Pattern Recognition Letters 36 (2014) 1–9

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Epipolar geometry estimation for wide baseline stereo by Clustering Pairing Consensus

Dazhi Zhang [a,d], Yongtao Wang [a], Wenbing Tao [a,b,c,*], Chengyi Xiong [b]

[a] Institute for Pattern Recognition and Artificial Intelligence, School of Automation, National Key Laboratory of Science and Technology on Multi-spectral Information Processing Technologies, Huazhong University of Science and Technology, Wuhan 430074, China
[b] College of Electronic and Information Engineering, Hubei Key Lab of Intelligent Wireless Communication, South-Central University for Nationalities, Wuhan 430074, China
[c] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
[d] Research Institute of Nuclear Power Operation, Wuhan, 430000, China

A R T I C L E I N F O

Article history:
Received 5 June 2012
Available online 30 August 2013

Communicated by A. Heyden

Keywords:
Wide baseline match
Fundamental matrix
Topological clustering
Clustering Pairing Consensus

A B S T R A C T

The problem of automatic robust estimation of the epipolar geometry for wide-baseline image pair is difficult because the putative correspondences include a low percentage of inlier correspondences, and it could become a severe problem when the veridical data are themselves degenerate or near-degenerate. In this paper, Clustering Pairing Consensus (CPC) algorithm is proposed to estimate the fundamental matrix. The CPC algorithm first produces the Matched Regions Clusters (MRCs) using topological clustering (TC) algorithm given a scale parameter. An estimation is produced from each valid pair of MRCs and is then provided to M-estimation to compute a fundamental matrix. Finally, the best one is chosen as the final model from all the estimation. The proposed CPC algorithm has been demonstrated to be able to effectively estimate fundamental matrix and avoid the degeneracy of the traditional method for some difficult image pairs.

1. Introduction

Finding epipolar geometry consistent with the largest number of tentative (local) correspondences is the final step of all wide-baseline algorithms, and is a fundamental problem in computer vision. Unlike classical short baseline stereo techniques, wide baseline stereo algorithms can tolerate a large change in viewpoint between the images. Although technically more difficult, wide baseline stereo has a number of advantages, including higher precision of depth measurement and smaller number of images needed to completely cover an object or scene. The price to be paid for these advantages is the need to cope with large geometric and radiometric distortions. The fundamental matrix is the algebraic representation of epipolar geometry, which is defined as the matrix satisfying the relation

$$x'^{T}Fx = 0 \quad \forall x', x, \tag{1}$$

where $x, x' \in R^3$ are the projections expressed in homogeneous coordinates of the same 3D point in the two images.

The estimation of fundamental matrix has received large attention in the last two decades. Random sampling and consensus (RANSAC) (Fischler and Bolles, 1981) is the most widely used robust estimator in computer vision today. Random sampling techniques aim to explore the search space of possible solutions well enough to have at least one candidate which is determined solely by inliers. However, the estimation of the epipolar geometry with wide baseline is difficult.

The first difficult situation is that these random sampling algorithms are not sufficient to cope with the high percentage of outliers which can frequently occur in wide baseline images. Traditional statistical estimators like LMedS or M-estimators have breakdown points that are no more than 50%, which is far from being satisfied with wide baseline (Torr and Murray, 1997). For RANSAC algorithm, there is no exact limit in the percentage of mismatches it can work with, but researchers agree it should not be used if beyond roughly 2/3 (67%) (Lowe, 2004; Wang et al., 2011).

For dealing with a large number of outliers, some improvement of basic RANSAC was proposed. The MAPSAC (Tordoff and Murray, 2002) is the generalization of RANSAC based on the same point selection strategy and the solution is to maximize likelihood. In Subbarao and Meer (2007), the projection based M-estimator (pbM) was proposed which can estimate model without a user specified threshold. LO-RANSAC (Chum et al., 2005) enhances RANSAC with a local optimization step. This optimization executes a new sampling procedure based on how well the measurements

* Corresponding author at: Institute for Pattern Recognition and Artificial Intelligence, School of Automation, National Key Laboratory of Science and Technology on Multi-spectral Information Processing Technologies, Huazhong University of Science and Technology, Wuhan 430074, China. Tel.: +86 27 87540164.
E-mail address: wenbingtao@hust.edu.cn (W. Tao).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.08.019

satisfy the current best hypothesis. Another frequently used strategy is replacing random sampling with guided sampling. In (Tordoff and Murray, 2002), the guidance of the sampling is based on the correlation score of the correspondences. In PROSAC (Chum et al., 2005), the correspondences are sorted by matching score and from this ordered list larger sets are progressively generated.

In wide baseline circumstances, the second difficult situation occurs when a large subset of inliers is consistent with degenerate epipolar geometry. In computer vision, the data are said to be degenerate if they are insufficient to determine a unique solution. As observed by Chum et al. (2005), RANSAC performs worse when the data are near degenerate. Moreover, as shown in Torr and Murray (1997), this is particularly severe when the veridical data are themselves degenerate or near-degenerate with respect to the model. To solve this problem, Torr and Zisserman propose The PLUNDER algorithm (Torr et al., 1998), which estimates multiple models separately and then performs model selection. In (Goshen and Shimshoni, 2008), BEEM algorithm is proposed for estimation of the epipolar geometry in difficult scenes. Besides dealing with degeneracies, the main feature of the BEEM algorithm is that it is able to generate an hypothesized fundamental matrix using only two pairs of matches.

In our previous work (Wang et al., 2008), a topological clustering (TC) algorithm was proposed and used to filter out mismatches. In fact, the algorithm can also be seen as a clustering algorithm, not only a filter. With this technique, false correspondences can be identified and rejected by checking the consistency of topological relationships between matched regions in image pair. Simultaneously, the resulting matched regions are partitioned into connected Matched Regions Clusters (MRCs). In this paper, the algorithm was extended and these MRCs are considered as an intermediate representation to facilitate fundamental matrix estimation. Since the original matched pairs include a number of mismatches which will degrade the performance of the fundamental matrix estimation, the topological clustering can be used to effectively eliminate the mismatch and reduce the number of the correspondences. Therefore, the estimation of fundamental matrix using the matched pairs produced by topological clustering can achieve better performance and higher computational efficiency than using the original matched pairs.

Besides eliminating the mismatch, by using the TC the second difficulty due to the degenerate epipolar geometry in fundamental matrix estimation can also be alleviated. The fundamental matrix has too many degrees of freedom which causes overfitting of the data when the scene includes a dominant plane in wide baseline image. By the topological clustering all the matched pairs in each cluster are topologically connected. Each correct MRC roughly corresponds to a plane in scene, and two planes are sufficient to determine the fundamental matrix **F** and avoid the degeneracy. Although it is possible that the arbitrarily selected two MRCs are included in a dominant plane in wide baseline image, we can consider all the two MRCs combination in the MRCs set to conduct the estimation and choose the best one. This mechanism can make the estimation effectively escape from degenerate models.

We first introduce the topological clustering algorithm more formally and profoundly using semi-local constraint, and define the topological relational matrix to illustrate the process of the topological clustering. Then the Clustering Pairing Consensus (CPC) algorithm is proposed to estimate the fundamental matrix. The CPC algorithm first produces the MRCs using TC algorithm given a scale parameter. An estimation is produced from each valid pair of MRCs and is then provided to M-estimation to compute a fundamental matrix. Finally, the best one is chosen as the final model from all the estimation.

The remainder of the paper is organized as follows. In Section 2 we first present the topological clustering algorithm more formally and profoundly using semi-local constraint. In Section 3 the Clustering Pairing Consensus (CPC) algorithm based on the topological clustering is proposed. In Section 4 we validate the algorithm by some experiments on public data set, followed by a brief conclusion section.

2. Topological clustering

2.1. Semi-local constraint

It is well known that natural images are not a random collection of independent pixels or blocks. The crucial observation is that most local feature region of image cover an approximately continuous surface of the scene. Intuitively, a surface is a two-dimensional manifold embedded in three-dimensional Euclidean space R^3, and the local region U of the surface is an open set in manifold. As shown in Fig. 1, φ and ψ are homeomorphisms from U to different two-dimensional Euclidean space separately. Following from topological theory, the transition map $\varphi \circ \psi^{-1}$ is bijective continuous map. It means that the connectedness or topological relationship among regions in a same surface is viewpoint-invariant, which is a semi-local constraint, and is the basis of topological clustering.

Define I as a given image in image domain Ω and $R_i, (i = 1, 2, \ldots, N)$ are regions in image domain Ω, where N is the

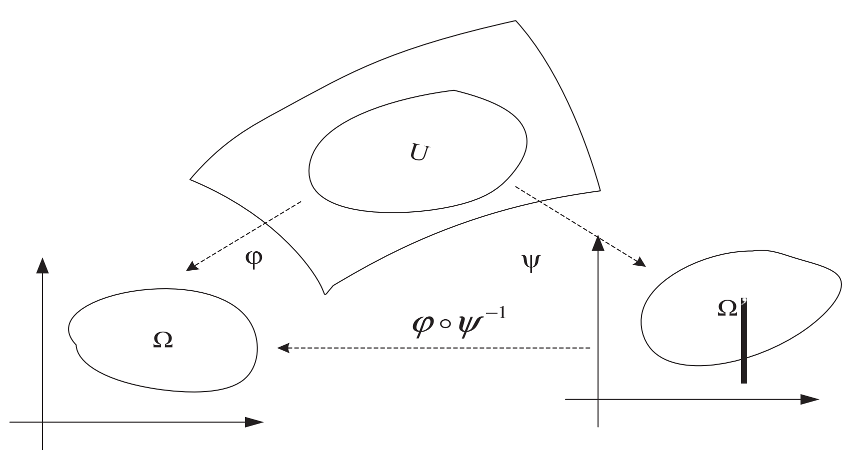

Fig. 1. Example of a two-dimensional manifold and its transition map.