Aircraft segmentation in SAR images based on improved active shape model

Xiaoqiang Zhang, Boli Xiong, Gangyao Kuang

Electronic Science College, National University of Defense Technology, Changsha, Hunan, China

Commission WG III/3

ABSTRACT:

In SAR image interpretation, aircrafts are the important targets arousing much attention. However, it is far from easy to segment an aircraft from the background completely and precisely in SAR images. Because of the complex structure, different kinds of electromagnetic scattering take place on the aircraft surfaces. As a result, aircraft targets usually appear to be inhomogeneous and disconnected. It is a good idea to extract an aircraft target by the active shape model (ASM), since combination of the geometric information controls variations of the shape during the contour evolution. However, linear dimensionality reduction, used in classic ACM, makes the model rigid. It brings much trouble to segment different types of aircrafts. Aiming at this problem, an improved ACM based on ISOMAP is proposed in this paper. ISOMAP algorithm is used to extract the shape information of the training set and make the model flexible enough to deal with different aircrafts. The experiments based on real SAR data shows that the proposed method achieves obvious improvement in accuracy.

1. INTRODUCTION

As the development of SAR technology, imaging technology becomes much more mature. High resolution SAR images bring much convenience to automatic target recognition (ATR), because more structure and geometry details can be displayed. This makes aircraft recognition possible. Aircraft targets have great value and always arouses much attention in image interpretation. There has been some effective methods of aircraft detection and recognition (F. Zhang, 2009) (G. Liu, 2013). But the related works based on SAR images are still limited. Due to the complex structure, it is hard to directly deal with the aircraft on the large image. A feasible way is that the detection algorithm is implemented firstly to generate ROIs which contain aircrafts, and then further processing will be carried out based the ROIs. In this procedure, aircraft segmentation is an important task, the base of feature extraction and target recognition. However, it is far from easy to segment the aircraft from the background completely and precisely in SAR images. Because of the complex structure, different kinds of electromagnetic scattering take place on the aircraft surfaces. As a result, aircraft targets usually appear to be inhomogeneous and disconnected in SAR images. It brings much trouble to target segmentation. Like CFAR algorithm (Novak, L.M., 1997), some classic methods, only utilizing image information, are able to extract strong scattering points of an aircraft according to intensity variations, but they fail to cluster the target and describe its shape accurately.

Aircrafts in SAR images usually appears to be different from those in optical images. Complicated scattering mechanisms on aircraft surfaces, caused by the complex structure, usually lead to inhomogeneity of local intensities. In the process of image segmentation, some dark regions of the aircraft will be easily classified as the background. As a results, the aircraft will be segmented into several isolated regions. As conventional approach, morphological operations are carried out for clustering. But it may make a great influence on the original structure of the aircraft.

The active contour model (ACM) is another successful segmentation method. Through minimization of the energy function, the active contour evolves towards boundaries. In SAR images, however, the multiplicative noise and intensity variations frustrate the evolution of the contour, and local minimums may be generated.

In order to improve the segmentation performance, it is necessary to introduce the shape prior knowledge. In this point of view, the active shape model (ASM) is a valuable tool. ASM is firstly proposed by Cootes et al (T.F. Cootes, 1995), and it uses a statistical approach to model shape statistics of samples. A compact representation from a set of training instances is extracted in this method. Thanks to the help of shape information, the contour is controlled in the procedure of evolution. ASM has achieved success in meaningful object segmentation of medical and optical images.

In the framework of ASM, dimensionality reduction is an important step in modelling shape statistics. Principal component analysis (PCA) has been widely used in many related works. This kind of linear learning techniques is usually based on a linearized shape space with limited deformation modes around a mean shape. It can keep good performance when shape samples in the training set are similar, otherwise it cannot embody the shape information of the samples fully. As for aircrafts, however, different types usually have different outlines. Two common types are taken as an example in Fig. 1 it can be seen that, the differences between (a) and (b) are obvious. When we implement ASM to deal with them, the training set should be consist of the shape samples of these aircrafts. Since PCA cannot extract full shape information of the training set when samples are not similar enough, the range of the contour deformation is limited, as shown in Fig. 2. This is not suitable for actual application. To solve this problem, a new aircraft segmentation method based on the non-linear shape statistics learning algorithm is proposed in this paper.
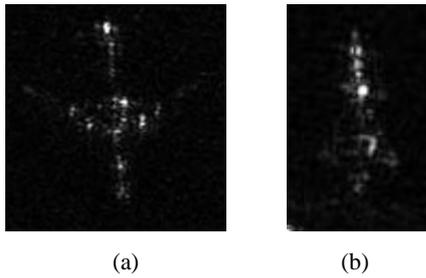
(a)   (b)

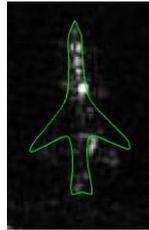Figure 1. Aircraft chips. (a) Type A. (b) Type B.

Figure 2. The limitation of ASM based on PCA.

2. ASM BASED ON ISOMAP

2.1 Describing a sample with Catmull–Rom Spline

Differentwith the traditional ASM, here we use CR spline to describe the shape samples, instead of the straight line, in order to decrease the control points and reduce computation. As for CR spline, presented by Catmull and Rom (E. Catmull, 1974), a curve segment can be defined by four control points at least. The control point is represented by $\mathbf{P}_i = [\mathrm{x}_i, \mathrm{y}_i]^\mathrm{T}$. A curve segment can be defined by control points $\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ and knot sequence t_0, t_1, t_2, t_3 as follow

$$\mathbf{C} = \frac{t_2 - t}{t_2 - t_1}\mathbf{B}_1 + \frac{t - t_1}{t_2 - t_1}\mathbf{B}_2 \tag{1}$$

where

$$\mathbf{B}_1 = \frac{t_2 - t}{t_2 - t_0}\mathbf{A}_1 + \frac{t - t_0}{t_2 - t_0}\mathbf{A}_2$$

$$\mathbf{B}_2 = \frac{t_3 - t}{t_3 - t_1}\mathbf{A}_2 + \frac{t - t_1}{t_3 - t_1}\mathbf{A}_3$$

$$\mathbf{A}_1 = \frac{t_1 - t}{t_1 - t_0}\mathbf{P}_0 + \frac{t - t_0}{t_1 - t_0}\mathbf{P}_1$$

$$\mathbf{A}_2 = \frac{t_2 - t}{t_2 - t_1}\mathbf{P}_1 + \frac{t - t_1}{t_2 - t_1}\mathbf{P}_2 \tag{2}$$

$$\mathbf{A}_3 = \frac{t_3 - t}{t_3 - t_2}\mathbf{P}_1 + \frac{t - t_2}{t_3 - t_2}\mathbf{P}_2$$

$$t_{i+1} = \left[\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}\right]^\alpha + t_i$$

in which $\alpha \in [0,1]$, $i = 1, 2, 3, 4$ and $t_0 = 0$. With different value of α , CR spline has variant forms. When $\alpha = 0$, the curve is standard. When $\alpha = 0.5$, the curve is centripetal CR spline. When $\alpha = 1$, the curve is chordal CR spline. According to the research of Yuksel et al (C. Yuksel, 2011), the centripetal CR spline has desirable advantages over the other types. Within

a curve segment, it does not form self- intersections and cusps. Therefore, we use the centripetal CR spline to describe the aircraft shape samples.

Since CR spline does not pass through the both end points, if we want the curve to go after $\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$, it is necessary to add $\mathbf{P}_{-1}$ and $\mathbf{P}_4$ to the ends respectively. They make a less influence on the cure, and can be define as

$$\mathbf{P}_0 - \mathbf{P}_{-1} = \beta(\mathbf{P}_1 - \mathbf{P}_0)$$
$$\mathbf{P}_4 - \mathbf{P}_3 = \beta(\mathbf{P}_3 - \mathbf{P}_2) \tag{3}$$

where $\beta = 0.5$. With this setting, the shortcoming can be rectified.

Due to our limited SAR data capacity of aircraft targets and the lack of ground truth, which are the common problems of many other researchers, the shape training set cannot be established based on SAR images. To solve this problem, a series of vertical views of aircrafts are collected from the Internet. We build the shape samples according to these images. As a result, 69 samples are obtained, in which 42 ones belong to Type A and 27 ones belong to Type B. The typical samples are shown in Fig.3. The shape model is consist of 23 control points that represent the stationary local features respectively.
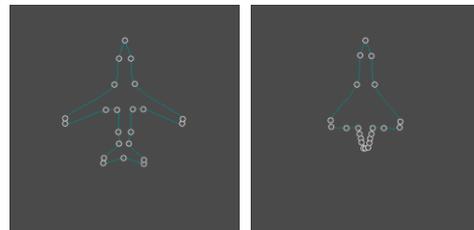

Figure 3. Typical shape samples.

Since the training set should contain the shape information only,

In order to build a shape space, where there is only the shape difference among samples, it is necessary to remove the other differences including translation, scale and rotation. That is why all the samples in the training set should be normalized and aligned. Cootes' method can be used for reference. Let control points $\mathbf{P}_1 = [x_1, \mathrm{y}_1]^T, \mathbf{P}_2 = [x_2, \mathrm{y}_2]^T, \dots, \mathbf{P}_m = [x_m, \mathrm{y}_m]^T$ form a shape denoted by $\mathbf{x} = [x_1, y_1, x_2, y_2, \dots x_i, y_i, \dots, x_m, y_m]^T$. The normalized points $[\hat{x}_i, \hat{y}_i]$ can be obtained by

$$\begin{cases} \hat{x}_i = (x_i - \overline{x}) / \max(\max|x_i - \overline{x}|, \max|y_i - \overline{y}|) \\ \hat{y}_i = (y_i - \overline{y}) / \max(\max|x_i - \overline{x}|, \max|y_i - \overline{y}|) \end{cases} \tag{4}$$

where $\overline{x}$ and $\overline{y}$ are the means of x_i and y_i .

To achieve the shape alignment between shapes $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$, a modified Procrustes algorithm is provided (J. C. Gower, 1975).

$$\begin{pmatrix} X_k & -Y_k & W & 0 \\ Y_k & X_k & 0 & W \\ Z & 0 & X_k & Y_k \\ 0 & Z & -Y_k & X_k \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} X_{k-1} \\ Y_{k-1} \\ C_1 \\ C_2 \end{pmatrix} \tag{5}$$

in which

$$a_x = s\cos\theta, \quad a_y = s\sin\theta$$

$$X_k = \sum_{i=1}^{m} w_i(x_i)_k, \quad Y_k = \sum_{i=1}^{m} w_i(y_i)_k$$

$$Z = \sum_{i=1}^{m} w_i[(x_i)_k^2 + (y_i)_k^2], \quad W = \sum_{i=1}^{m} w_i$$

$$C_1 = \sum_{i=1}^{m} w_i[(x_i)_{k-1}(x_i)_k + (y_i)_{k-1}(y_i)_k]$$

$$C_2 = \sum_{i=1}^{m} w_i[(y_i)_{k-1}(x_i)_k + (x_i)_{k-1}(y_i)_k]$$

where $[(x_i)_k, (y_i)_k]$ is the i th point of $\mathbf{x}_k$. w_i is the weight of each point. Higher weights will be allocated to the points with less movement. Let R_{ij} denote the distance between point i and j in a shape sample, and $V_{R_{ij}}$ denote the variance in the distance over the shape set. w_i is defined as

$$w_i = (\sum_{i=1}^{m} V_{R_{ij}})^{-1} \qquad (6)$$

The scheme of training set alignment is as follow

a) Align every shape with the first shape.
b) Obtain the mean shape of the aligned shapes.
c) Realign the shapes with this mean shape.
d) Repeat (b) and (c) until the process converges.

2.2 Dimensionality reduction by ISOMAP

Isometric feature mapping (ISOMAP) is the extension to multi-dimensional scaling (MDS), a classic method for embedding dissimilarity information into Euclidean space. ISOMAP returns an embedding, where the distance between points is approximately equal to the shortest path distance. And it obtains the estimate of the dimensionality of the underlying manifold. The specific steps are as follows.

a) Compute the L_2 distance between all pairs of points
b) Define the set of points comprising the neighbourhood by k-nearest neighbours (KNN).
c) Define a graph with a node for each input point and weighted undirected edges connecting each node to the nodes corresponding to the points.
d) Solve for all-pairs shortest paths on this sparse graph to calculate a complete pair-wise distance matrix.
e) Use MDS to find an embedding of these distances in Euclidean space.

The shape prior estimation based on ISOMAP is described as follows.

a) Project the training set shape samples S_i and the initial shape X_0 into the manifold subspace by ISOMAP. The corresponding projections are denoted by s_i and the initial shape x_0.
b) Construct the adjacent graph on the low-dimensional manifold. Let G denote a graph with n vertices. Find the the shortest paths on the graph between all pairs of vertices by Dijkstra's algorithm. The distance between the initial segmentation and the ith shape in the learning set is

$$d_i = g(s_i, x_0) \qquad (7)$$

where g is the length of the shortest path.

c) By using the learned manifold, the shape prior for an image to be segmented is estimated by choosing a certain amount of training shapes, which are closest to the

currently obtained shape. Let $N(x_0)$ describe the neighbourhood containing the KNNs of a shape x_0 on the learned manifold. And then the shape X_0 can be represented by a weighted convex combination of the shapes in $N(x_0)$. The reconstruction weight for each shape x_0 is given by

$$w_i = e^{\frac{-D_i}{\sigma}} \qquad (8)$$

where the normalization parameter σ is set in the experiments. The shape in the high-dimensional space is computed using high-dimensional features of the k nearest neighbours and the reconstructed weights (T.F. Cootes, 1995). Therefore, the estimated shape prior is

$$\hat{X}_0 = \sum_{i=1}^{k} w_i S_{N(x_0)} \qquad (9)$$

3. IMGAE MOTIVATION BY GVF SNAKE

In the proposed method, the image data motivation is realized by snake model based on edges.

3.1 Edge extraction by ROEWA operator

The gradient operator used in traditional snake and GAC models is the conventional edge detector. It does not work well in SAR images because of the multiplicative nature of speckle noise. Different from the preceding works, the ROEWA operator, an optimal multi-edge detector for SAR images (Fjørtoft R., 1998), is employed to reduce the influence of speckle. Although ROEWA operator can generate the edge strength satisfactorily, it cannot get the correct edge direction, which brings trouble to edge location and thinning. The watershed algorithm can give closed boundaries, but it easily leads to the over-segmentation. Though this phenomenon can be dealt with by some variant method (Sheng, G., 2012), manual markers are required.

To extract edges from the edge strength map, the multi-direction NMS is applied. In the absence of edge direction, the NMS is used in 8 different orientations (0, 45°, 90°, 135°, 180°, 225°, 270°, and 315°) to find the local maximas. The result is further filtered by a two level thresholding. The high threshold HTh and low threshold LTh are set as follow: the ratio of the number of non-edge pixel to the total amount of pixel N is HR. Pixels are counted from the lowest magnitude of the histogram of the edge strength map. When the number comes to $HR \cdot N$, HTh is set as the corresponding magnitude. Then $LTh = HTh \cdot LR$, where LR is a coefficient. If the edge strength magnitude of a pixel is higher than HTh, it belongs to the strong edge. Then if the magnitude between HTh and LTh, besides the pixel borders on strong edges, it is classified as the weak edge. Otherwise, the pixel does not belong to edges. Finally, a size filter is applied to eliminate isolated pixels.

3.2 GVF snake

A snake contour is an elastic curve that deforms to minimize the following energy functional.

$$E(C) = E_{int} + E_{ext} \qquad (10)$$

where the internal energy E_{int} derives from the contour itself and the external energy E_{ext} derives from the image. Since the contour evolution is time-consuming, especially when the contour is long, the spline curve is introduced to form an economical realization of snake. The internal energy is not required in the spline-based snake because of the restrictions of the spline. The B-spline snake, also called B-snake, has been widely used. In this paper, we choose the CR spline, which is easy to build the shape models.

On the other hand, since the capture range of the traditional snake is very limited, Xu and Prince improved the external force by introducing the gradient vector flow (GVF) field $V(x, y) = [u(x, y), v(x, y)]$ (Xu C., 1997), which minimizes the energy function

$$E_{ext} = \iint \left[\mu |\nabla V|^2 + |\nabla f|^2 (V - \nabla f)^2 \right] dxdy \quad (11)$$

where μ is a regularization parameter, and f is the edge map of the raw image.

4. METHODOLOGY

We want to move the points of the shape to the edges of the image while still satisfying the shape constraint of the model. A spline-based shape model s_i in the shape scale can be transformed into S_i in the image scale by affine transformation using the parameters including scale s, rotation θ and translation $\mathbf{t} = [t_x, t_y]$

$$x_i' = x_i \cdot s \cos \theta - y_i \cdot s \sin \theta + t_x$$
$$y_i' = x_i \cdot s \sin \theta + y_i \cdot s \cos \theta + t_y \quad (12)$$

Firstly the mean shape x_μ is transformed into X_μ by the parameters $s, \theta, \mathbf{t}$. Then X_μ is motivated by the GVF field V, and the deformed shape is X_{GVF}. Then the new shape X_{new} can be obtained according to X_{GVF}, via shape prior estimation based on ISOMAP. During this process, the parameters is updated and the new ones $s', \theta', \mathbf{t}'$ are equal to the values that align x_μ to X_{GVF}. The above operations are repeated, until the contour becomes stabilized. The contour evolves during iterations and always keeps similar to the shapes of the training set during the deformation. In the first iteration, the parameters should initialized.

5. EXPERIMENTS

To demonstrate the performance of the proposed segmentation method, it is compared with two ACMs including RSF and LGIF. The segmentation results are presented in Fig. 4. The aircraft chips are picked from the Ku-band airborne SAR images with 0.5 m resolution. Since the aircrafts usually appear to be inhomogeneous in SAR images, RSF and LGIF models only extract some isolated bright pixels and fail to find the real contours. It can be seen that the performance of them is unacceptable. By contrast, the proposed method extracts the complete contours and recovers the real shapes using the PDM. It is much better than the other two methods. What's more, the proposed method can deal with the two types properly. In this

framework, when we need to deal with other types of aircraft, just add the corresponding samples to the training set. The predictability is much improved.

(a)

(b)

(c)
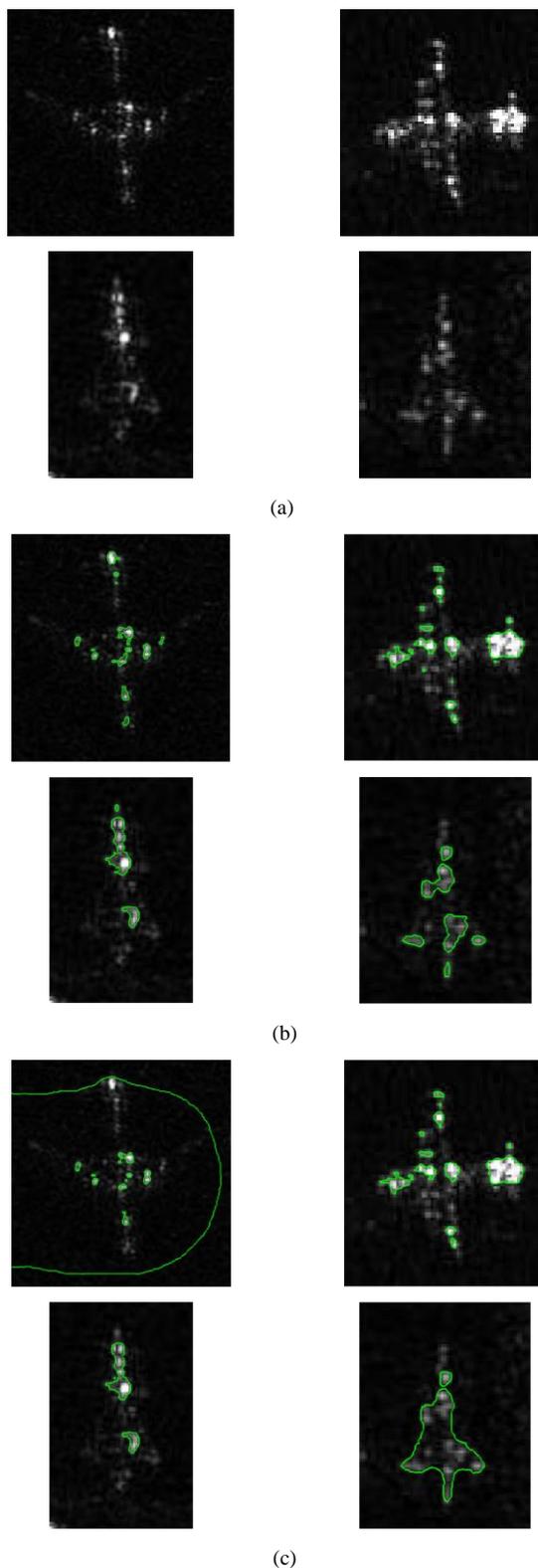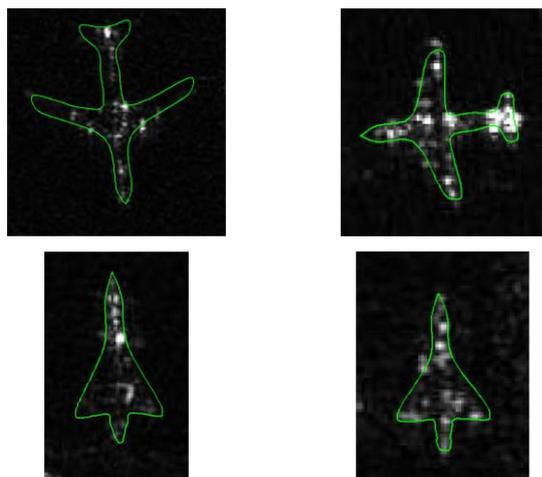
(d)

Figure 4. The experiment of comparison. (a) Original SAR chips. (b) LGIF. (c) RSF (d) The proposed method

T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, "Active shape models—their training and application," Comput. Vis. Image Underst. 61, 1995, pp. 38–59.

Xu, C. and Prince, J.L., "Gradient vector flow: A new external force for snakes," In IEEE Computer Vision and Pattern Recognition Proceeding, 1997, pp. 66-71.

6. CONCLUSION

A refined segmentation method for aircraft target in SAR images based on ISOMAP is proposed in this paper. It combines the image information and the shape prior to improve the segmentation accuracy. On the one hand, the image motivation is carried out by the GVF snake. On the other hand, the shape information is extracted by ISOMAP algorithm. The experiment based on real SAR data shows that the proposed method achieves obvious improvement in accuracy.

REFERENCES

C. Yuksel, S. Schaefer, J. Keyser, "Parameterization and applications of Catmull–Rom curves," Comput. Aided Des. 43, 2011, pp. 747–755.

E. Catmull, R. Rom, "A class of local interpolating splines, in: Computer Aided Geometric Design, Academic Press," 1974, pp. 317–326.

Fjørtoft R., Lopes A., Marthon P., "An Optimal Multiedge Detector for SAR Image Segmentation," IEEE Transactions on Geoscience and Remote Sensing, 36, 1998, pp. 793-802.

F. Zhang and S. Q. Liu, "Aircraft recognition in infrared image using wavelet moment invariants," Image Vis. Comput., vol. 27, no. 4, 2009, pp. 313−318.

G. Liu, X. Sun, K. Fu, H. Wang. "Aircraft Recognition in High-Resolution Satellite Images Using Coarse-to-Fine Shape Prior," IEEE Geoscience and Remote Sensing Letters, vol. 10, no. 3, 2013, pp. 573–577.

J. C. Gower, "Generalized Procrustes analysis," Psychomestrika 40, 1975, pp. 33-51.

Novak, L.M., Owirka, G.J., Brower, W.S., Weaver, A.L. "The automatic target-recognition system in SAIP, " Linc. Lab. J. 1997, 10, 187–202.

Sheng, G., Yang, W., Deng, X., He, C., 2012, Coastline Detection in Synthetic Aperture Radar (SAR) Images by Integrating Watershed Transformation and Controllable Gradient Vector Flow (GVF) Snake Model, IEEE Journal of Oceanic Engineering, 37, pp. 375-383s