Signal Processing: Image Communication 47 (2016) 193–206

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

A multi-scaled hierarchical structure model for multispectral image detection

Yi Ma, Jin Han, Jiaqi Liu, Yi Zhang, Lianfa Bai *

Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Article history:
Received 4 November 2015
Received in revised form
8 March 2016
Accepted 21 June 2016
Available online 22 June 2016

Keywords:
Multispectral images
LARK
Composite template set
Hierarchical structure
Similarity

ABSTRACT

In view of the complex information and low recognition efficiency of multispectral images, a multi-scaled hierarchical structure model based on locally adaptive regression kernels is proposed. First of all, we construct the 3D feature descriptors by fully using the spatial structure and the abundant spectral information of the template set and test images, in order to expand the target detection to the multi-spectral area. Then a template set with multi-scale and multi-pose is put forward to ensure that targets with different sizes can be identified by the proposed statistical principle of similarity. Meanwhile, the hierarchical structure model is established to reduce the time of recognition and to increase the accuracy. Experiment results show that this algorithm can gain good recognition effect in both colorful images and multispectral images of near infrared band, and it is also conductive to the improvements of recognition accuracy and efficiency.

1. Introduction

Focused on the testing samples or the templates, general target detection for a class of specific targets plays a vital part in the computer artificial intelligence, and the high simultaneity is the ultimate goal in this region. In the past 20 years, category recognition (classification) has been the mainstream approach in this field, such as probabilistic constellation [1] and parts-and-shape model [2]. These kind of methods rely heavily on classifiers, and divide into supervised recognition [3,4] (Neural networks [5]), unsupervised recognition [6] and semi-supervised recognition [7,8]. Those traditional classifier models, however, require a lot of training samples, which may lead to slow learning process. What's more, phenomena of over-fit training parameters may easily occur. Therefore, new ways of target detection without training have been developed in recent years.

Comaniciu [9] proposed the target tracking algorithm based on kernels in 2003. Several years later, Takeda et al. [10] conducted the nuclear regression model to extract the characteristics of local structure, so as to restore and reconstruct the test image. Later, some similar non-training algorithms [11,12] were put forward. On these bases, Seo et al. [13–16] proposed a new way (named Seo algorithm below) grounded on the locally adaptive regression kernels (LARK), which can be applied to face detection [14,15], human gesture detection [13,16], etc. But the template Seo used is single, which means the detection results are overly dependent on the template (Fig. 1). A test image is processed with the same parameters, except the template. (b) and (c) show different results with different sizes of the template, and (c) and (d) form different results under different templates. And the seo algorithm concentrates on the integral structure, thus it is powerless for targets with multi-pose. To solve these problems, Luo [17] took the local structure into consideration, and proposed a new algorithm called local similar structure statistical matching (LSSSM). Targets are detected by counting the number of similar structures between test images and templates. LSSSM improves the results of target detection, but cannot reach a win-win pattern in both “efficiency” and “accuracy”. Even though Luo expands the number of templates, false detection still exists, especially on images with multi-scaled targets. And, the similarity statistic of whole test image is time-consuming. These problems exert a bad effect on accuracy and integrity of target recognition.

Recently, some scholars have demonstrated that integrating spectral features into research could improve the performance in many fields such as image fusion [18,19] and image classification [20]. And Yuan et al. [21] exploit multi-spectral information to improve the performance of pedestrian detection. It follows that the multi-spectral information is a new view to learn the world more effectively. LSSSM, like Seo algorithm, just focuses on gray images, so the important spectral information of each image is lost. Take the plaster model as an example, it has the same spatial structure and different spectral information as the face. Thus, according to the similarity of spatial structure alone, it will form the recognition effect shown in Fig. 1. Due to the lack of spectral

* Corresponding author.
E-mail address: blf@njust.edu.cn (L. Bai).

http://dx.doi.org/10.1016/j.image.2016.06.010
0923-5965/

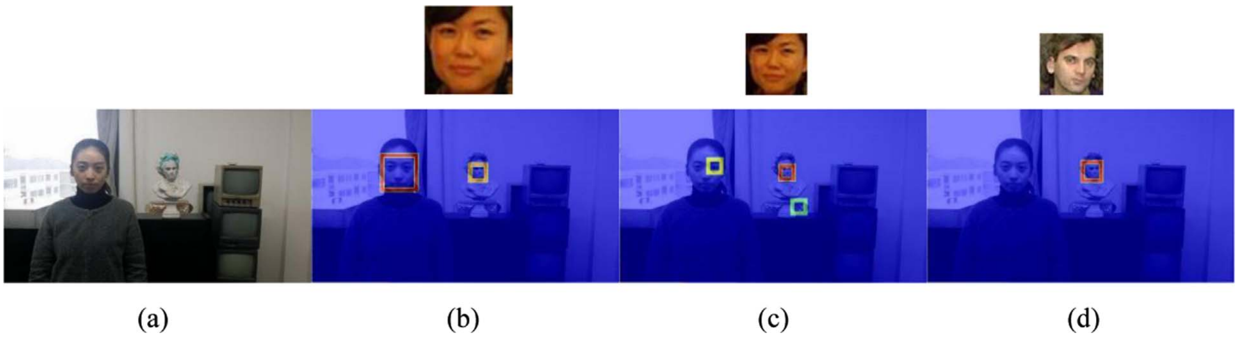

Fig. 1. The results of Seo algorithm: (a) the original image; (b), (c), (d) different templates and the corresponding identification results.

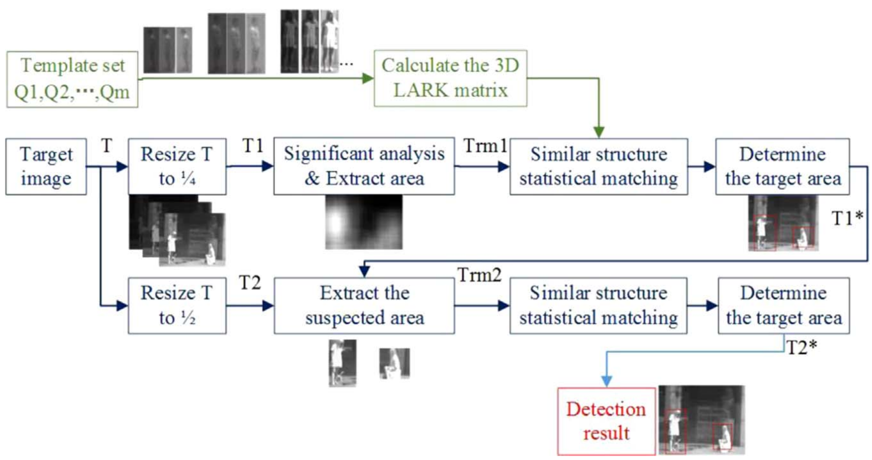

Fig. 2. The flow chart of the hierarchical structure model based on the 3D LARK.

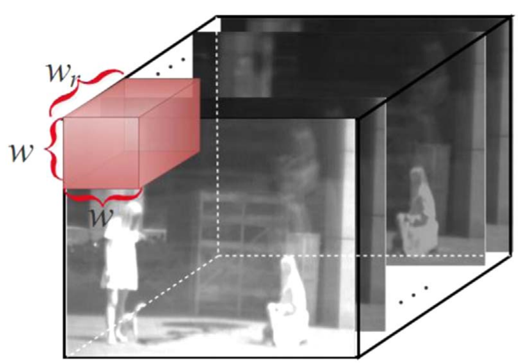

Fig. 3. The sketch of local window.

information, all of these three detection results regard the plaster model as the face.

In this paper, we propose a multi-scaled hierarchical structure model based on the 3D space and spectrum LARK. This algorithm has three features: multispectral image detection, multi-scaled template set and the hierarchical structure model. Firstly, through the plentiful spectral information in the template set and the test images, the 3D feature descriptors are constructed and the target detection is expanded to the multispectral area. Then, the composite template set ensures that targets with different sizes can be identified. Finally, with the hierarchical structure model, the recognition efficiency can be improved, accuracy as well.

2. Locally adaptive regression kernel

In Section 1, we have mentioned the LARK feature proposed by Seo. To make the concepts clear, the calculation of LARK is briefly described here.

In classic kernel regression algorithm, characteristics of images are extracted through kernel regression [10],

$$y_i = z(x_i) + \varepsilon_i,\ i{=}1,\ 2,\ \ldots,N, \tag{1}$$

where $z(\bullet)$ is the regression function, ε_i is the independent and identically distributed zero mean noise value, and N is the number

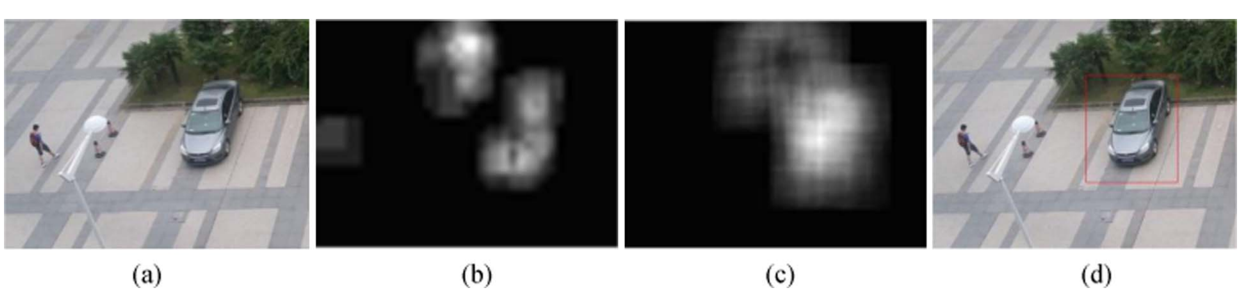

Fig. 4. (a) The original image; (b) R_m of the first layer; (c) R_m of the second layer; (d) detection result.