# Improving piecewise-linear registration through mesh optimization

Vicente Arévalo and Javier González

Dept. of System Engineering and Automation,
University of Málaga, Campus Teatinos, 29071, Málaga (Spain)
varevalo@ctima.uma.es, jgonzalez@ctima.uma.es

**Abstract.** Piecewise-linear methods accomplish the registration by dividing the images in corresponding triangular patches, which are individually mapped through affine transformations. For this process to be successful, every pair of corresponding patches must lie on projections of a 3D plane surface; otherwise, the registration may generate undesirable artifacts, such as broken lines, which diminish the registration quality. This paper presents a new technique for improving the registration consistency by automatically refining the topology of the corresponding triangular meshes used by this method. Our approach iteratively modifies the connectivity of the meshes by swapping edges. For detecting the edges to be swapped, we analyze the local registration consistency before and after applying the action, employing for that the mutual information ($MI$), a metric for registration consistency significantly more robust than other well-known metrics such as normalized cross correlation ($NCC$) or sum of square differences ($SSD$). The proposed method has been successfully tested with different sets of test images, both synthetic and real.

## 1 Introduction

Image registration is the process of overlapping two images of the same scene acquired on different dates, from differences point of views and/or using different sensors. In this process, one image remains fixed (*fixed* image) whereas the other (*moving* image) is spatially transformed until fitting with the first one. Image registration is a crucial step in many image analysis applications like image fusion, change detection, 3D scene reconstruction, etc. Traditionally, the registration process is dealt with in two stages. In the first one, the positions of a set of pairs of corresponding points (so-called correspondences) are identified in the images, and in the second stage, this set of correspondence pairs is exploited to robustly estimate a mapping function which is then used to transform all the pixels of the moving image onto the fixed one (some kind of interpolation is required in this step).

Different mapping functions have been reported in the literature for image registration, such as polynomial, radial basis, piecewise (linear or cubic), splines, etc. [1]. For registering images of polyhedral scenes (typical in indoor and urban environments), piecewise-linear functions are especially suitable, since they


3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07),
Part II, LNCS, vol. 4478, pp. 122-129, 2007.

divide the images into triangles which are individually registered through linear transformations that preserve the topology of the triangular mesh [2]. Of particular significance is the case where the perspective deformation of the images can be simplified by an affine transformation, since a triangle in the moving image must perfectly overlap onto the fixed one provided that it comes from the projection of a planar patch of the scene [3].

Given a set of corresponding point pairs in the images, isomorphic triangular meshes are typically generated onto them by using the Delaunay's triangulation method [4], which produces triangles of balanced size and shape, but which does not guarantee that the created topology is the best possible one for registering the images through a piecewise-linear method. For that purpose, it is clear that we should minimize the number of triangles covering on projections of different planar 3D patches, that is, those whose vertices are projections of 3D points of different planar patches (see fig. 1). This is the aim of this work: to improve the accuracy of piecewise-linear image registration by only applying edge swapping modifications to the mesh. This process can be seen as an optimization procedure that modifies the mesh connectivity, that is, without varying the number of vertices neither their coordinates. It is remarkable also that, the resulting optimized mesh is in compliance with the 3D scene structure up to the level that the mesh geometrical realization allows. To our knowledge, this is a novel approach for the image registration problem, since previous methods reported in the literature focus on optimization/simplification of 3D triangular meshes, requiring a complete knowledge of the scene geometry derived, for example, from a laser range finder [5][6] or calibrated images [7][8].

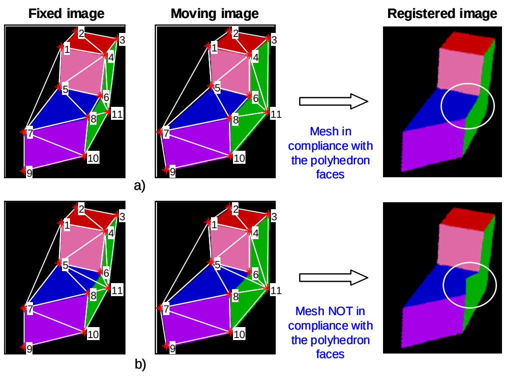


**Fig. 1.** For a piecewise-linear registration process to be successful, the triangles must be projections of one single polyhedral face of the scene as in (a), otherwise broken lines are produced and the registration of that triangle shows a clear inconsistency (b).

A key aspect in the proposed optimization method is that of determining when an edge swapping operation is necessary. Our solution consists of checking the local registration consistency of the two triangles involved (those that share the analyzed edge) before and after performing the swap. In this process, no threshold needs to be considered. Another novelty of this work is the usage



3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.

of the mutual information ($MI$) as a measurement of registration consistency [9] which, unlike other well-known metrics such as normalized cross correlation ($NCC$) or sum of square differences ($SSD$), is less sensitive to changes in lighting conditions or noise. The overall registration method has been successfully tested with a broad variety of test images (both synthetic and real) acquired under different lighting conditions and viewpoints.

The remainder of this paper is organized as follows. Section 2 contains several assumptions and definitions, as well as, the formulation used in subsequent sections. In section 3, we describe our method, the inconsistency estimation function and the optimization process. In section 4, we present and discuss some experimental results. Finally, some conclusions and future work are outlined.

## 2 Assumptions and definitions

In this work we assume that the 3D-to-2D camera projection can be modelled by a paraperspective transformation which basically means that parallel lines in space keep their parallelism in the image. This simplification is assumable in most computer vision setups and leads to a great reduction in complexity in many vision problems [10]. For image registration, this assumption implies that 3 correspondences (instead of the 4 correspondences required for its general form) suffice to estimate the affinity which transfers points from one image patch to another [3]. In other words, if a pair of corresponding faces are projections of a plane surface, the geometric transformation which maps the pixels of one to another is an affinity. Thus, after performing the mapping, both image patches should perfectly match; otherwise, the faces are not projections of a planar surface.

Next, we introduce the notation employed in this work as well as some useful definitions.

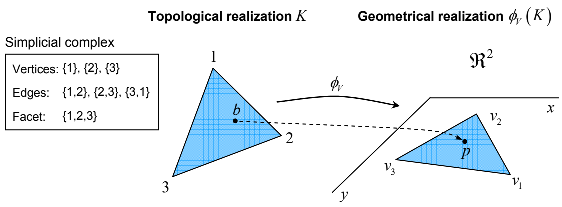


**Fig. 2.** Example of mesh representation: a mesh consisting of one face.

A mesh is a piecewise-linear surface, consisting of triangular faces put together along their edges. Formally, a mesh is a pair $M = (K, V)$, where $K$ is a structure, called simplicial complex [11], which determines the connectivity of the vertices, edges and faces (its *topological realization*), and $V = \{v_i | i = 1, \ldots, m\}$, $v_i \in \Re^2$ is a set of vertex positions which defines the shape of the mesh in $\Re^2$ (its *geometrical realization*) [6] (see fig. 2). To refer to any point within the mesh, we employ the notation $p \in \phi_V(s)$, where $s \subseteq K$, thus, we use $p \in \phi_V(t)$ to refer to one point within a triangular face $t = \{i, j, k\} \in K$; $p \in \phi_V(q)$ to refer to one point within a quadrilateral of $M$ consisting of two adjacent triangles $q = [\{i, j, k\}, \{i, j, l\}] \in K$, and so on.

In addition to the above general definition, we introduce the following ones, of interest for describing our method in the next section:

- An edge $\{i, j\} \in K$ is *external* or *boundary* if it is a subset of only one face, and *internal* or *shared* otherwise.
- An edge $\{i, j\} \in K$ is *3D-compatible* if it lies on a projection of a 3D plane surface, and *3D-incompatible* otherwise.
- Given a set of point correspondences $\{(v_i, v'_i) | i = 1, \ldots, n\}$, $v_i \in V$ and $v'_i \in V'$ identified in two images, two isomorphic triangular meshes $M = (K, V)$ and $M' = (K, V')$, and a simplicial complex $s \subseteq K$, we define the piecewise-linear function $\boldsymbol{f}$ which geometrically maps a point $p \in \phi_V(s)$ to another $p' \in \phi_{V'}(s)$ as follows:

$$p' = \boldsymbol{f}_{\phi_V(s)}(p) = \begin{cases} f_1(p) \;\, if \;\, p \in \phi_V(t_1) \\ \qquad \vdots \\ f_m(p) \;\, if \;\, p \in \phi_V(t_m) \end{cases} \tag{1}$$

where $t_i = \{j, k, l\} \in s$; $f_i$ is an affinity estimated from the geometrical realization of the vertices of $t_i$ in both meshes, namely the point pairs $(v_j, v'_j)$, $(v_k, v'_k)$, and $(v_l, v'_l)$; and $m$ is the number of triangular faces.

Notice that once the transformation has been applied $\phi_V(s) = \phi_{V'}(s)$, that is, the corresponding faces of both meshes must perfectly overlap.

## 3 Description of the proposed method

The method presented in this paper is aimed to improve the accuracy of piecewise-linear registration, especially when applied to images of polyhedral scenes. For this purpose, we iteratively modify the connectivity of the triangular meshes by swapping 3D-incompatible edges (see fig. 3(a)). To detect such edges our algorithm checks, before and after applying the swap, the registration consistency of the two triangles that share the analyzed edge: the edge is swapped if that operation leads to a registration improvement. Notice that this procedure only modifies the mesh connectivity, since the number of vertices and their coordinates remain without modification.

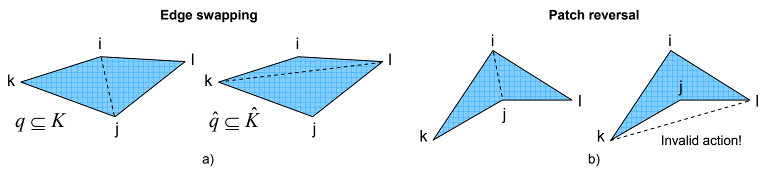


**Fig. 3.** The topological action of swapping an edge when a) all preconditions are verified and b) the action produces a patch reversal.



3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.

The employed metric for measuring the registration consistency is the mutual information ($MI$) [12]. From a statistical viewpoint, the $MI$ measures the statistical dependency or information redundancy of two random variables. Unlike other consistency measures such as the sum of square differences ($SSD$) or the normalized cross correlation ($NCC$) which assume a priori functional relationship between both image patches, the $MI$ postulates a statistical relationship which can be estimated from the joint entropy. The advantage of this metric is that it is more robust to image changes caused by different lighting conditions, observation angles, noise, etc. [13]. Mathematically, the $MI$ of two image patches $A$ and $B$ can be written as follows:

$$MI(A,B) = \sum_i \sum_j P_{A,B}\left(i,j\right) \log\left(\frac{P_{A,B}\left(i,j\right)}{P_A\left(i\right) P_B\left(j\right)}\right) \tag{2}$$

where $P_A(i)$, $P_B(j)$ and $P_{A,B}(i,j)$ are the probability functions estimated from the intensity joint histogram of $A$ and $B$ ($h_{A,B}$), that is:

$$\begin{aligned} &P_A\left(i\right) = \textstyle\sum_j h_{A,B}\left(i,j\right)/N, \\ &P_B\left(j\right) = \textstyle\sum_i h_{A,B}\left(i,j\right)/N, \text{ and} \\ &P_{A,B}\left(i,j\right) = \textstyle\sum_i \sum_j h_{A,B}\left(i,j\right)/N \end{aligned}$$

being $N$ the number of pixels.

We take advantage of the robustness of the $MI$ for effectively detecting 3D-incompatible edges. Thus, given two images $I$ and $I'$ to register and their corresponding meshes defined as $M = (K, V)$ and $M' = (K, V')$, we determine the 3D-compatibility of an edge $\{i, j\} \in K$ by measuring the improvement in consistency, before and after being swapped, through the following expression:

$$\omega\left(\{i,j\}\right) = MI\left(I\left(r\right), I'\left(\boldsymbol{f}_{\phi_V(\hat{q})}\left(r\right)\right)\right) - MI\left(I\left(r\right), I'\left(\boldsymbol{f}_{\phi_V(q)}\left(r\right)\right)\right) \tag{3}$$

where $r = \phi_V(q) \equiv \phi_V(\hat{q})$ are the pixels contained in $\phi_V(q)$ or $\phi_V(\hat{q})$, being $q = [\{i,j,k\}, \{i,j,l\}]$ and $\hat{q} = [\{l,k,j\}, \{l,k,i\}]$ the two adjacent faces, before and after the swapping, respectively. Thus, $I(r)$ represents the patch of the fixed image defined by $q$, and $I'(\boldsymbol{f}_{\phi_V(q)}(r))$ and $I'(\boldsymbol{f}_{\phi_V(\hat{q})}(r))$ the transformations of its moving counter parts according to the two possible topological configurations.

An edge is considered for swapping only if $\omega > 0$, otherwise, the topological realization of the meshes remains without modification. Also, before evaluating the 3D-compatibility of any edge $\{i, j\} \in K$, the edge should be checked to verify the following preconditions: 1) the edge $\{i, j\}$ is internal, 2) the resultant edge $\{k, l\} \notin K$, and 3) the action does not produce a patch reversal in $\hat{K}$ (see fig. 3(b)). It is important to notice that, in this process, (3) is used only for comparison, so no threshold needs to be applied in this procedure.

The overall optimization process is formulated as a *greedy* search [14], which starts with the two images $I$ and $I'$ to register, and the initial corresponding triangular meshes $M$ and $M'$ resulting of triangulating (by means of the Delaunay's method) a set of point pairs identified in both images. The process finishes


3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.

when the topological realization can not be longer improved by the greedy algorithm.

## 4 Experimental tests

In this section we show some experimental results which illustrate the performance of our approach. Most of the images considered in our experiments belongs to the ALOI library [15], which includes images of 1000 objects acquired under different viewpoints and lighting conditions. We have also evaluated our implementation with scenes more complex, where several different objects are put together.

Fig. 4 graphically illustrates the process described in section 3 when applied to two image pairs of polyhedral scenes. This figure shows the isomorphic meshes automatically generated from sets of corresponding points previously identified in each of the image pairs (see fig. 4(a)), and the optimized ones once the refinements have been accomplished (see fig. 4(b)). With the aim of showing the benefits of using the *MI*, we have repeated the experiments twice: firstly, employing (3), and secondly, replacing the *MI* by the *NCC*. The results of these experiments are summarized in table 1. They reveal the advantage of the *MI* against *NCC* for driving the optimization process, concretely: an improvement in the accuracy of the piecewise-linear registration process (see also fig. 4(c)) and a reduction in the computational time.

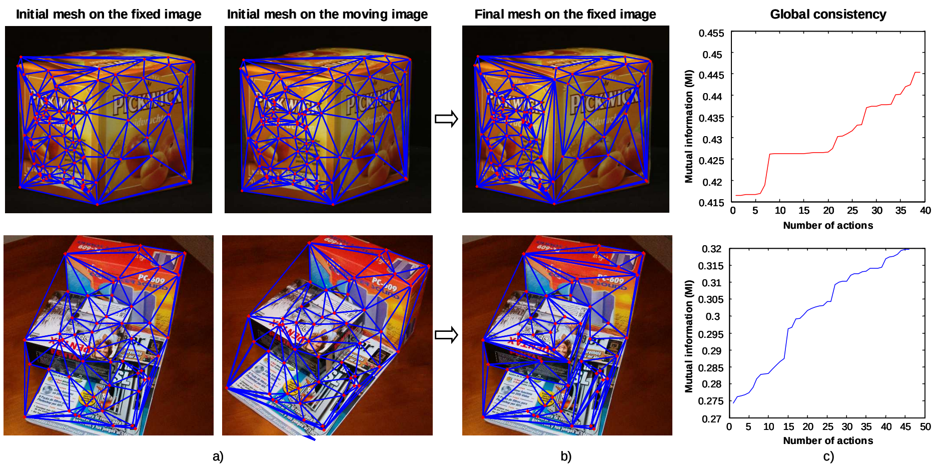


**Fig. 4.** (a) Real images of polyhedron scenes and their corresponding Delaunay triangular meshes. (b) Optimized triangular meshes provided by our method. Observe how the process swaps edges which go from one plane surface of the scene to another. (c) Overall registration consistency during the optimization process. The flat intervals mean that the actions performed there do not lead to significant improvements, though they carry out suitable topological changes that are exploited in subsequent iterations, as shown in the evolution of the curves.


3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.

**Table 1.** Experiment results.

| | *MI* | | *NCC* | |
|---|---|---|---|---|
| *Scene (# of edges)* | *Correctness*[1] (%) | *CPU time*[2] (sec.) | *Correctness* | *CPU time* |
| Cube (275) | 100 | 23.89 | 98.88 | 29.39 |
| Stacked boxes (140) | 99.28 | 18.56 | 93.23 | 21.34 |

Finally, with the purpose of showing that the optimization process ends up with meshes in compliance with the 3D scene structure (obviously, limited by the initial set of corresponding points), in figure 5, we have re-projected them into 3D space employing the factorization algorithm for affine reconstruction proposed in [3] (pag. 437). It can be clearly observed the undesirable artifacts which appear when the mesh contains 3D-incompatible edges.

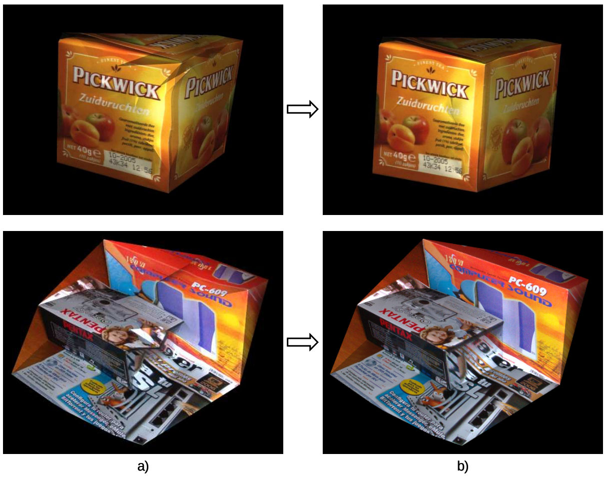


**Fig. 5.** 3D scene reconstructions generated from two meshes: (a) the initial mesh and (b) the refined one. In plots (a) we can observe some artifacts in those places where edges not in compliance with the 3D scene exist. These artifacts disappear when all edges are conveniently swapped, as showed in plots (b).

## 5 Conclusions and future work

In this paper we have proposed a new technique for automatically optimizing the triangular mesh employed by piecewise-linear registration process in order to improve the registration consistency. To achieve that, we iteratively modify the connectivity of both meshes through edge swapping actions. The function employed for evaluating the edge to be swapped is based on the *MI*, which is significantly more robust than other well-known metric such as *NCC*, since it is less sensitive to changes in lighting conditions or noise. The optimization

[1] Percentage of 3D-compatible edges, which are not boundary edges.
[2] We have employed Matlab on a Pentium 4 HT 2.6GHz for implementing the tests.


3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.

procedure is formulated as a greedy search which finishes when all mesh edges have been swapped. The proposed method has been successfully tested with different sets of test images acquired under different conditions (from different angles and lighting conditions) and sensors.

In spite of the achieved results, we have detected mesh configurations where the registration consistency can not be improved. Such configurations occur when the vertices of the mesh are not well-localized (i.e. in the central part of the faces). In these cases, additional actions should be considered, for example, edge splitting. Unlike edge swapping, it involves changes in both, the topological and geometrical realizations of the meshes, making the optimization process significantly more complex and time demanding, and generating new challenges such as, where the new vertices should be located or what is the best way of splitting an edge. This is one of our concerns for future work.

## References


1. Zitová, B., Flusser, J.: Image registration methods: A survey. Image and Vision Computing **21**(11) (2003) 977–1000
2. Goshtasby, A.: Piecewise linear mapping functions for image registration. Pattern Recognition **19**(6) (1986) 459–466
3. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, Cambridge, UK (2004)
4. Shewchuk, J.R.: Lecture notes on Delaunay mesh generation. Technical Report 3, University of California at Berkeley (1999)
5. Hoppe, H.: Progressive meshes. Computer Graphics **30**(Annual Conference Series) (1996) 99–108
6. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Mesh optimization. Computer Graphics **27**(Annual Conference Series) (1993) 19–26
7. Morris, D.D., Kanade, T.: Image-consistent surface triangulation. In: Computer Vision and Pattern Recognition (CVPR 2000). (2000) 332–338
8. Vogiatzis, G., Torr, P., Cipolla, R.: Bayesian stochastic mesh optimisation for 3D reconstruction. In: British Machine Vision Conference (BMVC 2003). (2003) 711–718
9. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. Computer Vision, International Journal of **24**(2) (1997) 137–154
10. Xu, G., Zhang, Z.: Epipolar Geometry in Stereo, Motion, and Object Recognition: A Unified Approach. First edn. Kluwer Academic Publishers, Norwell, MA, USA (1996)
11. Spanier, E.H.: Algebraic Topology. First edn. McGraw-Hill, New York, NY, USA (1966)
12. Cover, T.M., Thomas, J.A.: Elements of Information Theory. First edn. John Wiley & Sons, Inc., New York, NY, USA (1991)
13. Chen, H., Varshney, P., Arora, M.: Performance of mutual information similarity measure for registration of multitemporal remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on **41**(11) (2003) 2445–2454
14. Strang, G.: Introduction to Applied Mathematics. First edn. Wellesley-Cambridge Press, Wellesley, MA, USA (1986)
15. Geusebroek, J.M., Burghouts, G.J., Smeulders, A.W.M.: The Amsterdam library of object images. Computer Vision, International Journal of **61**(1) (2005) 103–112




3rd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'07), Part II, LNCS, vol. 4478, pp. 122-129, 2007.