Improving Piecewise Linear Registration of High-Resolution Satellite Images Through Mesh Optimization

Vicente Arévalo and Javier González, *Associate Member, IEEE*

***Abstract*—Piecewise linear transformation is a powerful technique for coping with the registration of images affected by local geometric distortions, as it is usually the case of high-resolution satellite images. A key point when applying this technique is to divide the images to register according to a suitable common triangular mesh. This comprises two different aspects: where to place the mesh vertices (i.e., the mesh geometrical realization) and to set an appropriate topology upon these vertices (i.e., the mesh topological realization). This paper focuses on the latter and presents a novel method that improves the registration of two images by an iterative optimization process that modifies the mesh connectivity by swapping edges. For detecting if an edge needs to be swapped or not, we evaluate the registration improvement of that action on the two triangles connected by the edge. Another contribution of our proposal is the use of the mutual information for measuring the registration consistency within the optimization process, which provides more robustness to image changes than other well-known metrics such as normalized cross-correlation or sum of square differences. The proposed method has been successfully tested with different pairs of panchromatic QuickBird images (0.6 m/pixel of spatial resolution) of a variety of land covers (urban, residential, and rural) acquired under different lighting conditions and viewpoints.**

***Index Terms*—High-resolution satellite images, mesh optimization, piecewise linear (PWL) registration.**

I. Introduction

IMAGE registration is an essential step in many remote sensing applications like image fusion, change detection, 3-D scene reconstruction, etc. In this process, one image remains fixed (the reference image), whereas the other (the input or moving image), which is acquired on a different date, from a different viewpoint and/or using a different sensor, is spatially transformed until fitting with the first one. Traditionally, the registration process is dealt with in two stages. In the first one, the positions of a set of pairs of corresponding points 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 applied to transform all the pixels of the input image onto the reference one (some kind of interpolation is required in this step) [1].

A variety of mapping functions have been reported in the literature for image registration, including polynomial [2], radial basis functions [3], piecewise linear (PWL) [4] or piecewise-cubic [5] functions, multiquadric functions [6], B-spline functions [7], etc. In remote sensing, global polynomial (POL) functions usually perform well with low and medium resolution images (National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer, EarthSat, Landsat, Indian Remote Sensing, etc.) but may not be powerful enough to register high-resolution ones (QuickBird, Ikonos, or OrbView), where geometric distortions may become important to attain short revisit time, and high-resolution satellites pitch along their orbit to observe the scene from off nadir; thus, two (temporal) images of a certain scene may have been acquired from quite different angles, which entails large local geometric differences, particularly in urban scenarios and high-relief terrain. For registering such images, PWL functions are particularly suitable (as revealed in [8]), because they divide the images into a mesh of triangular patches, which are individually registered through linear transformations.

In this registration method, it is of particular relevance—the case where the camera projection can be approximated by a paraperspective one. Under this assumption, provided that two corresponding image triangles come from the projection of a planar patch of the scene, they must perfectly overlap for a certain affine transformation [9] (see Fig. 1).

It is clear that, in order to meet such desirable condition for all mesh triangles, the mesh cannot be an arbitrary one, but it must fulfil some geometrical and topological requirements. Current approaches for PWL registration (including those commercially available in packages such as ERDAS, ENVI, PCI, etc.) create the triangular meshes for the two images from a set of correspondence (conjugate) pairs, which are localized either automatically or by hand. Typically, they are generated by using the Delaunay's triangulation method [10] (or other similar one), which produces triangles of balanced size and shape but are not optimal for covering as many planar patches as possible (see Fig. 2). The aim of this paper is to modify the topology (not the vertices) of a given initial mesh by iteratively swapping its edges in order to improve the global registration of a pair of images. This process can be seen as an optimization procedure that, at each step, focuses on a particular edge and improves the registration consistency of the quadrilateral formed by the

Manuscript received March 29, 2007; revised October 25, 2007. Current version published October 30, 2008. This work was supported in part by the Spanish Government under Research Contracts CICYT DPI-2005-01391 and AECI PCI-A-7286-06.

The authors are with the Department of System Engineering and Automation, University of Málaga, 29071 Málaga, Spain (e-mail: varevalo@ctima.uma.es; jgonzalez@ctima.uma.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2008.924003

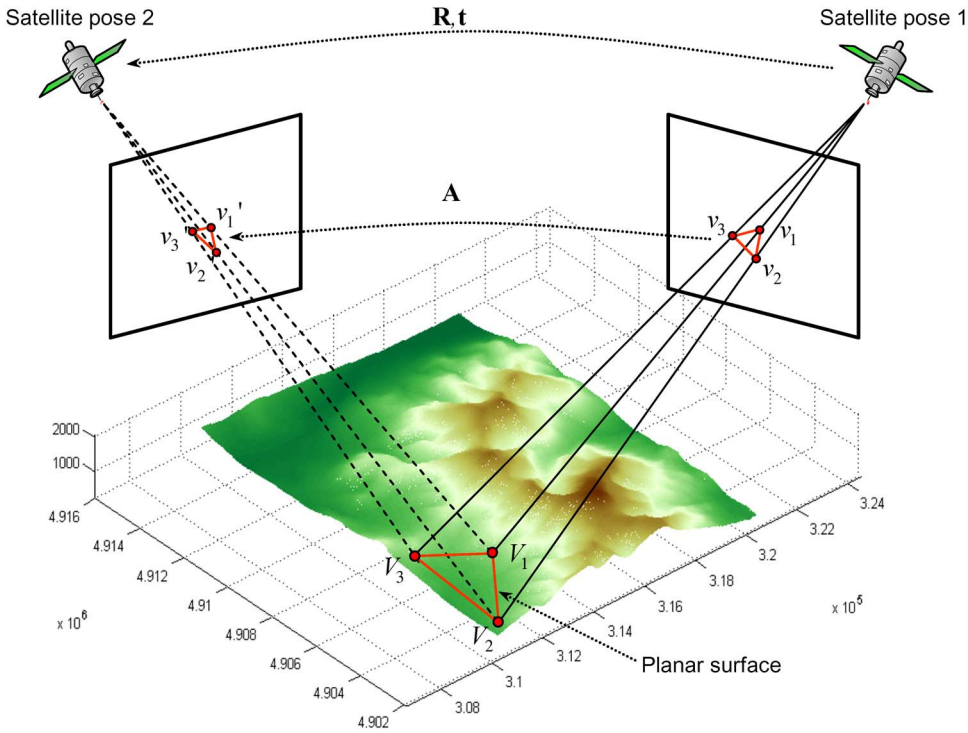

Fig. 1. Paraperspective camera projection. In satellite remote sensing, the perspective projection can be approximated by a paraperspective or parallel projection. For image registration, this simplification implies that three correspondences (instead of the four ones required for its general form) suffice to estimate the affinity ($\mathbf{A}$), which transfers points from one image patch to another.

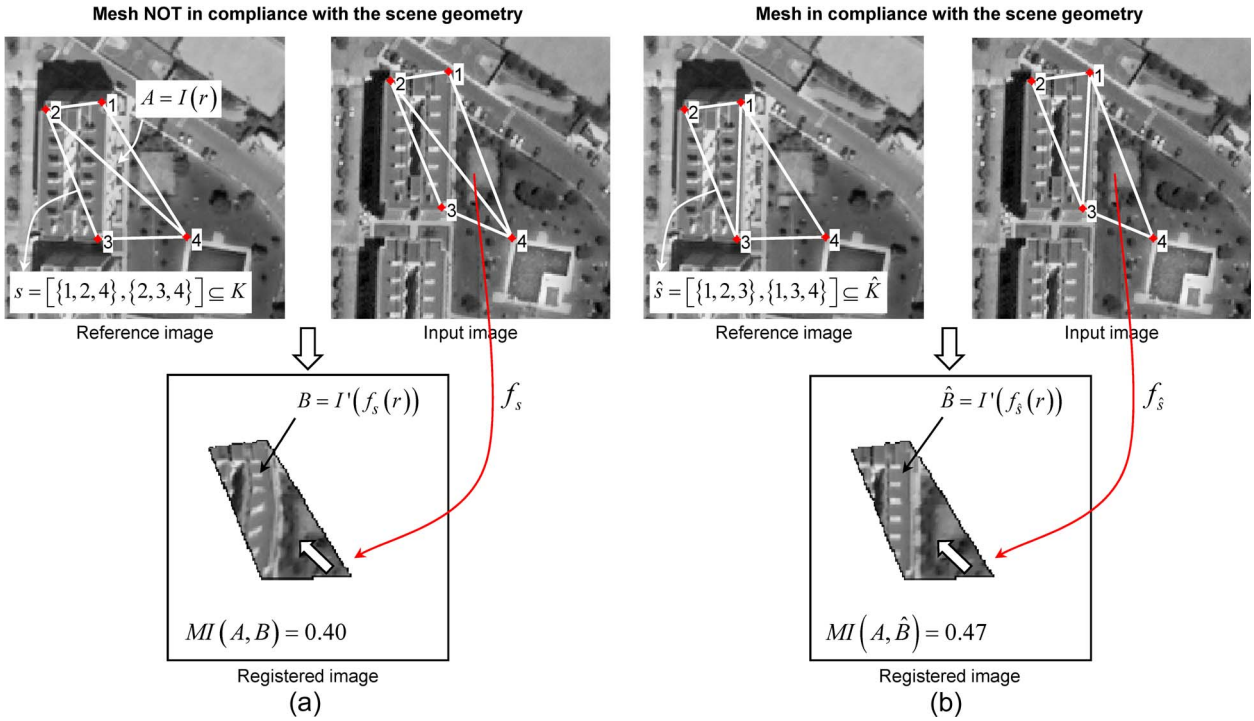

Fig. 2. For a PWL registration process to be successful, the triangles must be the projections of one single plane surface of the scene, as the triangle $\{1, 2, 3\}$ in (b); otherwise, broken lines are produced and the registration consistency decreases, as in (a).

two triangles sharing the edge. Because the resulting optimized mesh will be in compliance with the 3-D scene structure (up to the point that the geometrical realization allows), a by-product of the improvement in the registration is the possibility to project the triangles back to space and build an unscaled coherent 3-D model of the scene.

A key aspect of our proposal is that of measuring how good is the registration of a pair of conjugate quadrilateral image patches. In this paper, we propose to use the mutual information (MI) associated to the intensity values of the patches as a measure of their registration consistency [11], [12]. Unlike other well-known metrics such as normalized cross-correlation

(NCC) or sum of square differences (SSDs), the MI is less sensitive to changes in the images, as it will be reported later on.

The proposed swapping-based optimization process has been successfully applied for improving the PWL registration of panchromatic QuickBird images (0.6 m/pixel of spatial resolution) acquired under different lighting conditions and viewpoints.

The remainder of this paper is organized as follows. In Section II, we review some previous works that apply mesh optimization with particular emphasis on those aimed at improving image registration. In Section III, several assumptions and definitions, as well as the formulation used in subsequent sections, are presented. In Section IV, we describe our method, the inconsistency estimation function, and the optimization process. In Section V, we present and discuss some experimental results. Finally, some conclusions are outlined.

II. Related Works

Many techniques for generating optimal triangular meshes have been proposed in a variety of disciplines, including object modeling, image registration, 3-D scene reconstruction, image compression, etc. The goal of this optimization varies with the type of application, but for most of them, the optimization is aimed at generating 3-D triangular meshes that properly represent 3-D surface of objects (as in computer graphics) or gray-level intensities (as in image compression and reconstruction). In image registration, however, the goal is to generate 2-D triangular meshes that, used for a PWL mapping method, improve the image registration consistency between two images. Next, a review of the most representative techniques on mesh optimization is given, putting special emphasis in those proposed within the image registration field.

Mesh optimization can be classified according to the following issues: the mechanism used for modifying the mesh (i.e., type and scope of the actions), the measure for evaluating the goodness of a given mesh modification (energy or cost functions), and the procedure for accomplishing the mesh refinement.

According to the type and scope of the actions applied to modify the mesh, we have techniques where

1) Only the topological realization is modified by swapping edges [13]–[16].
2) Only the geometrical realization is modified by refining the vertex coordinates (approach typically employed in image registration) [17]–[20].
3) Both the topological and geometrical realizations are simultaneously refined by splitting/collapsing/swapping edges and refining the vertex coordinates [21]–[24].

Most of these techniques were developed in the field of geometric modeling with the objective of simplifying the number of faces of an initial very detailed 3-D mesh. This mesh is typically built upon a dense set of vertices provided by a 3-D sensor, for example, a laser range finder [21], [22]. Similar approaches have been also reported for PWL image representation under the denomination of data-dependent triangulation (DDT) [25], [26]. In this context, the idea of DDT is to consider the image as a 3-D discrete surface that is approximated by a triangular mesh obtained from an initial mesh that is modified in order to better fit the underling data (image intensities). This image approximation is of great interest, among others, for image reconstruction [16] and compression [27].

Unlike these approaches that rely on just one 3-D mesh, in image registration, instead, we are provided with two images and two initial conjugate 2-D meshes onto them, which must be modified to maximize the image registration consistency. One example of this is the work in [17], which relocates the mesh vertices (the mesh topology remains fixed) in order to compensate for the affine motion in video streaming. Vertex coordinate refinement, although being suitable for coping with smooth image distortions, does not provide enough correction to accommodate the possibly important geometric differences between the satellite images when they are acquired from off-nadir angles. This problem, particularly accentuated in high-resolution images, gives rise to unnatural image deformations in the registered image such as broken lines and artifacts (particularly in regions of high-relief terrain or tall buildings), as observed in Fig. 2.

Typically, mesh optimization techniques are formulated as minimization or maximization processes that range from random searches [13] to more complex procedures based on simulated annealing [28], Bayesian stochastic models [24], variational approaches[1] [29], etc. Whatever the applied optimization technique, one of the key points is that of defining a convenient cost or energy function to evaluate the enhancement in the refined mesh if certain action is carried out. Whereas, in object modeling and image DDT, the goodness of a mesh can be measured upon the available 3-D surface [16], [21], [22], [26], [27] in image registration, we must rely only on the radiometric relationship between the reference and moving images. So far, within the mesh optimization for image registration context, several metrics have been used for this purpose: NCC [15], SSDs [14], or some templates based on image differences [13]. However, they are affected by the nonfunctional radiometric differences in images due to different light conditions (shadows, specular reflections, etc.), land changes, different image sensors (i.e., different satellites), etc. [30], which becomes a serious drawback for the registration of satellite images. In this paper, instead, we propose a cost function based on the MI of image patches to drive the swapping actions over the mesh. Although MI has been used as a registration consistency metric in some works [30]–[34], this is the first time it has been integrated into a mesh optimization framework.

III. Assumptions and Definitions

In the remote sensing field, it is well known that, because of the camera scene configuration (i.e., the camera field of view is narrow, and the size of the sensed objects is small with respect to their distances to the camera), the scene projection onto the sensor can be approximated by a paraperspective transformation, which is also called affine or parallel projection. This simplification may lead to a great reduction in complexity in many

[1] In www.itk.org, we can find a broad variety of codes which implement numerous of these techniques, such as potential yields, elastic bodies, etc.

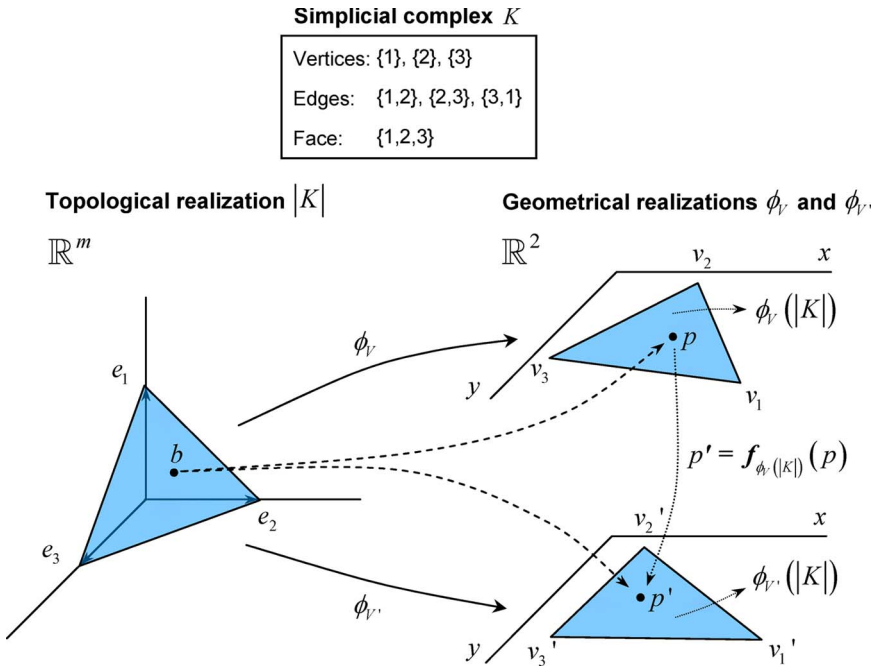

Fig. 3. Example of mesh representation based on simplicial complex. A mesh consisting of one face.

problems [35]. For image registration, this assumption implies that three correspondences (instead of the four ones required for its general form) suffice to estimate the homography (also called affinity under this assumption), which transfers points from one image patch to another [36] (see Fig. 1). Thus, after performing an affine mapping between two conjugate image triangles, they must perfectly match; otherwise, the triangles are projections of a nonplanar surface. Next, we introduce the notation employed in this paper, as well as some useful definitions.

A mesh is a PWL structure consisting of triangular faces put together along their edges. Formally, a mesh is defined as a pair $M = (K, V)$, where K is a simplicial complex[2] which determines the connectivity of the vertices, edges, and faces and $V = \{v_i | i = 1, \ldots, m\}$, where $v_i \in \mathbb{R}^2$, is a set of vertex positions, which defines the shape of the mesh in $\mathbb{R}^2$.

A geometrical realization of a mesh in $\mathbb{R}^2$ can be obtained as follows. For a given simplicial complex K, the topological realization $|K|$ in $\mathbb{R}^m$ results from identifying the vertices $\{1, \ldots, m\}$ with the standard basis vectors $\{e_1, \ldots, e_m\}$ of $\mathbb{R}^m$. Let $\phi : \mathbb{R}^m \mapsto \mathbb{R}^2$ be the linear mapping that sends the ith standard basis vector $e_i \in \mathbb{R}^m$ to $v_i \in \mathbb{R}^2$ (see Fig. 3). The geometrical realization of M is given by $\phi_V(|K|)$, where we write the subindex V in ϕ_V to emphasize that it is specified by that particular vertex set [22].

To refer to any point within a part s of the mesh, we employ the notation $p \in \phi_V(|s|)$, where $s \subseteq K$. Thus, for example, $p \in \phi_V(|t|)$ with $t = \{i, j, k\} \in K$ refers to one particular point within a triangular face, $p \in \phi_V(|q|)$ with $q = [\{i, j, l\}, \{i, j, k\}] \in K$ refers to a point within a quadrilateral of M consisting of two adjacent triangles, and so on.

[2]A simplicial complex K consists of a set of vertices $\{1, \ldots, m\}$ together with a set of nonempty subsets of the vertices, which are called the simplices of K, such that any set of exactly one vertex is a simplex in K and every nonempty subset of simplices in K is also a simplex in K. The zero simplices $\{i\} \in K$ are called vertices, the one simplices $\{i, j\} \in K$ are called edges, and the two simplices $\{i, j, k\} \in K$ are called faces (cf. Spanier [37]).

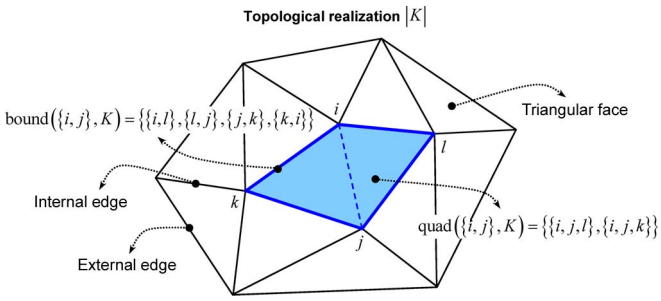

Fig. 4. Topology elements that take part in the PWL image registration process.

In addition to the aforementioned general definitions, we introduce the following particular ones, which are of interests for stating our method in the next section.

1) An edge $\{i, j\} \in K$ is said to be 3-D compatible if it lies on a projection of a 3-D plane surface and 3-D incompatible if otherwise.
2) An edge $\{i, j\} \in K$ is external or boundary if it is a subset of only one face and internal or shared if otherwise.
3) Given an internal edge $\{i, j\} \in K$, we define the following functions (see Fig. 4).
   a) quad$(\{i, j\}, K) = \{\{i, j, l\}, \{i, j, k\}\}$.
   b) bound$(\{i, j\}, K) = \{\{i, l\}, \{l, j\}, \{j, k\}, \{k, i\}\}$.
4) Given a set of conjugate points $\{(v_i, v_i') | i = 1, \ldots, n\}$, $v_i \in V$ 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 PWL function $\boldsymbol{f}$ which geometrically maps a point

$p=(x,y)^\top \in \phi_V(|s|)$ to another $p'=(x',y')^\top \in \phi_{V'}(|s|)$ as follows:

$$p' = \boldsymbol{f}_{\phi_V(|s|)}(p) = \begin{cases} f_1(p) & \text{if} \quad p \in \phi_V(|t_1|) \\ \quad \vdots & \\ f_m(p) & \text{if} \quad p \in \phi_V(|t_m|) \end{cases} \tag{1}$$

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

$$p' = f_i(p) \equiv \begin{cases} x' = a_{11} + a_{12}x + a_{13}y \\ y' = a_{21} + a_{22}x + a_{23}y \end{cases} \tag{2}$$

where a_{ij} ($i = 1, 2$ and $j = 1, 2, 3$) being the transformation coefficients.

Notice that once $\boldsymbol{f}_{\phi_V(|s|)}$ (for clarity, $\boldsymbol{f}_s$ from now on) has been applied, $\phi_V(|s|) = \phi_{V'}(|s|)$, i.e., the corresponding faces of both meshes must perfectly overlap (remember that $\phi_V(|s|)$ represents all the points—pixels—within the mesh given by the simplicial s).

IV. Description of the Proposed Method

The method presented in this paper is aimed to improve the accuracy of PWL registration when applied to high-resolution satellite images. For this purpose, we iteratively modify the connectivity of the conjugate triangular meshes by swapping 3-D-incompatible edges. To detect such edges, our algorithm checks the registration consistency of the two triangles that share the analyzed edge before and after applying the swap: The edge is swapped if that operation entails a registration improvement (as shown in the example in Fig. 2). Notice that this procedure only modifies the mesh connectivity because the number of vertices and their coordinates remain without modification.

A. Mutual Information as a Metric for Registration Consistency

The metric employed in this paper for measuring the registration consistency is the MI [38]. The MI measures the statistical dependence or information redundancy of two random variables. Unlike other similarity measures such as the SSDs or the NCC, which allow for a functional relationship between the gray levels of the image patches to register, the MI responses to their statistical relationship, which can be estimated from the joint histogram of the image patches. The advantage of this metric is that it is more robust to possible image radiometric differences that are difficult (or impossible) to model by a function, for example, shadows, intensity saturation, seasonal changes, land changes, etc. [30]. It is clear that, in remote sensing, this characteristic is highly desirable because such radiometric differences are rather common in multitemporal and multimodal analyses.

Mathematically, the MI of two equal-sized image patches A and B with intensity values of $a \in [0, M-1]$ and $b \in [0, M-1]$, respectively, is defined as

$$(A,B) = \sum_a \sum_b P_{A,B}(a,b) \log \left(\frac{P_{A,B}(a,b)}{P_A(a) P_B(b)} \right) \tag{3}$$

where $P_A(a)$, $P_B(b)$, and $P_{A,B}(a,b)$ are the probability distributions estimated from the intensity joint histogram h of A and B

$$P_{A,B}(a,b) = \frac{1}{N} h_{A,B}(a,b) \tag{4}$$

$$P_A(a) = \sum_b P_{A,B}(a,b) \tag{5}$$

$$P_B(b) = \sum_a P_{A,B}(a,b) \tag{6}$$

where N being the number of pixels in the image patches,[3] and $h_{A,B}(a,b)$ is the number of corresponding pairs having the intensity value a in the first image and the intensity value b in the second.

In the case of 8-bit gray-scale images, the original value of M is 256; however, in practice, it is convenient to use a lower value (e.g., 128, 64, 32, ...) for three reasons: first, to make the process less time consuming [less terms in the summation of (3)]; second, to provide the method robustness against intensity noise; and third, to make the estimation of the joint probability from the joint histogram more reliable, when N is not very large (i.e., the joint histogram needs to be representative enough). In our implementation, given that the size of the image patches is usually in the range between some hundreds and a few thousands (triangles that are too small are previously discarded from the initial mesh provided by the Delaunay triangulation), we have used 32 gray-level bins. An alternative way of overcoming the problem of small triangles is to use nonparametric estimation methods as the Parzen window, although it entails a higher computational cost (the reader may refer to [12] for more detail about this approach).

Observe in (3) that, when all the intensity values of the two images are independent one from another (i.e., without correlation), the argument of the logarithm becomes one, and the MI achieves its minimum at zero (MI is always nonnegative). To illustrate how the joint histogram captures the idea of statistical dependence, Fig. 5 shows the joint histograms of two pairs of equal-sized synthetic images of just four gray levels each. When one of them is rotated, the joint histogram exhibits more dispersion than when they perfectly overlap.

Finally, with the purpose of illustrating the performance of the MI in comparison with the NCC, we have conducted the experiments shown in Fig. 6. In the first one, the similarity between the two images used in the previous example is computed for a range of rotation angles, using both MI and NCC. MI outputs a maximum at the angle where they overlap (zero degree), despite the fact that the four gray levels of the images do not coincide one another. This statistical dependence goes

[3]Notice that the requirement that the two image patches have the same size always holds because the input image patch is mapped to the reference one through some interpolation technique, having then the same number of pixels.

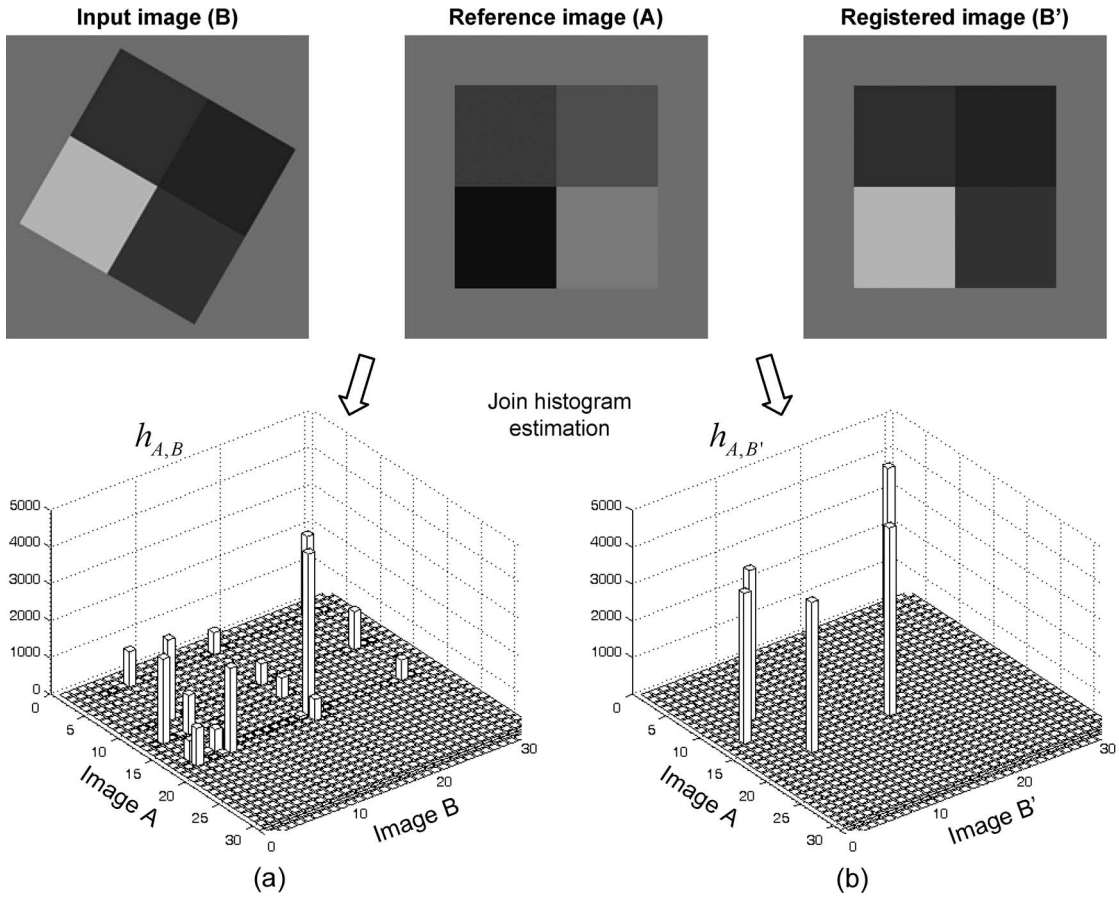

Fig. 5. Joint histograms of two pairs of synthetic images. (a) Misaligned pair. (b) Perfectly overlapped pair. Observe that the joint histogram (of 32 gray levels) presents less dispersion when the images are aligned. Notice that this fact is independent of the intensity values of the two images being identical or not.

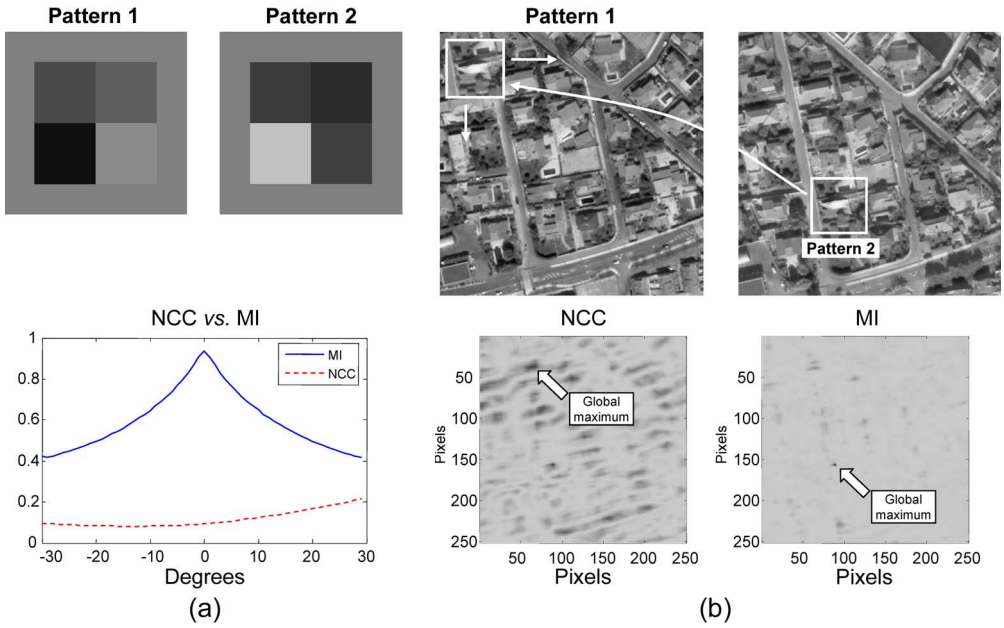

Fig. 6. Two experiments illustrating the suitability of the MI as a registration consistency measure for image registration. The plots show the MI in comparison with the NCC when (a) a synthetic pattern is rotated from -30° to 30° and (b) a small window of a QuickBird satellite image is translated through an image of the same scene acquired on a different date. Observe how the NCC fails in both experiments, giving no global maximum in (a) and an erroneous global maximum in (b).

completely unnoticed for the NCC. Another example is shown in Fig. 6(b), although in this case, a small QuickBird image is translated through a larger one, which is from the same scene but taken six months later (different seasons, daytime, objects in the scene, etc.). From these results, we can highlight the robustness and effectiveness of the MI when applied to image pairs

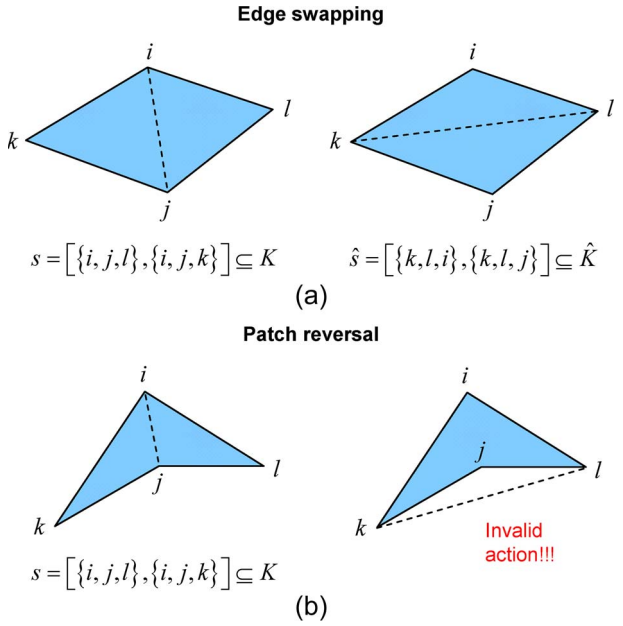

Fig. 7. Topological action of swapping an edge when (a) all preconditions, as explained in the text, are verified and (b) the action produces a patch reversal (because the quadrilateral is concave).

with nonfunctional radiometric changes, which, in contrast to NCC, can manage important image intensity differences but only if they follow a linear function (i.e., intensity shifts and/or contrast scaling).

B. Checking Edges for Swapping

We take advantage of the robustness of the MI for effectively detecting 3-D-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 3-D compatibility of an edge $\{i,j\}\in K$ by measuring the improvement in the registration consistency of the quadrilaterals defined by $s=\text{quad}(\{i,j\},K)$ and $\hat{s}=\text{quad}(\{k,l\},\hat{K})$, i.e., before and after the edge $\{i,j\}$ is swapped [see Fig. 7(a)]. Formally, that improvement is measured by

$$\omega\left(\{i,j\}\right)=MI\left(I(r),I'\left(\boldsymbol{f}_{\hat{s}}(r)\right)\right)-MI\left(I(r),I'\left(\boldsymbol{f}_{s}(r)\right)\right) \tag{7}$$

where $I(r)$ represents the quadrilateral region of the reference image defined by s, and $I'(\boldsymbol{f}_s(r))$ and $I'(\boldsymbol{f}_{\hat{s}}(r))$ are the transformations of its input-image counterparts according to the two possible topological configurations s and $\hat{s}$, respectively. Fig. 2 shows the examples of these two topological configurations and the MI associated to them.

Before evaluating the 3-D compatibility of any edge, for example, $\{i,j\}\in K$, it should be checked to verify the following preconditions.

1) $\{i,j\}$ is an internal edge.
2) The resultant swapped edge is a new one $\{k,l\}\notin K$.
3) The action does not produce a patch reversal[4] in $\hat{K}$ [see Fig. 7(b)].

[4]A mesh inconsistency produced when the shared edge of two adjacent faces, which make up a concave quadrilateral, is swapped.

```
OBJECTIVE
Given two images I and I' to register, and
two initial triangular meshes on them defined
by M = (K,V) and M' = (K,V'), determine
a new topological realization by iteratively
swapping edges which improve the consistency of
piecewise-linear image registration.

ALGORITHM
1. % Build a sorted list (indexed by edge) with
the ω of each edge
   w_list=[]
   FOR-EACH edge {i,j} ∈ K
      IF {i,j} verifies the preconditions THEN
         compute ω({i,j}) from expression (7)
         store ω({i,j}) in w_list
      END IF
   END FOR-EACH
   sort w_list in descending order[5]

2. % Iterate while there exist an edge swap that
improves the consistency
   WHILE the first element of w_list > δ
      swap its corresponding edge, say {i,j} ∈ K by
{k,l} ∈ K̂
      remove it from w_list

      % Update w_list with the ω of the adjacent
edges of {i,j}
      FOR-EACH edge {m,n} ∈ bound({i,j},K)
         IF {m,n} verifies the preconditions THEN
            compute ω({m,n}) from expression (7)
            store ω({m,n}) in w_list
         END IF
      END FOR-EACH

      update the topological realization K = K̂
      sort w_list in descending order
   END WHILE
```

[5]Notice that for properly sorting the list, the values of MI in (7) must be normalized, so, we normalize the expression (3) with the joint entropy of both image patches. Reader may refer to [11] for other possible MI normalized variants.

Fig. 8. Greedy algorithm proposed for improving the PWL registration of two images by applying edge swapping actions.

Once these preconditions are met, the swap of the edge is accepted if $\omega(\{i,j\})>\delta$, i.e., when it leads to an increase in the MI above a certain threshold. This threshold must be small but enough for preventing mesh modifications that entail a negligible improvement. One example of such cases occurs when the quadrilateral lies on a planar surface, and because of the image resampling, we may have small consistency differences for the two configurations, which should not be considered as true improvement. We have experimentally determined that δ values close to 0.01 solve this problem.

C. Mesh Optimization

The overall optimization process can be formulated as a greedy search [39], which starts with two corresponding triangular meshes M and M' that resulted, for example, by applying a Delaunay's triangulation method over a set of conjugate points identified in both images. The greedy search algorithm, which is shown in Fig. 8, starts by creating a list of the improvement in registration consistency ω for all the edges of the mesh. This list is computationally expensive to generate, but this is done just once, which is at the beginning. At each iteration of the optimization process, the first edge of the current list is

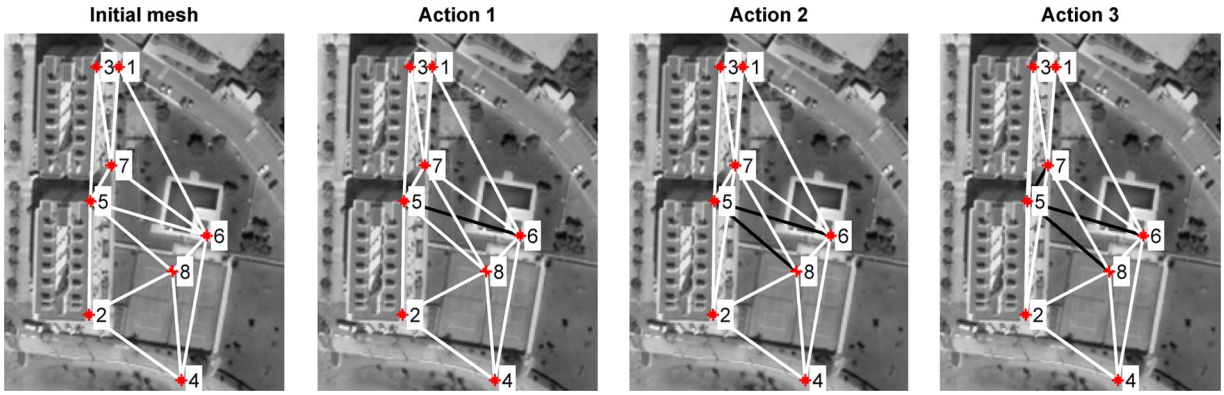

Fig. 9. Illustration of the optimization process. The initial mesh shows a topological configuration containing several 3-D-incompatible edges (edges $\{5,6\}$, $\{5,7\}$, and $\{5,8\}$). As depicted in the algorithm in Fig. 8, when the (first action) edge $\{5,6\}$ is swapped, its (second and third actions) adjacent edges, which are $\{5,7\}$ and $\{5,8\}$, are checked again for swapping.

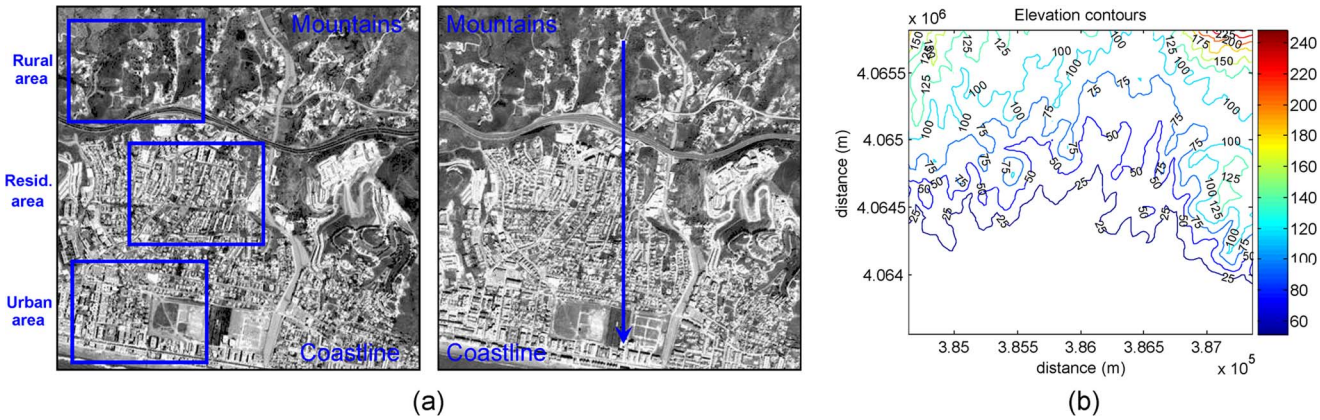

Fig. 10. (a) Test images of the coastal city of Rincón de la Victoria (Málaga, Spain) used in our experiments. Three regions of interest are highlighted in the reference image. (b) Elevation information considered to properly define the regions of study.

swapped, and the list is updated by checking only those edges affected by the swapping action (its bound). The algorithm stops when the list is empty, i.e., when all the mesh edges have been explored and no further improvement can be achieved by the swapping actions.

Unlike other optimization techniques employed in 3-D scene reconstruction such as [13] and [14], which are formulated as a random search, this procedure guarantees the iterative improvement of the image registration consistency up to the degree that the geometrical realization of the mesh allows. Notice that, without relocating the vertices and, possibly, introducing additional ones, the mesh may be not good enough to completely avoid 3-D-incompatible edges.

Fig. 9 shows the proposed optimization method, when applied to an initial topological configuration containing several 3-D-incompatible edges.

V. Experimental Results

This section presents some experimental results that show the performance of our approach for improving the PWL registration of high-resolution satellite images. A by-product of such improvement is the possibility of recovering a more accurate unscaled 3-D representation of the sensed scene by reprojecting the resultant meshes. This issue will also be illustrated in this section.

A. Data Sets and Methodology

We have employed two orthoready QuickBird images, of 3000×1500 pixels $\equiv 1800 \times 900$ m^2, of the coastal city of Rincón de la Victoria (in Málaga, Spain) (see Fig. 10). The images, which are acquired on spring and winter, present significant radiometric differences (brightness and contrast, shadows, changes on the land cover, etc.), and also geometrical differences induced by quite different off-nadir observation angles, concretely 32.1°. Apart from the whole image, we consider the following three different subimages (of 900×750 pixels) as areas of study:

1) urban area, which contains buildings of different heights on an almost plain terrain;
2) residential area, which is a high-relief terrain mostly containing low buildings (houses);
3) rural area, which is also an almost plain terrain but, in this case, without tall structures (except for some trees).

The purpose of selecting these zones is to test the method on different types of elevation profiles and image contents.

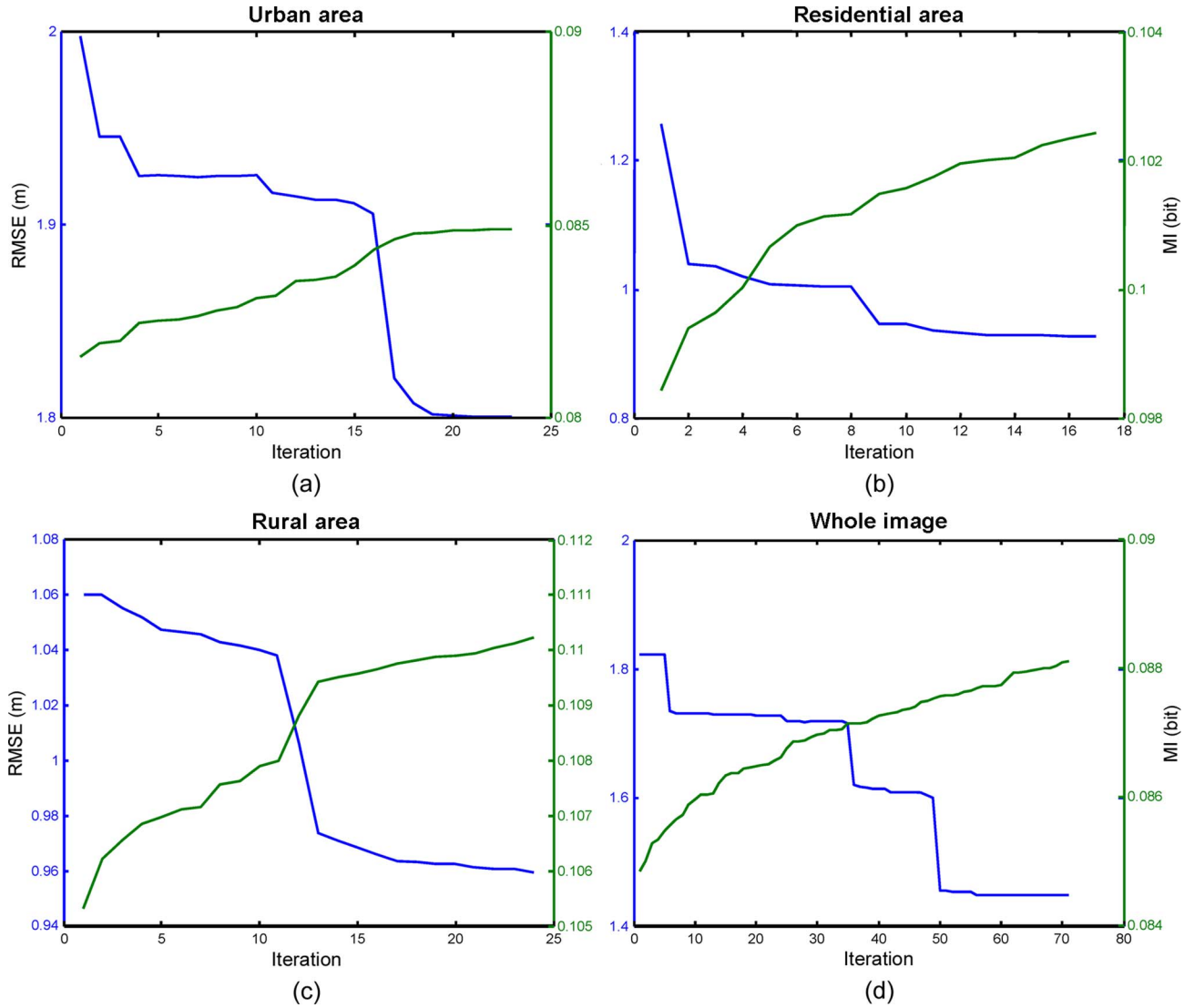

Fig. 11. RMSE (in meters) and MI (in bits) curves for each of the analyzed areas. (a) Urban. (b) Residential. (c) Rural. (d) Whole images shown in Fig. 10. The plots show a significant improvement in the registration consistency for (a) and (b) the urban and residential test images, whereas for (c) the rural case, whose images present less geometrical difference, the improvement is not so accentuated.

The conjugate control points (CPs) that define the geometrical realization of the mesh are automatically obtained through the following procedure (please refer to [8] for more details).

1) The Harris corner detector [40] identifies distinctive feature points in the reference image.
2) The Lucas–Kanade feature tracker [41] detects their corresponding points in the input image.
3) The affine epipolar geometry of the two images (i.e., the affine fundamental matrix) is robustly estimated by applying the RANdom SAmple Consensus algorithm [42] and employed for discarding those CPs which are not consistent with it (outliers) [43].

Despite the fact that this method provides consistently matched CPs, some of them may be identified in different scene planes (e.g., ground and top of the buildings), which gives rise to 3-D-incompatible edges once the Delaunay's triangulation algorithm [10] is applied (as in Fig. 2). The objective of our optimization method is to correct, as much as possible, those situations. Whether this improvement is achieved or not can be assessed in two different ways.

1) First is by evaluating the goodness of the image registration. We have employed two metrics for that: the MI of the complete images (although, to be precise, only the image region delimited by the convex hull of the mesh is evaluated), i.e.,

$$MI\left(I(m), I'\left(\boldsymbol{f}_K(m)\right)\right) \tag{8}$$

where $m = \phi_V(|K|)$ represents all image pixels within the mesh given by K, and the RMSE of a set of independent check points (ICPs) selected by hand in the images.
2) The second way is by checking if the unscaled 3-D scene reconstructed from the resulting mesh is more accurate than that obtained from the initial one. In the event that a precise digital terrain model (DTM) was available, it could be used for that evaluation. Otherwise, as it is the case here, we do it by visual inspection.

TABLE I
MESH DATA, NUMBER OF APPLIED ACTIONS, AND COMPUTATIONAL TIME FOR THE EXPERIMENTS DESCRIBED IN SECTION V-A

Scene	# Faces	# Edges	# Actions	Time* (sec.)
Urban	168	258	23	7.30
Residential	155	239	17	6.80
Rural	162	248	24	7.73
Whole image	502	759	71	39.04

*We have employed Matlab on a Pentium 4 HT 3.2GHz for implementing the tests.

B. Results

Fig. 11 shows the results obtained when running our method for the three areas of study and the whole image (mixed area).

The reader may observe that the registration consistency is improved in all cases and for the two metrics. Such improvement is more notorious for the urban and residential areas because, being the rural one almost plain terrain, the initial Delaunay's mesh is reasonably good and does not require many modifications. The number of iterations, and thus the computational time, for the algorithm to converge strongly depends on the number of edges (i.e., triangles) in the mesh. For the three first areas (a, b, and c), the algorithm requires less than 25 iterations, whereas for the whole image (which has a larger size), it takes 71 iterations. Table I summarizes the mesh data and the computational time for each of these experiments.

Notice that, during the optimization process, there are actions that improve the registration consistency, according to the MI (increasing its value), whereas the RMSE remains fixed (revealed as a small flat stretch in the curve). This "disagreement" between the two metrics is due to the fact that the triangles involved in the swapping actions do not contain any ICP inside, and consequently, the RMSE is not altered by the edge swap. Be aware that, even though the ICPs were placed (manually) in all the triangles of the initial mesh for having the maximum mesh coverage, it may happen that, along the course of the optimization, some of the new triangles generated have no ICPs lying on them.

For completeness, we have experimentally compared our approach with other well-known registration methods included in popular commercial packages of remote sensing like ERDAS, ENVI, and PCI, namely, POL (from second up to fourth order), PWL, piecewise cubic (PWC), and thin-plate spline (TPS). The results of this comparison are summarized in Table II.

From this comparison, we can highlight a variety of points: 1) POL functions, as expected, perform worse than local techniques because the images present significant distortions in the high-relief areas; 2) among the three local registration methods, TPS is the most precise one, as was already reported in [8]; and 3) by applying the mesh optimization method proposed in this paper, the optimized PWL solution achieves and even outperforms the accuracy results of the more precise one, which is the TPS method.

Because the iterative approach proposed in this paper requires a higher computational cost than PWL, PWC, or TPS, we believe that this is a handicap that may be negligible for many remote sensing applications, although it could be of significance in other fields such as robotics, medicine, etc., where the time requirement becomes a critical issue. Despite that, better times could be easily achieved with more efficient programming languages, such as C, optimized libraries, and even parallelizing some stages of the algorithm.

TABLE II
COMPARISON OF THE PROPOSED METHOD WITH OTHER REGISTRATION METHODS BROADLY USED IN THE REMOTE SENSING COMMUNITY: A POL METHOD (FROM SECOND TO FOURTH ORDER); TWO LOCAL METHODS (PIECEWISE LINEAR AND CUBIC); AND A HYBRID ONE (TPS)

		Areas of study			
Metric	**Method**	**Urban**	**Resid.**	**Rural**	**Whole**
RMSE (m.)	POL 2^{nd} order	3.20	2.71	2.18	8.12
	POL 3^{rd} order	3.27	2.05	1.82	7.67
	POL 4^{th} order	2.86	2.07	1.75	6.00
	PWL	2.00	1.26	1.06	1.82
	PWC	2.24	1.33	1.27	2.15
	TPS	2.02	0.92	1.03	1.66
	Optimized PWL	1.80	0.93	0.96	1.45
MI (bits)	POL 2^{nd} order	0.330	0.362	0.356	0.217
	POL 3^{rd} order	0.310	0.404	0.383	0.219
	POL 4^{th} order	0.338	0.406	0.393	0.243
	PWL	0.447	0.527	0.472	0.449
	PWC	0.425	0.494	0.444	0.435
	TPS	0.437	0.556	0.487	0.464
	Optimized PWL	0.465	0.549	0.494	0.467

Finally, we exploit the result of having a pair of registered conjugate images: We can reconstruct an unscaled 3-D surface of the sensed scene by projecting back the two meshes[5] (see Fig. 12). Apart from its undoubted interest for 3-D reconstruction, it allows us to check the effectiveness of the proposed procedure by contrasting the 3-D models associated to the initial and refined meshes. Unfortunately, we do not have an accurate DTM available of the area of study, neither the camera parameters, to carry out a rigorous error analysis from this viewpoint. Therefore, our analysis has been limited to visually check the 3-D inconsistencies.

VI. CONCLUSION

Image registration is an essential step in a broad variety of remote sensing applications, where the final result comes from the combination of several sources, for example, change detection, image fusion, 3-D scene reconstruction, etc.

In this paper, we have proposed a technique for automatically optimizing the conjugate triangular meshes employed by a PWL registration process: Having more suitable meshes means that the registration becomes more accurate. 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 particularly suitable for satellite images, because it is less sensitive to changes in the land cover, shadows, specular reflections, etc., than other well-known metrics such as NCC or SSD. The optimization procedure is formulated as a greedy search, which finishes when the mesh topology can no longer be refined, i.e., when all mesh edges have been successfully checked and no further improvement is possible.

[5]For example, by applying the factorization algorithm for affine reconstruction proposed in [9, p. 437].

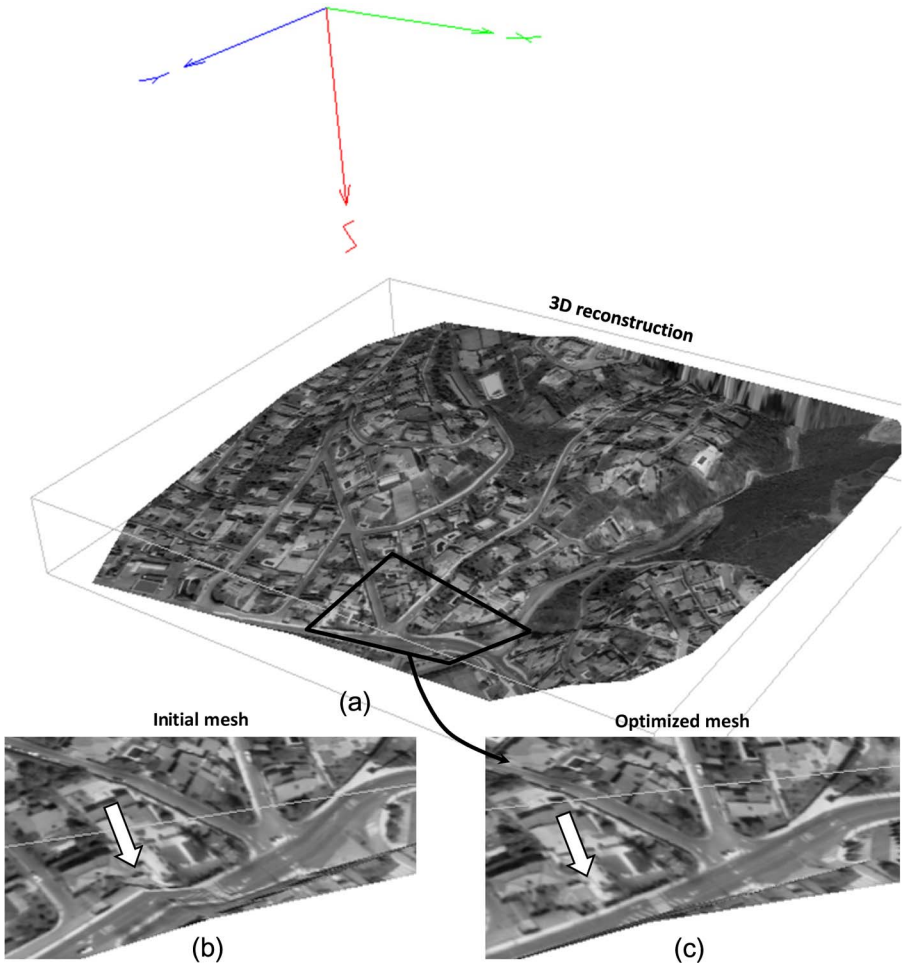

Fig. 12. (a) Three-Dimensional scene reconstruction generated from the refined mesh. Magnifications of the same area, (b) before and (c) after applying the proposed mesh optimization. Observe in (b) the presence of deformed roads because of an incorrect triangulation (Delaunay).

The proposed method has been successfully tested with different pairs of panchromatic QuickBird images (0.6 m/pixel of spatial resolution) acquired under different conditions (from different angles and lighting conditions), as well as compared with well-known registration methods employed by the remote sensing community.

ACKNOWLEDGMENT

QuickBird imagery used in this paper is distributed by Eurimage, S.p.A. (www.eurimage.com) and provided by Decasat Ingenieria S.L. (www.decasat.com).

REFERENCES

[1] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.
[2] J. González, G. Ambrosio, and V. Arévalo, "Automatic urban change detection from the IRS-1D PAN," in *Proc. IEEE-ISPRS Joint Workshop Remote Sens. Data Fusion Over Urban Areas*, Rome, Italy, Nov. 2001, pp. 320–323.
[3] F. Bookstein, "Principal warps: Thin-plate-splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
[4] A. Goshtasby, "Piecewise linear mapping functions for image registration," *Pattern Recognit.*, vol. 19, no. 6, pp. 459–466, 1986.
[5] A. Goshtasby, "Piecewise cubic mapping functions for image registration," *Pattern Recognit.*, vol. 20, no. 5, pp. 525–533, 1987.
[6] M. Ehlers and D. Fogel, "High-precision geometric correction of airborne remote sensing revisited: The multiquadric interpolation," in *Proc. SPIE—Image Signal Processing Remote Sensing*, Sep. 1994, vol. 2315, pp. 814–824.
[7] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1427–1442, Nov. 2003.
[8] V. Arévalo and J. Gonzalez, "An experimental evaluation of non-rigid registration techniques on QuickBird satellite imagery," *Int. J. Remote Sens.*, vol. 29, no. 2, pp. 513–527, Jan. 2008.
[9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[10] J. Shewchuk, *Lecture Notes on Delaunay Mesh Generation*. Berkeley, CA: Univ. California, 1999. Tech. Rep. 3.
[11] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.
[12] P. Viola and W. Wells, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, Sep. 1997.
[13] A. Nakatuji, Y. Sugaya, and K. Kanatani, "Mesh optimization using an inconsistency detection template," in *Proc. IEEE ICCV*, Beijing, China, Oct. 2005, vol. 2, pp. 1148–1153.
[14] D. Morris and T. Kanade, "Image-consistent surface triangulation," in *Proc. IEEE Int. Conf. CVPR*, Hilton Head, SC, Jun. 2000, pp. 332–338.
[15] J. Perrier, G. Agin, and P. Cohen, "Image-based view synthesis for enhanced perception in teleoperation," in *Proc. SPIE—Enhanced Synthetic Vision*, Apr. 2000, vol. 4023, pp. 332–338.
[16] X. Yu, B. Bryan, and T. Sederberg, "Image reconstruction using data-dependent triangulation," *IEEE Comput. Graph. Appl.*, vol. 21, no. 3, pp. 62–68, May/Jun. 2001.
[17] M. Servais, T. Vlachos, and T. Davies, "Bi-directional affine motion compensation using a content-based, non-connected, triangular mesh," in *Proc. CVMP*, London, U.K., Mar. 2004, pp. 49–58.

[18] B. Yip and J. Jin, "Image registration using triangular mesh," in *Proc. PCM—Advances Multimedia Information Processing*, 2004, vol. 3331, pp. 298–303.
[19] D. Schlesinger, B. Flach, and A. Shekhovtsov, "A higher order MRF-model for stereo-reconstruction," *Pattern Recognit.*, vol. 3175, pp. 440–446, 2004.
[20] B. Matuszewski, J. Shen, and L. Shark, "Elastic image matching with embedded rigid structures using spring-mass system," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2003, vol. 2, pp. 937–940.
[21] H. Hoppe, "Progressive meshes," in *Proc. Comput. Graph.*, 1996, vol. 30, pp. 99–108.
[22] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Mesh optimization," in *Proc. Comput. Graph.*, 1993, vol. 27, pp. 19–26.
[23] P. de Bruin, P. van Meeteren, F. Vos, A. Vossepoel, and F. Post, "Accurate and high quality triangle models from 3D grey scale images," in *Proc. MICCAI—Part II*, 2002, vol. 2489, pp. 348–355.
[24] G. Vogiatzis, P. Torr, and R. Cipolla, "Bayesian stochastic mesh optimization for 3D reconstruction," in *Proc. BMVC*, Norwich, U.K., Sep. 2003, pp. 711–718.
[25] N. Dyn, D. Levin, and S. Rippa, "Data dependent triangulations for piecewise linear interpolation," *IMA J. Numer. Anal.*, vol. 10, no. 1, pp. 137–154, 1990.
[26] E. Quak and L. Schumaker, "Cubic spline fitting using data dependent triangulations," *Comput. Aided Geom. Des.*, vol. 7, no. 1–4, pp. 293–301, Jun. 1990.
[27] B. Lehner, G. Umlauf, and B. Hamann, "Image compression using data-dependent triangulations," in *Proc. 3rd ISVC—Part I*, 2007, vol. 4841, pp. 351–362.
[28] L. Schumaker, "Computing optimal triangulations using simulated annealing," *Comput. Aided Geom. Des.*, vol. 10, no. 3/4, pp. 329–345, Aug. 1993.
[29] *National Library of Medicine Insight Segmentation and Registration Toolkit*. [Online]. Available: http://www.itk.org
[30] H. Chen, P. Varshney, and M. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2445–2454, Nov. 2003.
[31] A. Cole-Rhodes, K. Johnson, J. LeMoigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1495–1511, Dec. 2003.
[32] H. Xie, L. Pierce, and F. Ulaby, "Mutual information based registration of SAR images," in *Proc. IGARSS*, Jul. 2003, vol. 6, pp. 4028–4031.
[33] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.
[34] J. Kern and M. Pattichis, "Robust multispectral image registration using mutual-information models," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1494–1505, May 2007.
[35] G. Xu and Z. Zhang, *Epipolar Geometry in Stereo, Motion, and Object Recognition: A Unified Approach*, 1st ed. Norwell, MA: Kluwer, 1996.
[36] E. Mikhail, J. Bethel, and J. McGlone, *Modern Photogrammetry*, 1st ed. New York: Wiley, 2001.
[37] E. Spanier, *Algebraic Topology*, 1st ed. New York: McGraw-Hill, 1966.
[38] T. Cover and J. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 1991.
[39] G. Strang, *Introduction to Applied Mathematics*, 1st ed. Wellesley, MA: Wellesley-Cambridge, 1986.
[40] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, Manchester, U.K., 1988, vol. 4, pp. 147–151.
[41] J. Bouguet, "Pyramidal implementation of the Lucas-Kanade feature tracker. Description of the algorithm," Microprocessor Res., Intel Corp., Hillsboro, OR, 1999. Tech. Rep. (accessed 26/06/2006). [Online]. Available: http://robots.stanford.edu/cs223b04/algo_tracking.pdf
[42] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
[43] P. Torr, "A structure and motion toolkit in Matlab: Interactive adventures in S and M," Microsoft Res., Cambridge, U.K., Tech. Rep. MSR-TR-2002-56, 2002. (accessed 27/03/2007). [Online]. Available: ftp://ftp.research.microsoft.com/pub/TR/TR-2002-56.ps

Vicente Arévalo received the M.S. and Ph.D. degrees in computer science from the University of Málaga, Málaga, Spain, in 2001 and 2008, respectively.

Since 2003, he has been with the Department of System Engineering and Automation, University of Málaga. He was granted by the Spanish Government as a Predoctoral Researcher. His research mainly focuses on mobile robots, computer vision, and remote sensing, being the author of nearly 15 journal and conference papers.

Javier González (S'92–A'93) received the M.S. degree in electrical engineering from the University of Seville, Seville, Spain, in 1987 and the Ph.D. degree from the University of Málaga, Málaga, Spain, in 1993.

From 1990 to 1991, he was with the Field Robotics Center, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, working on perception for mobile robots. He joined the System Engineering and Automation Department, University of Málaga, in 1988, where he is currently an Associate Professor and has led national and European Union projects. His research interests focus on mobile robot autonomous navigation, computer vision, and remote sensing. In these areas, he is the author and coauthor of three books and more than 80 papers in international conferences and journals.