Neurocomputing 134 (2014) 100–110

Contents lists available at ScienceDirect

Neurocomputing

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

Sparse representation with geometric configuration constraint for line segment matching

Qing Wang*, Tingwang Chen, Lipeng Si

School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China

ARTICLE INFO

Article history:
Received 10 June 2012
Received in revised form
13 November 2012
Accepted 31 December 2012
Available online 30 January 2014

Keywords:
Line segment matching
Line grouping
Sparse coding
Feature representation
Geometric configuration constraint

ABSTRACT

In the paper we propose a novel line segment matching method over multiple views based on sparse representation with geometric configuration constraint. The significant idea of the paper is that we transfer the issue of line correspondence into a sparsity based line recognition. At first, line segments are detected by a LSD (line segment detector) and clustered according to spatial proximity to form completed lines. For each point within a line, SIFT is extracted to represent the attribute of point and PHOG is also considered to describe the appearance of the patch centered at the point. SIFT and PHOG are simply concatenated as a single feature vector and then all these point features are put together by a max pooling function to form a distinctive line signature. Then, all line features extracted from training images are trained into a dictionary using sparse coding. Lines with the same similarity may fall together in the high-dimensional feature space. Finally, line segments in a test view are matched to their counterparts in other views by seeking maximal pulses from the coefficient vector. Under our framework, line segments are trained once and matched across all other views. Experimental results have validated the effectiveness of the approach for planar structured scenes under various transformations and degradation, such as viewpoint change, illumination, blur and compression corruption.

1. Introduction

Meta-pattern (point, line, contour, region, etc.) matching over multiple views is one of the most fundamental tasks in computer vision, which benefits in various applications, such as 3D structure reconstruction, object recognition, landmark alignment and image registration. Although a great number of approaches have been proposed for point correspondence, only a few methods of automatic line segment matching are reported in the literature so far. Due to the limitations of middle level feature representation of line segment, effective similarity cross lines and optimization methods, line correspondence is still a challenge issue till now.

Existing approaches to line matching can roughly be divided into two categories: local appearance based approaches and topological layout based methods. As the first appearance based type, Schmid and Zisserman [1] proposed to use graylevel information for matching individual line segments between images. Wang et al. [2] proposed a new descriptor named MSLD (Mean-Standard deviation Line Descriptor) for line matching. A pixel support region is defined for each pixel within a line, and a histogram of image gradient is accumulated in the support region. The final descriptor is constituted of the mean and standard deviation of these histograms. Appearance based line feature is less distinctive than local feature descriptors since line segments lack rich textures in their local neighborhood. Moreover, this method may fail when encountering the repeated textures, such as an urban scene with buildings and facades. On the other hand, topological layout based approaches are twofold: disambiguation derived from more geometric information, and the increased complexity. Lourakis et al. [3] proposed to use two points and two lines to estimate a projective invariant for matching planar surfaces with points and lines. However, this method is rather complicated and can hardly be handled in non-planar scenes. Wang et al. [4] proposed a line signature to match lines. The angles and length ratios between lines computed by the endpoints of lines are used to describe a pair of line segments, and then line matching is completed on the basis of pairs of line segments. Since the descriptor of a pair of line segment relies on the endpoints of line segments, this method may fail when the location of endpoint is not accurate enough. Fan et al. [5] explored an affine invariant from two points and one line, and utilized them to match lines even with different endpoints. But it relies on points matching seriously and fails when point matching goes down. Bay et al. [6] proposed a hybrid method to match lines based on the appearance and topological layout. Since the matching propagation is an iterative process, its computational cost is extremely huge and unacceptable in real world applications.

* Corresponding author.
E-mail address: qwang@nwpu.edu.cn (Q. Wang).

0925-2312/$ - see front matter
http://dx.doi.org/10.1016/j.neucom.2012.12.079

Epipolar geometry constraints are introduced to reduce the disambiguation and computational cost 1,7,8. Hartley [9] proposed to use the trifocal tensor to match lines across views. Plane sweep methods [10] are also employed to reduce combinatorial expansion, but these methods do not perform well when the endpoints of 2D line segments are not consistent in different images of the same scene. Another way to increase the matching efficiency is to use color histogram based on feature descriptors for 2D line segments [6], but this method assumes that colors only undergo slight changes (Lambert surface) and the image does not contain any specular highlights. Recently, Chen and Wang [11] proposed a weak matching model for 2D line segment based on the reconstructed 3D point clouds. The computational time is reduced at a great scale but it needs 3D points from SfM (shape from motion). Eden and Cooper [12] proposed to obtain the line correspondence within a 3D ROI (region of interest), which divided the space region into smaller cubes and solved the matching problem for line segments that lied inside each cube. This method could reduce the searching scope of line matching in a one-off process, but the total computational complexity is not changed at all.

We propose a novel method combined distinctive line segment representation with geometric configuration constraints as [6]. Here, we use an efficient sparse representation of line segment feature after pooling the concatenation vector of the SIFT [15] and PHOG [17] descriptors within the line. Sparse signal representation has been proven to be an extremely powerful tool for acquiring and representing signals, which is widely used in computer vision and pattern recognition. This success is mainly due to the fact that important classes of signals such as audio and images have naturally sparse representations with respect to fixed bases (*i.e.*, Fourier, Wavelet), or concatenations of such bases. Moreover, efficient and provably effective algorithms based on convex optimization or greedy pursuits are available for computing such representations with high fidelity [13]. In line segment matching, we have to learn a task-specific (often under-complete) dictionary from given sample lines. So we need to extend the existing theory and algorithms of sparse representation to this new scenario. Our algorithm is based on the assumption that if the matched lines correspond to the same line in a 3D space, they may fall into a same distribution in the high-dimensional feature space. After line segment collection of representative samples is found for this distribution, we expect a typical sample that has a very sparse representation with respect to such (possibly learned) bases. Such a sparse representation, if computed correctly, could naturally encode the semantic information of the line segments over views. Extensive experiments have demonstrated that the proposed method can match lines with high accuracy under various conditions, such as scaling, rotation, illumination changes as well as some extent of viewpoint changes. The contributions of our approach are as follows:

- Line segment is normalized with the max pooling function to form a distinctive descriptor.
- All lines from training images are put together and trained as a dictionary. Lines in the test image are matched with those in the training images by the pulse function of sparse coefficient vector.
- Geometric configuration constraints are used to remove the false matching.

The remainder of the paper is organized as follows. The overview of our approach is introduced in Section 2. The line segment detection and grouping are discussed in Section 3. The process of feature extraction and representation and line matching over multiple views by sparse coding are presented in Section 4. We further pose geometric configuration constraints to eliminate some false matched lines in Section 5. Experimental results of line matching of multi-view images are reported and evaluated in Section 6. Finally, the concluding remarks are drawn in Section 7.

2. Overview of our approach

On the survey of the state-of-art methods and analysis of shortcomings of line matching algorithms, we propose a novel approach based on sparse coding for automatic line matching over multi-view images. Fig. 1 depicts an overview of the proposed approach. The left of Fig. 1 is the input of a photo collection of a benchmark data set (Hall), and the right is the output of line segment matching across three views. In the middle of Fig. 1, the

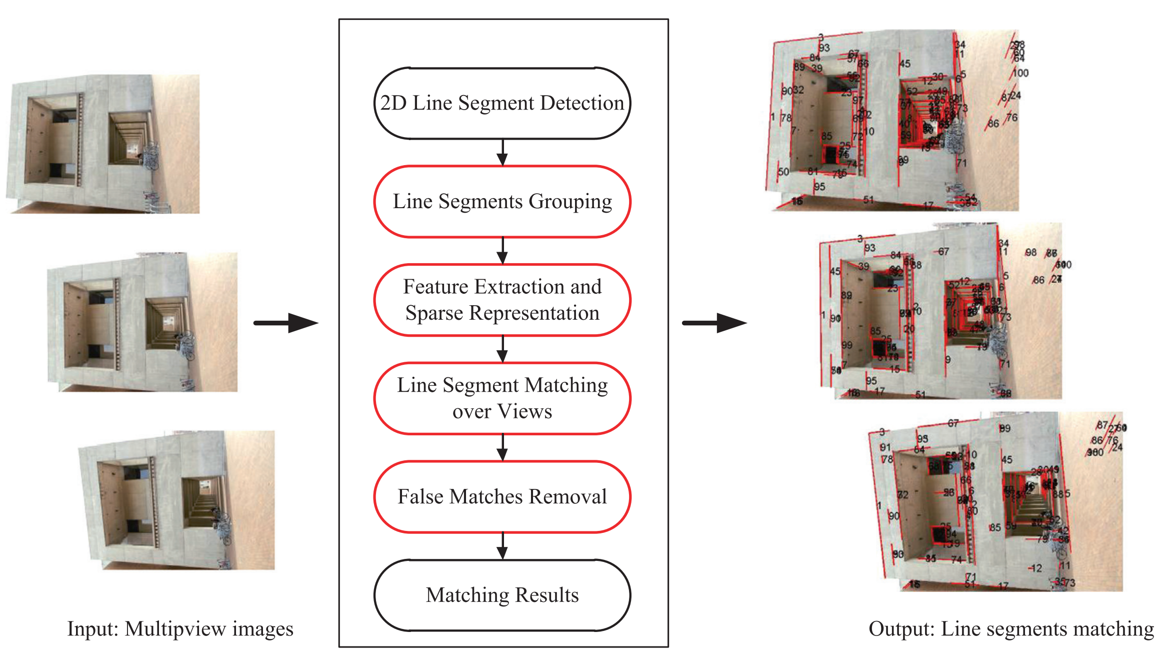

Fig. 1. Overview of the proposed multi-view line matching approach.