Pattern Recognition 47 (2014) 3006–3018

Contents lists available at ScienceDirect

Pattern Recognition

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

Scale-invariant template matching using histogram of dominant gradients

Jisung Yoo [a,*], Sung Soo Hwang [a], Seong Dae Kim [a], Myung Seok Ki [b], Jihun Cha [b]

[a] Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Guseong-dong, Yuseong-gu, Daejeon, Republic of Korea
[b] Broadcasting & Telecommunications Convergence Media Research Department, Electronics and Telecommunications Research Institute (ETRI), Gajeongno, Yuseong-gu, Daejeon, Republic of Korea

ARTICLE INFO

Article history:
Received 3 January 2013
Received in revised form
22 January 2014
Accepted 26 February 2014
Available online 21 March 2014

Keywords:
Template matching
Dominant gradient enhancement
Scaling-tolerant histogram vector

ABSTRACT

This paper presents a histogram-based template matching method that copes with the large scale difference between target and template images. Most of the previous template matching methods are sensitive to the scale difference between target and template images because the features extracted from the images are changed according to the scale of the images. To overcome this limitation, we introduce the concept of dominant gradients and describe an image as the feature that is tolerant to scale changes. To this end, we first extract the dominant gradients of a template image and represent the template image as the grids of histograms of the dominant gradients. Then, the arbitrary regions of a target image with various locations and scales are matched with the template image via histogram matching. Experimental results show that the proposed method is more robust to scale difference than previous template matching techniques.

1. Introduction

Template matching is a well-known problem in the field of computer vision. Once a template image and a target image are given, it determines the matched region of the target image that is similar to the template image. Template matching has a wide application area such as object detection, video compression, automatic inspection and so on.

Most previous studies have focused on the template matching algorithms invariant to rotation and illumination changes. Normalized cross correlation (NCC)-based method is one of the most popular methods for illumination-invariant matching, and various schemes for efficient computation of NCC have been proposed [1–4]. For rotation-invariance, Ref. [5] introduced the concept of the orientation codes similar to the histogram of oriented gradients (HOG) and Ref. [6] exploited the Fourier coefficients of radial projections. Recently several algorithms invariant to both rotation and illumination changes have been proposed using circular projection and Zernike moments [7], ring-projection transform [8], and NCC combined with Fourier–Mellin Transform [9].

On the other hand, template matching methods that are robust to large scale difference have rarely been considered. As image capturing technologies have been developed extensively, users have been easily able to acquire high-resolution images and to control their resolution. For example, one may think of a multi-camera surveillance system that uses a camera for monitoring wide area while details of a specific region are monitored by another camera. (the demo video for our implementation is in Refs. [16,17]). As a result, the need for developing template matching method for images with large scale difference has increased. While template matching methods that are rotation- and illumination-invariant have been widely studied, only a few studies have been conducted on scale-invariance. And even though several proposed algorithms are scale-invariant [10,11], they are unable to handle large scale difference between template and target images. Local descriptor-based scale-invariant image matching methods, such as SIFT, could be a solution to matching images with scale difference [12,18]. However, as shown in Fig. 1, even local descriptor-based image matching can be inaccurate when a template image and a target image have large scale difference. Moreover, the SIFT-based matching can suffer from a lack of local feature points if the image is very small, and its performance can vary according to the content of images.

In this paper, we propose a template matching method that is robust to large scale difference between template image and target images. To ensure scale-invariance, *dominant gradients*, which appear consistently regardless of scale changes of an image, are

* Correspondence to: Korea Advanced Institute of Science and Technology, Room #3107, LG Semicon Hall, Guseong-dong, Yuseong-gu, Daejeon, Republic of Korea. Tel.: +82 42 350 5430; fax: +82 42 350 8030.
E-mail addresses: yjs1018@kaist.ac.kr (J. Yoo), dreamerjoe@kaist.ac.kr (S.S. Hwang), sdkim@kaist.ac.kr (S.D. Kim), serdong@etri.re.kr (M.S. Ki), jihun@etri.re.kr (J. Cha).

http://dx.doi.org/10.1016/j.patcog.2014.02.016
0031-3203/

Fig. 1. An example of failure cases of local descriptor-based image matching. The left image is a template image and the right image is a target image. Scale of the template image is 6 times larger than one of the target image. As shown in Fig. 1-(a), almost matched points between these two images are invalid pairs. Fig. 1-(b) shows that our proposed method matches these two images successfully despite of the large scale difference between them. (a) The result of SIFT-based image matching and (b) the result of the proposed method.

extracted from each image. We analyze the structure of a template image using its dominant gradients and represent the image as 2D grids of histograms of dominant gradients (HDGs). Sub-images of a target image with various locations and scales are also represented as the HDG grids, and each sub-image is matched with the template image via histogram matching. As our algorithm performs a repetitive histogram matching procedure that requires a large amount of computation, the integral histogram [13] is used to speed up the procedure.

The remaining of this paper is organized as follows: in Section 2, we first briefly describe the related works of our research. In Section 3, we describe our proposed algorithm for scale-invariant template matching. Experimental results and performance evaluations are shown in Section 4, and we conclude in Section 5.

2. Related works

The template matching process involves cross-correlating a template image with a target image using a sliding-window method based on various similarity measures. As basic similarity measures, the sum of squared or absolute differences (SSD/SAD) and NCC have been widely used. SSD/SAD uses pixel-by-pixel differences of intensities between two windows. Hence, SSD/SAD has a low computational cost but is sensitive to illumination changes. Since NCC computes the normalized cross correlation of intensity values between two windows, it is tolerant to illumination changes but has a high computational cost. To overcome this limitation of NCC, various studies on accelerating the computation of NCC have been proposed [1–4]. In Ref. [1], NCC is computed rapidly from unnormalized cross correlation using integral images. Ref. [2] combined adaptive multilevel partition with the winner update scheme to skip unnecessary NCC calculations. Refs. [3,4] also proposed the skipping algorithms that reduce the search space by measuring the transitivity of correlation and employing a coarse-to-fine scheme, respectively. However, a simple sliding-window cross-correlation method is not invariant to rotation and scaling of the template image.

In order to ensure the rotation invariance of template matching, many algorithms have been proposed [5–9]. Refs. [5,7] utilized the histogram of oriented gradients (HOG) and the Zernike moments for estimating the rotation angle of a template image respectively. In Ref. [6], complex coefficients of the discrete Fourier transform of the radial projections are utilized as new rotation-invariant local features. Ref. [8] used the ring-projection transform process to convert a 2D template image into a rotation-invariant 1D vector, and Ref. [9] combined NCC and phase correlation based on the Fourier–Mellin transform to compute the rotation and scaling factors of a template image.

Recently several template matching methods were proposed that are robust to scale difference. In Ref. [10], by replacing image functions with complex gray-level edge maps, log-polar Fourier representations are derived to estimate the scaling and rotation of images. They also introduced the normalized gradient correlation for robust matching. In Ref. [11], a set of coefficients in the wavelet transform domain are selected as features of a template image first and a new adaptive transform is designed to eliminate the selected features. By using both non-adaptive and adaptive transforms, the matched region of a target image that is the most similar to the template image is determined. Even though their methods work properly under small scale difference between template and target images, they could not deal with the large scale difference of over four times.