix}
 a_1 & a_2 \\
 a_3 & a_4
 \end{pmatrix} \begin{pmatrix}
 u_1(k) \\
 v_1(k)
 \end{pmatrix} \tag{8}
\]

So we have to estimate \( a_1, a_2, a_3 \) and \( a_4 \). This system can be represented as following

\[
 \begin{cases}
 u_2(k_1) = \alpha e^{jkl_0} a_1 u_1(k_1) + \alpha e^{jkl_0} a_2 v_1(k_1) \\
 v_2(k_1) = \alpha e^{jkl_0} a_3 u_1(k_1) + \alpha e^{jkl_0} a_4 v_1(k_1) \\
 \vdots \\
 u_2(k_n) = \alpha e^{jkl_0} a_1 u_1(k_n) + \alpha e^{jkl_0} a_2 v_1(k_n) \\
 v_2(k_n) = \alpha e^{jkl_0} a_3 u_1(k_n) + \alpha e^{jkl_0} a_4 v_1(k_n)
 \end{cases} \tag{9}
\]

Then to estimate matrix \( A \)’s parameters (eq. 7), we have to resolve a system with \( 2N \) equations and \( 4 \) unknown parameters that can be written as

\[
 K_{2n \times 4} A_4 = U_{2n} \tag{10}
\]

The solution of the system (eq. 9) can be obtained by

\[
 KA - U = e \tag{11}
\]

so we have to minimize the quadratic error \( e \) by using the pseudo-inverse of \( K \). We show by the following figure an example of affine motion estimation between two synthetic curves after contours re-sampling. For more affine contours matching results using the method explained above, the reader can be referred to [19, 20, 21, 22]. The final result of curves alignment after affine motion estimation are presented by (Figure.2). The estimated motion parameters are presented by Table 1.
In presence of many templates, we have to choose the most suitable one according to the evolving curve. Let $\alpha$ and $\beta$ be positive real numbers, and $k_0$, $k_1$, $k_2$ and $k_3$ four positives integers. Let $C^x_n$ and $C^y_n$ be the complex Fourier coefficients of the coordinates $(u, v)$, $\Delta$ denotes the determinant.

$$\Delta^m_n = \Delta \begin{vmatrix} C^x_n & C^y_n \\ C'^x_n & C'^y_n \end{vmatrix}$$  \hfill (13)

In [23], the author introduced two sets of invariant descriptors $I$ and $J$ which are respectively given by (eq. 14) and (eq. 15).

$$I_{k_1} = |\Delta_{k_1,k_0}|, \quad I_{k_2} = |\Delta_{k_2,k_0}|, \quad I_k = \frac{\Delta_{k_1,k_0} \Delta_{k_2-k_1} \Delta_{k_1-k_0} \Delta_{k_2,k_0}}{\Delta_{k_1,k_0} \Delta_{k_2-k_1} \Delta_{k_2,k_0}}$$
$$\quad \text{for all } k \in N^* - \{k_0, k_1, k_2\}$$

$$J_{k_1} = |\Delta_{k_1,k_3}|, \quad J_{k_2} = |\Delta_{k_2,k_3}|, \quad J_k = \frac{\Delta_{k_1-k_3} \Delta_{k_2-k_3} \Delta_{k_1,k_3} \Delta_{k_2,k_3}}{\Delta_{k_1-k_3} \Delta_{k_2,k_3}}$$
$$\quad \text{for all } k \in N^* - \{k_1, k_2, k_3\}$$

In [19], [20], authors have shown experimentally that such descriptors are complete and stable. The completeness guarantee the uniqueness of matching, the stability gives a robustness under non linear shape distortions and numerical errors. In [23], the author demonstrates that the shape space $S$ can be considered as a metric space with a set of metrics. Hence, the Euclidean distance (eq.16) between the set of the presented invariants can be used to compare the evolving curve and the available templates.

$$d_{\alpha}(F, H) = ||I_{\alpha}(f) - I_{\alpha}(h)||_p = (\sum |I_{\alpha}(f) - I_{\alpha}(h)|^\alpha) ^ {\frac{1}{\alpha}}$$ \hfill (16)

for any real number $\alpha$ such that $\alpha > 1$. Where $f$ and $h$ are two normalized affine arc length parametrization of two objects having respectively the shapes $F$ and $H$. The shape having the minimum distance according to the evolving active contour is used as template.

3. SHAPE PRIORS FOR GEOMETRIC ACTIVE CONTOURS

Geometric active contours are iterative segmentation methods which use the Level Set approach [2] to determine the evolving front at each iteration. Several models have been proposed in literature that we can classify into edge-based or region-based active contours. In [4], the level set approach is used to model the shape of objects using an evolving front. The evolution’s equation of the level set function $\phi$ is

$$\phi_t + F |\nabla \phi| = 0,$$  \hfill (17)

$F$ is a speed function of the form $F = F_0 + F_1(K)$ where $F_0$ is a constant advection term equals to $(\pm 1)$ depends of the object inside or outside the initial contour. The second term...
The total discrete evolution’s equation that we propose is as follows

\[ g(x, y) = \frac{1}{1 + |\nabla G_\sigma f(x, y)|}, \]

where \( f \) is the image and \( G_\sigma \) is a Gaussian filter with a deviation equals to \( \sigma \). This stopping function has values that are closer to zero in regions of high image gradient and values that are closer to unity in regions with relatively constant intensity. Hence, the discrete evolution equation is

\[ \frac{\phi^{n+1}(i, j) - \phi^n(i, j)}{\Delta t} = -g(i, j) F(i, j) |\nabla \phi^n(i, j)|, \]

It’s obvious that the evolution is based on the stopping function \( g \) which depends on the image gradient. That’s why this model leads to unsatisfactory results in presence of occlusions, low contrast and even noise. To make the level set function evolves in the regions of variability between the shape of reference and the target shape, we propose the new stopping function as follow

\[ g_{\text{shape}}(x, y) = \begin{cases} 0, & \text{if } \phi_{\text{prod}}(x, y) \geq 0, \\ \text{sign}(\phi_{\text{ref}}(x, y)), & \text{else,} \end{cases} \]

where \( \phi_{\text{prod}}(x, y) = \phi(x, y) \cdot \phi_{\text{ref}}(x, y) \), \( \phi \) is the level set function associated to the evolving contour, while \( \phi_{\text{ref}} \) is the level set function associated to the shape of reference after alignment. As it can be seen, the new proposed stopping function only allows for updating the level set function in the regions of variability between shapes. In these regions \( g_{\text{shape}} \) is either 1 or -1 because in the case of partial occlusions, the function is equals to 1 in order to push the edge inward (deflate) and in case of missing parts, this function is equals to -1 to push the contour towards the outside (inflate). This property recalls the Balloon snake’s model proposed by Cohen in [24] in which the direction of evolution (inflated or deflate) should be precised from the beginning. In our work, the direction of evolution is handled automatically based on the sign of \( \phi_{\text{ref}} \). The total discrete evolution’s equation that we propose is as follows

\[ \frac{\phi^{n+1}(i, j) - \phi^n(i, j)}{\Delta t} = -(w g(i, j) + (1 - w) g_{\text{shape}}(i, j)) F(i, j) |\nabla \phi^n(i, j)|, \]

\[ w \] is a weighting factor between the image-based force and knowledge-driven force. See [16] for our proposed shape prior for a region-based active contours.

4. EXPERIMENTAL RESULTS

4.1. Robustness of the proposed shape priors

We present in Fig. 3 an example of successive evolutions between several shapes of different topologies under the proposed shape priors only \((w = 0)\). This simulation shows that the proposed shape priors can well constrain an active contour to take a given shape (known as reference) and handling non trivial geometric shapes with holes and complex topologies.

![Fig. 3. Curve evolution under the proposed shape prior only.](image)

4.2. Application to object detection

We will devote this section to present some results of object detection obtained by the proposed model in case of partially occluded object under affine transformation. We first evolve the active contour without shape prior until convergence (i.e. \( \epsilon = 1 \)) to reduce the computational complexity and to have a good estimation of the parameters of the rigid transformation as in [11, 8]. This first result provides an initialization for the model with prior knowledge. More weight is assigned to prior knowledge (generally \( w \leq 0.5 \)) to promote convergence toward the target shape.

Consider for this first experiment the spider object, image (a), obtained from MCD data base \(^1\). The (b) image is the object of interest which is obtained from (a) after affine transformation and partial occlusion. As a first step, we perform object (b) segmentation without prior knowledge. The obtained contour is then aligned with the contour of the (a) object to determine the regions of occlusions. We present in (c) the result of shape alignment. The red sampled contour corresponds to the (a) shape after the affine transformation estimation with (b) and the blue one corresponds to the (b) shape. The estimated values are \( \alpha = 2, \beta_0 = 0, \gamma = [0.5, 0.2, 0.2] \). By Fig.5, we present the obtained results without and with shape prior.

We considered in this experiment four templates (a spider, a chopper, a device and a bird), see Fig.6. We were based on the set of the presented invariant descriptors to choose the suitable reference according to the occluded spider (Image (b) of Fig.6).

Table 2 presents the obtained Euclidean distance between the target shape and the available templates. We notice that the minimum distance can be easily identified because the suitable form in this experiment is distinguishable. In the

---

\(^1\)http://vision.ece.ucsb.edu/ zuliani/Research/MCD/MCD.shtml
Fig. 4. Spider’s shapes alignment.

Fig. 5. Detection without, image (a), and with, image (b), affine prior knowledge.

Fig. 6. Image (a): The available templates, Image (b): the transformed object.

Table 2. Distances between the occluded spider and the used templates.

<table>
<thead>
<tr>
<th></th>
<th>Spider</th>
<th>Chopper</th>
<th>Device</th>
<th>Bird</th>
</tr>
</thead>
<tbody>
<tr>
<td>The occluded spider</td>
<td>0.037</td>
<td>0.46</td>
<td>0.44</td>
<td>0.53</td>
</tr>
</tbody>
</table>

next experiment, a mosaic application is considered. The mosaic images that are considered are taken from the Bardo Museum of Tunisia which contains the biggest collection of mosaic images in the world. In mosaic images, objects are composed of tessellas and are often partially occluded as it is shown by images (c) and (d), Fig.7. So given that in mosaic images the object are often repeated and in order to study the robustness of our method, we try to find to true contour of an occluded object based on another one having the same shape. We have approximated the perspective projection to an affine transformation which is often used in literature according to the acquisition conditions. Let’s consider the (c) image. As it’s shown, this image contains many forms. Among all forms available in these two images, (c) and (d), taken from two sides, we use the Euclidean distance between the Affine Invariant Fourier Descriptors to localize two lions. The left one is partially occluded. We will use the right lion of the (d) image as template in order to have a better segmentation. In the last figure, images (a) and (b), we present in red the used curves to perform shape alignment and shape prior computation. The estimated values are $\alpha = 1, l_0 = -0.52, A = [-0.002, -0.003, 2.003, -0.35]$. We present by the (c) image the obtained curves alignment result and by the (d) image the segmentation result based on the proposed shape prior. Such a results is particularly interesting since it can be used for mosaic images restoration under partial occlusions and missing parts. The underlying idea is to extract similar forms using minimal distance between descriptors in order to define prior knowledge for such occluded or cluttered objects based on the presented affine shape alignment method for the purpose of a good segmentation result.

Fig. 7. Some mosaic images from the Bardo Museum of Tunisia.

Fig. 8. Robust object detection in mosaic image.
5. CONCLUSION

We presented in this paper an alternative approach to incorporate prior knowledge into a level set based active contours in order to have robust object detection in case of large shape distortion that can be analyzed by the class of affine transformations. We presented also a geometric solution to choose the reference shape in case of many available templates given that the statistical approach needs a training set and PCA. Then the application of a given classifier like Bayesian classifier to determine the appropriate reference shape like the work of Fang et Chan [25]. Given that the proposed approach invoke only pixels of the regions of variability between the shape of reference and the object of interest in the process of curve evolution and based on the fast Fourier transform for affine motion estimation and invariants computation, the method is faster compared to [15] and [26] where at each iteration shape descriptors are calculated for a given order that has to be set empirically. The obtained results are promising in the case of real and simulated data and the method can be used for the restoration of mosaic images in the archeological field. As future perspectives, we are working on integrating our model in the context of 3D object reconstruction from silhouettes sequence in order to refine the obtained 3D model.

6. REFERENCES


