

Abstract

This paper proposes to recognize an aircraft in satellite image using template matching for accurate detection and tracking. High resolution multispectral satellite images with multi-angular look capability have tremendous potential applications. Here the system involves an object tracking algorithm with three-step processing that includes moving object estimation, target modeling, and target matching. Potentially moving objects are first identified on the time-series images. The target is then modeled by extracting both spectral and spatial features. In the target matching procedure, template will be used as matching model to recognize with each frame by frame for accurate detection. Here, normalized cross correlation and spatial features are used as features model for recognition. This recognition model will be continued for all sequence of satellite images. Final simulated will be demonstrated the capability of object tracking using change detection algorithm in a complex environment with the help of high resolution multispectral satellite imagery.

Keywords: Aircraft recognition, image segmentation, level set method, shape prior, template matching.

Introduction

Aircraft recognition is an important issue of target recognition in satellite images and has many important applications in practice such as airfield dynamic surveillance. As the resolution of satellite images gets higher, more abundant color, texture, and spatial information are provided. Such information offers good opportunity to recognize aircraft that has a very complex structure. However, automatic aircraft recognition is not a simple problem. Besides the complex structure, different aircraft differ in size, shape, and color, and even for one kind of aircraft, the texture and intensity are usually dissimilar in different scenarios. Moreover, recognition often

suffers from various disturbances such as clutter, different contrasts, and intensity in homogeneity. Thus, the robustness and resistance to disturbance are highly required for the method. We illustrate some typical satellite aircraft images in Fig. 1 to show the difficulties. A lot of work have been done for aircraft recognition, such as moment invariant features [1], Fourier descriptor [2], and so on [3], [4]. These methods use shape features on either binary image or object contour and suppose that the object region or edges can be well obtained which is often difficult in practice.

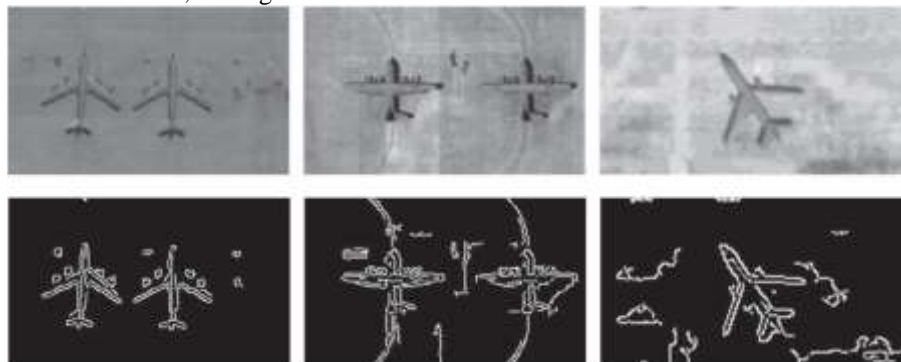


Fig. 1. Aircraft images and Canny edge detections with thresholds 60 and 150.

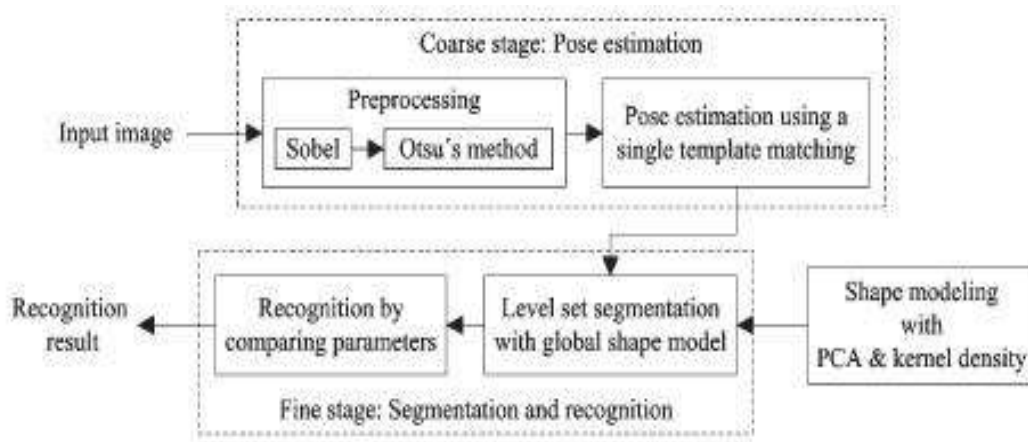


Fig: 2 proposed aircraft recognition framework

Xu and Duan [5] apply an artificial bee colony algorithm with an edge potential function to the recognition task for low altitude aircraft, whereas in satellite images, object boundary is easily blurred and edges usually cannot be well extracted due to poor contrast as shown in Fig. 1. In [6], principal component analysis (PCA) is applied on the binary image obtained by segmentation with Otsu's method [7] to estimate the main axis, and recognition is then achieved by template matching. Hsieh *et al.* [8] use a symmetry-based method to find the optimal axis direction after image binarization using minimum within-group variance dynamic threshold; then, four features are extracted and combined to recognize the object. Shape features are used in the recognition stage for both [6] and [8], while their segmentations are both only pixel based. Thus, the segmentations are easily affected by various disturbances, which, as a result, will affect the final recognition performance.

The contribution of our work is threefold. First, unlike traditional methods, shape prior is integrated to segmentation process and utilized coarsely to finely in our recognition framework as shown in Fig. 2. The similarity and difference of shapes for aircraft samples are respectively exploited in the coarse and fine stages. Second, pose is first approximately estimated before segmentation in the coarse stage, using a single template matching with a defined score criterion. The template is regarded as describing the similarity of shapes for aircraft, and the estimated pose result provides an initialization for the following segmentation. The refined pose information is integrated in our contour evolution process in the fine stage and obtained when the segmentation is done. Third, we model shapes globally with PCA and kernel density function that have good effects on both dimension reduction and sample space description,

and a new energy function is proposed to embed the shape model to segmentation using a level set method. The final obtained model coefficients in segmentation are directly used to recognize the type. Experiments on Quick Bird images show that our method is more robust with respect to various disturbances compared with other methods.

Materials and methods

Coarse Stage : Pose Estimation

Pose estimation is essential in aircraft recognition and is usually done after segmentation in conventional methods [6], [8]. In our method, shape is integrated in the segmentation with a parametric model, so the refined pose information is contained while segmenting, and will be obtained when the segmentation is accomplished. On the other hand, the integrated shape model needs an initial pose. Due to the special complex structure of aircraft, segmentation with a random pose initialization easily runs into local minimum. Thus, in this stage, we roughly estimate the pose information, which coarsely addresses the problem of translation, rotation, and scaling. Considering the common cross structure of aircraft, we adopt the method of coarse template matching with the average shape of test aircraft.

A. Preprocessing

Template matching is usually edge based or region based. In our method, edge information is used. As mentioned before, edge extraction often suffers from image blurring and poor contrast. However, the edges obtained still preserve structure information for aircraft to some extent. In addition, aircraft is usually on parking aprons, most of which are often flat and smooth, so edges are mostly

produced by the contrast between object and background. On the contrary, color feature is not steady. Intensity in homogeneity on aircraft often occurs due to illumination, and sometimes, the intensities of pixels on aircraft look even the same as background. The preprocessing step aims to extract the edge information. Considering the existence of various contrasts for optical satellite images, we use Sobel operator to calculate the gradient magnitude instead of traditional edge detectors, e.g., Canny for which fixed thresholds are needed. In detail, we first apply Sobel operator on the input image; then, Otsu's method is used on the output from Sobel operator to calculate a global adaptive threshold to remove low magnitudes, most of which are produced within background. The resultant gradient image is used as the input of the template matching.

B. Pose Estimation

Aircraft have the common feature of cross structure. To roughly estimate the pose of an aircraft,

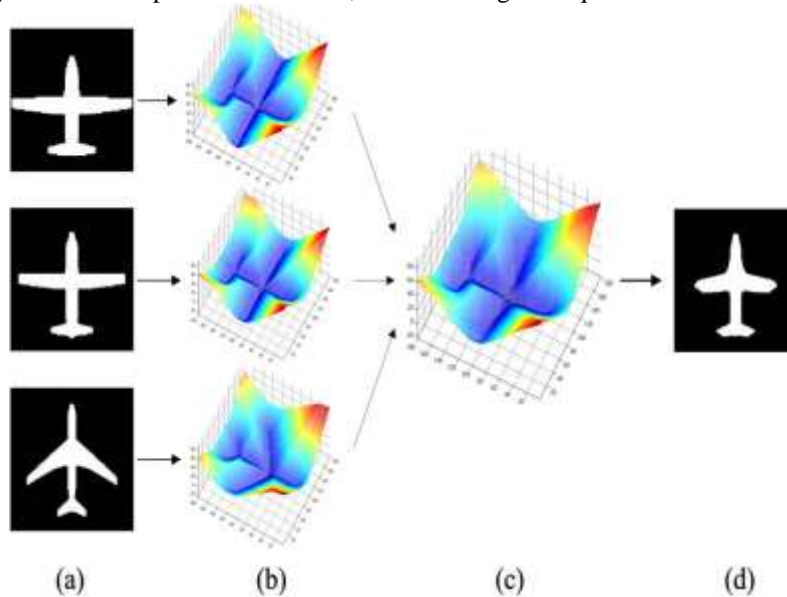


Fig. 3. Process of making average template. (a) Aligned test shapes. (b) SDFs. (c) The average SDF. (d) Binarization of the average SDF.

Fig. 3 shows how to get this average template. The aligned test aircraft shapes are first represented by signed distance functions (SDFs), and then, the average SDF is calculated. This is also one of the steps in shape modeling process introduced in the next section. The template is obtained by binarization of this average SDF. It is noticed that the template is solid. There are two reasons we do not use edge template. First, we align aircraft shapes through region correlation introduced in the shape modeling, so this solid template represents the common feature of aircraft. Second, in the

this common structure can be adopted. In our pose estimation, the binary shape template of test aircraft is used to match the preprocessed gradient image coarsely on translation r , rotation θ , and scaling s with the following defined score criterion:

$$\text{Score}(r, \theta, s) = \frac{\sum_x I(x)}{(D(x, T(r, \theta, s)) + 1) * s^\zeta (1)}$$

Where $I(x)$ denote the value of the preprocessed image,

$T(r, \theta, s)$ represents the template with translation, rotation, and scaling, and $D(x, T(r, \theta, s))$ is the distance from pixel x to the template which can be obtained by distance transformation. ζ is a constant, and s^ζ is a scaling regular term of Score. This equation gives the score criterion, which is very important factor for rough estimation. By using the above equation we can find binary shape template of test aircraft to match the preprocessed gradient image on translation. The average template is obtained by using this equation.

experiments, we find that solid template works better for the large difference of wings of various aircraft. Since the distance $D(x, T(r, \theta, s))$ becomes smaller when the scaling s is larger or the size of this solid template is bigger, the regular term s^ζ is introduced in the denominator in (1). In addition, in our experiments, we get the best result when ζ is set to 1.1.

Fine Stage: Segmentation And Recognition

In the fine stage, shapes of aircraft are globally modeled with a parametric representation and kernel

density estimation. This model is then integrated into the popular energy-based level set method which is convenient to combine various priors. Curve evolution is driven by the image region-based energy and simultaneously regulated by the shape model. The last obtained coefficients of shape model are directly used in the recognition.

A. Shape Modeling

A lot of methods have been developed to model a class of objects, e.g., [9]–[11]. We choose PCA method here due to its efficiency and good effect on dimension reduction. First, shape templates in the training set are aligned with method in [10]; then, the boundary is embedded as the zero level of an SDF. PCA is applied to these SDFs to get the mean shape ϕ_0 and K eigenvectors ϕ_k called eigen shapes. Then, a shape can be parametrically represented as

$$K\phi(y) = \phi_0(T \cdot y) + \sum_{k=1}^K w_k \phi_k(T \cdot y)$$

$$K = \sum_{k=0}^K w_k \phi_k(T \cdot y), \text{ with } w_0 = 1 \quad (2)$$

Where $W = \{w_1, \dots, w_k, \dots\}$ denotes the weights of eigen shapes and T is the transformation matrix that includes translation, scaling, and rotation

$$T = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

In practice, the M ($M < K$) largest eigen shapes measured by the corresponding eigen values are used to reconstruct the shape. In addition, the average reconstruction accuracy is the sum of the M largest eigen values divided by the sum of all the eigen values. It is noticed that, in the aforementioned linear representation, a uniform distribution is actually assumed for the coefficients W or for the shapes in the PCA subspace. However, the types of aircraft are finite in practice. To better capture the shape distribution, a kernel density function for W is introduced here

$$p(W) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|W - W_i|^2}{2\sigma^2}\right) \quad (4)$$

Where, W_i represents the PCA coefficients of samples and N is the number of samples. Like [11], we set σ^2 to be the mean squared nearest neighbor distance.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N \min_{j \neq i} |W_i - W_j|^2. \quad (5)$$

B. Shape-Based Segmentation

Proposed by Osher and Sethian [12], level set method has become a very popular framework for segmentation [13]–[15]. Compared with other methods such as graph cut, it is more convenient to combine with shape prior. The basic idea of level set is the evolution of contour driven by some energy terms. In our method, a new energy function integrating a shape prior is proposed to guide the segmentation.

$$E = \sum_{i=1}^2 \int \int K_\tau(x - y) |I(y) - f_i(x)|^2 M_i(\phi(y)) dy dx - \alpha \frac{1}{N} \sum_{i=1}^N \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|W - W_i|^2}{2\sigma^2}\right). \quad (6)$$

The first term of E is a region-scalable fitting (RSF) term [16], and the second is the kernel density function introduced earlier. In addition, α is the weight for these two terms which is positive. In the first term, integrals over y and x are both taken on the image domain; K_τ is a Gaussian kernel with scale parameter τ ; $I(y)$ denotes the image intensity at point y ; $f_1(x)$ and $f_2(x)$ are two values that approximate local intensities of images

$$f_i(x) = \frac{K_\tau(x) * [M_i(\phi(x))I(x)]}{K_\tau(x) * M_i(\phi(x))}, \quad i = 1, 2 \quad (7)$$

Where, $M_1(\phi) = H(\phi)$, $M_2(\phi) = 1 - H(\phi)$, and H is the Heaviside function approximated smoothly with a constant ϵ as follows:

$$H(x) = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan\left(\frac{x}{\epsilon}\right) \right]. \quad (8)$$

RSF is a region-based level set method proposed by Li *et al.* [16]. Compared with traditional piecewise constant method [17], it has good tolerance to intensity in homogeneity which often occurs in satellite images. It is noticed that, in (6), ϕ is not random at the beginning. It is the PCA representation in (2) with initial PCA coefficients W usually set to zero and initial pose coefficients T obtained from the coarse stage. Thus, the contour evolution is updating of the shape parameters W and pose parameters T . Since the energy in (6) is parametrically represented with W and T , its minimization can be easily achieved by the gradient descent approach. The gradients of E , taken with respect to W and T , are given by the following equations:

$$\nabla_{w_i} E = \int \int g(x, y) \delta(\phi(y)) \phi_i(T \cdot y) dy dx - \frac{\alpha}{2\pi\sigma^2} \sum_{j=1}^N \exp\left\{-\frac{|W - W_j|^2}{2\sigma^2}\right\} \frac{(w_i - w_{ji})}{\sigma^2} \quad (9)$$

$$\nabla_{T_i} E = \int \int g(x, y) \delta(\phi(y)) \frac{\partial \phi(y)}{\partial T_i} dy dx \quad (10)$$

where,

$$g(x, y) = K_T(x - y)(f_2(x) - f_1(x))(2I(y) - f_1(x) - f_2(x)) \quad (11)$$

$$\frac{\partial \phi}{\partial a} = \sum_{i=0}^k w_i \frac{\partial \phi_i}{\partial x} \quad (12)$$

$$\frac{\partial \phi}{\partial b} = \sum_{i=0}^k w_i \frac{\partial \phi_i}{\partial y} \quad (13)$$

$$\frac{\partial \phi}{\partial h} = \sum_{i=0}^k w_i \left\{ \frac{\partial \phi_i}{\partial x} (x \cos \theta - y \sin \theta) + \frac{\partial \phi_i}{\partial y} (x \sin \theta + y \cos \theta) \right\} \quad (14)$$

$$\frac{\partial \phi}{\partial \theta} = \sum_{i=0}^k w_i h \left\{ \frac{\partial \phi_i}{\partial x} (-x \sin \theta - y \cos \theta) + \frac{\partial \phi_i}{\partial y} (x \cos \theta - y \sin \theta) \right\} \quad (15)$$

The shape parameter W and pose parameters T are then updated in iteration by

$$w_i^{t+1} = w_i^t - \Delta t w_i \nabla_{w_i} E \quad (16)$$

$$T_i^{t+1} = T_i^t - \Delta t T_i \nabla_{T_i} E \quad (17)$$

Where, $\Delta t w_i$ and $\Delta t T_i$ are positive step-size parameters.

C. Aircraft Recognition

After segmentation, the last obtained PCA coefficients are directly used in the recognition process. Since the resolution of satellite image is known, it is easy to calculate the size of aircraft from the final W and T . Considering that there is no obvious relation between aircraft size and PCA coefficients, we adopt a two-step k -nearest neighbor (KNN) recognition method. First, we select K types from the training set, the sizes of which are most close to the test sample. We define the closeness of size as follows:

$$clo = 1 - \frac{\sqrt{(l - L)^2 + (w - W)^2}}{\sqrt{l^2 + w^2}} \quad (18)$$

where l and w are the calculated length and width, respectively, and L and W correspond to the training

samples. The closeness may be negative when the difference of size is too big. However, this has no influence since bigger clo means that sizes are closer. Second, among the K candidates, we verify the type of aircraft as the closest one by measuring the $L2$ distance of PCA coefficients.

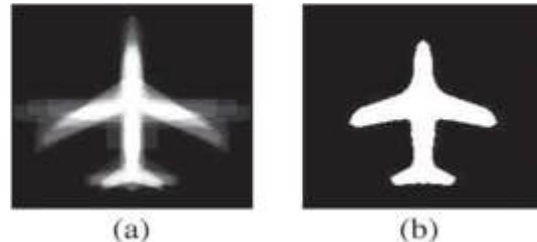


Fig: 4 Shape modeling. (a) Shape alignment. (b) Average shape template.

Results and discussion

In our experiments, due to the lack of standard data sets of high-resolution satellite images for aircraft recognition, 300 gray images, 30 per type, including ten kinds of airplanes, are collected from Quick Bird with the resolution of 0.61 m to evaluate the proposed method. The engines of most of the aircraft are removed since they are small structures and have little common features. In our shape modeling, first, we align all the templates with the region correlation method in [10]. Then, PCA is done for the SDF representations. We take four eigen shapes to parametrically represent the shape of the aircraft, and the average reconstruction accuracy is 96.31%. The aligned shapes and the average shape template used in (1) are shown in Fig. 4.

A. Pose Estimation

In the pose estimation, since the center of aircraft is usually located near to the middle of the image in practice, we restrict our search region for translation r in (1) just around the middle, which is a 50×50 square considering the resolution of image and the sizes of samples. The steps of translation r , rotation θ , and scaling s are set to 5 pixels, 15° , and 0.2, respectively. The length and width of template that we used are 48 and 46 m, respectively, and the scaling s ranges from 0.6 to 1.4. Fig. 6

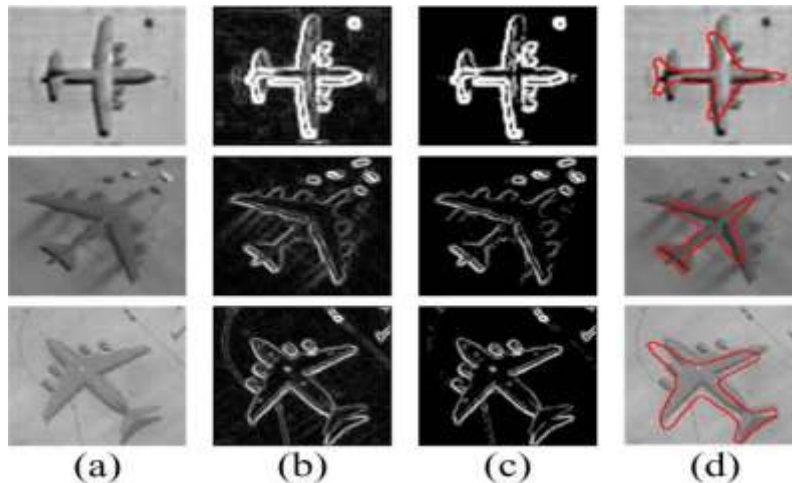


Fig. 5. Pose estimation. (a) Test image. (b) Sobel image. (c) Threshold with Otsu's method. (d) Pose estimation result.

TABLE I
 PRECISION ON POSE ESTIMATION OF DIFFERENT METHODS

Method	Method in [6]	Method in [8]	Our method
Precision	88.7%	86.7%	94.3%

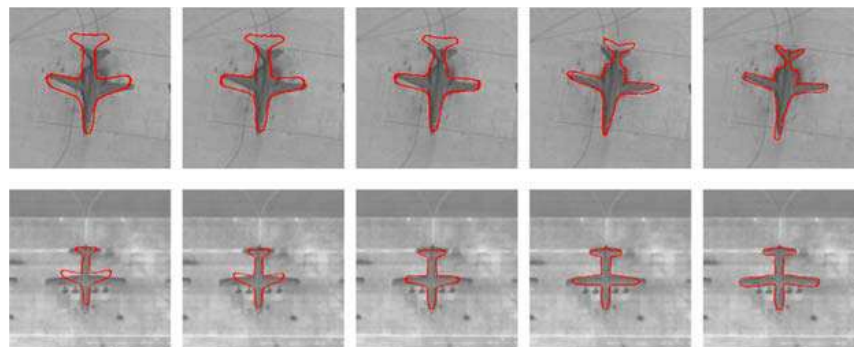


Fig. 6. Curve evolutions in segmentation.

shows some matching results. From these images, we can see that, although there exist different contrasts and background disturbances, the structure of aircraft is still kept in the preprocessed gradient image, and the Score in (1) would get the maximum at the desired pose. Thus, the average template can well coarsely estimate the pose information. We compare our method with those in [6] and [8]. Since they only settle the problem of estimating main axis of aircraft and the requirements of accuracy are not exactly the same, we suppose that a correct pose is gotten when estimated axis direction is within an error of 15°. Table II lists the accuracies of these methods. Our method obtains much better performance than the other two methods. It is mainly because that their pixelbased segmentations before pose estimation are

easily affected by various disturbances in practice, which, as a result, affects the accuracy of PCA and symmetric analysis in pose estimation.

B. Shape-Based Segmentation

We apply the proposed energy function to segmentation. The parameters are chosen empirically. The scale τ of Gaussian kernel is set to three. The ε in Heaviside function is set to five. We set $\alpha = 1 \times 10^4$ in (6). The step-size parameters of W in (16) are set to 2.0, 1.2, 0.8, and 0.6, respectively, and those of a , b , h , and θ are set to 3×10^{-4} , 3×10^{-4} , 1×10^{-8} , and 2×10^{-8} , respectively. We fix the number of iteration to 200 to stop the minimization process. Fig. 7 shows some examples of contour evolution. We can see that, with the integration of the shape prior, the contours of segmentation always look regular and

good results are obtained at last. In segmentation, to illustrate the effect of our pose estimation, we compare our method with the segmentation using a random initialization. Fig. 8 shows some comparisons. It can be seen that segmentation using a

random initialization easily runs into a local minimum, while with our pose estimation, this problem can be better avoided. We also compare our method with Otsu's method and RSF that is the first energy term in (6) without shape prior.

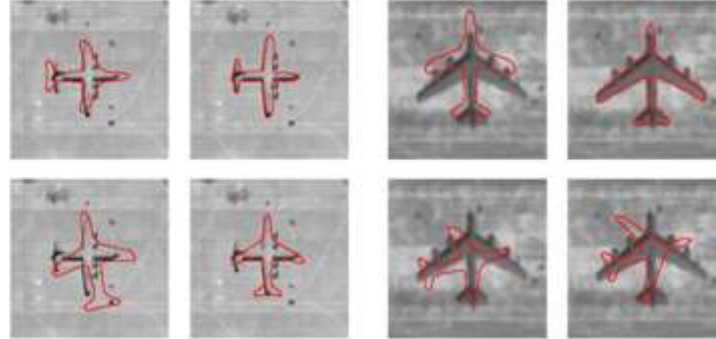


Fig. 7. Comparison of segmentation with different initializations. (Top row) With our pose estimation. (Bottom row) With a random initialization.

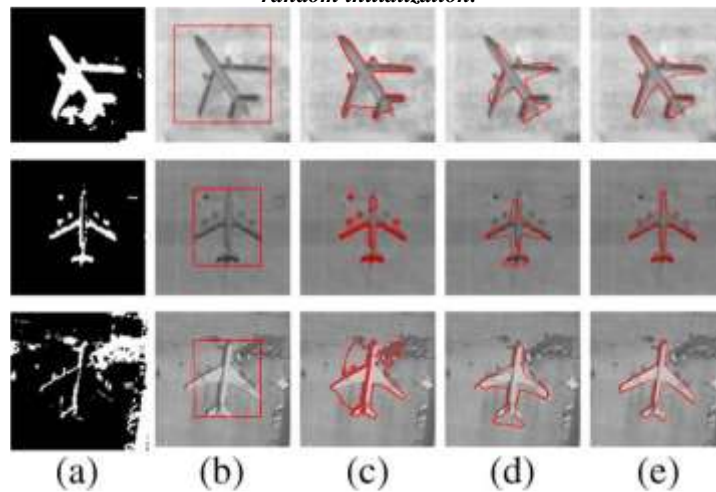


Fig. 8. Comparison of segmentation. (a) Otsu's method. (b) Initialization of RSF. (c) Result of RSF. (d) Our pose estimation. (e) Our segmentation.

Table II
 precision on recognition of different methods

Method	Method in [6]	Method in [8]	Our method
Precision	85%	80.3%	89.3%

Fig. 9 shows some results. In the first row, parts of the airplane's wing and empennage have poor contrast with background; in the middle row, many pixels of the target have intensities similar with that of the with that of the background; and in the third row, there exist so many background disturbances. From these examples, it can be seen that methods using only image appearance information are not capable of separating target from background and are easily affected by bad image quality and background disturbances. Besides these, their results are often so irregular which makes the following recognitions

difficult. On the contrary, since we model shapes of aircraft globally as well as restrict the solution in the PCA linear space, the result is more regular and more robust with respect to various disturbances.

C. Recognition Results

In recognition, considering the number of aircraft samples used in our experiment, we set K to three in the first KNN stage. The length and width of the object are calculated with the obtained parameters W and T . Table III shows the recognition accuracies of our method and the other two methods. Due to the better performance of pose estimation and

segmentation of our method, we get better recognition accuracy than the other two approaches. From the accuracies of pose estimation and recognition, it can also be seen that template matching [6] for recognition is very accurate when pose is rightly estimated or segmentation is well done

in other words. However, it has no dimension reduction, so that it is not scalable when the number of aircraft increases. Experimental results show the advantage of our shape-model-based recognition method.

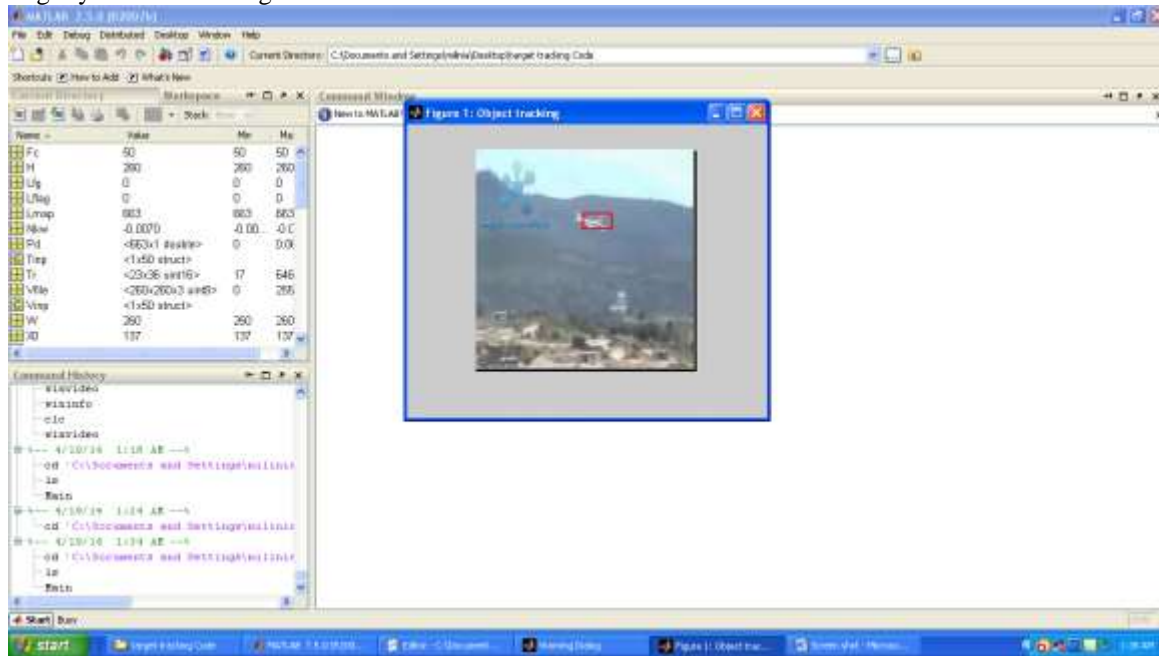


Fig 9.Screenshot of the output

Conclusion

In this letter, we have proposed a coarse-to-fine aircraft recognition method. The similarity and difference of shapes for aircraft are explored in the coarse and fine stages, respectively. A single template with a defined score is adopted to estimate the pose roughly in the coarse stage. In the fine stage, PCA and kernel density function are used to model shapes to get good effects on dimension reduction and sample space description. A new energy function is proposed to embed the shape model to segmentation. In addition, the obtained size and PCA coefficients are used directly to recognition. Experiments show that our proposed method is robust with respect to various disturbances.

References

1. F. Zhang and S. Q. Liu, "Aircraft recognition in infrared image using wavelet moment invariants," *Image Vis. Comput.*, vol. 27, no. 4, pp. 313–318, Mar. 2009.
2. C. C. Lin and R. Chellappa, "Classification of partial 2-D shapes using Fourier descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI- 9, no. 5, pp. 686–690, Sep. 1987.
3. G. J. Scott, M. N. Klaric, C. H. Davis, and C.-R. Shyu, "Entropy-balanced bitmap tree for shape-based object retrieval from large-scale satellite imagery databases," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1603–1616, May 2011.
4. S. C. Tien, T. L. Chia, and Y. Lu, "Using cross-ratios to model curve data for aircraft recognition," *Pattern Recognit. Lett.*, vol. 24, no. 12, pp. 2047–2060, Aug. 2003.
5. C. F. Xu and H. B. Duan, "Artificial bee colony (ABC) optimized edge potential function (EPF) approach to target recognition for low altitude aircraft," *Pattern Recognit. Lett.*, vol. 31, no. 13, pp. 1759–1772, Oct. 2010.
6. D. P. Shao, Y. N. Zhang, and W. Wei, "An aircraft recognition method based on principal component analysis and image model matching," *Chin. J. Stereol. Image Anal.*, vol. 14, no. 3, pp. 261–265, Sep. 2009.

7. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
8. J. W. Hsieh, J. M. Chen, C. H. Chuang, and K.-C. Fan, "Aircraft type recognition in satellite images," *Proc. Inst. Elect. Eng.—Vis. Image Signal Process*, vol. 152, no. 3, pp. 307–315, Jun. 2005.
9. D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *Int. J. Comput. Vis.*, vol. 69, no. 3, pp. 335–351, Sep. 2006.
10. A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, Feb. 2003.
11. M. Rousson and D. Cremers, "Efficient kernel density estimation of shape and intensity priors for level set segmentation," in *Proc. MICCAI, 2005*, vol. 1, pp. 757–764.
12. S. J. Osher and J. A. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on Hamilton–Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, Nov. 1988.
13. Y. Shuai, H. Sun, and G. Xu, "SAR image segmentation based on level set with stationary global minimum," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 644–648, Oct. 2008.
14. H. Ma and Y. Yang, "Two specific multiple-level-set models for high resolution remote-sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 558–561, Jul. 2009.
15. Q. Xu, B. Li, Z. He, and C. Ma, "Multiscale contour extraction using a level set method in optical satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 854–858, Sep. 2011.
16. C. M. Li, C. Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.
17. T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

Author Bibliography

	<p>Author Name here Arathy Krishnan G.U ME:Avionics PSNCET,Tirunelveli arathykrishnangu1@gmail.com</p>
--	---