第 39 卷第 7 期
2010 年 7 月

光 子 学 报
ACTA PHOTONICA SINICA

Vol. 39 No. 7
July 2010

Article ID: 1004-4213(2010)07-1278-6

3D Target Recognition Based on Fast Skeleton Extraction

WANG Song-wei, LI Yan-jun, ZHANG Ke

(*College of Astronautics Northwestern Polytechnical University*, *Xi'an* 710072, *China*)

Abstract: Aiming at the IR 3D target recognition question, a method based on fast skeleton extraction is proposed. Based on the topology structure of target reflected by the skeleton, the correspondence between different targets' structural portions are established. Within the corresponding curve segments, curves are matched based on curve alignment, and the similarity is evaluated according to the sum of all the segments' cost. This method performs well in the present of commonly occurring visual transformations. Moreover, incorrect correspondence introduced by curve match's drawback, that is, curve matching tends to match local parts, and ignores the global view of how these parts are spatial arranged with respect to each other, can be avoided. Experimental results show that this algorithm has a rather good effect on 3D target recognition.

Key words: 3D target recognition; Target shape; Skeleton; Curve matching

CLCN: V556, TP391. 4 **Document Code**: A **doi**: 10. 3788/gzxb20103907. 1278

0 Introduction

When prototype-views method[1-2] is used in the recognition of 3D target, similarities between the target and prototype-views will be measured to identify the category of target. In this method, only several prototype-views are used to represent 3D target. So, compared with the traditional recognition ways[3-6], algorithm must work well in the presence of modest amounts of view-point variation. Moreover, to improve algorithm's robustness, performance under visual transformation, affine transformation, some articulation and deformation must be well.

Many useful information of targets lie in its shape, and the representations of the shape of targets can have a significant impact on the effectiveness of a recognition strategy. To satisfy above demands of 3D target recognition, curve-based and skeleton-based shape representations are the common selection.

One type of shape representation models the shape outline as curves and matched[7-8]. The major advantage of curve matching is its computational efficiency. This method can give intuitive matches in the presence of commonly occurring visual transformations including modest amounts of articulation, viewpoint variation, partial occlusion, and some affine transformation. But curve matching tends to match local parts, and ignores the global view of how these parts are spatial arranged with respect to each other. In the presence of changes in the part structure, curve matching fails.

Based on skeletons directly, many approaches have been developed for shape matching[9-10]. The benefits of applying skeleton-based methods are its natural consistency with human intuition and capability to descibe the local geometrical features, allowing the performance of articulated matching. But traditional skeleton extraction algorithm is time-consuming, and sensitive to the boundary noise. Small boundary disturb can create great skeleton branches, all of which will have a bad effect on the recognition result of skeleton-based method.

In this paper, a fast target skeleton extraction algorithm is proposed, and the resulting skeleton reflects the main topology structure of the target. Instead of pursuing the exact matching of skeleton graphs, which would introduce extra complexity and instability into the matching process, we divide the contour of the target into several segments based on the endings of skeleton, one of these segments corresponding to a structural portion of target. Within segments, we utilize the curve matching, which works well in the present of visual transformations of limited extent, to

* Supported by the National Nature Science Foundation of China (60575013), the Aeronautics Science Foundation (04I53067) of China and Astronautics Science Innovation Foundation (05C53005) of China

Tel: 029-82375057 Email: wangsongwei79@163.com

Received date: 2008-11-03 Revised date: 2009-05-05

calculate the matching cost of the two curves and identify the category of target. In this way, the drawbacks of curve-based method, that is, curve matching tends to match local parts, and ignores the global view of how these parts are spatial arranged with respect to each other, can be overcome. And the experiment shows that the proposed method has a good recognition results compared with the traditional way.

1 Fast skeleton extraction algorithm

An adjacent skeleton point determination criterion of knowing skeleton is presented in this paper. According to this criterion, the adjacent skeleton point of an initial skeleton point can be determined. In this way, the whole skeleton structure of target can be acquired with rather low complexity. Combined with the discrete curve evolution model[11], the redundant skeleton branches are eliminated. In this way, we can decrease the sensitivity of skeleton to boundaries noises greatly and remain the topology structure of target completely.

1.1 Skeleton based on the maximal disk and related definitions

Suppose that the contour C of a target in an image is represented by a continuous closed curve. Inside the contour C, the planar shape F represents the content of the object. The corresponding skeleton S can be determined, as shown in Fig. 1. According to Blum′s definition[12], the skeleton S of a planar shape is defined as the locus of the centers of the maximal disks contained inside the planar shape F.

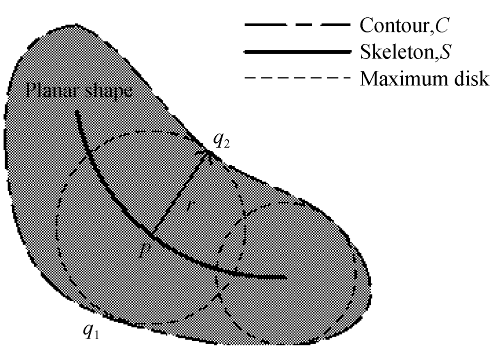


Fig. 1 The definition of a skeleton

It can be observed that almost all skeleton points are associated with at least two boundary points whose respective distances to the skeleton point are the shortest except the endings of the skeleton. In Fig. 1, with a maximal disk centered at p, the object′s contour and the maximal disk touch each other at points q_1 and q_2.

In order to efficiently find the maximal disk enclosed in an object, a distance map should be generated before locating the center of the maximal disk, that is, the Distance Transform $DT(F)$. For each pixel in binary image, the distance transform is assigned by a number that is the distance between that pixel and the nearest nonzero pixel of BW. Not only is the shortest distance to the object′s boundary determined, but the position of the corresponding point on the boundary is also provided.

1.2 Adjacent skeleton point determination criterion

Suppose the maximum disk corresponding to the skeleton point p and the contour touch each other at the points q_1, q_2, …, q_n. These n points divide the contour C into n segments $\boldsymbol{q_1q_2}$ $\boldsymbol{q_2q_3}$, … $\boldsymbol{q_nq_1}$. Combing pq_1, pq_2, …, pq_n, these n segments constitute n connected domains, as shown in Fig. 2. For every connected domain, we have the following theorem.

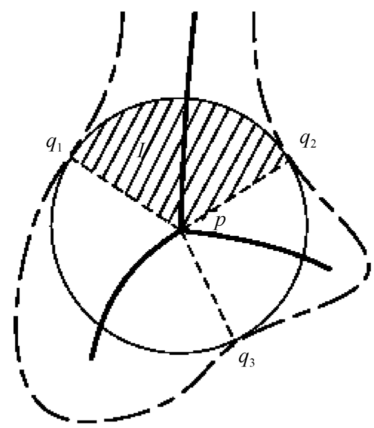


Fig. 2 The tangent points between the maximum disk and target boundary

Theorem 1: If the maximum disk corresponding to the skeleton point p and the contour touch each other at points q_1, q_2, …, q_n, there exists one and only one adjacent skeleton point of p in connection region $q_iq_{i+1}pq_i$.

Proof: Consider region $q_1q_2pq_1$. Because there is two tangent points between this region and the maximum disk centered at p, there must exist a maximum disk which has at least 2 tangent points with this region. Without loss of generality, suppose the maximum disk centre at p_k, and three tangent points of all their tangent points are q_1^k, q_2^k, q_3^k, as shown in Fig. 3. There must exist a series of maximum disks between maximum disk centered at p and maximum disk centered at p_k, and these maximum disk has two tangent point with region $q_1q_2pq_1$, which lie in arc $q_1q_1^k$ and $q_2q_2^k$, respectively. As shown in Fig. 3, the maximum disk center at p_1, with the tangent points q_1^1 and q_2^1. The centers of all these maximum disks constitute a curve which starts at point p and ends at point p_k. So there exists one and only one adjacent skeleton point of p in region $q_1q_2pq_1$.

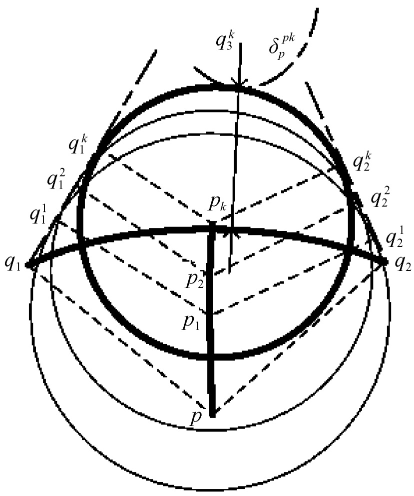


Fig. 3 The illustration for proving Theorem 1

Theorem 1 restricts the positions of the adjacent skeleton points of the known skeleton point p into certain scope.

Utilizing the Distance Transform, we can obtain the nearest neighborhood boundary points of the known skeleton point p and its 8 adjacent points, q_i ($i=1, \cdots, 9$). Cluster these boundary points, and regard the cluster center locations as the tangent points between the maximum disk centered at p and the contour. Lines between p and the tangent points divide the 8 adjacent points of p into several regions, as shown in Fig. 4, in which A and B are the tangent points. According to theorem 1, there exists one skeleton point in every region. Skeleton point is the medial axis point, so point in every region with the closest distance to the medial axis is the skeleton point.

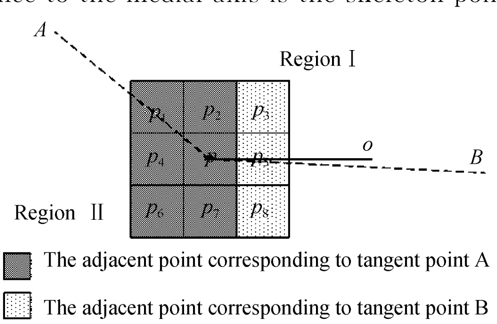


Fig. 4 The tangent points of skeleton point p

As shown in Fig. 4, two adjacent points corresponding to different tangent points can be found in every region, for example, p_2 and p_3 in region I, and p_7 and p_8 in region II. These two points are points closest to the medial axis in region obviously and can be taken as the candidate skeleton points. According to the distance transform, q_2, q_3 are the nearest neighborhood boundary point of p_2, p_3, respectively. For p_2, set $r_1=\text{dis}(p_2, q_2)$ and $r_2=\text{dis}(p_2, q_3)$; for p_3, set $r_1=\text{dis}(p_3, q_3)$ and $r_2=\text{dis}(p_3, q_2)$, calculate $\delta=|r_1^2-r_2^2|$ respectively. The litter δ, the closer is the point to the medial axis. And we set this point as the skeleton point in this region.

1.3 Skeleton extraction algorithm

The above criterion can be used to determine the adjacent skeleton point of a known skeleton, so we must set the initial skeleton point. According to[11], we have the following theorem.

Theorem 2 For planar shape F, point $p\in F$, if $DT(p)=\max(DT(F))$, $p\in SK(F)$.

According to theorem 2, point in planar shape with the maximum distance transform value is the initial skeleton point. Utilizing the above criterion, its adjacent skeleton points can be obtain, and push them into stack. Then we pop the top element of stack, and calculate its adjacent skeleton points which will be push into stack. Continue this process, till the stack is null. In this processing, point which has been pushed into stack can not be pushed again. So the computing complexity of this algorithm is $O(s)$, where s is the number of skeleton points.

In order to decrease the sensitivity of this algorithm toward boundary noise and disturbs, during the skeleton growing process, the redundant skeleton branches are eliminated by the discrete curve evolution model[11] and only the main visual branches remains, as shown in Fig. 5.

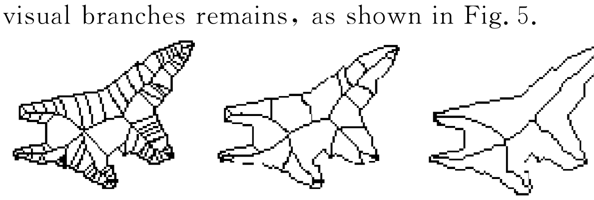


Fig. 5 Hierarchical skeleton by discrete curve evolution

2 Segment matching

As shown in Fig. 5(c), we get the skeleton which reflect the main topology structure of the target, each important structural piece of the shape leads to a corresponding branch on the medial axis and the terminals of these branches, i. e. the endings, naturally correspond to the maxima of positive curvature along the shape curve[13] and thus provide a good capability for indexing these structural portions[14]. Therefore, in this work, we use these endings divide target contour into several segments, one of which corresponding to a structural portions of target. Then, we will establish the correspondences between two segments of contours. To describe the features of these endings, we utilize the same context descriptor as in Ref. [15].

Given the contour points $C=\{c_i\}_{i=1}^{L}$ of a shape F and the endings $E=\{e_k\}_{k=1}^{K}$ on its medial axis, the context feature of each $e_k \in E$ is defined as a log-polar histogram h_k

$$h_k(b)=\#\{c\neq e_k:(c-e_k)\in \mathrm{bin}(b)\},\quad b=1,\cdots,B$$

where B denotes the number of bins used to describe the point set. The resulting coarse histogram records the relative locations of the sampling points referring to e_k. In this paper we use five bins for distance measure $\log r$ and 12 bins for angle θ. Thus, there are totally 60 bins in the log-polar space. Using this feature, the optimal matching of two sets of endings can be estimated by minimizing the total matching cost $\sum_1^K U(e_k, e_{\pi k})$, where πk is the index of correspondence of e_k, and $U(e_k, e_{\pi k})$ is the difference between the context features of e_k and $e_{\pi k}$ defined as

$$U(e_k, e_{\pi k})=\frac{1}{2}\sum_{b=1}^{B}\frac{[h_k(b)-h_{\pi k}(b)]^2}{h_k(b)+h_{\pi k}(b)} \qquad (1)$$

If the two sets of endings have different sizes, dummy points may be added to each set. The resulting square cost matrix is then input to an assignment problem (AP), and the matching can be solved in $O(n^3)$ by using the Hungarian method[16]. Fig. 7 shows examples of matching the segment between two target contours. With the aid of the topology structure, we get a correct identification of the corresponding parts of the shapes.

3 Curve matching within segment

Once the segment matching is done, we can precede the curve match within segment. Let the curves segments to be matched, T and $\hat{T}$; be parameterized by arc length s and $\hat{s}$; respectively. Let $\theta(s)$ and $\hat{\theta}(\hat{s})$ be the tangent orientations of T and $\hat{T}$, respectively. The cost of deforming entire curves can be expressed as the sum of the costs of deforming infinitesimal sub-segments, which in turn is penalized by length and curvature difference. Specifically, we define the cost of matching infinitesimal curve segments as $d\mu = |d\hat{s}-ds| + R\,|d\hat{\theta}-d\theta|$; where R is a scale-dependent constant. The problem is then cast as minimizing an energy functional over all possible alignments between the two curves. The notion of an alignment curve is introduced to represent the alignment which ensures symmetric treatment of the two curves. Specifically, the alignment curve α, $\alpha(\zeta)=(s(\zeta), \hat{s}(\zeta))$ represents the points on the two curves $T(s(\zeta))$ and $\hat{T}(\hat{s}(\hat{\zeta}))$ that correspond. The optimal alignment curve α^* is found by minimizing the functional

$$\mu(\alpha)=\int_0^{\tilde{L}}\left[\left|\frac{d\hat{s}}{d\zeta}-\frac{ds}{d\zeta}\right|+R\left|\frac{d\hat{\theta}(\hat{s})}{d\zeta}-\frac{d\theta(s)}{d\zeta}\right|\right]d\zeta \qquad (2)$$

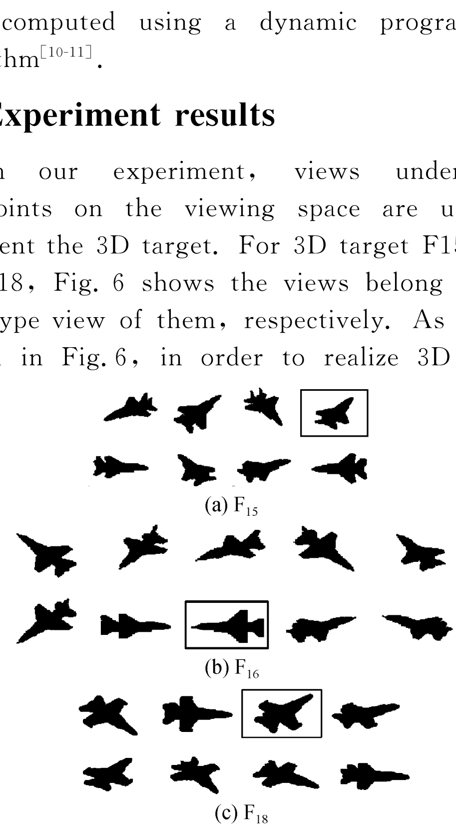
where $\tilde{L}$ is the length of the alignment curve. A key advantage of using the alignment curve is that the alignment can be expressed in term of a single function. We choose the angle between the tangent of the alignment curve and the x-axis, $\psi(\zeta)$: Then, the functional in (1) can be rewritten in terms of ψ as

$$\mu[\psi]=\int_0^{\tilde{L}}[|\cos(\psi)-\sin(\psi)|+R|\kappa\cos(\psi)-\hat{\kappa}\sin(\psi)|]\,d\zeta \qquad (3)$$

where κ and $\hat{\kappa}$ are curvatures along the curves T and $\hat{T}$, respectively. The optimal alignment curve α is computed using a dynamic programming algorithm[10-11].

4 Experiment results

In our experiment, views under 100 viewpoints on the viewing space are used to represent the 3D target. For 3D target F15, F16 and F18, Fig. 6 shows the views belong to one prototype view of them, respectively. As can be shown in Fig. 6, in order to realize 3D target



(a) F_{15} (b) F_{16} (c) F_{18}

Fig. 6 The views belong to a prototype view of F15, F16 and F18 recognition, algorithm must work well in the presence of modest amounts of view-point variation and deformation of parts.

Utilizing algorithm presented in this paper, skeleton of real-time target is extracted. And we establish the correspondence between contour segments of real-time target and the prototype views, as shown in Fig. 7, observe that the correspondence are intuitive.

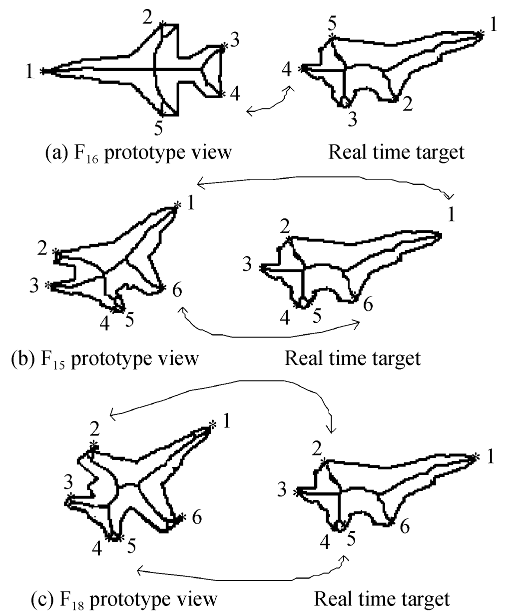


Fig. 7 Correspondences between prototype views and real time target

In the corresponding segments, curve matching costs are calculated. The sum of every matching cost of every segment is the cost of the total matching cost, that is, 0. 345 2, 0. 272 5 and 0. 447 8. And we can draw the conclusion that the real-time target is F15.

We apply the algorithm presented in this paper, algorithm based on curve representation and algorithm based on the skeleton representation to the planes shown in Fig. 7. In this experiment, each shape was matched to three prototype views and the matching costs are recorded. The minimum cost corresponds target category.

Our results for this data set are 23/23, while the results based on curve representation are 18/23 and results based on skeleton are 21/23, respectively.

In addition, the proposed technique and skeleton-based method work well in the presence of occlusion. As shown in Fig. 8 (a), intuitive correspondences are obtains. The matching costs between target and F15, F16, F18 prototype views are 0. 512, 0. 598, and 0. 622, respectively. Correct recognition result acquired. While in Fig. 8 (b), for the part structure of the target changes, method based on curve representation fails.

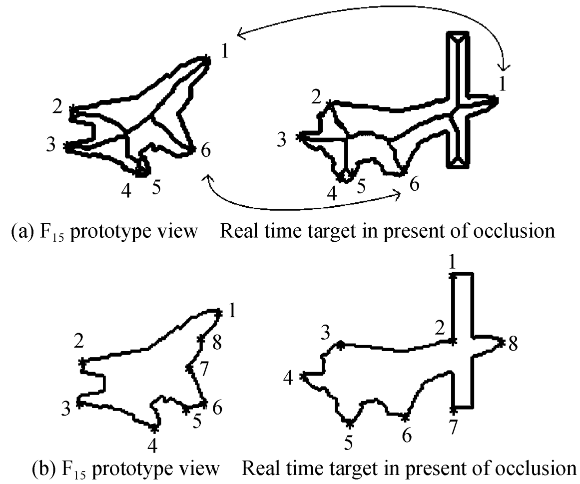


Fig. 8 Correspondences between prototype and real target in present of occlusion

5 Conclusion

Aiming at the 3D target recognition question, a method based on fast skeleton extraction is proposed. Target skeleton reflects its topology structure. So with the help of the skeleton of target, we can establish the correspondence of structural portions of different targets. Within these structural segments, utilizing the curve-based method which is symmetric in its treatment of the two curves, insensitivity to sampling, and invariant to rotation and scaling, the matching cost of different targets can be obtained. Experiment results show that this new algorithm combined target skeleton overcomes the drawbacks of curve-based method.

References

[1] CYR C, KIMIA B. 3D object recognition using shape similiarity-based aspect graph [C]. Proceedings of the International Conference on Computer Vision, Vancouver, Canada, 2001: 254-261.

[2] KLEIN P, SEBASTIAN T, KIMIA B. Shape matching using edit-distance: an implementation[C]. 20th Annual ACMSIAM Symposium on Discrete Algorithms, Washington, D C, 2001: 781-790.

[3] YANG Jing, QIU Jiang. Research of line moments and its application in image boundaries recognition [J]. *Acta Photonica Sinica*, 2003, **32**(3): 336-339.

[4] QIU Jiang, LIU Bo, YANG Jing, *et al*. Dentification algorithm of multi-object image using high-level Hu's moment [J]. *Acta Photonica Sinica*, 2001, **30**(9): 1141-1145.

[5] QIU Jiang, YANG Jing, LIU Bo. Identification of images of objects using wavelet and moment[J]. *Acta Photonica Sinica*, 2001, **30**(7): 836-843.

[6] LI Ying-hua, GUO Zheng-hong, FENG Liang, *et al*. Research on recognition of rotating targets in optical correlation using Karhunen-Loeve transform [J]. *Acta Photonica Sinica*, 2008, **37**(4): 750-753.

[7] SEBASTIAN T B, KLEIN P N, KIMIA B B. Alignment-

based recognition of shape outlines[C]. Proceedings of the International Workshop on Visual Form, Springer, Berlin, 2001: 606-618.
[8] SEBASTIAN T B, KLEIN P N, KIMIA B B. On aligning curves[J]. *Pattern Anal Machine Intell*, 2003, **25**(1): 116-125.
[9] GOLLAND P, GRIMSON W. Fixed topology skeletons[C]. Proceedings of the IEEE Computer Vision and Pattern Recognition, vol. I, Hilton Head Island, SC, 2001: 10-17.
[10] SEBASTIAN T B, KLEIN P N, KIMIA B B. Recognition of shapes by editing their shock graphs[J]. *IEEE Trans Pattern Anal Mach Intell*, 2004, **26**(5): 550-571.
[11] LIU W Y, BAI Xiang, ZHU Guang-Xi. A skeleton- growing algorithm based on boundary curve evolution[J]. *Journal of Automatic*, 2006, **32**(2): 255-262.
[12] BLUM H, NAGEL R N. Shape description using weighted symmetric axis features[J]. *Pattern Recognition*, 1978, **10**(2): 167-180.
[13] LEYMARIE F, LEVINE M D. Simulating the grass fire transform using an active contour model[J]. *IEEE Trans Pattern Anal Mach Intell*, 1992, **14**(1): 56-75.
[14] KIMMEL R, SHAKED D, KIRYATI N. Skeletonization via distance maps and level sets[J]. *Comput Ler Vision Image Understanding*, 1995, **62**(3): 382-391.
[15] JUN Xie, HENG P A, SHAH M. Shape matching and modeling using skeletal context[J]. *Pattern Recognition*, 2008, **41**(5): 1756-1767.
[16] PAPADIMITRIOU C H, STEIGLITZ K. Combinatorial optimization: algorithms and complexity[M]. Englewood Cliffs: Prentice-Hall, 1982: 12-16.

基于快速骨架提取的三维目标识别

王松伟,李言俊,张科
(西北工业大学 航天学院,西安 710072)

摘　要:针对三维目标识别问题,提出了一种基于快速骨架提取的方法。根据骨架所反映的目标拓扑结构,建立了不同目标局部结构之间的对应关系;而在对应的局部曲线段上,采用基于曲线配准的方法进行匹配;以各个局部匹配的成本之和评估不同目标的相似性.这种方法在目标出现一定程度的视觉变形时仍具有较好的识别效果,同时避免了基于曲线方法的匹配目标的局部,而忽略局部之间相互的空间组织的缺点所造成的误匹配.算例结果表明这种算法对于三维目标有较好的识别效果.
关键词:三维目标识别;目标形状;骨架;曲线匹配

WANG Song-wei was born in 1979. He is a Ph. D. degree candidate, and his research interests include 3D target recognition and tracking, multi-scale time series analysis.

LI Yan-jun was born in 1944. He received the Ph. D. degree in Northwestern Polytechnical University. He is the professor in the college of astronautics in polytechnical university. His research interests include adaptive control, Fault diagnostics and fault tolerant control.