Signal Processing: *Image Communication* 41 (2016) 61–71

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

DFOB: Detecting and describing features by octagon filter bank for fast image matching

Zhenyu Xu, Yiguang Liu *, Shuangli Du, Pengfei Wu, Jie Li

SiChuan University, College of Computer Science, Chendu 610065, China

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form
1 December 2015
Accepted 3 December 2015
Available online 14 December 2015

Keywords:
Feature detection
Feature description
Feature orientation
Local feature

ABSTRACT

Feature correspondence is vital in image processing and computer vision. To find corresponding pairs efficiently, in this paper it is proposed that feature detector and descriptor are constructed from the same octagon filter bank (DFOB). The DFOB method is a novel method for the detection, orientation computation, and description of feature points, and is very efficient as computationally implemented by integral images. The matching capability of DFOB is close to the prevalent methods such as SIFT and SURF, because they all detect blob-like image structures as interest features and describe these features using histogram of oriented gradients. Experimental results on benchmark datasets demonstrate that the matching performance of DFOB is comparable with the SIFT and SURF algorithms, while the computational cost is much lower, especially the proposed descriptor is about 50 times faster than SURF descriptor.

1. Introduction

Image matching is fundamental for many computer vision applications such as object recognition, 3D structure reconstruction, image stitching, visual mapping, and target tracking. Extracting stable and repeatable features and encoding these features into robust descriptors with high discriminability are two key steps in image matching. During the last decade, researchers were racing to find faster and better approaches to address this issue. Among them, SIFT [1] and SURF [2] are two of the most famous approaches due to their high quality with respect to the detector repeatability and the descriptor discriminability under a variety of different image geometric and photometric deformations. However, the two methods are memory consuming and computationally expensive, which are not suitable for real-time tasks or applications on mobile devices with limited computational capability. Recently, high computational speed and low memory consumption are achieved by using the combination of the variant of the FAST [3] detector and variant of the BRIEF [4] descriptor, such as ORB [5], BRISK [6] and FREAK [7]. Compared to these methods, the proposed method is more efficient, this is because the descriptor and the detector in proposed method are created from the same filter bank and this makes a dynamic programming strategy applicable to accelerating the computation. The main contribution of our work are highlighted in the following three trickles:

- A novel binary descriptor based on histogram of oriented gradients, which is about 50 times faster than SURF descriptor with about the same robustness and discriminability for matching.
- A stable method for feature point orientation computation, which is efficient and accurate.
- A variant of the CenSurE detector for feature point detection, which allows the detection to run faster and cover more scales than the original.

* Corresponding author.
E-mail addresses: sanxu@outlook.com (Z. Xu), lygpapers@aliyun.com (Y. Liu), dushuangli@163.com (S. Du), wupengfeinihao@hotmail.com (P. Wu), ljcd86@gmail.com (J. Li).

http://dx.doi.org/10.1016/j.image.2015.12.001
0923-5965/

The DFOB method has been tested on standard benchmarks and the experimental results demonstrate that the matching performance of DFOB is close to SIFT and SURF which are famous for the high quality in matching, and the time efficiency is higher than BRISK and ORB which are known for their high computation speed. Another advantage of our method is that both the detector and the descriptor are constructed by octagon filters without pre-processing steps such as Gaussian-smoothing or data training. So, the implementation of our method is very simple, and has quite good efficiency.

2. Related works

In this section we briefly introduce several popular feature point detectors and descriptors related to DFOB method.

2.1. Feature point detectors

We roughly categorise the feature point detectors into two classes in terms of the structure they detect: the corner and the blob. The corner detectors can be traced back to the work of Moravec [8], which calculates the sum-of-squared-differences (SSD) to identify whether the candidate is a corner. Harris and Stephens [9] improved the Moravec detector and designed the well-known *Harris corner detector* with higher repeatability under image deformations. The Harris corner is widely used for its efficiency and accuracy. Scale invariance is achieved in the Harris-Laplace and the Hessian-Laplace detector [10,11] by using Harris method along with scale-space techniques. Recently, the most popular corner detectors may be the FAST and its variants [12,3,13]. The FAST family detectors compare the center pixel to those in a circular ring about the center. The center pixel is identified to be a corner, when there are more than N contiguous pixels all brighter or all darker than the center pixel by at least t. Training on specific scenes in pre-process allows these methods to run at very high speed.

As for blob detectors, the Laplacian of Gaussian (LoG) detector may be one of the first blob detectors. Lowe [14] approximated the LoG by Difference of Gaussian (DoG) which runs faster. Pei and Horng [15] made a even simpler approximation of the LoG detector by two circular mean filters: the bi-level Laplacian of Gaussian (BLoG). Agrawal et al. [16] proposed to use octagon filter to approximate the circular filter for its efficiency in computation. Another family of the blob detectors are resulted from the determinant of the Hessian matrix, and the most famous one is the SURF [2] detector. The maximally stable extremum regions (MSER) [17] can also detect some blob-like structures in the image, which is more robust under perspective deformation. More recently, Alcantarilla et al. proposed the KAZE features [18], in which the detector is able to detect features in a nonlinear scale space by using nonlinear diffusion filtering, thus this detector is edge preserving compared to Gaussian filter based detector.

2.2. Feature point descriptors

Many feature point descriptors rely on *histogram of oriented gradients*. The SIFT descriptor [1], proposed a decade ago, is highly discriminative and robust under a variety of image deformations, and is used as a reference descriptor in descriptor performance evaluations. However, it is a 128 dimensional floating point vector which is prohibitively slow and memory-consuming. The PCA-SIFT [19] reduce the sift descriptor to 36 dimensions, but sacrifice some of the discriminability. The GLOH descriptor [20] created by Mikolajczyk and Schmid is even more discriminative than SIFT descriptor, but it is also more computationally expensive. Inspired by SIFT and GLOH, Tola E, Lepetit V, et al. proposed the DAISY descriptor [21] for dense matching. The signature based SURF descriptor [2], which is faster than SIFT descriptor and has similar discriminability, also has affinity with histogram of oriented gradients, since the first order Haar wavelet responses in x and y direction computed in SURF descriptor is similar to gradients computed in a block pattern.

Another class of feature point descriptors describes feature point using binary string for memory saving and computational efficiency. The BRIEF and its variant descriptors [4,6,7,5] are binary strings composed of binary bits resulted from image intensity comparisons. Yang and Cheng [22] proposed a more discriminative binary descriptor by gradient comparisons rather than intensity comparisons. Machine learning techniques can be applied in these methods to train a pattern for intensity or gradient comparisons, which is helpful to make these descriptors compact and discriminative. Strecha, Christoph, et al. created their binary descriptor by mapping descriptor vectors into the Hamming space [23]. More recently, Lepetit et al. proposed a binary descriptor called BinBoost [24], in which the bits are obtained from learning boosted hash functions. The BinBoost is very robust to viewpoint and illumination changes, but also more computational expensive [25]. Baroffio et al. proposed a more computational efficient binary descriptor than BinBoost by using asymmetric pairwise boosting and box filters [25]. The binary descriptors are matched by computing the Hamming distance, which is extremely fast on modern CPUs which usually provide specific instructions to perform XOR and bit count operations.

3. DFOB: the method

The DFOB method mainly consists of 3 parts: feature point detection, feature point orientation computation and feature point description. Each method in the 3 parts can be used in combination with the methods in other popular image matching algorithms, such as SIFT and SURF. It's worth noting that different to other methods building the detector and descriptor separately, DFOB method builds them together. That is, the descriptor is built using the intermediate result produced in the feature detection step. The using of intermediate result can largely reduce the computational cost.