J. Vis. Commun. Image R. 30 (2015) 299–311

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A topology-based approach to computing neighborhood-of-interest points using the Morse complex ☆

Ricardo Dutra da Silva [a,*], William Robson Schwartz [b], Helio Pedrini [c], Jesus Pulido [d], Bernd Hamann [d]

[a] Informatics Department, Federal University of Technology, Curitiba, PR 80230-901, Brazil
[b] Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-010, Brazil
[c] Institute of Computing, University of Campinas, Campinas, SP 13083-852, Brazil
[d] Institute for Data Analysis and Visualization, Department of Computer Science, University of California, Davis, CA 95616-8562, USA

ARTICLE INFO

Article history:
Received 13 August 2014
Accepted 3 May 2015
Available online 22 May 2015

Keywords:
Interest points
Topological neighborhood
Matching
Correspondence
Morse complex
Local context
Image descriptors
Geometric similarity

ABSTRACT

A central problem in image processing and computer vision is the computation of corresponding interest points in a given set of images. Usually, interest points are considered as independent elements described by some local information. Due to the limitations of such an approach, many incorrect correspondences can be obtained. A specific contribution of this paper is the proposition of a topological operator, called Local Morse Context (LMC), computed over Morse complexes, introduced as a way of efficiently computing neighborhoods of interest points to explore the structural information in images. The LMC is used in the development of a matching algorithm, that helps reducing the number of incorrect matches, and obtaining a confidence measure of whether a correspondence is correct or incorrect. The approach is designed and tested for the correspondence of narrow-baseline synthetic and specially challenging underwater stereo pairs of images, for which traditional methods present difficulties for finding correct correspondences.

1. Introduction

The correspondence or matching of image interest points is a basic step in computer vision that is used to find corresponding locations in different images for tasks such as image stitching, image registration, scene reconstruction, object detection and recognition [3,27,31,34,37,47,53]. A well explored advantage of the correspondence of interest points is that it allows matching in the presence of occlusions and changes in scale and orientation [46]. Furthermore, more reliable matching of images can be computed [11] and might be used to compute denser correspondences [30]. However, the approach may fail depending on the image acquisition method, noise corruption, and transformations between images [38].

This paper investigates the correspondence of interest points as an auxiliary step for establishing a denser set of correspondences to further construct 3D models and merge point clouds produced from images acquired at different time steps (video frames). The considered images are challenging, such that we should avoid incorrect matches due to multiple regions with similar characteristics. Furthermore, the confidence of a match being correct should be measured.

The basic stages for corresponding interest points between images include: (1) interest point detection, (2) description, and (3) matching. After these steps, (4) filtration (as SIFT-ratio) and (5) model fitting (RANSAC) are commonly used. We define these steps as the basic framework.

Interest points are usually considered as independent elements described by some limited local information, basically, the appearance of patches of pixels surrounding the point location [46]. The limited local information clearly is not able to discriminate between interest points in some cases. The difficulty arises from patches related to two interest points not having enough dissimilarity between them. Examples of such cases occur in smooth regions, repeated patterns and symmetries. Therefore, images with different regions of high similarity often incur in problems of discrimination due to the limited information to compute descriptors.

Besides such difficulties, many other may occur. For example, for the underwater images considered in this paper, besides the morphological nature of the regions in the images, which may create highly similar structures, noise may be present due to particles in the water and the acquisition may produce distortions and illumination differences.

☆ This paper has been recommended for acceptance by "Yehoshua Zeevi".

* Corresponding author.
E-mail address: ricardodutr@gmail.com (R.D. da Silva).

http://dx.doi.org/10.1016/j.jvcir.2015.05.001
1047-3203/

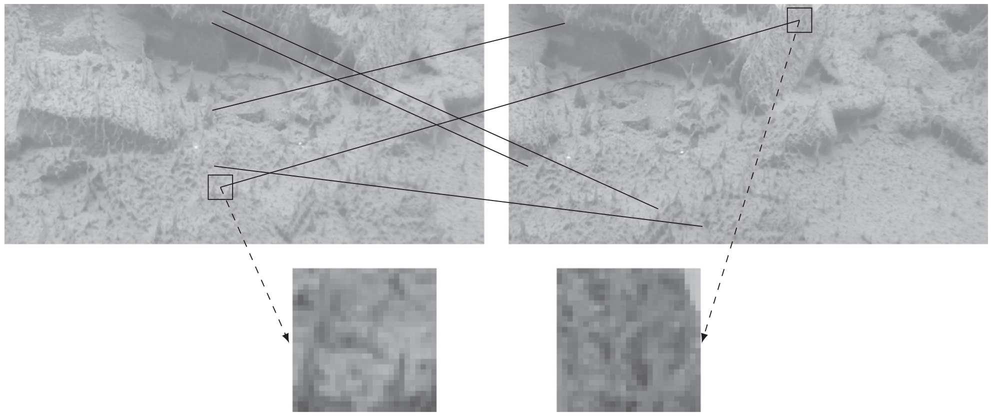

Fig. 1. Examples of incorrect matches computed due to regions of high similarity. The descriptor by itself is not able to discriminate between some regions. The zoomed regions show the pixel level texture similarities that produce close descriptors and consequently difficulties for the correspondence of interest points.

Commonly, the correspondences of images (step 3 of the framework) are computed by pairwise comparison of their interest point local descriptors, such as performed by *k*-nearest neighbor algorithm [34,46]. This approach is computationally efficient and suitable for real-time applications or for problems that deal with massive amounts of data. However, the difficulties arising from point detection and description are propagated and not properly handled at the matching stage. Fig. 1 shows some examples of incorrectly matched points.

The neighborhood relation proposed in this paper is mainly evaluated at the matching level. We use the neighborhood relation to produce local restrictions to the set over which correspondences are searched for. The idea is that, for the narrow-baseline images considered in this paper, spatially close interest points in one image should be related to spatially close putative correspondences in a second image.

Cases of incorrect correspondences are very common and there are some different levels in which it is possible to consider ways of eliminating them. At the early detection step, it is possible to threshold the detected interest points by using a measure of importance or of how salient the interest point is. In such a manner, it is possible to avoid detecting interest points in highly homogeneous regions.

At the matching level, approaches such as the one used in the Scale-Invariant Feature Transform (SIFT), the SIFT-ratio [32], perform a comparison of similarity between the *k*-closer matches of an interest point (usually *k* is 2) and, therefore, when they are too similar in terms of descriptors, the matches are removed. At another level, approaches such as RANdom SAmple Consensus (RANSAC) [17] use a model fitting to find a transformation such that outliers are removed from the correspondence set. It is clear that all these approaches reduce the number of correspondences. In fact, the initial number of correspondences invariantly suffers a drastic reduction. Fig. 2 shows an example of correspondences obtained through the application of interest point thresholding and SIFT-ratio. A set of good matches is acquired and can be useful in many applications, such as registration. However, the correspondences are usually very sparse.

A region growing approach is presented in [6] for correspondence verification. The method outperforms the selection of SIFT. However, it is still a method that produces sparse number of correspondences. A small number of correspondences may not be enough for obtaining denser matches for some images, such as the ones considered in this paper.

Alternatively, the structural relations between interest points can be introduced to obtain more global information for the matching process. Such relations are commonly modeled by using graphs and allow the exploration of structural arrangements to better discriminate regions in images. Successful applications to find or discriminate sparse number of interest points [7,22,25,51] have been reported. However, finding correspondences within dense sets of points using graphs is still a challenging problem. Algorithms for this purpose are computationally expensive and sensitive to noise in the data [28,40,52]. Much of the efforts to solve correspondences based on graphs consider such structures to connect all interest points [4,10,14,19,40,45], which makes the limitations of graph matching even more severe when the number of interest points is dense.

Computational topology is a field gaining importance for analyzing images at qualitative, structural and abstract levels [2,5,9,15,16,29,39]. In this work, we present an approach based on the topology of functions, given by Morse complexes, to defining locally meaningful connectivity of interest points. We have devised and tested the method for obtaining correspondences in images to demonstrate its contributions:

- A general neighborhood relation which can be adapted to different applications.

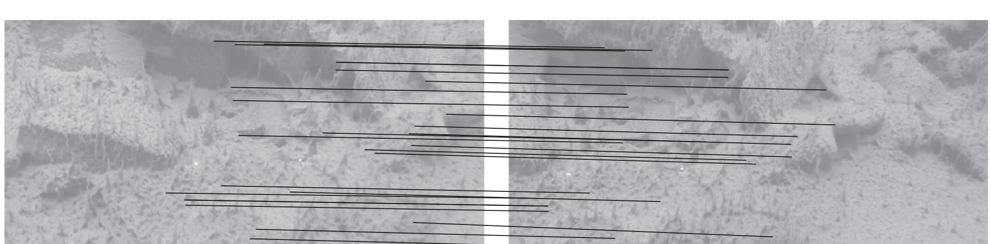

Fig. 2. Examples of output matches obtained through SIFT method. Incorrect correspondences are filtered out, however, the number of correspondences is drastically reduced and many regions do not contain any paired points.