Pattern Recognition Letters 98 (2017) 123–129

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Lattice-Support repetitive local feature detection for visual search

Dipu Manandhar*, Kim-Hui Yap, Zhenwei Miao, Lap-Pui Chau

School of Electrical and Electronics Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

ARTICLE INFO

Article history:
Received 30 August 2016
Available online 12 September 2017

Keywords:
Repetitive pattern detection
Visual burstiness
Image search and retrieval

ABSTRACT

Repetitive patterns such as building facades, floor tiles, vegetation, and wallpapers are commonly found in sceneries and images. The presence of such repetitive patterns in images often leads to visual burstiness and geometric ambiguity, which poses challenge for state-of-the-art visual search technologies. To alleviate these problems, we propose a new lattice-support repetitive local feature detection method to detect repetitive patterns, estimate the underlying lattice structure, and enhance descriptors used for subsequent visual image search. Existing methods for repetitive pattern detection are commonly based on determining the underlying lattice structures. However, these structures do not correspond directly to robust features that are scale- and rotation-invariant. This paper proposes a new lattice-support repetitive local feature (LS-RLF) detection method that aims to integrate lattice information into repeated local feature detection and extraction. The advantage of the proposed method is that the detected features can be directly used by current visual search technologies. The LS-RLF method estimates the undetected repeated features in the lattice structure using Hough transform-based feature estimation. Further, in order to handle the visual burstiness issue, a new LS-RLF based image retrieval framework is developed. Experiments performed on benchmark datasets show that the proposed method outperforms the state-of-the-art methods by mean Average Precisions (mAP) of 4.5%, 5.5% and 3.2% on Oxford, Paris, and INRIA holidays datasets respectively. This demonstrates the effectiveness of the proposed method in performing visual search for images which contain wide range of repeated patterns.

1. Introduction

The presence of repetitive patterns provides pleasing aesthetic to image, and are commonly found in many images. Repetitive patterns are frequently encountered in images from natural scene as well as man-made structures. For example, in the city landscape, images often contain repetitive patterns/structures such as building facades, windows, tiles, etc. Repeated structures can also present in background objects such as trees and bushes, which result in negative impact on the performance of visual recognition [15]. Sample images that contain repetitive patterns are shown in Fig. 1. In general, repetitive patterns are often considered as a challenge that hinders visual search and recognition in computer vision.

In recent year, local feature-based image representations such as Bag-of-Words (BoW) [8,21,22,25] have been widely used in image search systems. The BoW model finds wide applications in visual recognition tasks such as place and landmark recognition [2,3,26]. The model represents the local feature as discrete visual words (VWs), and then represents an image as a histogram of visual words. The visual words are computed by using unsupervised algorithms such as k-means clustering in high dimensional feature space. Variants of k-means like Approximate *K*-Means (AKM) [21] and Hierarchical *K*-Means (HKM) [19] have been proposed to achieve efficient computation of visual vocabulary. Eventually, local features from the database images are quantized using the visual vocabulary and indexed using inverted file structure. To retrieve similar images for a query image, database images are scored and ranked using weighting functions such as *term frequency-inverse document frequency* (tf-idf). To further enhance the performance, post-processing strategies such as geometric verification [21,25] and query expansion [5,6] are used to re-rank the images.

Nevertheless, the similarity scoring and geometric verification in the BoW framework are adversely affected by the presence of repetitive patterns in images. The repetitive patterns pose significant challenges to image matching due to the following observations:

(i) ***Violation of Feature Independence Assumption*** **:** The repeated patterns in images violate the i.i.d. (independent and identically distribution) assumption required by state-of-the-art algorithms in image matching, retrieval and reconstruction [7,13,24].

* Corresponding author.
E-mail addresses: dipu002@e.ntu.edu.sg, dips4717@gmail.com (D. Manandhar), ekhyap@ntu.edu.sg (K.-H. Yap), zwmiao@ntu.edu.sg (Z. Miao), elpchau@ntu.edu.sg (L.-P. Chau).

http://dx.doi.org/10.1016/j.patrec.2017.09.021

Fig. 1. Sample images with different types of repetitive patterns in man-made structures and natural sceneries.

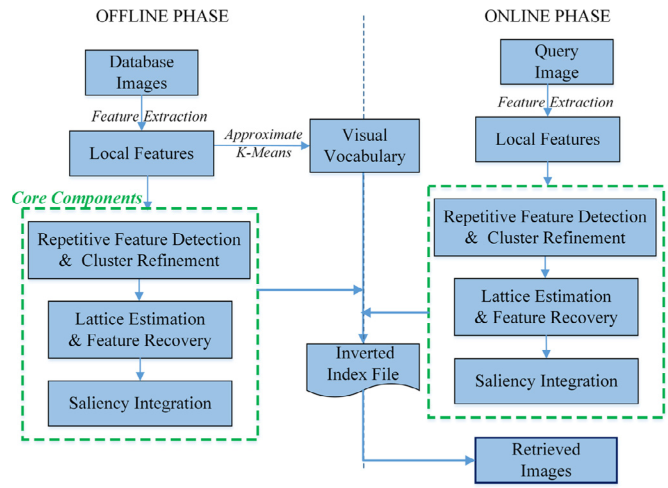

Fig. 2. Overview of the proposed LS-RLF framework in visual search.

Hence, the presence of repetitive patterns results in inferior retrieval performance of the algorithms.

(ii) ***Burstiness of Visual Words*** **:** Features from the same repetitive patterns are likely to be quantized to the same visual words, thus causing visual words from these repeated patterns to overwhelm the non-repeated ones. This phenomenon is known as visual burstiness [13]. The similarity score is hence dominated by visual words with large bin counts that correspond to repetitive features. The burstiness of visual words will often lead to incorrect retrievals and degrades performance of visual search systems.

(iii) ***Ambiguity for Geometric Matching*** **:** The spatial verification of images are performed on the putative correspondences established during visual word matching. Hence, multiple matches of features from the same repetitive patterns will cause ambiguity in determining the true correspondences during geometric verification.

To handle the above-mentioned challenges, detection of repeated patterns and incorporation of this information into the retrieval framework is needed. In recent years, a number of repetitive pattern detection methods have been studied. Leung et al.[16] used local patches to find similar elements in the neighborhood. Park et al. [20] and Hays et al. [11] modeled the repetitive patterns with deformed 2D lattice. They used normalized cross correlation (NCC) on lattice template and vectors to find the dominant lattice structure in the image. However, these lattice-based detection methods [11,16,20] are not able to handle different types of repetitive patterns. Further, not all the repeated patterns in images can be modeled using quadrilateral lattice elements. Most importantly, the detected lattice elements do not correspond to discriminative features that can be used to characterize the image for visual search/recognition directly.

In image search and retrieval, repetitive building facades have been studied in [23] by measuring the similarity of motifs using NCC. Similarly, Doubek et al. [9] employed shift invariant descriptor for lattice representation. However, both Doubek et al. [9] and Schindler et al. [23] can only handle small dataset. This is because their methods do not produce discriminative features that are not scalable with respect to the dataset dimension.

For scalable retrieval, the study on the effect of repetitive patterns has been done in [13]. The authors different weighting functions such as square-root operator $\sqrt{(\cdot)}$, and $\log(\cdot)$ to down-weigh the contribution of repetitive features. However, the method does not exploit repetitive pattern detection. This leads to sub-optimal solution to the burstiness issue. Recent work on repttile method [26] used local features to detect the presence of repetitive patterns. The method relies on the top K nearest visual words to measure the feature similarity, and hence suffers from feature quantization errors. Moreover, Torii's repttile method only makes coarse estimation of repetitive patterns in images without utilizing the underlying lattice structure.

In view of this, this paper presents a new Lattice-Support Repetitive Local Feature (LS-RLF) detection method to detect repetitive patterns, estimate the underlying lattice structure, and extract descriptors that can used directly for visual image search. By using both geometric and descriptor information of local features, the proposed LS-RLF method is able to accurately detect different types of repetitive patterns at various scales. An advantage of this approach is that since the detected local features are invariant to scale change and rotation, the detected pattern can be used directly for image search and matching. Moreover, as the local descriptors can handle a certain amount of view, geometric and photometric variations, the proposed method is more robust than lattice patch approaches [9,11,20,23]. As compared to [26], the proposed LS-RLF method integrates both feature information and underlying lattice structure to estimate more accurate repetitive patterns. Undetected features in the patterns can also be recovered using Hough transform-based interpolation. The detected patterns can be readily used to address the visual burstiness issue. Since the bursty features from the image background such as trees, bushes, flora, etc. have negative impact on retrieval, feature saliency information is also extracted and incorporated into the framework to de-emphasize these background features. Experiments show that the proposed method can outperform the state-of-the-art method by an mAP of 4–5%.

2. Overview of the proposed method

The flowchart of the proposed LS-RLF framework is shown in Fig. 2. It consists of offline and online phases. In the offline phase, local features extracted from the database images are used for vocabulary construction and repetitive feature detection/extraction. The repetitiveness of features are analyzed based on their attributes, and the features are clustered using connected component analysis. As the initial clustering may suffer from misclassification due to variations in features of repeated patterns, cluster refinement is performed to improve the clustering outcome. Next, for each cluster, the repetitive properties and its underlying lattice structures are estimated. Missing feature in the lattice structures are identified, and recovered using Hough transformed-based interpolation. Finally, the saliency information are incorporated to emphasize the foreground regions of interest, and weaken the burstiness of repeated patterns in the backgrounds.