Computer Vision and Image Understanding 142 (2016) 50–64

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Consensus of multiple correspondences between sets of elements☆,☆☆

Carlos Francisco Moreno-García, Francesc Serratosa*

Universitat Rovirai Virgili, Spain

ARTICLE INFO

Article history:
Received 26 February 2015
Accepted 20 August 2015
Available online 31 August 2015

Keywords:
Consensus
Points' correspondence
Feature extraction
Linear solver
Hamming distance
Image registration

ABSTRACT

In many pattern recognition and computer vision problems, it is often necessary to compare multiple sets of elements that are completely or partially overlapping and possibly corrupted by noise. Finding a correspondence between elements from the different sets is one of the crucial tasks that several computer vision, robotics or image registration methods have to cope with. The aim of this paper is to find a consensus correspondence between two sets of points, given several initial correspondences between these two sets. We present three different methods: iterative, voting and agglomerative. If the noise randomly affects the original data, we suppose that, while using the deducted correspondence, the process obtains better results than each individual correspondence. The different correspondences between two sets of points are obtained through different feature extractors or matching algorithms. Experimental validation shows the runtime and accuracy for the three methodologies. The agglomerative method obtains the highest accuracy compared to the other consensus methods and also the individual ones, while obtaining an acceptable runtime.

1. Introduction

Suppose we have several correspondences between sets and there is some level of intersection between them. If we want to establish a consensus correspondence between the whole correspondences, we face two main problems. First, there are discrepancies between the element mappings. If our scenario is based on an automatic method, these differences are gauged by the features or the weights on these features. Contrarily, if the scenario is based on a human-machine interaction (for instance semi-automatic medical or forensic recognition), the strategy is based on the experience of the specialist. If such elements represent regions of segmented images, some subjects may think that the area is more important than the colour, but other specialists may think differently. Second, the intersection between sets is not null although some elements belong to only one or few sets.

Several applications can benefit from our proposed solution to finding consensus correspondences. For instance, suppose some human specialists manually extracted the minutiae of a pair of fingerprints and deducted the correspondence between these minutiae. Some discrepancies can appear in these correspondences due to different localizations of the minutiae and also different mappings between minutiae. Our method can help present a final minutiae correspondence and therefore a final distance between these fingerprints.

Our method could also be applied to pattern recognition problems in medical images. In this case, the localization of some parts in an image and the correspondence between these local parts between images is based on the experience of the specialist. Again, our method could be used to deduct a final correspondence between local parts of two images. Finally, we could also use our method in an automatic framework. Suppose we want to solve the automatic image registration problem. In this case, local descriptors can be extracted from images and also different matching algorithms can be used to find correspondence between images. In this case, we could use our method to deduct a consensus correspondence with the aim of increasing the quality of image registration. In the experimental section, we have applied our method to automatic image registration.

Image registration is the process of transforming different sets of data into one coordinate system. Data may be multiple pictures, multiple views or data from different sensors or times. It is used in computer vision, medical imaging, analysing images in general and data from satellites. Registration is necessary in order to be able to compare or integrate the data obtained from these different measurements. Interesting image registration surveys are [1] and [2], which explain the problematic of this goal. Given two images to be aligned, the image registration process is usually composed of three main steps [3]. First, a set of salient points is extracted from each of the two images. Second, a correspondence between the extracted points is deducted. Third, an alignment, for instance a homography, is deducted with the initial correspondence. In this process, it is usual to deduct a final correspondence adapted to the homography.

☆ This paper has been recommended for acceptance by Xiaolei Huang.
☆☆ This research is supported by the Spanish CICYT project DPI2013-42458-P, by project TIN2013-47245-C2-2-R and by Consejo Nacional de Ciencia y Tecnología (CONACyT Mexico).
* Corresponding author.
E-mail addresses: carlosfrancisco.moreno@estudiants.urv.cat (C.F. Moreno-García), francesc.serratosa@urv.cat, francesc.serratosa@gmail.com (F. Serratosa).

http://dx.doi.org/10.1016/j.cviu.2015.08.008
1077-3142/

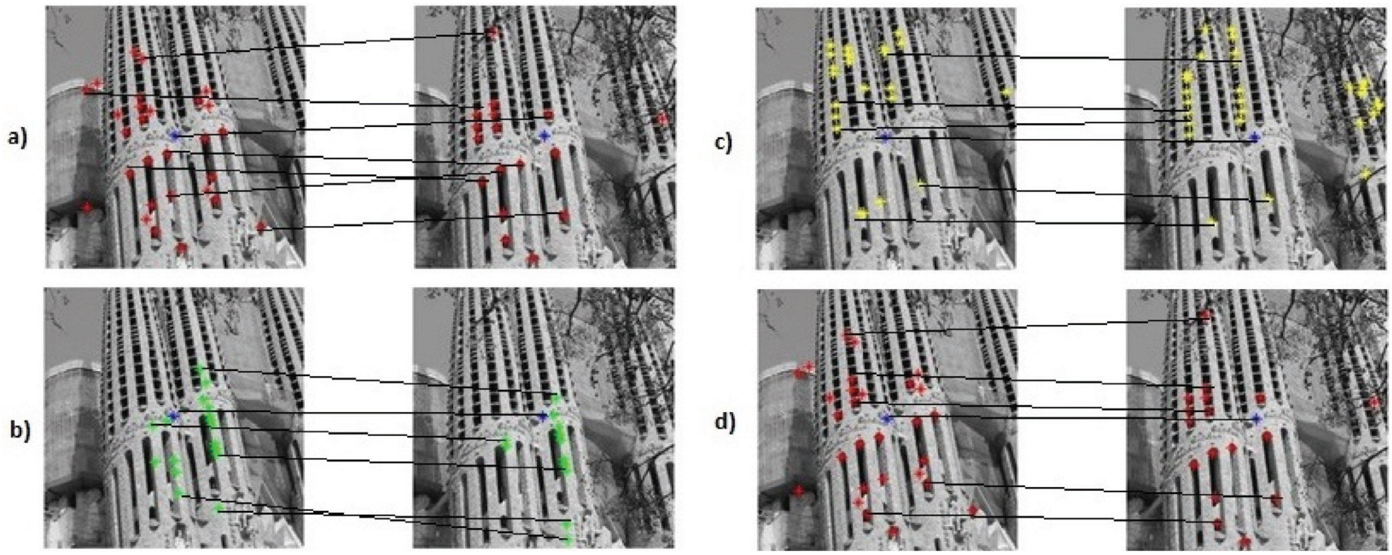

Fig. 1. Four possible salient point correspondences using different combinations of feature extractors and matching methods.

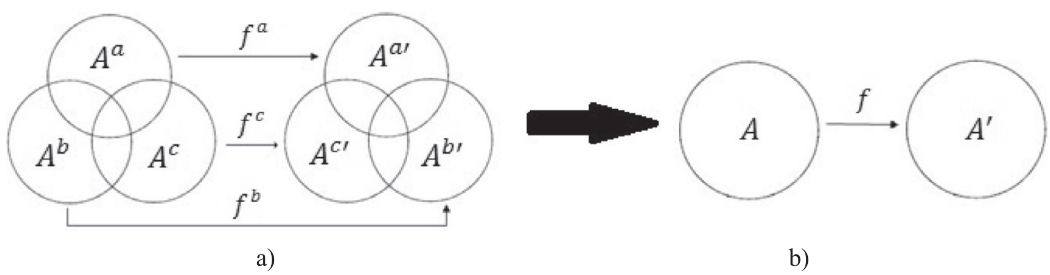

Fig. 2. (a) Input of our problem: Some correspondences between partially disjoin sets. (b) Output: Only one correspondence between two sets.

In the first step, several methods have appeared to select salient points in images [3], for example SIFT [4], Harris corners [5] or SURF [6]. These methods are based on assigning some local features (for instance, a vector of 128 features) to each extracted point or pixel of the image. Each local feature usually depends on the information on the image given a radius and an angle. The second step is based on finding a correspondence between the extracted salient points. In the second step, matching algorithms have been used with outlier rejection. That is, explicitly considering some points can be generated due to noise and so, they do not have to be mapped with elements of the other set. For instance, Bipartite (BP) [7], or a new version called Fast Bipartite (FBP) [8,9], is one of the algorithms used to find a correspondence between points or between graphs if the second order relations between points are considered. This algorithm obtains the point correspondences but it does not deduct the homography and it uses the features located at each point (for instance SIFTs or SURFs) or the second order features located at the relations between points. It is based on reducing the problem into a linear assignation problem and applying a linear solver such as the Hungarian method [10]. In the third step, the homography is extracted that transforms the coordinate system of one of the images to the other one given an initial correspondence. Iterative Closest Point (ICP) [11] is an algorithm employed to minimize the difference between two clouds of points. ICP is often used to reconstruct 2D or 3D surfaces from different scans. It only uses the position of the points but not the local features and an initial correspondence. It is usual to use ICP together with RANSAC [12], which is a method to discard points that do not fit on the deducted homography and correspondences and so eliminates the spurious correspondences. That is, points to be considered that have appeared due to noise in the images or sensors. Other more sophisticated algorithms have appeared that consider the features of each point and also the homographies such as [13], which is based on the Expectation Maximisation algorithm. In [14], they propose a method to deduct the vector field given two images and also the best correspondence between salient points. Finally, the Hough transform [15–17] is a technique used to find imperfect instances of objects represented by sub-sets of salient points within an image by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform.

Fig. 1 shows two images in which four different sets of points and correspondences have been found. The salient points extractors and matching algorithms are: 1.(a) SIFT extractor [4] and Hungarian method [10]. 1.(b) SURF extractor [6] and Hungarian method [10]. 1.(c) Harris corners [5] and *matchFeatures* function from Matlab [18]. 1.(d) SIFT extractor [4] and PF-Registration [19,20].

Several mapping combinations have been formed and all of them containing mistaken mappings. Nevertheless, due to the noisy nature of the errors, mistaken mappings tend to be non-repetitive. For this reason, if a consensus correspondence is defined, the final correspondence tends to have less mapping errors than the original ones. Moreover, the final sets of points are the union of the points in all the sets. Therefore, the consensus correspondence has the advantage of being composed by a larger set of point correspondences. With a larger set of correct mappings, the image registration process (for instance ICP + RANSAC) tends to obtain a better image alignment. Fig. 2 schematically shows the consensus method. In this case, we suppose there are three different correspondences f^a, f^b and f^c with their pairs of sets. The intersection of sets is not null. Our method deducts the sets A and A' and the consensus correspondence f.

A method to deduct a correspondence consensus given only two correspondences was presented in [21] and [22] for sets and in [23] for graphs. In the current paper, we present a method to deduct a