Neurocomputing 73 (2010) 3114–3122

Contents lists available at ScienceDirect

Neurocomputing

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

New similarity measures of localities for a two-layer matching scheme and estimation of fundamental matrices

Qi Li *, Guangming Xing

Department of Mathematics and Computer Science, Western Kentucky University, United States

ARTICLE INFO

Article history:
Received 20 November 2009
Received in revised form
23 April 2010
Accepted 23 June 2010
Communicated by D. Tao
Available online 29 July 2010

Keywords:
Fundamental matrix
Interest point
Cross correlation
Locality

ABSTRACT

Estimation of fundamental matrices is important in 3D computer vision. It is well known that the estimation of fundamental matrices is sensitive to outliers—even a few of imprecise point correspondences may result in an estimated fundamental matrix inconsistent with the geometry setup of input images. In terms of interest points with localities, we have proposed a two-layer matching scheme that is a generalization of conventional normalized cross correlation (NCC), aiming to improve the precision of point correspondence. A *locality* of interest points means a set of interest points that are contiguous to each other in terms of 8-connectivity. The first layer of the matching scheme establishes locality correspondence, and the second layer refines point correspondence within matching localities. In this paper, we analyze a limitation of the similarity measure of localities proposed in our previous work and then we propose two new similarity measures to address the limitation. We test the two-layer matching scheme on images in Middlebury stereo datasets, with known fundamental matrices as the ground truth. From the experimental results, we observe the improvement of new measures of locality similarity, and we also observe that the estimated fundamental matrices are consistently close to the ground truth.

1. Introduction

The fundamental matrix is the algebraic representation of epipolar geometry [9]. Estimation of the fundamental matrix has wide applications in 3D computer vision [2,9,22]. Given a pair of images, a popular way to estimate the fundamental matrix is to apply the (normalized) eight point algorithm [8], assuming that correct and precise corresponding points are provided. In the context of automatic estimation, however, point correspondence built by automatic matching methods (such as cross correlation of descriptors [24,15]) may contain outliers, either mismatching points or imprecise correspondence [24,9]. Randomized methods, such as RANSAC [6], or LMS [17], were introduced to remove mismatching-type outliers, and extensive studies have shown the robustness of these methods [23,24]. Recently, a linear mixed-effect modeling method [25] was proposed to estimate the fundamental matrix, aiming to relax the assumption of Gaussian distribution used in previous approach.

Imprecise point correspondence can severely degrade the accuracy of the fundamental matrix [24]. Imprecise point correspondence are caused by the localization error of image points detected by an interest (feature) point detector [20]. Localization error can be modeled as a noise distribution [24]. Non-linear optimization methods were introduced to search the optimal parameters of the fundamental matrix by minimizing its residue [5]. (Residue is defined as the average distance between a point and the epipolar line of the corresponding point.) Previous studies showed that non-linear optimizations focused on the parameters of the fundamental matrix seem not to bring significant improvement of the accuracy of the fundamental matrix. A more sophisticated scheme is the bundle adjustment [1] that extends the search space to the locations of corresponding points (besides the unknown parameters of the fundamental matrix). Studies showed that the bundle adjustment is effective to improve the accuracy of the fundamental matrix, but its expensive computational cost, due to the high-dimensional search space, may limit its applications. From a systematic perspective, model-selection was introduced to handle the uncertainty of fundamental matrix due to the choice of a particular feature detector, the choice of the matching algorithm, the motion model, iterative hypothesis generation and verification paradigms [22].

Another strategy to improve the accuracy of the fundamental matrix is to circumvent the localization error arising in interest point detection, and apply optical-flow techniques to build point correspondence [3,21]. In [3], the authors argue that "more emphasis should be placed on improving the quality of data used for estimation rather than focusing the attention of the development of very sophisticated estimation procedure", and proposed to apply optical-flow to improve the point correspondence. The

---

* Corresponding author.
E-mail addresses: qi.li@wku.edu (Q. Li), guangming.xing@wku.edu (G. Xing).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.06.015

key contribution of the paper [3] is the confidence measure to select a sparse set of reliable point correspondence using Hough transform. In [21], the authors proposed a pseudo-warping transform functions to support optical-flow techniques for the estimation of the fundamental matrix. The limitation of the optical-flow based estimation is the assumption of small motions (disparities).

The motivation of this paper is consistent with the argument of improving the quality of data [3]. However, we insist the feature based estimation scheme due to its flexibility in applications. We analyze that the localization error of conventional interest points likely comes from the *non-maximum suppression*, a popular candidate selection scheme used in interest point detectors [16,15]. (Roughly speaking, a detector consists of two steps: (i) interest strength assignment, and (ii) candidate selection [14].) Non-maximum suppression applies a small suppression window (e.g., 3×3) to select local maxima as good candidates of interest points. Non-maximum suppression is effective to force interest points to scatter in the entire image regions, which can benefit many higher level applications, including the estimation of the fundamental matrix. However non-maximum suppression also degrades the localities of interest points—In this paper, we define a set of interest points that are neighboring to each other in terms of 8-connectivity as a *locality*. It is easy to see that it is impossible for interest points detected by non-maximum suppression to be contiguous to some others in terms of 8-connectivity. (Note that optical-flow based methods do not involve with non-maximum suppression, and thus they may circumvent the localization issue.)

In our previous work [14], we proposed *imbalance oriented selection*, as an alternative of non-maximum suppression, to preserve good candidates for interest point detection in sparsely textured images. The intuition on imbalance is to distinguish different nature of interest points and edge points. Specifically, imbalance is determined by a non-negative integer, called *index of maximum difference*. (More details will be reviewed in the beginning of Section 2.) Interest points detected by imbalance oriented selection have non-trivial localities. With the interpretation of the index of maximum difference as a quantified imbalance degree, the diversity of imbalance degree can be used as an interest assignment scheme to enhance the distinctiveness of localities of interest points. The rationale is that interest points of diverse imbalance are expected to have complex local appearances, and thus help contribute distinctive descriptors to reduce mismatching rates in (correlation based) point correspondence.

Based on interest points with non-trivial localities, we generalized the normalized cross correlation (NCC) single-layer matching scheme to a two-layer matching scheme [11,13]. The first layer of the matching scheme aims to establish locality correspondence, and the second layer aims to refine point correspondence within each pair of matching localities. Relatively speaking, our study of interest points with localities was from the lower-level perspective than previous studies on groups of interest points. For example, aiming to construct (transformation) invariant descriptors, Brown and Lowe [4] proposed to group interest points in terms of nearest neighbor criterion.

The similarity measure of localities is the key part of the two-layer matching scheme. The similarity of localities can be basically formulated as the similarity of two sets of descriptors. The similarity measure proposed in our previous work [11] accumulates overall best contribution from every descriptor. The intuition behind this measure is that the overall contribution is more robust than a certain singleton contribution. However, there is a limitation of this measure, which is, if an interest point in one locality has no correspondence in the other locality (i.e., outliers within corresponding localities), its contribution becomes an uncertainty, which degrades the reliability of the similarity measure of localities. In this paper, we propose two new similarity measures of localities to address this limitation. One measure introduces a threshold to discard un-qualified individual contribution during the accumulation of point-wise similarities. The other measure is based on the best point-wise similarity.

We test the two-layer matching scheme with new similarity measures of localities on images in Middlebury stereo datasets [18,19,10], with the known fundamental matrix as the ground truth. Our experimental results show that new similarity measures result in a larger number of precise point correspondence, and the estimated fundamental matrices by the two-layer matching scheme are consistently close to the ground truth.

The rest of the paper is organized as follows: In Section 2, we present interest points with localities. In Section 3, we first review the two-layer matching scheme, and then propose two new similarity measures. In Section 4, we present experiments results. Finally, we conclude and give future work in Section 5.

2. Interest points with localities

In this section, we first give a brief description of imbalance oriented selection, and then we present an interest strength assignment based on the diversity of imbalance degrees to enhance the distinctiveness of the localities of interest points. More details of diversity of imbalance can be found in [12].

2.1. Imbalance oriented selection

Imbalance oriented selection was proposed to minimize the occurrences of edge points [14]. Since edge points have similar local appearances (i.e., not distinctive to each other), they increase the chance of mismatching in the higher-level applications. Edge points can be characterized as points of balanced local appearances. As shown in Fig. 1(a), intensities are supposed to change slightly at the same side of an edge while they change significantly across an edge. Here, $n = 8$ directions are considered. A long (short) arrow indicates a strong (weak) intensity change along the associated direction. The number of long arrows is equal to the number of short arrows, which indicates the balance nature of an edge point. Fig. 1(b) shows a

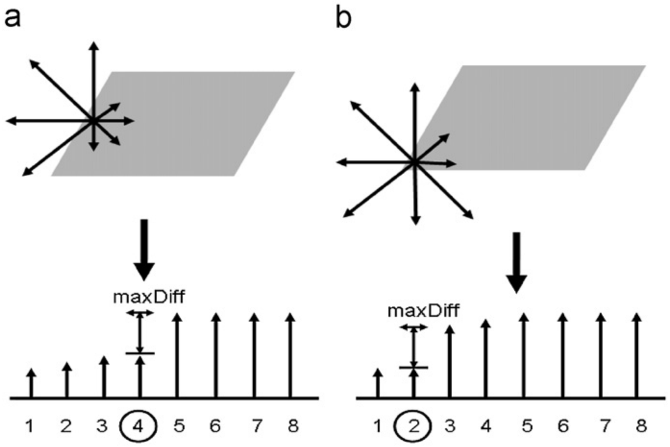

Fig. 1. (a) An edge point of balanced local appearance, where the index of maximum difference is 4 (half of 8 directions); (b) a candidate of imbalanced point, where the index of maximum difference is 2. An arrow starting from a point denotes the magnitude of the first-order directional derivative of the point.