

Contour Segment Feature Weighting in Infrared Automatic Target Recognition

Shi-Wei Chen, Sheng-Xiu Zhang, Xiao-Gang Yang and Nai-Xin Qi

Hi-Tech Institute of Xi'an, Xi'an 710025, P. R. China

cshw3876@tom.com

Keywords: Contour segment feature; Target recognition; Feature weighting; Bag-of-Words

Abstract. We study automatic target recognition (ATR) in infrared (IR) imagery by applying two recent computer vision techniques, Hierarchical description of contour and Bag-of-Words (BoW). We propose the idea of contour segment (CS) features which are extracted from the outer contour of IR target image and we developed a new weighted scalable vocabulary tree (SVT) that is learned from a set of training images to support efficient and scalable Bow-based ATR. We develop a weighted BoW model to improve the ATR performance by the feature weighting. Different from traditional BoW model, the saliency information is incorporated into feature weighting to enhance the voting confidence in Bow-based classification. The proposed CS features weighting ATR algorithm is evaluated against recent relevance grouping of vocabulary (RGV) approaches that reportedly outperform traditional methods. Experimental results on dedicated IR dataset demonstrate the advantages of the newly proposed algorithm outperform the baseline BoW method and the recent RGV approach.

Introduction

Automatic target recognition (ATR), which is useful and important in many civilian and military applications, has always been a challenging problem in pattern recognition and machine learning for decades. The Infrared (IR) imagery usually has poor quality compared with the visible one and the appearance of IR target may change dramatically under different working conditions and ambient environment [1]. So IR ATR is a more challenging problem. Cognitive psychology experts believe that people are more sensitive to the shape of the target than the color and texture features. Due to the simple yet robust nature of shape feature, scholars have put forward many target recognition algorithm based on contour shape [2]. But the description of target contour feature is not complete in these algorithms. In order to solve the improper contour description, hierarchical description algorithm of contour description is proposed [3]. The representation method is rich in information, which is very strong to the description of the target contour, but the method is not strong enough to suppress the noise, and the recognition accuracy is not high. However, directly computing the similarity between descriptors is very time consuming in the algorithm. Recently, Khan et al. [4] develops a new relevance grouping of vocabulary (RGV) technique to address the issue and improve the image recognition efficiency. It was reported in [4] that the RGV-based algorithm outperforms others. This algorithm will be used as the major competing methods in our research. In this work, our research is motivated by recent computer vision advancements in image classification, contour extraction and image retrieval.

Related Works

In the section, we will introduce the idea of contour segment features extraction. Then we will present a new weighted BoW model to improve the traditional BoW-based classification. We also provide the pseudo code of the proposed algorithm.

Contour Segment Feature Extraction. In this work, we propose a contour segment feature descriptor based on the valuation scale of contour segmentation. The method firstly extracts the real

corner form the target contour by the curvature. And then the target contour is segmented into feature segments and non-feature segments according to the valuation scale. This segment features have obvious concave and convex characteristics and enhance the integrity and effectiveness of target contour description, and improves the target recognition rate.

Corner Extraction. Typeset sub-subheadings in medium face italic and capitalize the first letter of the first word only. Section numbers to be in roman.

The corner is the important feature of image contour. It is simplicity and robustness to noise, occlusion and variations of views and illumination [5]. Generally speaking, the corner is thought of as the curvature maxim points on the image contour curve. So it is a key question to calculate the curvature of contour points. In the traditional curvature calculation method, the curvature is the reciprocal of fitting circle radius [6]. The fitting circle is made up of the contour point and its two neighboring points. But the result may be easily corrupted by noise. Fully considering the distribution of near contour points, the curvature of contour point $K(i)$ is defined as follows [7],

$$K(i) = \frac{4\sqrt{\left(\left|P_i P_{i+k}\right| + \left|P_{i-k} P_i\right|\right)^2 - \left|P_{i-k} P_{i+k}\right|^2}}{\left(\left|P_i P_{i+k}\right| + \left|P_{i-k} P_i\right|\right)^2} \quad (1)$$

P_i is a point on the contour and P_{i-k} , P_{i+k} are two points on the contour that are K distance from P_i .

Because the distributions of the curvature values of contour points approximately conform to the law of normal distribution, we propose a sparse representation-based extraction method to choose the corners from the contour points. We can find the real corners P_i as follows:

$$P_i = \left\{ K(i) = \max_{|i-j|<k} K(j), K(j) < 2\sigma \right\} \quad (2)$$

$K(j)$ is one of the curvature values of the contour points which are extracted by the sparse representation-based method, and $K(i)$ is the curvature value of the real corner. We choose k as 5-7.

Contour Segment Features Extraction. From the perspective of human cognitive psychology, the concave and convex curve is more easily captured by human visual [8]. We can use the value scale to measure the concave and convex of the contour segments.

In our work, the segment features will be normalized. Firstly, each feature segment is expressed as an ordered set of points after equal interval sampling. And then each contour point will be transform from its coordinate to another one according to certain rules.

SVT For Weighted BoW-based Classification. Because the contour segments are highly overlapped, there are significant redundancies among the contour segment features. These features can be quantized in a feature space to form a collection of visual words. Thus a target contour can be represented by a BoW that encodes the occurrences of different visual words [9]. Each word contains some attributes pertaining to certain target type. To support scalable and fast ATR, we construct a weighted SVT via hierarchical K-means clustering of all contour segment features extracted from training images. In SVT algorithm, each cluster is recursively partitioned up to a certain level which forms a hierarchical vocabulary tree. Each leaf of the tree or lowest level cluster center represents a visual word. The traditional SVT-based classification treats each visual word independently and equivalently in the process of building vocabulary tree. However, this approach results in high computational complexity. In this work, we will consider samples with different weights. The value scales of contour segment features are calculated to obtain the weights $W = \{W_i\}, i = 1, 2, \dots, M$. All of the visual words $V = \{V_i\}, i = 1, 2, \dots, M$ and their corresponding weights are used in the weighted SVT construction.

For an image consisting of m descriptors $V = \{V_1, V_2, \dots, V_m\}$, each descriptor V_i is passed through the weighted vocabulary tree at each level to choose its closest clusters. The distance between a contour segment H_t in training image and a contour segment H_q in query image can be computed as eq. (3), where $C(p, q)$ is the Shape Context (SC) distance in [3].

$$D(H_t, H_q) = \frac{1}{n_1} \sum_{p \in H_t} \arg \min_{q \in H_q} C(p, q) + \frac{1}{n_2} \sum_{q \in H_q} \arg \min_{p \in H_t} C(p, q) \quad (4)$$

Where p a point of the contour is segment H_t , and q is a point of the contour segment H_q .

Experiments

In this section, we evaluate the proposed CS features weighting ATR algorithm by comparing it with the baseline BoW classification method and the RGV algorithm. We use the dedicated IR dataset provided by the Army Research Lab for algorithm evaluation, which involve 1884 training images and 123 query images for ten target types. Sample images are shown in Fig.1. The training images are high quality images taken under good condition. On the contrary, the query images are captured in different conditions, including illumination variation, partial occlusion, viewpoints and scale.

Although there is a tremendous redundancy in the CS features extraction, the weighted SVT can efficiently quantize them into a more compact set of visual words. We use different vocabulary sizes under different branch factors (K) and tree levels (L). Table 1 shows that the proposed method outperforms RGV methods in computing time. And the proposed method is able to achieve near perfect results for both recognition rate and computing time which outperforms the original BoW-based classification.

Table 1: Target recognition performance (%) and computing time(s) using the proposed method, RGV and baseline BoW-based methods at different vocabulary sizes.

Vocabulary Tree Parameters	Vocabulary Size	BoW	RGV	The proposed method
K=5 L=5	3,125	80.36	86.45	87.94
K=5 L=6	15,625	85.60	92.58	92.55
K=6 L=6	46,656	90.28	97.42	98.18
Average Computing Time		0.97854	1.15479	0.84215

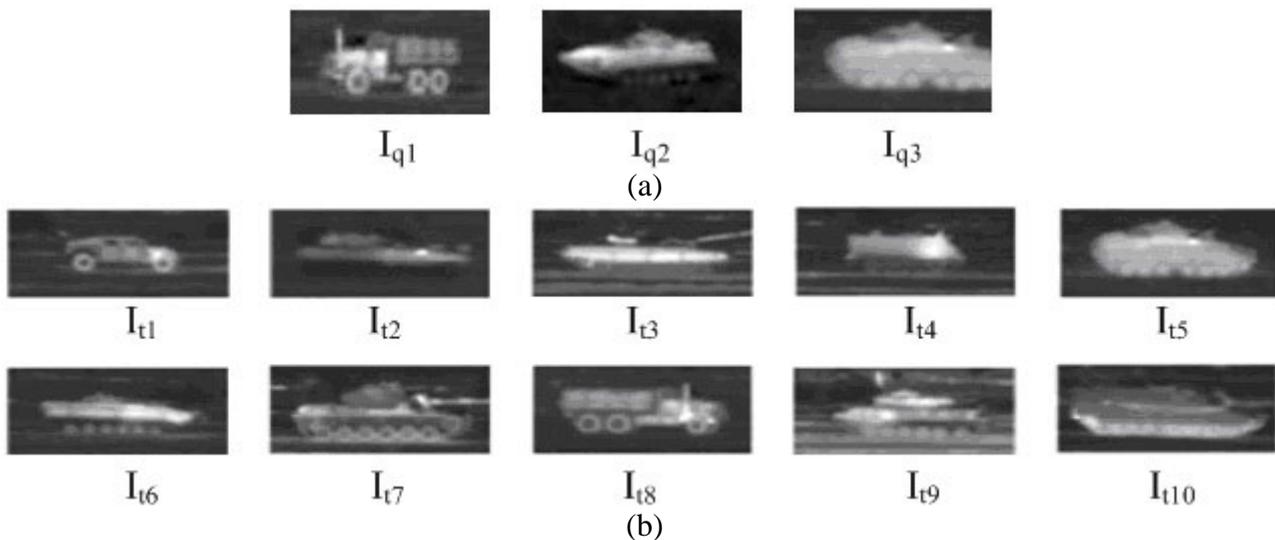


Figure 1: Ten different types training images and three different types query images from the dedicated IR dataset: (a) Query images; (b) Training image.

Conclusion

In this work, we study the challenging IR ATR problem. Our research is built on the recent advancements in the fields of image classification, contour extraction and image retrieval. First, it is shown that the CS feature is an effective approach to the ATR task where IR images usually have limited spatial resolution. Second, the weighted SVT is useful to reduce the redundancy of CS features by creating a compact set of visual words for scalable object recognition. Third, the weighted BoW recognition is efficient and robust in the context of IR ATR, although the spatial information of local features is totally ignored. Nevertheless, the proposed CS features weighting ATR algorithm further improves the ATR performance by boosting up the weights of coexistent visual words belonging to the same cluster.

Acknowledgements

This work was financially supported by National Natural Science Foundation of China (61401470).

References

- [1] Yilmaz A, Javed O, Shah M. Object tracking: A survey[J]. *Acm computing surveys (CSUR)*, 2006, 38(4): 13.
- [2] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2002, 24(4): 509-522.
- [3] Rodríguez-Sánchez A J, Tsotsos J K. The roles of endstopped and curvature tuned computations in a hierarchical representation of 2D shape[J]. *PLoS One*, 2012, 7(8): e42058.
- [4] Khan M N A, Fan G, Heisterkamp D R, et al. Automatic target recognition in infrared imagery using dense HOG features and relevance grouping of vocabulary[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2014: 293-298.
- [5] Bhat K S, Goldenthal R, Ye Y, et al. High fidelity facial animation capture and retargeting with contours[C]//*Proceedings of the 12th ACM SIGGRAPH/eurographics symposium on computer animation*. ACM, 2013: 7-14.
- [6] Prasicek G, Otto J C, Montgomery D R, et al. Multi-scale curvature for automated identification of glaciated mountain landscapes [J]. *Geomorphology*, 2014, 209: 53-65.
- [7] Petrakis E G M, Diplaros A, Milios E. Matching and retrieval of distorted and occluded shapes using dynamic programming[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(11): 1501-1516.
- [8] Xu X Z, Ding S F, Shi Z Z, et al. New theories and methods of image segmentation[J]. *Acta Electronica Sinica*, 2010, 38(2A): 76-82.
- [9] Wallach H M. Topic modeling: beyond bag-of-words[C]//*Proceedings of the 23rd international conference on Machine learning*. ACM, 2006: 977-984.