Signal Processing: Image Communication 68 (2018) 58–67

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Multispectral target detection based on the space–spectrum structure constraint with the multi-scale hierarchical model☆

Zhuang Zhao [1], Yuwei Zhang [2], Lianfa Bai [3], Yi Zhang [4], Jing Han *,[5]

XiaoLingWei Street 200, Nanjing, China

ARTICLE INFO

MSC:
00-01
99-00

Keywords:
Multispectral target detection
space–spectrum Differential Structure
Local and Neighbor Binary Pattern
Multiscale Hierarchical model

ABSTRACT

Aimed at the interference of scene and the local overly split, we fully excavate the information of spatial and spectral domain to construct the feature descriptor called local and neighbor multiscale, and construct a hierarchical structure model for multispectral target detection. Based on multidimensional differential distance measure, we put forward Space and Spectrum Differential Structure (SSDS) operator to extract small scale fine structure and robust feature of multispectral targets. Based on the idea of binary patterns, we propose Local and Neighbor Binary Pattern (LNBP) operator for large scale neighborhood structure extraction of multispectral targets. Finally, we construct pyramid hierarchical structure model (MH-LS, Multi-scale Hierarchical-LS), design a multiscale and multilevel computing architecture, and fully unite the characterization and separability of two operators on the multidimensional multiscale structure. Also, MH-LS can make robust matching based on small sample coming true and improve the detection accuracy and efficiency. Experiments show that MH-LS does not need a large number of sample training, and can effectively detect multispectral objects in different scenes, postures, views and scales based on a small number of template sets.

1. Introduction

Target detection methods can be roughly divided into two types: one is based on background modeling. For example, through background modeling, Sobral et al. [1] and Benezeth et al. [2] used the training-based method to detect multispectral moving targets. The other one is the foreground target extraction. And it can be divided into two parts: target recognition [3] and target detection [4–8]. The former is to classify a given object into one of the several given categories, while the latter is to separate the interested object from the background in a test image. The traditional classifier models rely heavily on the classifiers and require a large number of training samples, which may cause a slow learning process. And the phenomena of over-fit training parameters may easily occur. However, many application scenarios can only obtain a small number of classification samples, and how to obtain good classification and detection results in small samples is a difficult problem for target detection. In recent years, many scholars have proposed non-training detection algorithms [9,10]. In this class of non-training methods the object analysis is mainly based on the well represented image descriptors, such as local self-similarity [11,12], SIFT [13] and more, and some optimized matching schemes. Their advantage is that there is no need for a large number of sample learning, and have certain robustness and generalization ability for both rigid and non-rigid targets. In addition to these non-training methods, there are many multispectral or hyperspectral target detection models which is based on training [14–16], and they are mainly used for pixel-level small targets detection in the application scenario of remote sensing.

In 2003, Comaniciu et al. proposed the kernel as a descriptor for target tracking [9]. In 2007, Takeda et al. employed the classic kernel regression to recover the high frequency information of image, so as to restore and reconstruct the test image [10]. In 2009, Milanfar et al. studied the adaptive kernel regression method to remove noise, enhance image details and target detection [11]. Seo et al. made further efforts and proposed the Locally Adaptive Regression Kernel method (LARK),

☆ This document is a collaborative effort.

* Corresponding author.
E-mail addresses: zhaozhuang3126@gmail.com (Z. Zhao), 1836773123@qq.com (Y. Zhang), mrblf@163.com (L. Bai), zhy441@sina.com (Y. Zhang), eohj@njust.edu.cn (J. Han).

[1] Since 1990.
[2] Since 1993.
[3] Since 1965.
[4] Since 1966.
[5] Since 1987.

https://doi.org/10.1016/j.image.2018.06.014
Received 31 January 2018; Received in revised form 7 May 2018; Accepted 23 June 2018
Available online 4 July 2018

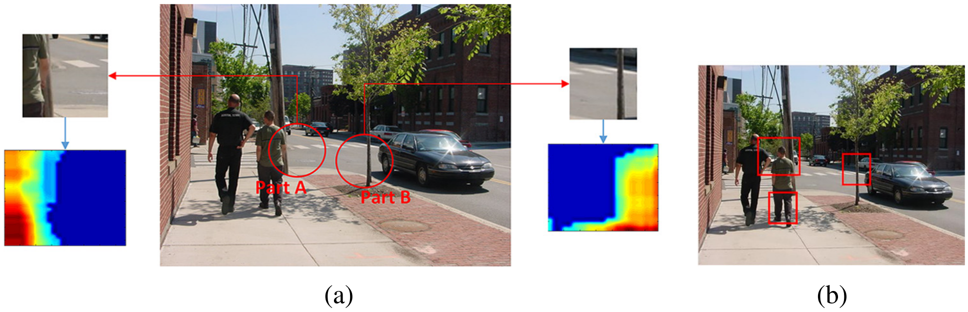

Fig. 1. (a) Comparison between different Image patches; (b) Detection result of the method [18]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

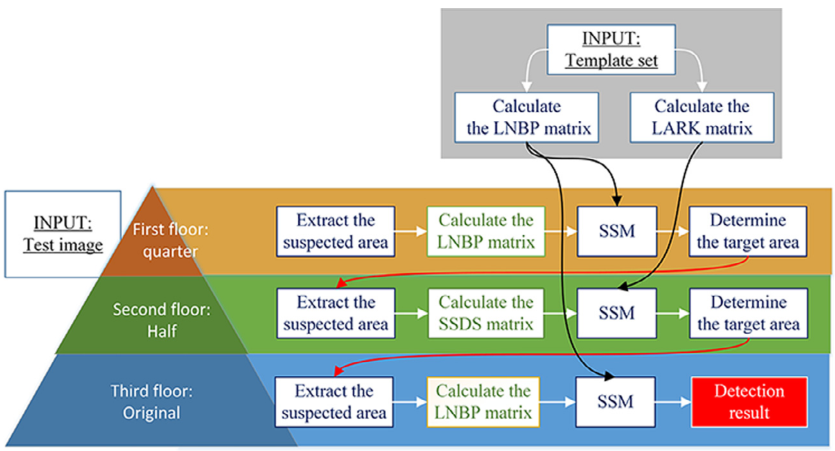

Fig. 2. Target detection model based on local and neighborhood structure feature constraints.

which is a non-parameter method based on a single template and can be applied to face detection [5,6], human gesture detection [4,7], etc. In 2015, Luo et al. took the local structure into consideration, instead of the integral structure, and proposed a LSSSM algorithm for non-compact targets with multi-pose [8]. The both two algorithms only focus on gray images, so the important spectral information of images is lost. In [17], RGB color information is considered, but the Liu method is time-consuming.

Besides, LSSSM, Seo and Liu methods are all aimed at image patch (the relationship between center pixel and the surrounding pixels), and ignore the structure relationship among image patches. Take Fig. 1 for example, both of the human's arm and the tree branch have high similarity with the query (Red represents higher resemblance), but they have different surrounding structures. So the false detection rate may be high if we just pay attention to the internal structure of image patch and neglect the relationship between patches. Meanwhile, the overall matching of the template can also lead to the leakage of walking change attitude. The Neighbor Structure Reconstruction Matching algorithm(NSRM) is proposed by Xue, although the neighborhood structure is introduced based on nonnegative linear reconstruction [18]. However, most natural date is non-linear, and the neighborhood reconstruction has no inclusion relation with LARK on window scale in NSRM.

How to adapt to the multi-scale matching of local and overall structure, the realization of "robustness" and "efficiency" are the core difficulties of non-learning methods. We propose a pyramid hierarchical structure model (MH-LS Multi-scale Hierarchical-LS) based on LNBP (Local and Neighbor Binary Pattern) and SSDS(Space and Spectrum Differential Structure). Fig. 2 shows an overview of the model. It has two feature descriptors LNBP (Local and Neighbor Binary Pattern) and SSDS (Space and Spectrum Differential Structure), which combine the internal structure feature of image patch and the relationship between patches to achieve multi-scale matching effect, so as to reduce the interference of background noise and solve the problem of local feature overly splitting. In addition, both the spatial and spectral information is taken into account to improve the accuracy. Moreover, the template set is expanded from the view of multi-scale and multi-pose. Finally, we design a pyramid hierarchical structure model to conduct the alternating detection with two feature descriptors, which can further increase the efficiency and improve the accuracy.

2. Pyramid hierarchical structure model (MH-LS) based on LNBP and SSDS

The MH-LS model consists of two feature operators, Local Neighbor Binary Pattern (LNBP) and Space-Spectral Differential Structure (SSDS), which can realize multi-scale and multi-criterion image robust structure. Take the two-dimensional image in Fig. 3 for example, first set a local window (5×5) as the center window and then expand its eight neighbors. Calculate the SSDS feature matrix of each local window and the LNBP matrix of the neighboring windows. Then, the features inside and among the local windows can be combined, which can improve the detection accuracy, as is described in Fig. 3. At the same time, the pyramid hierarchical structure model is constructed to conduct the detection with the alternate use of the two feature descriptors, as in Fig. 2.

2.1. Space and Spectrum Differential Structure (SSDS) descriptor based on 3D LARK

Classic kernel regression algorithm extracts the relevant features of the image by kernel regression. Based on the pixels' difference between