Neurocomputing 184 (2016) 207–220

Contents lists available at ScienceDirect

Neurocomputing

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

Hierarchical line matching based on Line–Junction–Line structure descriptor and local homography estimation

Kai Li [a], Jian Yao [a,*], Xiaohu Lu [a], Li Li [a], Zhichao Zhang [a,b]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei, PR China
[b] Collaborative Innovation Center of Geospatial Technology, Wuhan, Hubei, PR China

ARTICLE INFO

Article history:
Received 1 January 2015
Received in revised form
18 May 2015
Accepted 27 July 2015
Available online 13 January 2016

Keywords:
Line matching
Hierarchical matching
Local feature descriptor
Homography estimation

ABSTRACT

This paper presents a hierarchical method for matching line segments from two images. Line segments are matched first in groups and then in individuals. While matched in groups, the line segments lying adjacently are intersected to form junctions. At the places of the junctions, the structures are constructed called Line–Junction–Line (LJL), which consists of two adjacent line segments and their intersecting junction. To reliably deal with the possible scale change between the two images to be matched, we propose to describe LJLs by a robust descriptor in the multi-scale pyramids of images constructed from two original images. By evaluating the description vector distances of LJLs from the two images, some candidate LJL matches are obtained, which are then refined and expanded by an effective match-propagation strategy. The line segments used to construct LJLs are matched when the LJLs they formed are matched. For those left unmatched line segments, we match them in individuals by utilizing the local homographies estimated from their neighboring matched LJLs. Experiments demonstrate the superiorities of the proposed method to the state-of-the-art methods for its robustness in more difficult situations, the larger amounts of correct matches, and the higher accuracy in most cases.

1. Introduction

As a low-level vision task, image matching is fundamental for many applications which require recovering the 3D scene structure from 2D images, like robotic navigation, structure from motion, 3D reconstruction, and scene interpretation. The majority of image matching methods are feature point based [1–6] which commence the extraction of feature points from images, followed by the utilization of the photometric information adjacently associated with the extracted points to match them. Objects in real scenes, however, can be easily outlined by line segments, especially for man-made scenes. This indicates that it is better to recover 3D scene structures based on line segments than that on feature points, at least for some scenes [7–12]. For example, for poorly textured scenes, where feature points are hard to be detected and matched, recovering their 3D structures from line matches seems the only choice because their structures can be easily outlined by several edge line segments [13]. Despite these advantages, both the lack of point-to-point correspondence and the loss of connectivity and completeness of the extracted line segments make line segment matching a tough task, which also partly explains why line segment matching has been less investigated.

Line matching methods in existing literatures can generally be classified into two categories: the methods that match line segments in individuals and those in groups. Some methods matching line segments in individuals exploit the photometric information associated with individual line segments, such as intensity [14,15], gradient [16–18], and color [19] in the local regions around line segments. All these methods underlie the assumption that there are considerable overlaps between corresponding line segments. This assumption leads to the failure of these methods in situations where corresponding line segments share insufficient corresponding parts. Besides, in regions with repetitive textures, these methods tend to produce false matches since the lack of variations in the photometric information for some line segments. Other methods matching line segments in individuals leverage point matches for line matching [20–23]. These methods first find a large group of point matches using the existing point matching methods, and then exploit the invariants between coplanar points and line(s) under certain image transformations to evaluate the correspondence of the line segments from two images. The line segments which meet the invariants are regarded to be in correspondence. These methods utilize geometric relationship between line segments and points, rather than photometric information in the local regions around line segments, and are thus robust even

* Corresponding author.
E-mail address: jian.yao@whu.edu.cn (J. Yao).
URL: http://www.cvrs.whu.edu.cn/ (J. Yao).

http://dx.doi.org/10.1016/j.neucom.2015.07.137
0925-2312/

when local shape distortions are severe. However, these methods share a common disadvantage that they are incapable of processing images in which the scenes captured are poorly textured since feature points are hard to be detected and matched in this kind of scenes, which consequently disables the use of point matches for line segment matching.

Matching line segments in groups is more complex, but more constraints are available for disambiguation. Most of these methods [13,24–26,32] first use some strategies to intersect line segments to form junction points and then utilize features associated with the generated junction points for line segment matching. These methods transfer line matching to point matching, a widely investigated problem which many effective algorithms target to solve. But junction points contain more information than feature points detected by some detectors [27–30]. They are the results of intersecting pairs of adjacent line segments and the relationship between junction points and line segments forming them is additional and important information that can be exploited for matching them. How to effectively exploit features associated with junction points to help match them is still an open issue. In [31], rather than exploiting features of junctions for line segment matching, the stability of the relative positions of the endpoints of a group of line segments in a local region under various image transformations is exploited. This method first divides line segments into groups and then generates a description of the configuration of the line segments in each group by calculating the relative positions of these line segments. Since the configuration of a group of line segments in a local region is fairly stable under most image transformations, the description of the configurations of two groups of line segments in correspondence should be similar. In this way, groups of line segments can be matched. This method is robust in many challenging situations, but the dependence on the approximately corresponding relationship between the endpoints of corresponding line segments leads to the tendency of this method to produce false matches when substantial disparity exists in the locations of the endpoints.

Our proposed line matching method in this paper is a combination of the two categories methods described above. It matches line segments both in groups and in individuals under a hierarchical framework. The framework is composed of three stages where line segments are matched in groups in the first two stages while in individuals in the third stage. The three-stage flowchart of the proposed line matching algorithm is illustrated in Fig. 1. For two sets of line segments extracted from two images, the first stage commences intersecting neighboring line segments to form junction points. At the places of the formed junction points, we form the structures called Line–Junction–Line (LJL), utilizing two adjacent line segments and their intersecting junction. To greatly reduce the effect of the scale change between the two images, we propose to build Gaussian image pyramids for the original images and adjust the LJLs constructed in the original images to fit each image in the image pyramids and described them there by the proposed LJL descriptor. Some initial LJL matches can be found by evaluating the description vector distances of LJLs from the two images. These LJL matches are then refined and expanded in the second stage, where we propagate LJL matches by iteratively adding new matches while deleting possibly false ones. In the above two stages, the line segments lying closely with each other from the two images are matched along with their constructed and matched LJLs. For those line segments lying far away from others and are not used to constructed LJLs, we match them in individuals in the third stage by utilizing the local homographies estimated from their neighboring matched LJLs.

This work is an extension of our work presented in [32]. Compared with the previous one, this work makes promotions in the following aspects. First, a more general way is utilized to generate junctions using adjacent line segments. In [32], sets of line segments extracted by some line segment detectors in a image are processed in advance before they are used for matching by a series of procedures to obtain a new line segment set where some line segments are extended to be longer and are connected with others. In this work, the line segments extracted by line detectors are not required to be refined in advance, and can be used to generate junctions directly based on the local spatial proximity. This promotion helps generate more junctions and contributes to better matching results. Second, a more robust descriptor is proposed to describe the structure (called RPR in [32], while LJL in this paper) formed by two adjacent line segments and their intersecting junction. Third, we propose a more reasonable strategy to deal with the possible scale change between images. To match line segments from two images with scale change, in [32], the global scale change factor between the two original images is estimated and one of the two images is adjusted to the same scale as the other one. This strategy is reasonable only when the scale change between the two images is a global one. When scale changes between the two images vary with local regions (often introduced by viewpoint changes), this strategy is unable to reliably work. This disadvantage is solved in this paper and the proposed strategy

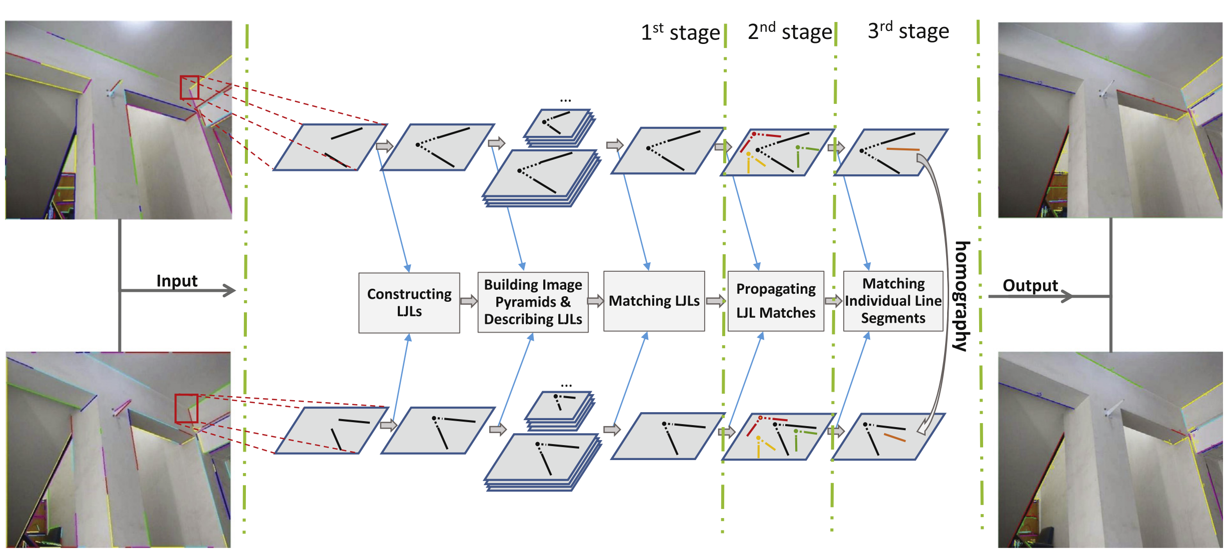

Fig. 1. The flowchart of the proposed line matching algorithm.