Pattern Recognition Letters 94 (2017) 46–54

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A feature descriptor based on the local patch clustering distribution for illumination-robust image matching

Han Wang[a,d], Sang Min Yoon[b], David K. Han[c], Hanseok Ko[d,*]

[a] School of Transportation, Nantong University, No.9 Seyuan Road, Nantong City, 226019, China
[b] HCI Lab. School of Computer Science, Kookmin University, 77 Jeongneng-ro, Sungbuk-gu, Seoul 02707, Republic of Korea
[c] Office of Naval Research, Arlington, VA, 22203 USA
[d] School of Electrical Engineering, Korea University, 145 Anam-ro, Sungbuk-gu, Seoul, 02841, Republic of Korea

A R T I C L E I N F O

Article history:
Received 21 December 2015
Available online 12 May 2017

Keywords:
Local patch clustering distribution
Feature descriptor illumination change,
Image matching

A B S T R A C T

This paper proposes a feature descriptor based on the local patch clustering distribution (LPCD), which preserves the salient features of a given image following changes in illumination. To mitigate the effects of illumination change, the proposed LPCD methodology consists of two steps. First, a local patch clustering assignment map is constructed by pairing the source image with a reference image. To resolve the quantization problem caused by an illumination change, a dual-codebook clustering method is employed so that an effective local patch clustering feature space can be constructed. Second, in the feature encoding process, the impact of the informative local patches that contain textural information is enhanced when using a saliency detection response as a method of weighting every local patch when the histogram feature is extracted. Experimental results show that the proposed local patch clustering space is more robust than the conventional intensity order-based space in response to changes in illumination.

1. Introduction

The goal when building a novel feature descriptor is to develop an effective feature extraction method that is robust to various image transformations or environmental changes. Many local feature descriptors have been successfully used in computer vision to achieve robust performance for different applications. For example, the Harris corner detector [1] and the difference of Gaussian (DoG) detector [2] are adopted for the detection of interesting points, while the Harris–Affine [3] and Hessian–Affine [4] detectors are employed for the detection of affine covariant regions. Other local feature descriptors such as the scale invariant feature transform (SIFT) [5], local binary pattern (LBP) [16], and similar strategies [6–9] are widely applied for object detection [10], object tracking [11,12], and object recognition [13].

Previous successful applications have proven that the local feature-based pattern recognition methods are very effective at mitigating the geometric variation and distortion. However, a lingering issue in computer vision is that how to effectively handle complex changes in illumination, a common phenomenon in the real world. Therefore, illumination-robust feature descriptors have had a very important effect on the performance of many computer vision applications. In the past 5 years, local features based on the intensity order rather than on the raw intensity have become more popular for illumination-robust image matching because the distribution of intensity order or the local relationship between pixels is invariant to monotonic changes in the pixels. Tang et al. [14] proposed the ordinal spatial intensity distribution (OSID), in which the relative order of pixel intensity in an entire image patch is used to generate illumination-robust features. However, ordering pixels by discrete intensity yields significantly different order distributions when the illumination changes, thereby leading to changes in the nonlinear intensity. To address this problem, an exact order method was used in the feature descriptor EOD (Exact Order based Descriptor) by Kim et al. [15].

In addition to the global intensity order distribution of the image patches, the local relationship between the pixels has also been employed as a design feature descriptor. LBP was proposed by Ojala et al. [16]. To reduce the dimensionality of the data, Heikkila et al. [17] proposed the center-symmetric LBP (CS-LBP). They compared the center-symmetric pairs of pixels instead of comparing all the pixels neighboring the center pixel. Thus, the original LBP-based 256-dimensional feature was reduced to 16 dimensions. Similarly, the center-symmetric local ternary pattern (CS-LTP) was proposed by Gupta et al. [18]. In this strategy, a 9-dimensional feature was generated by comparing only the diagonal pairs of a given pixel, thereby creating a binary code. In contrast to LBP-like fea-

* Corresponding author.
E-mail addresses: hsko@korea.ac.kr, hanseok.ko.34@gmail.com (H. Ko).

http://dx.doi.org/10.1016/j.patrec.2017.05.010
0167-8655/

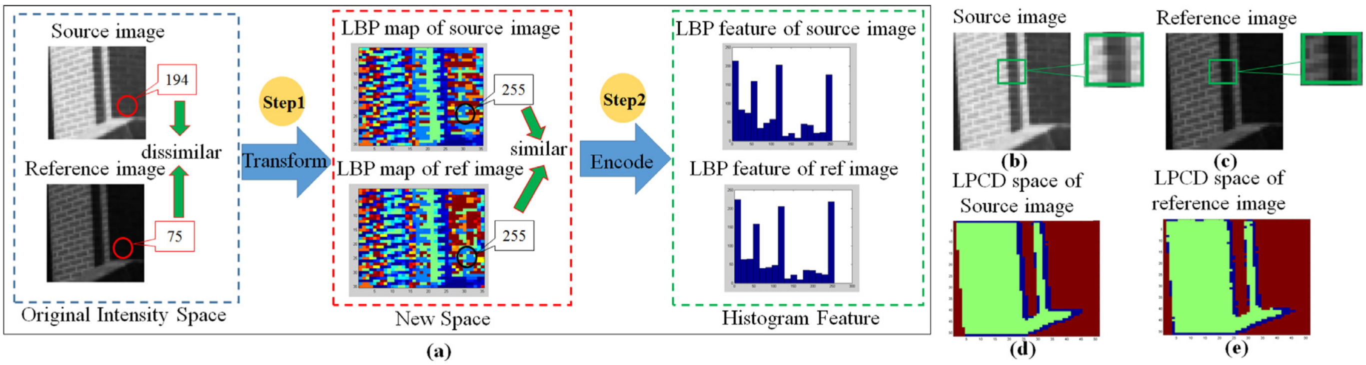

Fig. 1. Illumination-robust feature descriptor extraction process. (a) Establishment of a general feature descriptor. (b) Source image patch. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

tures [16–18], Wang et al. [19] adopted a nondecreasing method to encode different permutations of pixel intensity values and proposed the local intensity order pattern (LIOP). In this case, a spatial binning method based on the global intensity ordinal information led to a rotation-invariant performance.

As evident from the intensity order-based local feature descriptors listed above, the feature descriptor extraction process can be divided into two steps as depicted in Fig. 1(a). First, the raw intensity of the pixels is transformed into a new space by imposing certain rules [14,16–19]. A histogram feature is then extracted from the transformed space. For example, the image intensity of a pixel changes from 195 to 75 following a change in illumination as shown in Fig. 1(a), while in the transformed space (LBP space [16]), the identical feature value of 255 is generated when the illumination changes. This example illustrates that the new transformed space can be less sensitive to changes in illumination than the raw intensity space alone. Therefore, to build an illumination-robust feature descriptor, an effective approach confining the feature values of the transformed space in response to changes in illumination needs to be developed.

In this paper, unlike previous pixel operation-based methods, a feature descriptor based on the local patch clustering distribution (LPCD) is proposed. A dual codebook-based clustering method is then applied to resolve incorrect quantization caused by illumination change. In the encoding process, to enhance the impact of informative patches, a weight value based on a saliency detection response is assigned to every local patch when the histogram feature is extracted.

The rest of this paper is organized as follows. Section 2 describes the details of the proposed feature extraction procedure. The experimental results are presented in Section 3. Finally, we provide the concluding remarks in Section 4.

2. A descriptor based on LPCD

The basic idea of the proposed LPCD descriptor is inspired by the fact that the internal geometric layout of an image does not vary when the intensity changes [20]. This is illustrated by the two small green rectangle patches in Fig. 1(b) and (c). Although the intensity values are obviously different when the illumination changes, the internal geometric layout within these two local patches remains invariant while the local structure information is always represented by a local patch. Therefore, we can conclude that the results of the assignment of local patch clustering do not vary with changes in illumination. Fig. 1(d) and (e) presents the local patch clustering for Fig. 1(b) and (c), respectively, where the patch class number for the local patch is 3. This example clearly supports our proposition.

2.1. Generation of LPCD descriptor space

The effectiveness of an illumination-robust feature descriptor depends on how the new feature space is constructed after the illumination changes. Codebook-based quantization is a common approach for creating a local patch clustering assignment map,. However, incorrect quantization may result when a source image codebook is used to quantize the reference image patches in the generation of a clustering assignment map because the intensities of the source and reference images differ due to the change in illumination.

To resolve this quantization problem, a dual codebook-based clustering method that follows five steps is proposed (Fig. 2). In *Step 1*, a source and a reference local feature vector sets are extracted using a sliding window under the same spatial order using all the local patches in each image:

$$F_{Sour} = \{P_1, P_2, \ldots, P_N\} \tag{1}$$

$$F_{Ref} = \{Q_1, Q_2, \ldots, Q_N\}, \tag{2}$$

where F_{Sour} and F_{Ref} are the local feature vector sets of the source and reference images, respectively. P_i and Q_i are the intensity feature vectors of the ith corresponding local image patches from the source and reference images, respectively.

In *Step 2* and *Step 3*, the local feature vector set from the source image is clustered by the K-means algorithm [21] to extract the clustering centers as the codebook for the source image Θ_{Sour}.

$$\Theta_{Sour} = \left\{P_{J_1}, P_{J_2}, \ldots, P_{J_m}\right\}. \tag{3}$$

Because the source and reference feature vector sets are extracted using the same spatial order, the spatial order information from the source image codebook can be directly shared with the reference image in *Step 4*. Thus, the codebook feature vector of the reference image is extracted according to the corresponding spatial order of the source image codebook in *Step 5* as illustrated in Fig. 2.

$$\Theta_{Ref} = \left\{Q_{K_1}, Q_{K_2}, \ldots, Q_{K_m}\right\} \tag{4}$$

$$K_n = J_n \ (n = 1, 2, \ldots, m), \tag{5}$$

where Θ_{Sour} and Θ_{Ref} are the dual codebooks extracted from the source and reference local feature vector sets described by (1) and (2), respectively. J_n is the codebook feature vector index of the source image that contains the positional information of the corresponding local patches in the source local feature set. The relationship established by (5) depicts that the source and reference codebooks share the same clustering index. Therefore, although the codebooks P_i and Q_i have different intensities, they represent the