Computer Vision and Image Understanding 118 (2014) 84–96

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Bayesian perspective for the registration of multiple 3D views ☆

X. Mateo [a,c,*], X. Orriols [b], X. Binefa [c]

[a] Department of Computer Science, Universitat Autònoma de Barcelona, Barcelona, Spain
[b] Microsoft Search Technology Center Asia, Beijing, China
[c] Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

ARTICLE INFO

Article history:
Received 19 September 2011
Accepted 10 September 2013
Available online 20 September 2013

Keywords:
3D registration
Pairwise correspondences
Expectation Maximization algorithm

ABSTRACT

The registration of multiple 3D structures in order to obtain a full-side representation of a scene is a long-time studied subject. Even if the multiple pairwise registrations are almost correct, usually the concatenation of them along a cycle produces a non-satisfactory result at the end of the process due to the accumulation of the small errors. Obviously, the situation can still be worse if, in addition, we have incorrect pairwise correspondences between the views. In this paper, we embed the problem of global multiple views registration into a Bayesian framework, by means of an Expectation–Maximization (EM) algorithm, where pairwise correspondences are treated as missing data and, therefore, inferred through a maximum a posteriori (MAP) process. The presented formulation simultaneously considers uncertainty on pairwise correspondences and noise, allowing a final result which outperforms, in terms of accuracy and robustness, other state-of-the-art algorithms. Experimental results show a reliability analysis of the presented algorithm with respect to the percentage of a priori incorrect correspondences and their consequent effect on the global registration estimation. This analysis compares current state-of-the-art global registration methods with our formulation revealing that the introduction of a Bayesian formulation allows reaching configurations with a lower minimum of the global cost function.

1. Introduction

Current technology allows to obtain 3D representations of real objects or scenarios by using different methods. From usual methods like stereoscopy to complex devices like LADAR or time-of-flight cameras, the possibilities are growing continuously. However, acquisition techniques usually have problems with occluded surfaces or the limited field of view, so it is usually necessary to combine different views of the same object or scenario in order to obtain a full representation. Using this process another problem then arises: the registration of these individual 3D views which will enable, at the end of the process, to a whole 3D reconstruction of the desired object or scene.

First step for this objective is the so-called pairwise registration, where the 3D representations are registered pair to pair. Depending if the two structures have overlap, pairwise registration methods will give as a result a transformation which registers the first 3D view to the second one.

Multiview registration is the second step of this process, and is usually a more complex situation. Assuming that pairwise registrations are correct, their concatenation along a cycle will probably result in a non-satisfactory multiview registration because of the accumulation of the different pairwise errors. An example of this effect can be seen in Fig. 1.

In addition, there could exist another situation which produces more problems. Even if two structures register perfectly in the pairwise registration process, their transformation could be incorrect in a global environment. This could happen specially if we are working with objects with symmetries, planes or repetitive patterns. In these cases, most of the current multiview registration algorithms will fail because they are not ready to deal with this kind of error.

The main contribution of this paper, in opposition to other state-of-the-art papers, is the possibility of detecting these incorrect registrations between different views and therefore minimize their impact in the global registration process. This feature is achieved thanks to the use of different weights which encode the reliability we have in the correspondences between the views.

The structure of the document is as follows: a review of different registration algorithms is presented in Section 2, followed by an introduction to the main problematic of multiview registration in Section 3. The proposed algorithm is explained in detail in Section 4, and the obtained experimental results are shown in Section 5. Finally, conclusions and possible future improvements are explained at the end of the paper in Section 6.

☆ This paper has been recommended for acceptance by K.L. Boyer

* Corresponding author at: Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain.

E-mail addresses: javier.mateo@uab.es, javier.mateo@upf.edu (X. Mateo), xaorriol@microsoft.com (X. Orriols), xavier.binefa@upf.edu (X. Binefa).

1077-3142/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2013.09.003

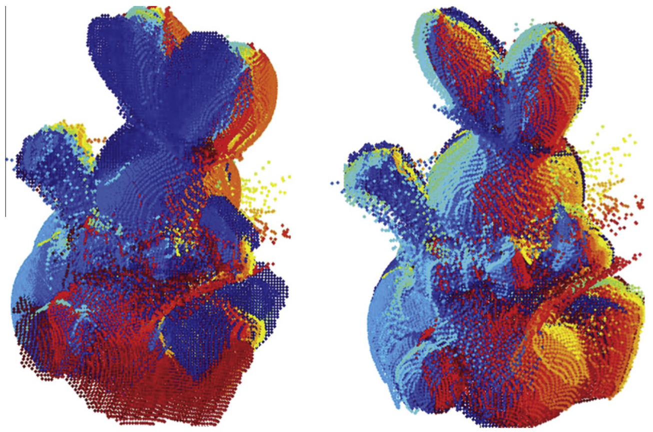

Fig. 1. Left side: result after applying the pairwise information for all the views of an object (each color represents a different view). Right side: desired result, where only the noise produced by the sensor can be appreciated.

2. State of the art

As previously stated the first step in order to achieve the registration of multiple 3D representations is the computation of the pairwise registrations. The most usual method consists in the establishment of correspondences between points of the two views, which can be found by using descriptors like the Spin-Images of Johnson and Hebert [1] or covariance descriptors like the one presented by Fehr et al. in [2]. The combination of correspondences between the two views determines an approximated transformation between them, having a minimum of 3 point correspondences in order to obtain it. This step is known as coarse pairwise registration. After this approximation, the transformation is refined by the so-called fine pairwise registration methods, which minimize the distance between all the points of both point clouds. The most representative examples of fine coarse registration methods are the Iterative Closest Point [3] presented by Besl and McKay in 1992 and the registration algorithm developed by Chen and Medioni [4]. At the end of this pairwise registration execution we will obtain a rigid transformation, composed by a rotation matrix and a translation vector, which registers the first view to the second one.

With respect to the multiview registration algorithms, they can be basically divided into two families: sequential registration methods and simultaneous registration methods.

The first ones consist in the sequential pairwise registration of additional views to an aggregated view which is continuously growing. This aggregated view is also denominated metaview in some literature, as stated in Matabosch et al. [5]. This method was initially proposed in the aforesaid paper of Chen and Medioni [4] and afterwards improved in the works presented by Pulli [6] and Nuchter et al. [7]. This kind of multiview registration methods have the advantage that they do not need to previously obtain all the views implicated in the registration process, so they do not suffer problems of memory in their execution and can also be useful for applications where the 3D structures are acquired at the same time as the registration is produced, achieving a performance which could be considered near to real-time. However, their main problem is the non-possibility of modifying the already registered views, producing a final result which is usually not as ideal as desired.

On the other side, simultaneous registration methods make use of all the information of pairwise registrations at the same time. A good example of this kind of method is the work presented by Eggert et al. [8], where the authors present an iterative algorithm which simultaneously updates the transformations of all the views by using the information of position and normal of the correspondence points. In Silva et al. [9] the multiview registration is achieved thanks to the use of genetic algorithms and a metric defined by the authors called Surface Interpenetration Measure, which indicates the level of confidence on a registration according to the multiple crossings of the surfaces between themselves (i.e. the interpenetration between themselves).

Inside the category of simultaneous registration methods a high number of papers base their algorithm on the use of a registration graph which encodes the different pairwise relations between the 3D views. Each node on a registration graph represents an specific view of the object, while each edge encodes the rigid transformation (rotation and translation) between two views. As can be seen in the registration graph of Fig. 2, each node is connected, at least, with the previous and the following node of the sequence of views and, in addition, with a variable number of additional nodes. The establishment of these additional edges to non-contiguous nodes may depend on the result of the automatic pairwise registration between the views (if the registration quality is lower than a pre-defined threshold, this registration is discarded) or also on the preferences of the user.

In order to obtain the multiview registration some properties of the registration graph are usually applied, like the property that rotations and translations along a loop of the registration graph should be null. This is the basis for the work presented by Sharp et al. [10], where, for the different basic cycles of the registration graph, the error in rotation and translation is distributed along all the edges. Also the concept of graphs minimization is used in Shih et al. [11] where, using concepts of Lie algebra and circuit theory, the authors develop an algorithm which achieves good results in standard databases, also with the advantage of having a low computational cost.

Other papers do not mention explicitly the concept of graph but use it intrinsically, using the same properties and ideas. A good example are the two papers presented by Krishnan et al. [12,13] which are based on the notion that, concatenating translations and rotations along a path, the same 3D view should be obtained independently on the direction we use to arrive to it. One particularity of this method is that it directly works with the correspondences itself and not with the rigid transformations obtained from the pairwise registration methods, so the algorithm uses all the data without losing any information. This fact causes an apparent complexity of the formulation, but using some concepts of Lie algebra and manifolds the authors achieve a compact representation of the problem. These two papers of Krishnan et al. are used as the basis of our algorithm due to the interesting conceptualization of the problem and the good results obtained after their implementation.

Independently of the election of using a sequential registration method or a simultaneous registration method, one of the main lacks of the aforesaid papers and the majority of other state-of-the-art publications is that they do not take into account the possibility of having absolutely bad registrations or bad correspondences. One of the few papers that deal with this possibility is the work presented by Hubert and Hebert in 2003 [14], where using the joint distribution probabilities the incorrect correspondences are eliminated and, in consequence, the graph can be separated into different splitted sub-graphs which can afterwards be studied individually. As developed in the following sections of the present paper, our algorithm will also serve for the detection of incorrect correspondences between the views, allowing to minimize the impact of this incorrect information into the global registration process.

3. Problematic issues in the multiview registration process

The most usual problem related with the multiview registration process is the minimization of the global error which is produced