International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-3 Issue-3, August 2013

Target Detection and Tracking in High-Resolution Aerial Images using Contour-Based Spatial Model and Correlation Tracker

G.Kishore, P.Srinivasulu

Abstract: ***This project aims at developing a contour-based spatial model which can detect geospatial targets accurately in high-resolution Aerial images. The detected targets are tracked using target tracking Correlation Algorithm. To detect the geospatial targets with complex structures, each image was partitioned into pieces as target candidate regions using multiple segmentations at first. Then, the automatic identification of target seed regions is achieved by computing the similarity of the contour information with the target template using dynamic programming. Finally, the contour-based similarity was further updated and combined with spatial relationships to figure out the missing parts. In this way, a more accurate target detection result can be achieved. The detected target further has to be monitored for its movements, this is achieved by implementing a correlation based tracking algorithm which efficiently tracks the target movements in successive image frames and its 2-D coordinates are plotted. The 2-D coordinates gives the observer a view of the Target movements and intentions.***

Keywords-

I. INTRODUCTION

Target detection is one of the most challenging tasks in the high-resolution remote sensing image analysis. With the development of remote sensing imaging, the high spatial resolution can provide abundant spatial and contextual information for target detection [1] . The use of such information offers the opportunity to detect geospatial targets with complex structures accurately, such as aircraft. While the fields of target detection and segmentation have been researched widely in recent years, it is common for many authors to consider these two tasks separately. Target detection has been achieved using, for example, constellation models [2] and the deformable shape model [3] . However, the target detection results by these methods are not precise enough. Some research studies which combined object detection and segmentation have been achieved, such as visual category filter [4] and implicit shape model [5]. Due to the lack of spatial and structure information, some methods can only get successful applications with limited data. Moreover, it is difficult to find a good partition method that works well for all objects in a complex scene, such as the high-resolution remote sensing image. Multiple segmentations received significant attention because some regions in some of the segmentations are almost correct for targets in one image.

Manuscript published on 30 August 2013.
*****Correspondence Author(s)
G.Kishore, M.Tech Scholar Deptt.of Electronics and Communication Engineering Sri Kalahasteeswara Institute of Technology, Sri kalahasti-517640, India.
P.Srinivasulu, Assistant professor Deptt.of Electronics and Communication Engineering Sri Kalahasteeswara Institute of Technology, Sri kalahasti-517640, India.

Meaningful and interesting objects are extracted using multiple segmentations [6],[7]. The question is how to pick out the meaningful regions from the pool of image segments [8]-[10].

The model is developed and simulated using Mat lab software. The results for images obtained from UAV imagery will be used for testing the system.

A threefold method is proposed which can detect geospatial targets present in high-resolution remote sensing images and delineate the target boundaries precisely. First, a contour-based spatial (CBS) model is proposed for segmenting and detecting target instances simultaneously [12]. Second, the contours of segments, instead of the extracted edges, are combined with the similarity calculation for identifying the meaningful target regions. This is a natural and simple way after the segmentations and can avoid clutter and help in extracting good edges. Third, in our model, spatial relationship functions are utilized to represent spatial context of different image pieces which can help in achieving accurate detection results [11].

The detected target is templated (reference image) and the target is tracked in successive image sequences acquired by the remote sensing device using image tracking algorithms such as correlators this tracker tracks the target accurately in the geospatial imagery and gives the details of target movement in 2D-X,Y coordinates as well as correlation coefficient information. with further processing this tracker information can be used to track the target movements.

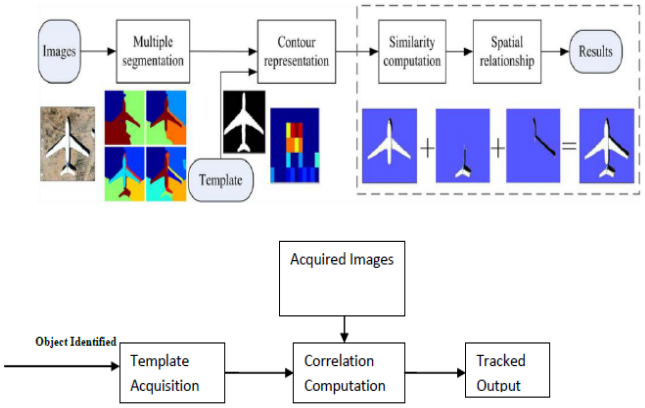

Fig:1 Overview of the process of our target detection and tracking algorithm. A single binary mask is given as the target template.

II. SEED REGION IDENTIFICATION

The problem can be decomposed into three stages. First, multiple segmentations are used to extract candidate regions from the given image.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The result is a set of image pieces. After this, the contour information of each piece is obtained using the shape descriptor. Finally, the similarity of each image contour with the target template is calculated by dynamic programming, and the similar pieces are picked up as target seed regions. The template is made according to the shape information of the target.

A. Multiple Segmentations

The first step aims at obtaining sufficient candidate regions of the image to have a high chance of obtaining the correct target seed regions. The choice of segmentation algorithm is not critical since not all segmentations are expected to be correct. The normalized cut framework [13] is used to produce the candidate regions, because it aims to find a global segmentation with large segments that have a chance to be targets. To produce multiple segmentations, we vary one key parameter of the segmentation algorithm, the number of segments.

According to the size of the image, we run the segmentation algorithm with the parameter set

$$\{ci : ci+1=ci+2,\ \lceil\alpha\sqrt{W\times H}\rceil \le ci \le \lceil\beta\sqrt{W\times H}\rceil\}$$

where W and H are the pixel numbers of image width and height, respectively, and $\lceil\ \rceil$ is the ceiling function. α and β are the segment coefficients ($\alpha = \beta$ means a single segmentation and $\alpha < \beta$ means the multiple segmentations).

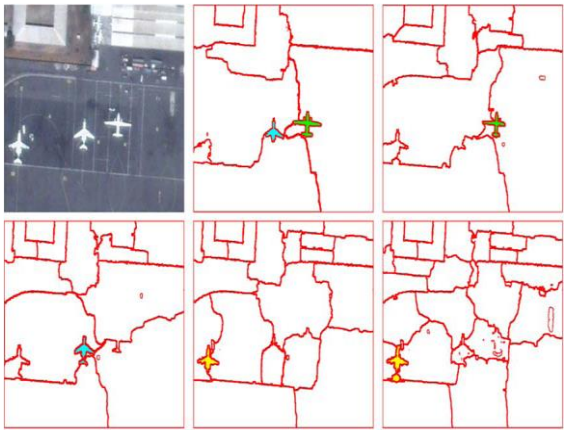

Fig.2.Image and its multiple segmentation sampling with α = 0.015 and β = 0.05. The selected image pieces are shown in colours.

B. Segment Information Description

The regions obtained from segmentation are represented with their contour information using the shape descriptor. The shape descriptor is expected to be robust and has invariance properties in translation, scale, rotation, and slight shape variation .The shape context (SC) descriptor is a comprehensive choice. The idea of SC was first proposed by Belongie *et al.* It describes the relative spatial distribution of landmark points around the edge point.

Given a segment P specified by a sequence of edge points $\{pi, i = 1, 2, \ldots, n\}$ sampled from its contour, the SC at point pi is defined as a histogram hi of the relative coordinates of remaining $n - 1$ points. The bins of the histogram are uniform in log−polar2 space

$$hi(k) = \#\{pj : j \_= i, pj - pi \in bin(k), 1 \le k \le K\}$$

Where # means the cardinality of the set, $bin(k)$ is the kth bin, $K = S \times T$ is the number of bins, and S and T denote the numbers of distance bins and orientation bins, respectively. It takes $O(n2)$ time to compute the histograms for n sample points.

C. Segment Selection

After forming the descriptor for each segment contour, our aim is to search for the similar segments as target seed regions. Similarity calculation is carried out by contour matching, which is to look for a set of correspondences between two contours. Belongie *et al.* used a framework combining SC and thin plate splines [15] and showed that it is very effective for shape matching tasks. Thayananthan *et al [16].* suggested an efficient dynamic programming scheme for SC matching including a figural continuity constraint. Our approach includes a similar constraint by assuming that contour points are ordered and using dynamic programming for matching the SC at contour sample points.

Let p and q be two points on two different contours and their SCs are corresponding to Hp and Hq, respectively. The similarity of the two points is defined as the distance between their Descriptors. The distance between two histograms is defined using $\chi 2$ statistic as in

$$dist(p, q) = \chi 2(hp, hq) = 1/2 \sum_{k=1}^{k} \frac{[hp(k) - hq(k)]2}{hp(k) + hq(k)}$$

where K is the number of SC bins. $hp(k)$ and $hq(k)$ are the normalized shape descriptors at points p and q, respectively. Given two contours P and Q, i.e., $P = \{pi, i = 1, \ldots, Np\}$ and $Q = \{qj, j = 1, \cdots, Nq\}$, the match cost matrix between them is denoted as

$$Dist(P,Q) =$$

$$\begin{matrix} dist(p1, q1) & dist(p1, q2) & \cdots & dist(p1, qNq) \\ dist(p2, q1) & dist(p2, q2) & \cdots & dist(p2, qNq) \\ & \ldots & & \\ & \ldots & & \\ & & \cdots & \\ & \ldots & & \\ dist(pNp, q1) & dist(pNp, q1) & \cdots & dist(pNp, qNq) \end{matrix}$$

The matching cost can be used to measure the similarity between a segment P and the template Q. Dynamic programming is utilized for this matching, and the dynamic programming [17] table equals the match cost matrix $Dist(P,Q)$. The dynamic programming procedure runs in $O(n3)$ time. In practice, we use 100 points in each contour. The segment selection takes about 0.3 s on a standard personal computer Intel Core i5, running at 3.10 GHz with 2-GB RAM.

III. MODELING SPATIAL RELATIONSHIPS

After the images are segmented and the target seed regions are identified, the next step in our method is modeling of spatial relationships. Spatial relationships of region pairs can be defined as context classes [18].

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

According to their relative location, the class membership is denoted as ψc with $c \in \{Dis, Bor, Int, Sur\}$ corresponding to disjoined, bordering, intersected, and surrounded, respectively. the spatial relationships of region pairs are defined as relationship classes, and the value ψc indicates the degree of region pairs to class c.

The spatial relationship between all regions in an image can be represented by a region relationship matrix. For a pair of regions, we first compute the following

1) areas of each region ςi and ςj ;
2) common area between two regions ςij computed as the number of shared pixels between two regions;
3) ratio of the common area to the minimum area $\eta ij = \varsigma ij / \min(\varsigma i, \varsigma j)$. $i, j \in \{1, \ldots, n\}$, and n is the number of regions in the image.

For the spatial relationship, the area ratio ηij is used in the following trapezoid membership functions:

1) Disjoined

$$\Psi Dis(\eta ij) = \begin{cases} 1, & \text{if } \eta ij = 0 \\ 0, & \text{otherwise.} \end{cases}$$

2) Bordering

$$\Psi Bor(\eta ij) = \begin{cases} 1, & \text{if } 0 < \eta ij \le 0.04 \\ \frac{-50}{49}\eta ij + \frac{51}{49}, & \text{if } 0.04 < \eta ij \le 1 \\ 0, & \text{otherwise} \end{cases}$$

3) Intersected

$$\Psi Int(\eta ij) = \begin{cases} \frac{50}{9}\eta ij - \frac{1}{9}, & \text{if } 0.02 \le \eta ij < 0.20 \\ 1, & \text{if } 0.20 \le \eta ij < 0.80 \\ -\frac{15}{4}\eta ij + 4, & \text{if } 0.80 \le \eta ij < 1 \\ 0, & \text{otherwise} \end{cases}$$

4) Surrounded

$$\Psi Sur(\eta ij) = \begin{cases} 10\eta ij - 8, & \text{if } 0.80 \le \eta ij < 0.90 \\ 1, & \text{if } 0.90 \le \eta ij \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

Each region pair can be assigned a degree of their spatial relationship using these spatial relationship functions. The functions and specific examples are shown

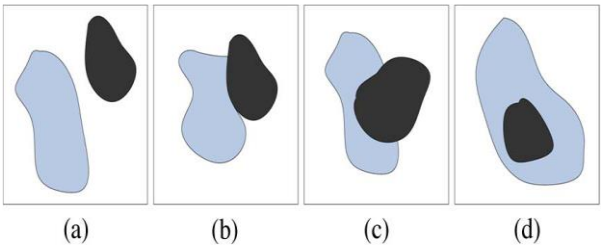

Fig:3Spatial relationships of region pairs. (a) Disjoined. (b) Bordering. (c) Intersected. (d) Surrounded

IV. TARGET DETECTION

Suppose that S denotes the target seed region and R denotes the set of image pieces. The combination scheme alternates between the following two steps. Since there are the lack of spatial contextual information in segmentations and background interference, the seed region pieces ordinarily are not whole targets and miss some parts of the targets. The missing parts of targets are extracted according to their spatial relationships in the image

In the missing part extraction process, we select the missing parts by a criterion function

$$\lambda = ws - \tau$$

where w is the weight according to spatial relationship, s is the shape similarity, and τ denotes the threshold. The value of criterion function λ is compared with zero. Positive value means that the piece is the target missing part. The threshold is initialized as the similarity of seed regions and adjusted during the process

1)Search step: Fix the threshold τ , and search the image pieces R according to the criterion function. For every image piece, calculate its response value of the criterion function. Extract the image pieces $r$target whose response is positive. This routine can be implemented efficiently using dynamic programming.

2) Update step: Fix the image pieces that belong to the target, compose them together forming the new seed region $S = S \cup r\text{target}$ and $R = R - r\text{target}$, and update the threshold. The algorithm ends when no missing piece is obtained at the search step. This target piece selection procedure can be shown to converge to a local maximum of contour similarity.

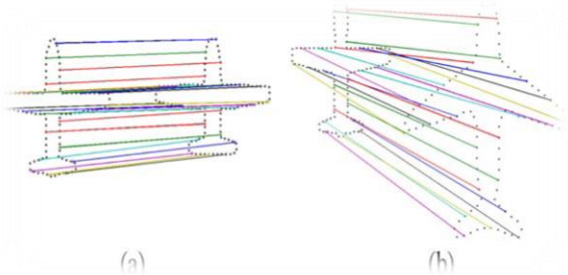

Fig: 4 Aircraft silhouette matching. (a)Matching between similar silhouettes. (b) Matching between dissimilar silhouettes. Note that the airfoils are correctly matched. Only quarter of the matched pairs are shown for illustration.

V. CORRELATION TRACKER

The standard way of tracking features between two images is by template matching. This approach involves taking a given pattern in one image and shifting a template containing the same pattern in another image until the best comparison is found. The most common and effective way of doing this task is by Normalized Cross Correlation (NCC) methods, which have a significant advantage over standard cross correlation (CC) methods, in that these methods are robust to different lighting conditions across an image and less sensitive to noise [19] . However, both methods can be computationally intense, especially for large imag

A.*Correlation Computation*

R : Reference image
S : Search image
i,j : i & jth pixel

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$c(i,j)=\frac{M^2\sum_{i=1}^{m}\sum_{j=1}^{n}RS_{ij}-\sum_{i=1}^{m}\sum_{j=1}^{n}R^*\sum_{i=1}^{m}\sum_{j=1}^{n}S_{ij}}{\left[\left\{M^2\sum_{i=1}^{m}\sum_{j=1}^{n}R^2-\left(\sum_{i=1}^{m}\sum_{j=1}^{n}R\right)^2\right\}*\left\{M^2\sum_{i=1}^{m}\sum_{j=1}^{n}S_{ij}^2-\left(\sum_{i=1}^{m}\sum_{j=1}^{n}S_{ij}\right)^2\right\}\right]^{\frac{1}{2}}}$$

N : No. of search locations
M2: Reference Image Size : 75 pixels x 55 lines
S2 :Search Image Size : 115 pixels x 95 lines

VI. RESULTS AND EVALUATION

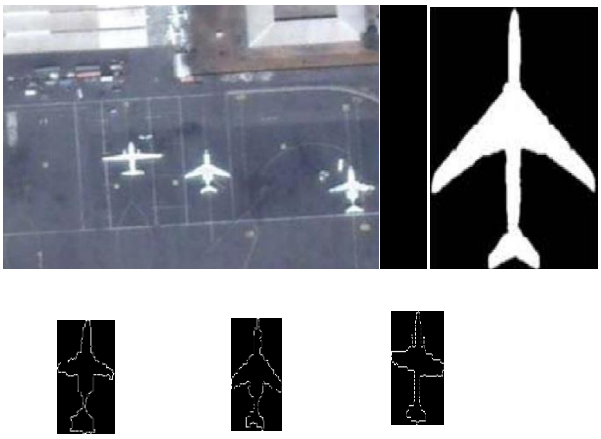

Fig:5 Results of our algorithm for aircraft accurate detection. The first rows include search and reference images. The second row shows detected results

Fig: Tracking Results

VII. CONCLUSION

In this letter a CBS model and correlation tracker has been proposed to solve the problem of detecting and tracking geospatial targets present in high resolution aerial images. The CBS model is a global shape model, which captures the global characteristic of targets. Therefore, the proposed method can perform well on targets with relatively fixed shape. The detected target is tracked by using correlation tracker, this is done by using the correlation computation between the search image and reference image. Here for correlation tracker the object that is detected by CBS model acts as template. An improved work of video tracking in real time is our future scope

REFERENCES

1. D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," IEEE Geosci. Remote Sens., vol. 47, no. 11, pp. 3866–3879, Nov. 2009.
2. R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in Proc. CVPR, Jun. 2003, vol.2, pp. II-164–II-271.
3. A. C. Berg, T. L. Berg, and J. Malik, "Shape matching and object recognition using low distortion correspondences," in Proc. CVPR, Jun. 2005, vol. 1, pp. 26–33.
4. R. Fergus, P. Perona, and A. Zisserman, "A visual category filter for google images," in Proc. ECCV, 2004, vol. 3021, pp. 242–256.
5. B. Leibe, A. Leonardis, and B. Schiele, "Combined object categorization and segmentation with an implicit shape model," in Proc. ECCV, 2004, pp. 17–32.
6. H. Akcay and S. Aksoy, "Automatic detection of geospatial objects using multiple segmentations," IEEE Geosci. Remote Sens., vol. 46, no. 7, pp. 2097–2111, Jul. 2008.
7. X. Sun, H. Wang, and K. Fu, "Automatic detection of geospatial objects using taxonomic semantics," IEEE Geosci. Remote Sens. Lett., vol. 7, no. 1, pp. 23–27, Jan. 2010.
8. V. Ferrari, L. Fevier, F. Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 1, pp. 36–51, Jan. 2008.
9. J. Shotton, A. Blake, and R. Cipolla, "Multiscale categorical object recognition using contour fragments," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 7, pp. 1270–1281, Jul. 2008.
10. A. Opelt, A. Pinz, and A. Zisserman, "A boundary-fragment-model for object detection," in Proc. ECCV, 2006, vol. 2, pp. 575–588. [11] D. M. Gavrila, "A Bayesian, exemplar-based approach to hierarchical shape matching," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 8, pp. 1408–1421, Aug. 2007.
11. X. Bai, Q. Li, L. J. Latecki,W. Liu, and Z. Tu, "Shape band: A deformable object detection approach," in Proc. CVPR, 2009, pp. 1335–1342.
12. J. Shi and J. Malik, "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, Aug. 2000.
13. M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," Pattern Recognit., pp. 43–90, Nov. 2008.
14. S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 4, pp. 509–522, Apr. 2002.
15. A. Thayananthan, B. Stenger, P. H. S. Torr, and R. Cipolla, "Shape context and chamfer matching in cluttered scenes," in Proc. CVPR, 2003, vol. 1, pp. I-127–I-133.
16. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, 2nd ed. Cambridge, MA: MIT Press, 2001.
17. S. Santini, Exploratory Image Databases: Content-Based Retrieval. New York: Academic Press, 2001.
18. M. Raffel, C. Willert, and J. Kompenhans, Particle Image Velocimetry: A Practical Guide. Berlin: Springer-Verlag, 1998.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)