HAIRIS: A Method for Automatic Image Registration Through Histogram-Based Image Segmentation

Hernâni Gonçalves, José Alberto Gonçalves, and Luís Corte-Real, *Member, IEEE*

***Abstract*—Automatic image registration is still an actual challenge in several fields. Although several methods for automatic image registration have been proposed in the last few years, it is still far from a broad use in several applications, such as in remote sensing. In this paper, a method for automatic image registration through histogram-based image segmentation (HAIRIS) is proposed. This new approach mainly consists in combining several segmentations of the pair of images to be registered, according to a relaxation parameter on the histogram modes delineation (which itself is a new approach), followed by a consistent characterization of the extracted objects—through the objects area, ratio between the axis of the adjust ellipse, perimeter and fractal dimension—and a robust statistical based procedure for objects matching. The application of the proposed methodology is illustrated to simulated rotation and translation. The first dataset consists in a photograph and a rotated and shifted version of the same photograph, with different levels of added noise. It was also applied to a pair of satellite images with different spectral content and simulated translation, and to real remote sensing examples comprising different viewing angles, different acquisition dates and different sensors. An accuracy below 1° for rotation and at the subpixel level for translation were obtained, for the most part of the considered situations. HAIRIS allows for the registration of pairs of images (multitemporal and multisensor) with differences in rotation and translation, with small differences in the spectral content, leading to a subpixel accuracy.**

***Index Terms*—Histogram, image registration, image segmentation, matching, Wiener filtering.**

I. Introduction

AUTOMATIC image registration (AIR) is still a present challenge regarding image processing related applications. Reviews concerning image registration methods can be found in [3] and [45]. Remote sensing applications is one of the fields where further research on AIR methods is required. Under this scope, there are particular difficulties so that AIR methods suitable for many computer vision applications will present limited performance.

The rigid-body model under the scope of automatic image registration methods is still a present subject of research [22], in particular under the scope of remote sensing applications [2], [40]. The problem of registering remote sensing images can roughly be locally seen as the determination of translations and a small rotation. Under the scope of computer vision applications, the rigid-body transformation may seem a simple problem to solve with many existing methods. However, under the scope of remote sensing applications, one of the major problems is related to the radiometric content (due to multisensor or multispectral pairs of images). Moreover, scale is frequently known, as most satellite images are provided with sufficiently accurate scale information, being the exception aerial photographs when the aircraft is flying across a region with significant differences in the terrain elevation. Similar comments may be applicable regarding differences in the view-angle, as satellite images are roughly acquired at an altitude of 600 km, being the affine or homography mostly required for aerial photographs under the circumstances previously mentioned for aerial photographs.

Over the last 30 years, a large amount of articles has been published regarding the topic image segmentation, either related to methods themselves or their application to several fields such as medicine, remote sensing, among others [10], [32], [33], [38], [43]. However, image segmentation (IS) is still an actual field of research, regarding automatic methods of image processing. IS is generally defined as the process that partitions an image into regions, each of them fulfilling a given criteria, which can be from the image domain and/or feature space. From image segmentation methods, we expect the extraction of a set of objects present on an image, as we visually detect them. In other words, it is expected that a segmentation method acts as artifical intelligence on the identification of objects on a scene. However, the objective of the segmentation may be quite subjective, depending upon the detail and features we are expecting. For instance, on the segmentation of an image of a human body, one may be interested in delineating the whole body as a single object, or its constituent parts, which may become itself quite subjective. Regarding remote sensing applications, this aspect may become even more complex [28].

Image segmentation comprises a wide variety of methods [8], [33], either for monochrome or color images (or to a single or multiple bands of satellite images). It can be employed to image or video applications, in the latter to each frame individually. Most image segmentation methods can be classified according to their nature: histogram thresholding, feature

Manuscript received April 30, 2010; revised August 25, 2010; accepted August 26, 2010. Date of publication September 13, 2010; date of current version February 18, 2011. The work of H. Gonçalves was supported by Fundação para a Ciência e a Tecnologia, Portugal. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Brian D. Rigling.

H. Gonçalves is with the Departamento de Geociências, Ambiente e Ordenamento do Território, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal. He is also with the Centro de Investigação em Ciências Geo-Espaciais, Universidade do Porto (e-mail: hernani.goncalves@fc.up.pt).

J. A. Gonçalves is with the Departamento de Geociências, Ambiente e Ordenamento do Território, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal (e-mail: jagoncal@fc.up.pt).

L. Corte-Real is with the Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia, Universidade do Porto and with INESC Porto, Instituto de Engenharia de Sistemas e Computadores do Porto, 4200-465 Porto, Portugal (e-mail: lreal@inescporto.pt).

Digital Object Identifier 10.1109/TIP.2010.2076298

space clustering, region-based approaches, edge detection approaches, fuzzy approaches, neural networks, physics-based approaches and any combination of these [8]. Any of these generally intends to transform any image to a binary image: objects and background.

Regarding histogram thresholding, several methods have been reported. Typically, an histogram-based image segmentation comprises three stages: recognizing the modes of the histogram, finding the valleys between the identified modes and finally apply thresholds to the image based upon the valleys [7]. Some works published in this field cover the peaks detection on the histogram curve based upon homogeneity criteria [7], propose recursive thresholding techniques based upon discriminant analysis [9], maximum correlation criterion for bilevel thresholding [42], entropy-based [5], [21], using fuzzy sets [41], among several others.

Traditionally, this class of methods look for a peak (mode or local modes) which is associated to an homogeneous region clearly visible on an image. Most of the known methods are devoted to unimodal, bimodal or multimodal histogram distributions, where it is assumed that each mode correspond to a certain class of objects. However, these are not appropriate for other applications such as remote sensing, where such distinction is not so evident after a simple visual inspection. Fuzzy c-means (FCM) and hard c-mean (HCM) account for the presence of different classes of objects on an image [34]. However, these fuzzy-based methods require *a priori* information, in particular the need for a previous definition of the number of clusters. Furthermore, the presented results illustrate its ability in delineating very large regions, which is not sufficient for applications such as automatic image registration requiring several "control points." Despite a multimodal gray-level histogram modelling and decomposition has also been reported [4], once again the focus relies on the most prominent modes, ignoring the remaining parts of the histogram. Although a work which adresses the segmentation of bright targets using wavelets and adaptive thresholding has been reported, in practice it is a method which mainly look for a mode on a bimodal distribution [44].

The use of image segmentation as a step in image registration had been scarcely explored [12], [18], [24]. The work by Liang *et al.* [24] is able to handle with a rigid-body transformation, comprising a new region-based similarity measure applied to the segmentation obtained by a multiscale segmentation algorithm, followed by the Powell direction set optimization method to find the peak on a similarity surface. The given examples do not include additive noise neither remote sensing images, which are widely known to be associated with particular difficulties in the registration process. Furthermore, it is based upon a search approach and consequently being computationally expensive. The work from Goshtasby *et al.* presents as major contribution a region refinement procedure so that optimally similar corresponding regions are obtained. As in [24], the images content presented in [18] are more prone to lead to a clearer segmentation result, since the MSS and TMS images presented exhibit a quite good contrast leading to a much simpler segmentation and consequent registration. Additionally, the simulated noise in [18] is quite smooth, as can be observed by larger regions in some examples which are not affected even by the larger levels of noise. The work by Dare and Dowman [12] is another example where image segmentation is used as a step in image registration. Although the authors state that their method is fully automatic, in practice it requires an initial manual alignment, in order to remove gross differences in scale and rotation [12]. Moreover, the proposed segmentation methods are more adequate to images containing sufficiently large and homogeneous objects (such as large water bodies), presenting limited performance for other datasets such as satellite images in urban context.

In this paper, a method for automatic image registration through histogram-based image segmentation (HAIRIS) is proposed, which allows for a more detailed histogram-based segmentation, rather than the traditional methods, and consequently to an accurate image registration. HAIRIS is able to estimate the rotation and/or translation between two images—which may be multitemporal or multisensor—with small differences in the spectral content.

II. HAIRIS Description

Finding the relationship between two coordinate systems using pairs of measurements of the coordinates of a number of points in both systems is a classic photogrammetric task [19]. The transformation between two Cartesian coordinate systems can be thought of as the result of a rigid-body motion and can, thus, be decomposed into a rotation and a translation [19].

Suppose that we have a pair of images in the same scale (same pixel size with respect to the scene), and the existence of a translation and/or rotation difference between the two images, where one of them is "static" (image 1) and the other (image 2) is to be registered onto the "static" image. Assuming that (X, Y) are the coordinates of the "static" image and (P, L) are the (Pixel,Line) pair of the image to be registered. The transformation may be written as [3]:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} P \\ L \end{bmatrix} + \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \tag{1}$$

where the origin is considered to be the upper left corner of the "static" image, θ is the orientation difference, and (δ_x, δ_y) is the shift between the two images. The proposed methodology of automatic image registration is schematicaly represented in Fig. 1. It begins with a preprocessing stage in order to reduce unnecessary detail on the images content, important for the subsequent histogram-based image segmentation phase (which includes a relaxation parameter α). The objects extracted from the segmentation stage are characterized and matched according to some related properties, which finally allows for the statistically-based rotation and translation parameters estimation. In the following, the several steps involved in HAIRIS are explained, where the main objective is to estimate θ, δ_x, and δ_y.

A. Preprocessing

Too much detail on the pixel domain may lead to undesirable segmentation results. Therefore, it is advisable an image

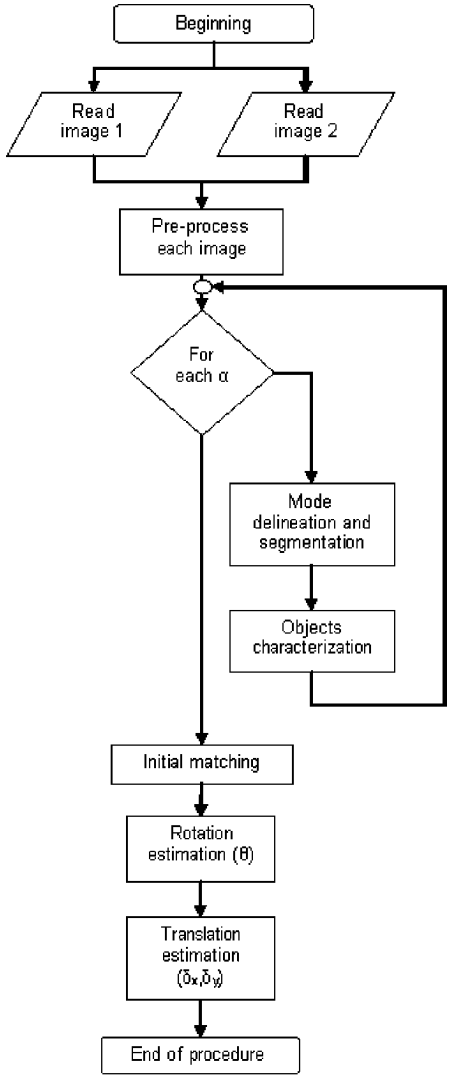

Fig. 1. Flowchart of the main steps of HAIRIS.

enhancement step prior further processing. By image enhancement (which is itself a largely subjective process), it is intended to obtain an image with less detail than the original version, nearest to the "object" identification which is performed by the human eye. Although typically more mathematical and complex, restoration algorithms may provide the exploitation of the detailed characteristics of an image and its degradation [25]. Despite the main purpose of image restoration methods is to model and remove the degradation, these methods may also be used with other purposes. Since with image segmentation it is intended to extract objects (in particular their boundaries), one may view the image objects which have some texture as a kind of degradation. Therefore, it is intended to remove that degradation, which is assumed to be additive random noise.

The Wiener filter [25] is one of the most used filters under the scope of image restoration methods [17]. However, it may also be used for image enhancement, with the aim of reducing the detail on an image, since it is typically a lowpass filter and consequently induces a significant blurring effect. Its frequency response $H(\omega_1, \omega_2)$ is given by

$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_v(\omega_1, \omega_2)} \tag{2}$$

where $P_f(\omega_1, \omega_2)$ stands for the original image power spectra and $P_v(\omega_1, \omega_2)$ for the additive random noise power spectra. Although one might be interested in inducing some blurring on the image, the use of a fixed (space-invariant) filter throughout the image may decrease the clearness of the objects boundaries. Therefore, an adaptive image restoration is required. A "pixel-by-pixel processing" approach may become quite computationally expensive, in opposition to a "subimage-by-subimage processing" where we divide the image into a certain number of tiles. The latter is typically considered for subimages with size between 8×8 and 32×32 pixels [25]. However, since the objective of the Wiener filter employment is different from restoration, it is advisable to consider the conservative smallest possible square tile (beyond a single pixel) size of 3×3 pixels. Although this latter approach might induce the so-called "blocking effect," it may be ignored for images with low SNR.

Recalling the additive random noise, we generally have no *a priori* information, which may be overcome by considering measured features such as the local variance, providing the determination of the presence of significant high-frequency details. Since we are interested in delineating the objects boundaries, an edge-sensitive adaptive image restoration version of the Wiener filter [25] is adequate. This method is based upon the idea of reducing more noise near edges without additional edge blurring through a cascade of 1-D adaptive filters. Let $T_i[\cdot](1 \leq i \leq N)$ represent a 1-D (space variant) filter, where N is four and represents the four directions corresponding to the angles of 0°, 45°, 90°, and 135°. Then, these operators are sequentially applied, which lead to an improvement over the performance of some 2-D adaptive restoration algorithms, and to typically less computational requirements [25].

Additionally, in order to overcome significant differences between the histograms of the images to be registered, an histogram equalization of image 2 using the histogram counts of image 1 is performed, prior to the application of the Wiener filter. In this way, the Wiener filtering on image 2 allows both for the reduction of the image detail, as well as to the smoothing of the histogram, which becomes spiky due to the histogram equalization step.

B. Histogram-Based Segmentation: Mode Delineation and Image Segmentation

The method used for mode delineation is based upon the analysis of the consecutive slopes of the histogram. Let $x(m)$ be the image histogram counts, $m = 0, \ldots, M$ and

$$y(n) = x(n) - x(n-1), \quad n = 1, \ldots, M \tag{3}$$

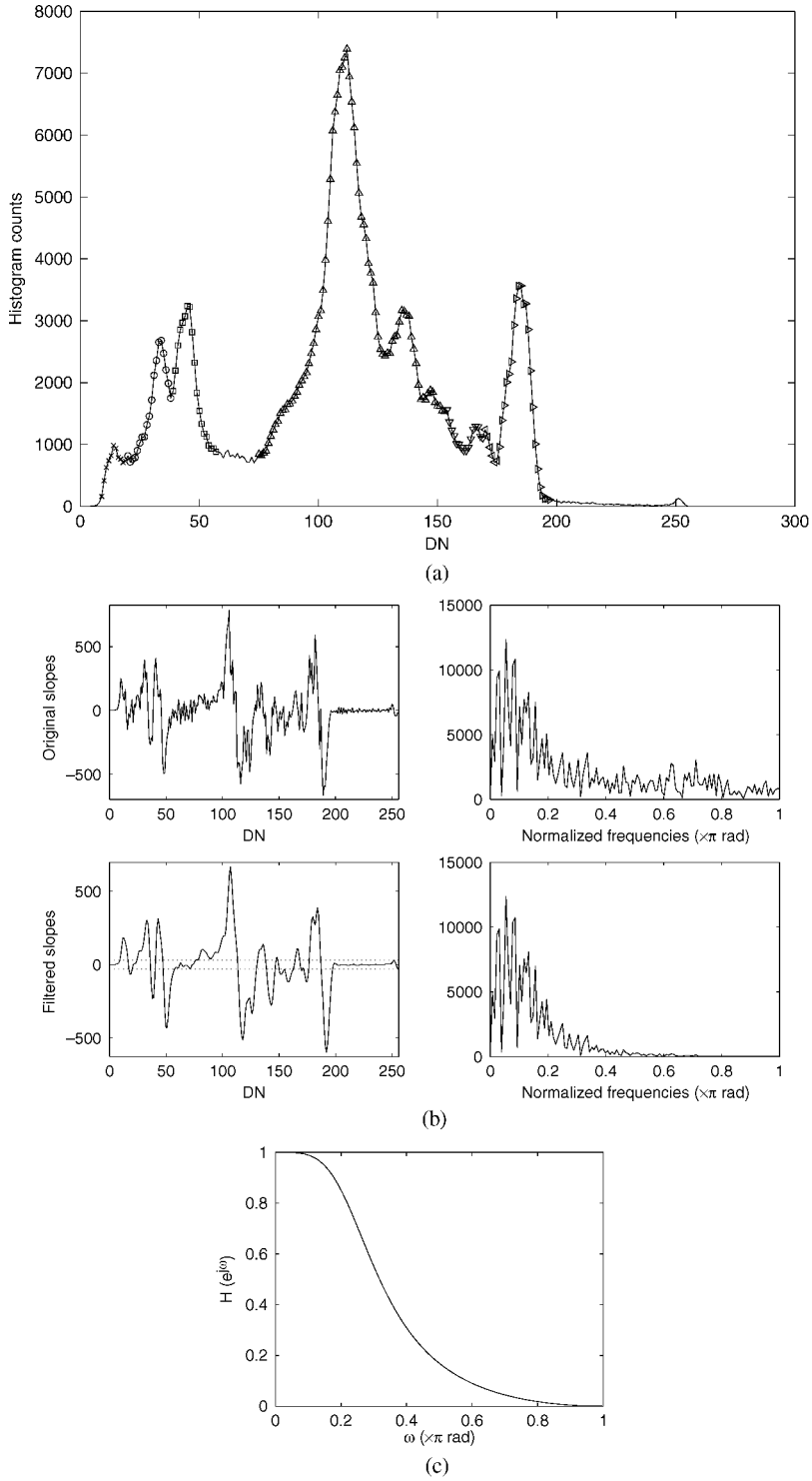

Fig. 2. (a) Histogram of the image presented in Fig. 4(a), where the different superimposed symbols represent the seven identified modes. (b) Slopes sequence (on the left) corresponding to the histogram in (a), before (upper plots) and after (lower plots) the application of the low-pass filter represented in (c), and their representation in the frequency domain (on the right). The limits (horizontal dashed lines) superimposed on the filtered slopes sequence correspond to the 99% confidence interval (further details in the text). (c) Frequency response $(H(e^{j\omega}))$ of the low-pass filter applied to the original slopes sequence in (b).

the sequence of the consecutive slopes, where $M + 1$ is the number of histogram levels ($M = 255$ for an 8-b image). The idea behind this approach is to choose an adequate threshold for considering whether or not one is in the presence of a mode, which is characterized by a significant increase and/or decrease on the slopes sequence. As illustrated by the example in Fig. 2, a mode is characterized by extreme positive and negative values of the slopes sequence. Therefore, one obvious and functional solution for delineating a mode is to obtain a confidence interval for the slopes sequence, where the presence of a mode is detected by the slopes which are outside the 99% confidence interval [Fig. 2(b)]. The t-Student based confidence interval for the mean is adequate for the most common 8-b images, since it corresponds to 256 points which is sufficiently higher than

the common advisable minimum of 30 points [13]. In order to achieve a robust detection through the statistical approach of the confidence interval, a preprocessing of the slopes is required, in order to smooth the slopes sequence irregularity outside the presence of a mode [Fig. 2(b)], which may induce the detection of false modes. This preprocessing is performed through a second-order low-pass Butterworth filter, with a normalized cutoff frequency at 0.25, which frequency response is represented in Fig. 2(c).

In Fig. 2, the histogram of the image in Fig. 4(a) is shown, where roughly seven modes are visible. Also in the same figure, the consecutive slopes of the histogram are shown, where each global transition from positive slopes to negative slopes is associated to a mode on the histogram.

A relaxation parameter (α) is considered on the mode delineation, which in theory is a continuous parameter, defined on the space [0,1], and has to pass through a discretization process in practice. The inclusion of this parameter leads to the obtention of several different segmentation results, which allows for the subsequent stages of the proposed methodology to be more robust. Furthermore, for situations where only a single and sharp mode is detected on the histogram, the inclusion of this parameter reveals to be of great importance, as illustrated in Section III for the satellite images pair. The relaxation parameter (α) corresponds to the proportion of the height of the histogram—considered to correspond to the highest mode—for which below this value, the mode is to be considered as a "flat" region. In this way, the method becomes much more useful and adaptable to a large variety of situations. The common methods of merely identifying a mode, correspond to selecting the null value for α.

Once the modes and "flat" regions are identified for each value of α, the image may be segmented considering each interval (either a mode or a flat region) as a class. Then, for each class, those pixels which are 4-connected are considering as belonging to the same object, resulting in the final segmentation of the image.

In some cases, a significant amount of no data ($\mathrm{DN} = 0$ or $\mathrm{DN} = 255$ for unsigned 8-b images) may be present, which can mask in some way the histogram shape. Therefore, those corresponding pixels should be discarded prior to the representation of the image histogram.

C. Characterization of the Extracted Objects

The extracted objects at the segmentation stage are characterized by four attributes which allow for their adequate morphological description: area (A_{rea}), perimeter (P_{erim}), axis ratio (AR_{at}) and fractal dimension (D_b).

The attribute area (A_{rea}) is merely obtained by the number of pixels which form an object, whereas the perimeter (P_{erim}) is obtained by calculating the distance between each adjoining pair of pixels around the border of the region. These two attributes allow for the evaluation of an object with respect to its size and compactness, respectively.

The major axis length corresponds to the major axis length of the ellipse that has the same normalized second central moments as the object, from which it may be also obtained the minor axis length. The ratio between the major and the minor axis length lead to the attribute AR_{at} (axis ratio). The AR_{at} attribute allows for the characterization of the object according to its narrow or wide nature.

Althoug the three previously described measures comprises the most general aspects of an object characterization, several uncertainties would still persist on the objects matching stage. Therefore, there is the need for considering a complementary attribute which considers the particular complexity nature of an object shape: the fractal dimension. Fractal dimension is one among several notions of dimension proposed by mathematicians. In this work, the box-counting dimension (D_b)—one of the special forms of Mandelbrot's fractal dimension—was the considered fractal dimension [27]. It generally consists on the slope of a straight line, fitted to a scatter plot with $\log(N(s))$ and $\log(1/s)$ on the vertical and horizontal axis, respectively, where s is the mesh size of the grid overlaid on the object and N is the number of grid boxes which contain pixels of the object.

The four previously described attributes are used for the later stage of objects matching. These four attributes are expected to be similar for corresponding objects. As stated at the beggining of HAIRIS description, it is assumed that the pair of images differ with respect to rotation (θ) and translation (δ_x and δ_y). Therefore, in order to allow for the registration of the pair of images, there is the need to have an orientation and positioning indicators. Regarding orientation, for each extracted object, the angle between the x-axis and the major axis of the ellipse that has the same second-moments as the object is stored. With respect to translation, the centroid of each object (the center of mass of the object) is also stored.

D. Matching

The matching step begins with the evaluation of a cost function, between every possible two-by-two combination of objects obtained by the segmentation of the two images, for every possible combination of the α values considered for both images. This leads to a matrix with n_1 rows and n_2 columns, where n_1 and n_2 correspond to the number of extracted objects from images 1 and 2, respectively. The cost function (γ), evaluated for the values of the properties of the objects from images 1 and 2, is defined as follows:

$$\gamma = \frac{(A_{\mathrm{rea1}} - A_{\mathrm{rea2}})}{\overline{A_{\mathrm{rea}}}} + \frac{(AR_{at1} - AR_{at2})}{\overline{AR_{\mathrm{at}}}} + \frac{(P_{\mathrm{erim1}} - P_{\mathrm{erim2}})}{\overline{P_{\mathrm{erim}}}} + \frac{(D_{b1} - D_{b2})}{\overline{D_b}} \quad (4)$$

where $\overline{A_{\mathrm{rea}}}$, $\overline{AR_{\mathrm{at}}}$, $\overline{P_{\mathrm{erim}}}$ and $\overline{D_b}$ are the average of each property for images 1 and 2 values. With γ defined in this manner, it allows for a normalized sum of the considered properties, and consequently to the sum of them.

Then, the γ values are represented in the form of boxplots, with the image which led to the lower number of segmented objects corresponding to the horizontal ("categorical") axis. A valid matching between two objects should lead to the lower values of γ, sufficiently far from the majority. This can be statistically evaluated through the outlier detection criterion used in the boxplots representation, where a point is considered an outlier (regarding the smaller values) if it is smaller than $q_1 - k \times (q_3 - q_1)$, where q_1 and q_3 are the first and third quartiles, respectively [15]. Although k is typically considered as

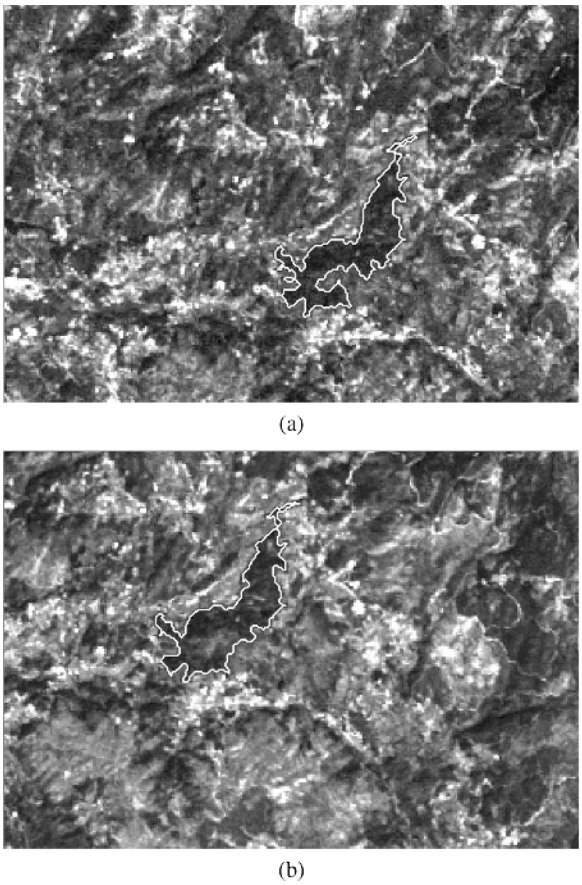

Fig. 3. Segment of the images shown in Fig. 7, where (a) corresponds to band 1 and (b) to band 2 with the simulated shift. Superimposed (in white) are the boundaries of two objects obtained from the segmentation stage, which have been matched. The attributes values $(A_{\mathrm{rea}}, AR_{\mathrm{at}}, P_{\mathrm{erim}}, D_b)$ from the object in (a) are (1759, 2.66, 647.7, 1.12) and from the object in (b) are (1685, 2.67, 555.2, 1.09).

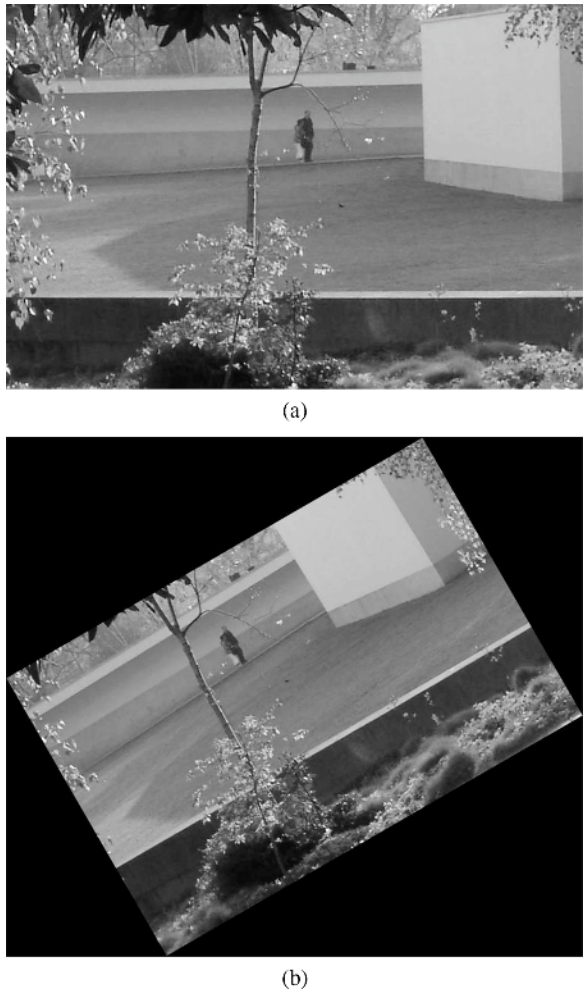

Fig. 4. (a) 8-b image with size 768 × 512 pixels, extracted from a larger scene acquired by a digital camera. (b) Image extracted from the same larger scene as (a), with a shift of 60 and 40 pixels on the horizontal and vertical axis, respectively, followed by a rotation of $\theta = 30^\circ$.

1.5, in this step the more flexible value of 1 is required (also commonly used in practice), in order to reduce the loss of eventual matching candidates. This procedure of outlier detection is applied to each object of the image in the horizontal axis. Fig. 3 illustrates an example of two regions which have been matched, and the values of the four attributes.

E. Rotation Estimation: $\hat{\theta}$

The rotation and translation are determined on a statistical basis. Given the set of matching objects candidates, the histogram of the extracted objects orientation differences is represented. This allows for the detection of a modal class, restricting the set of possible values for rotation. Then, among these rotation candidates, a robust strategy of finding $\hat{\theta}$ is performed, through considering the frequencies of the rotation candidates, and finding the rotation value which absolute frequency corresponds to the higher outlier, according to the procedure of boxplot outliers detection previously described. This procedure lead to a robust estimation of θ.

F. Translation Estimation: $(\hat{\delta_x}, \hat{\delta_y})$

Once $\hat{\theta}$ is obtained, only the initial matching candidates which correspond to the obtained rotation are considered. Then, a similar procedure as that followed in the rotation estimation is considered for obtaining $\hat{\delta_x}$ and $\hat{\delta_y}$. This statistically based procedure also leads to a robust estimation of δ_x and δ_y, as illustrated in the following section.

III. Application of HAIRIS

In this section, the assessment of HAIRIS performance is performed in the following situations: simulated rotation and translation, the same simulated rotation and translation with different levels of Gaussian white noise added to the image to be registered, as well as in simulated and real examples of remote sensing applications. For each image, the values considered for parameter α ranged from 0% to 100%, through steps of 20%, 10%, and 5%, in order to allow for a sensitivity analysis of the parameter α. The results were obtained on a computer with an Intel Core 2 6400 2.13 GHz processor and 2.0 Gbytes of physical memory, using MATLAB Release 2009b.

A. Simulated Rotation and Translation

An 8-b image with size 768 × 512 pixels, acquired by a digital camera was considered for the illustration of HAIRIS appli-

TABLE I
VALUES OF $\hat{\theta}$, $\hat{\delta_x}$ AND $\hat{\delta_y}$ (AND ASSOCIATED COMPUTATIONAL TIME) OBTAINED THROUGH HAIRIS APPLICATION IN THE REGISTRATION OF THE EXAMPLES DESCRIBED IN SECTIONS III-A, III-B AND III-C, FOR α IN STEPS OF 20%, 10%, AND 5%. REFERENCE $(\theta, \delta_x, \delta_y)$ VALUES FOR III-A AND III-B ARE (30°, 252.0, 292.7). REFERENCE $(\theta, \delta_x, \delta_y)$ VALUES FOR PAIRS 1, 2, AND 3 OF III-C ARE RESPECTIVELY (0°, 60, 40), $(0°, 3.88, -0.39)$ AND $(0°, -20.75 \pm 2.77, -3.38 \pm 0.70)$. * EXCLUDING THE TIME ASSOCIATED TO THE SEGMENTATION OF THE REFERENCE IMAGE ALREADY PERFORMED FOR III-A

α steps		III-A	III-B				III-C		
			10%	20%	50%	100%	Pair 1	Pair 2	Pair 3
20%	$\hat{\theta}$	30°	30°	30°	29°	30°	0°	0°	-1°
	$\hat{\delta_x}$	252.93	252.85	252.81	246.35	242.02	59.83	3.75	-12.15
	$\hat{\delta_y}$	293.17	293.15	293.09	286.97	293.00	-40.23	-0.84	-29.67
	Time (sec)	167	71*	75*	61*	58*	169	198	223
10%	$\hat{\theta}$	30°	30°	30°	30°	30°	0°	0°	2°
	$\hat{\delta_x}$	252.88	252.83	252.87	252.87	243.08	59.73	3.90	-27.71
	$\hat{\delta_y}$	293.16	293.08	293.10	293.04	293.06	-40.03	-0.94	0.32
	Time (sec)	291	133*	140*	115*	109*	358	394	428
5%	$\hat{\theta}$	30°	30°	30°	30°	30°	0°	0°	2°
	$\hat{\delta_x}$	252.95	252.91	252.88	252.90	243.08	59.76	3.80	-27.73
	$\hat{\delta_y}$	293.12	293.09	293.11	293.06	293.08	-40.01	-0.96	0.66
	Time (sec)	565	269*	270*	234*	216*	722	785	831

cation [Fig. 4(a)]. In order to have a known basis for the evaluation of HAIRIS performance, a translation of 60 and 40 pixels on the horizontal and vertical axis, respectively, followed by a rotation of $\theta = 30°$ were simulated [Fig. 4(b)]. As illustrated in the Appendix, when the rotated image is rotated backward by the same angle θ, there is an additional translation effect on the image, where $a = w\sin^2(\theta)$ and $b = w\cos(\theta)\sin(\theta)$. Therefore, the values of δ_x and δ_y the method should achieve are 252.0 and 292.7, respectively.

After the application of HAIRIS, we have obtained the correct value for θ for the three considered resolutions of α (20%, 10%, and 5%), and also a subpixel accuracy for the same three considered resolutions regarding δ_x and δ_y (Table I).

B. Simulated Rotation and Translation With Added Noise

In order to evaluate the robustness of HAIRIS in the presence of noise, several noise levels were considered, through the simulation of Gaussian white noise added to the uncorrected image considered in the previous subsection [Fig. 4(b)]. Given that the standard deviation of the image in Fig. 4(b) is 45 (excluding the background area induced by the simulated rotation), Gaussian white noise with standard deviation of 10%, 20%, 50%, and 100% of the image standard deviation were simulated (Fig. 5). The obtained values for $\hat{\theta}$, $\hat{\delta_x}$ and $\hat{\delta_y}$ through HAIRIS for the four considered noise levels are presented in Table I. HAIRIS has correctly estimated the rotation value (with the exception of a small error in one situation, Table I), estimated δ_y at the subpixel value with an error of less than 0.5 pixels for practically all situations, and to a smaller extent but still accurately estimated δ_x, for the three considered resolutions of α (Table I). The registration of the image with a noise level of 50% for an α resolution of 10% is illustrated in Fig. 6.

C. Remote Sensing Examples: Simulated and Real Situations

The first remote sensing example consists on a pair of satellite images with size 512 × 512 pixels of the northwest region of Portugal, acquired from the sensor Landsat/ETM+ (Fig. 7). The first was obtained from selecting a part of a larger scene from band 1 (0.45–0.515μm), whereas the second image was obtained from the same larger scene from band 2 (0.525–0.605μm), with a known shift of 60 and 40 pixels, on the horizontal and vertical axis, respectively. From the results presented in Table I, it can be seen that HAIRIS correctly estimated θ, and led to an accuracy higher than 0.30 pixels (subpixel level), for both δ_x and δ_y.

The second example consists on a segment of 512 × 512 pixels of the green band of an aerial photograph of the northwest region of Portugal (with an initial spatial resolution of 0.5 m resampled to 2.5 m), and on a segment of 512 × 512 pixels of a panchromatic image from the same region, acquired by the sensor ALOS-PRISM image (Fig. 8), with an initial approximate geometric correction part of the product (with a spatial resolution of 2.5 m). The displacement between these two images was manually obtained, through the identification of six conjugate points, leading to a null difference in rotation and to a difference of 3.88 pixels and −0.39 pixels on the horizontal and vertical axis, respectively. Also in this example, HAIRIS has correctly estimated θ, leading to a subpixel accuracy for both δ_x and δ_y (Table I).

The third example consists on a segment with 512 × 512 pixels of the green band of an orthophotograph covering part of the city of Porto (with an initial spatial resolution of 0.5 m resampled to 1 m), and on a segment with 512 × 512 pixels of

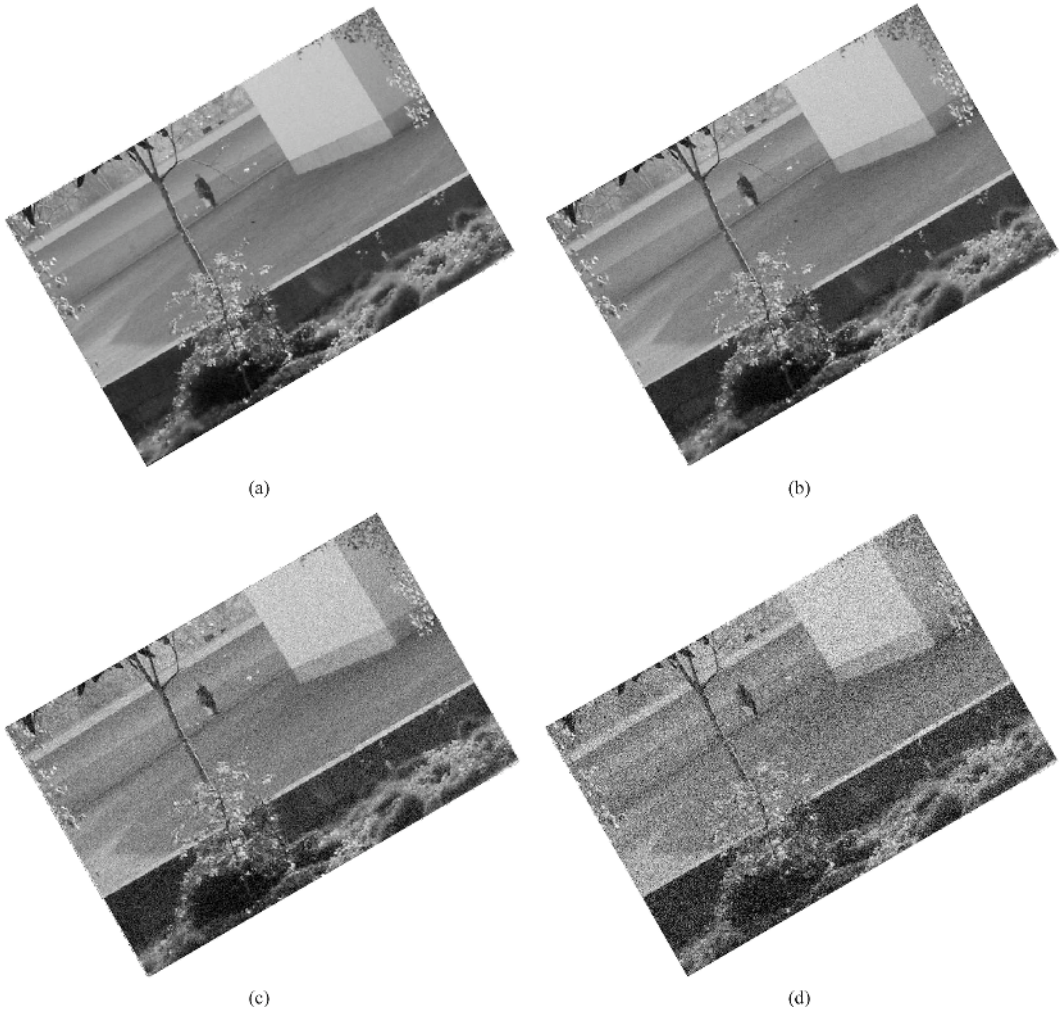

Fig. 5. Image represented in Fig. 4(b), after the addition of white Gaussian noise with σ equal to (a) 10%, (b) 20%, (c) 50%, and (d) 100% of the image standard deviation.

a panchromatic image from the same region, acquired by the sensor IKONOS (Fig. 9), with an initial approximate geometric correction part of the product (with a spatial resolution of 1 m). The temporal difference between these images is around one year, where significant changes such as new buildings (left part of the image) can be observed during this period. Additionally, beyond the difference between the sensors, it can be clearly seen the effect of different viewing angles (IKONOS image was acquired at a viewing angle of around 20°), which is associated to significant shadow effects (Fig. 9). All these differences in this pair of images makes this example a quite difficult pair to register (Fig. 9). The displacement between these two images was manually obtained, through a difficult identification of six conjugate points. The horizontal shifts between these conjugate points ranged between -16.25 and -23.25 (average of -20.75 and standard deviation of 2.77), whereas the vertical shifts ranged between -2.25 and -4.25 (average of -3.38 and standard deviation of 0.70). These differences are explained by the set of significant differences previously mentioned. Through the least squares method a null rotation was found between this pair of images and considering the six identified conjungate points. Also in this example, HAIRIS led to acceptable results considering the average and standard deviation of the reference values, despite all the difficulties contained in this pair of images (Table I).

D. Comparison of HAIRIS With Other Methods

In order to establish a baseline for comparison of HAIRIS with other AIR methods, two popular methods of AIR have been considered: scale invariant feature transform (SIFT) [26] and a contour-based approach [14], [23]. Since the former method (implementation of the author) provides a set of conjugate points, the parameters θ, δ_x and δ_y were estimated through the least squares method [29] (once redundancy is guaranteed) and according to (1). We considered the distance ratio equal to 0.6, as the value 0.8 suggested in [26] led to worst results in both photographs and remote sensing images. The implementation of the later method allows for considering the RST transformation (and not merely RT), and the default parameters have been considered. The results obtained from applying these two methods to the previously considered pairs of images are presented in Table II, including the associated computational time.

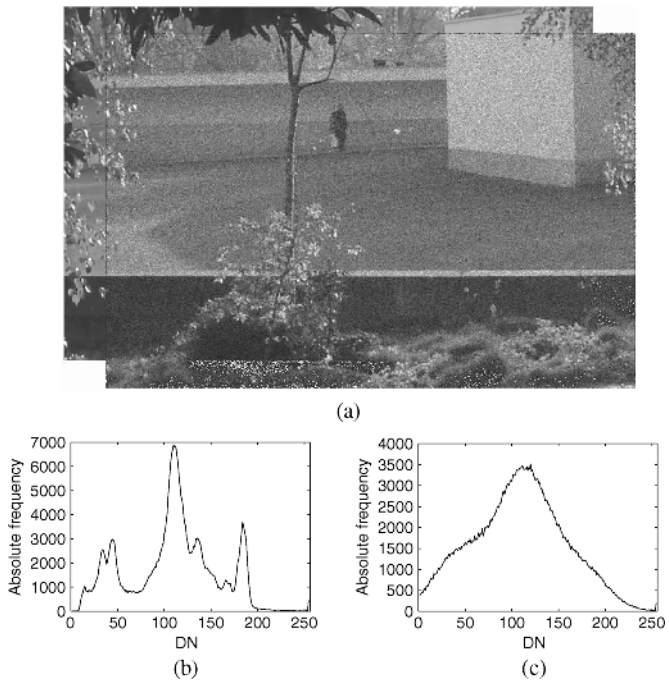

Fig. 6. (a) Registration of the original image in Fig. 4(a) with the simulated rotation and translation version with added noise at the level of 50% [Fig. 5(c)], for an α resolution of 10%. (b) Histogram of the original image in Fig. 4(a). (c) Histogram of the image in Fig. 5(c), ommiting $\mathrm{DN} = 0$ to allow for the visualization of the remaining histogram.

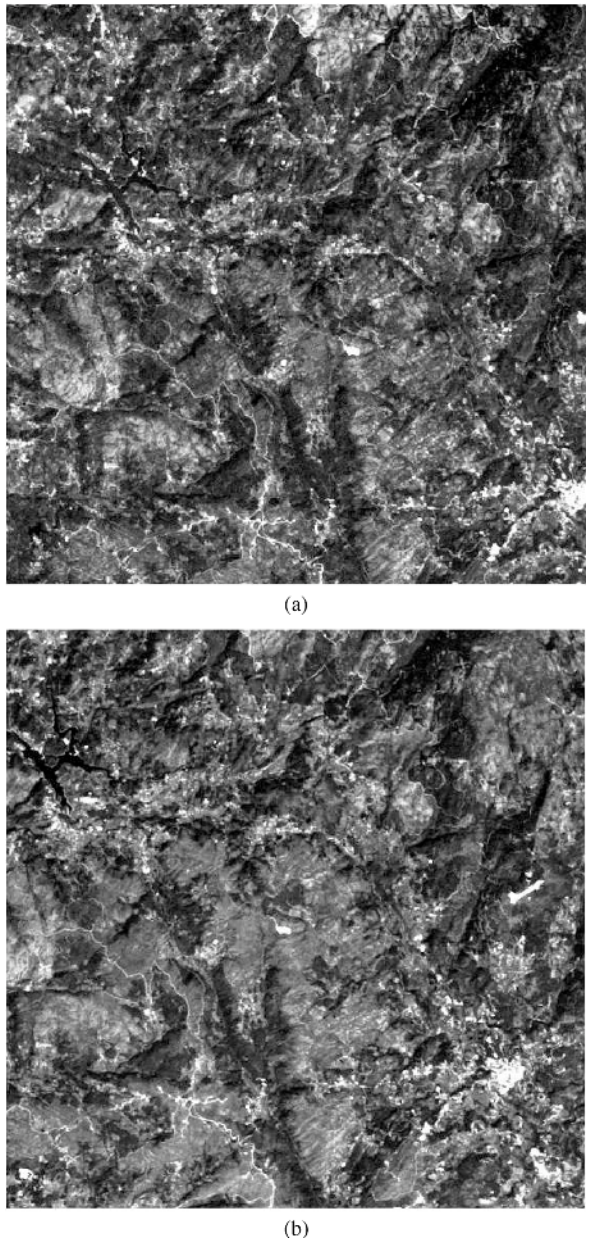

Fig. 7. (a) 8-b image with size 512 × 512 pixels, extracted from band 1 of a larger scene acquired by the sensor ETM+ from Landsat 7 satellite. (b) 8-b image with size 512 × 512 pixels, extracted from band 2 of the same larger scene as (a), with a shift of 60 and 40 pixels on the horizontal and vertical axis, respectively. The presented images were contrast stretched in order to allow for a better visualization.

It can be observed that HAIRIS generally outperformed SIFT and the countour-based approach, in particular for the remote sensing examples. Although SIFT is a powerful method under the scope of computer vision applications, its performance is quite limited for remote sensing images [30], [39]. Furthermore, its dependence upon the parameter associated to the ratio of distances is an undesirable property under the scope of fully AIR methods [26]. The contour-based approach [14], [23] exhibited in general the worst performance. Although it is a quite fast method, it is not able to lead with more complex situations such as significant presence of noise or multisensor pairs of images (Table II), whereas HAIRIS had accurately registered these complex pairs of images (Table I). Although higher computational times were associated with HAIRIS (mainly related to the segmentation stage), it is worth to mention that we implemented HAIRIS in MATLAB (the contour-based approach implementation is based upon C++ which is faster) and the code is far from being optimized, since several outputs are produced given the experimental state of the method. Moreover, it showed to be able to register pairs of images covering a wide range of situations (multitemporal, multisensor, and in the presence of noise).

IV. Discussion and Conclusion

A large variety of automatic image registration methods can be found in [3] and [45]. In several applications, the registration model only assumes rotation and translation [22], where the registration of satellite images is an example [2], [40]. In this paper, a new approach for automatic image registration through histogram-based image segmentation (HAIRIS) is proposed, with clear advantages by joining these two main areas of image processing. Since HAIRIS does not require any search interval either for rotation or translation, it is a fully automatic procedure.

With the filtering step—an important preprocessing stage of the proposed methodology—the objective is to transform the original image in order to take advantage of the psychophysical aspects of the human visual system. The Wiener filter is one of the solutions among several other possible alternatives. The proposed Wiener filtering, although with its associated disadvantages, showed to be adequate for the variety of examples considered in this manuscript. The only parameter associated to

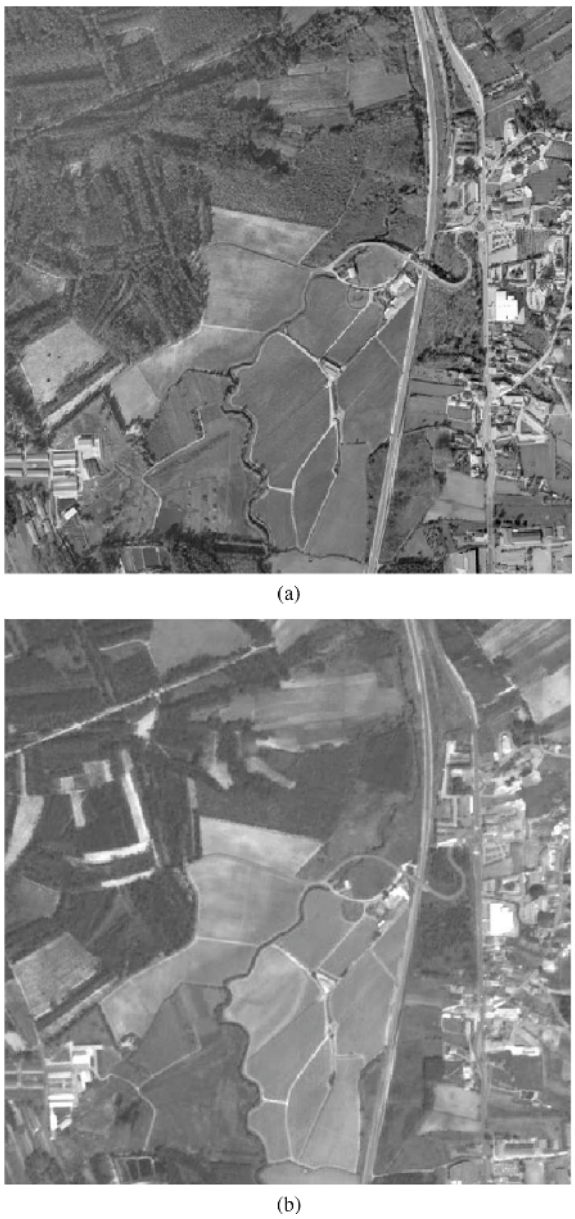

Fig. 8. (a) Green band of an aerial photograph with size 512×512 pixels (with spatial resolution of 2.5 m). (b) Panchromatic 8-b image with size 512×512 pixels (with spatial resolution of 2.5 m), extracted from a larger scene acquired by the ALOS-PRISM sensor of the same larger scene as (a), with a shift of about 3.88 and -0.39 pixels on the horizontal and vertical axis, respectively, obtained through the average of six conjugate points manually identified.

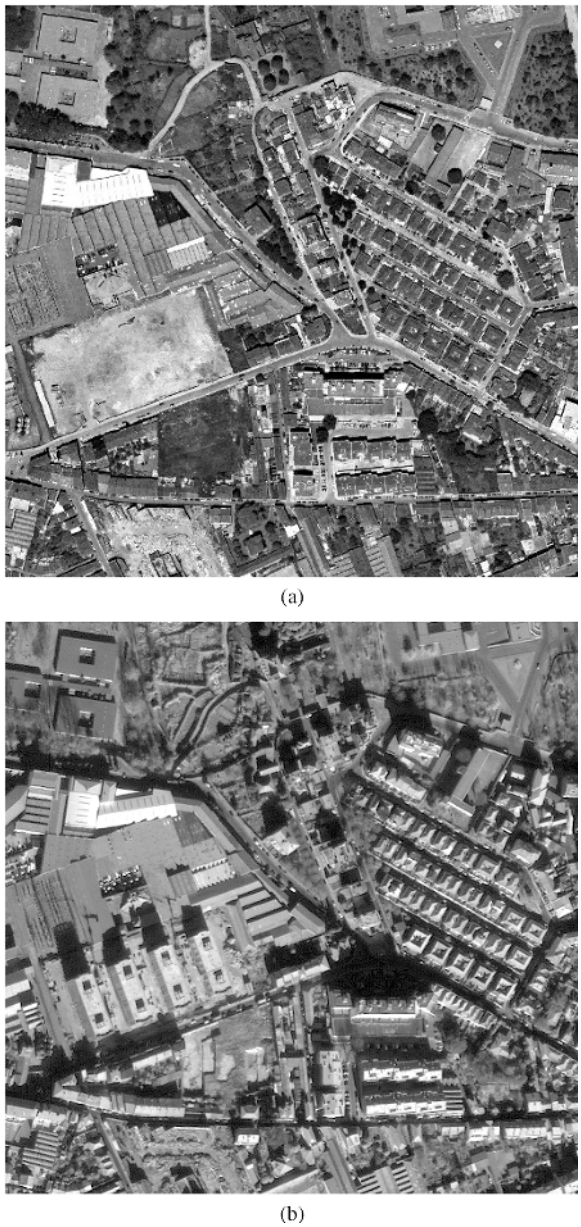

Fig. 9. (a) Green band of an orthophoto with size 512×512 pixels (with spatial resolution of 1 m). (b) Panchromatic 8-b image with size 512×512 pixels (with spatial resolution of 1 m), extracted from a larger scene acquired by the IKONOS sensor of the same larger scene as (a), with a shift of about -20.75 ± 2.77 and -3.38 ± 0.70 pixels on the horizontal and vertical axis, respectively, obtained through the average of six conjugate points manually identified.

the proposed Wiener filtering is the size of the local neighborhood, where our choice revealed to be adequate for a considerable variety of examples. Althgouh adaptive restoration algorithms require considerably more computation than nonadaptive algorithms, they often perform better than nonadaptive restoration algorithms.

Several other filters had also been tested, from different filter categories. The median filter, one of the most used filters, produces a blurred version of an original image. As a consequence, beyond the intended detail reduction, the distinction between objects on the image is sometimes lost, since it induces a significant smoothing on the entire image, including the objects limits. We have also exhaustively tested the application of the anisotropic diffusion for the considered examples [35]. However, anisotropic diffusion showed to be sensitive to the parameters choice (conduction coefficient, number of iterations and function generating the scale-space). In particular, it was not possible to find a set of parameters (or a clear rule) which would be adequate for all the considered examples. As HAIRIS is a fully automatic image registration method, it is not compatible with a preprocessing step sensitive to a wrong selection of parameters. The use of other methods, such as wavelets, have not been considered given their even higher dependence upon several parameters for different situations, which is not compatible with automatic image registration methods.

TABLE II
VALUES OF $\hat{\theta}$, $\hat{\delta_x}$ AND $\hat{\delta_y}$ (AND ASSOCIATED COMPUTATIONAL TIME) OBTAINED THROUGH THE APPLICATION OF SIFT [26] AND A CONTOUR-BASED APPROACH [14], [23] IN THE REGISTRATION OF THE EXAMPLES DESCRIBED IN SECTIONS III-A, III-B, AND III-C. REFERENCE $(\theta, \delta_x, \delta_y)$ VALUES FOR III-A AND III-B ARE $(30°, 252.0, 292.7)$. REFERENCE $(\theta, \delta_x, \delta_y)$ VALUES FOR PAIRS 1, 2, AND 3 OF III-C ARE RESPECTIVELY $(0°, 60, 40)$, $(0°, 3.88, -0.39)$ AND $(0°, -20.75 \pm 2.77, -3.38 \pm 0.70)$. * IT WAS NOT FOUND ANY CONJUGATE POINT. ** INDICATION OF BAD FIT BY THE METHOD IMPLEMENTATION

		III-A	III-B				III-C		
			10%	20%	50%	100%	Pair 1	Pair 2	Pair 3
SIFT [26]	$\hat{\theta}$	30°	30°	30°	30°	30°	0°	*	*
	Scale	1.00	1.00	1.00	1.00	0.94	0.90	*	*
	$\hat{\delta_x}$	252.15	250.83	251.57	248.94	249.87	51.29	*	*
	$\hat{\delta_y}$	292.72	291.72	290.88	287.04	260.64	-80.94	*	*
	Time (sec)	16	16	15	13	12	2	10	16
Contour-based approach [14], [23]	$\hat{\theta}$	30°	28°	37°	**	**	0°	-11°	83°
	Scale	1.00	1.03	0.78	**	**	1.00	0.77	0.43
	$\hat{\delta_x}$	251.27	232.52	333.26	**	**	60.00	45.70	140.37
	$\hat{\delta_y}$	293.49	305.14	106.59	**	**	-40.00	187.47	308.77
	Time (sec)	1	2	2	1	1	1	1	2

Although a large variety of histogram shapes can be found in practice, single or multimode histograms often appear, which reinforces the importance of the mode delineation. The approach of mode delineation proposed in this paper, using the histogram consecutive slopes sequence, is a robust method regarding an accurate delineation of the modes present on an histogram. The considered nonparametric approach—mainly based upon the histogram consecutive slopes analysis—has shown to be efficient. In some rare cases, it is possible that no mode is present on the histogram or it is almost undetectable. In these particular cases, the solution may pass through an histogram transformation prior to the mode delineation. An objective identification of these cases may be performed through statistical methods such as the kurtosis measure [13] for instance.

Several papers are devoted to the selection of the appopriate threshold regarding single or multiple modes on the histogram, through a large variety of methods such as entropy [37], hierarchical cluster analysis [1], among others. Despite most of the known segmentation methods work well for certain applications, with images containing less evident objects—such as remote sensing applications—they will exhibit a lower performance. This aspect leads to the interest of considering the relaxation parameter (α) proposed in this paper, which allows for a more efficient histogram-based segmentation, having in mind a posterior image registration procedure. This alternative revealed to be adequate, avoiding alternatives which are only appropriate to particular histogram shapes, such as assuming one [36] or two univariate populations, even devoting attention to the common incomplete assignment of voxels in the image, by indicator kriking for instance [31].

Image segmentation methods based upon histogram thresholding are among the simpler methods. Although this class of methods may present some limitations, it has the advantage of being less dependent upon the choice of parameters, which is an utmost requirement under the scope of fully automatic procedures such as HAIRIS. In particular, the proposed approach allows for reducing in some way some of the limitations of using histogram thresholding. In fact, this is one of the strengths of HAIRIS, as supported by the presented examples where a pair of images with completely different histograms is accurately registered. The histogram thresholding approach which is proposed cannot be considered as the traditional histogram thresholding. Through the proposed "Mode delineation and image segmentation" stage, several levels of segmentation based upon the several modes which may be detected lead to a set of extracted objects, which when combined at the subsequent stages of HAIRIS, lead to a robust and accurate registration for a large variety of real situations (supported by the presented results).

The proposed method is based upon detecting closed similar regions in both images. Taking into account that the pair of images to be registered presented limited differences regarding their spectral content, it will usually be possible to detect similar regions in both images, even for regions with low contrast (as Fig. 3 is an example). Furthermore, one important characteristic of HAIRIS is the segmentation which is produced at different levels, by considering a range of values for the relaxation parameter α. This allows for the obtention of "several segmentations" and consequently to a more robust subsequent stage of initial matching. Moreover, it should be noticed that a closed region is a subjective concept. For instance, the segmentation of a river may be seen as a line, but may also be considered a closed region depending upon its width.

A sensitivity analysis of the parameter α has been performed, which indicated that, as expected, a higher resolution of the parameter space led to better results. One of the reasons behind this achievement is that it inevitably leads to a more robust determination of the rotation and translation parameters, since they are determined on a statistical basis. This may also justify some "saturation" on the results accuracy when comparing the results

obtained between a resolution of 10% and a resolution of 5%. Therefore, a resolution of 10% for parameter α may be adequate for several situations.

The determination of the rotation and translation values, which is a statistically based procedure, is a new approach regarding image registration methods, which revealed to be a robust and meaningful process. This was possible since it was combined with the also new approach for histogram-based image segmentation, which together leads to consistent results, as shown by the presented results in the different situations (different levels of noise and controlled differences of spectral content).

Although a set of measures for an objective evaluation of the geometric correction process quality have been recently proposed [16]—with particular importance under the scope of the registration of satellite images—since HAIRIS final result is the value of the rotation and translation difference between the two images, rather than a set of conjugate points, the referred set of measures would not bring additional value to the evaluation of HAIRIS performance. Regarding the evaluation of automatic image registration methods performance, it still lacks a dataset similar to the Berkeley dataset, widely used regarding image segmentation methods. With such a dataset, it would facilitate a universal comparison and convey the limitations and abilities of each image registration method.

For larger images than those included in this work, the division of the images into tiles may be appropriate, since too much differences on the image content may difficult the application of the proposed methodology. The limit to consider the division of the image into tiles may depend upon several factors, such as the image content, among others. Moreover, the division of larger scenes into tiles, and considering the centers of each tile as matching points, it may become possible to correct for stronger distortions than the rigid-body transformation.

Most of the known image registration methods work well for well defined pairs of points, mainly on noise free images and/or with objects clearly distinguishable from the background. With the proposed methodology, we are able to deal with these difficulties at a certain level, as it was demonstrated with the presented results.

The histogram shape may greatly vary depending upon the context of the application. Therefore, the concept of segmentation with different values for the relaxation parameter itself, may become useful in the context of several applications of pure image segmentation applications, since it allows for the obtention of several sets of objects, which may lead to more meaningful results than the "rigid" histogram-based segmentation methods.

The most popular methods of image registration are those based upon the correlation coefficient [3], [20], [45], in which a smaller part of an image is used as a window on the other image, looking for a maximum of correlation. Although they are appropriate for some situations, when there is a difference in orientation between the pair of images to be registered (for larger images, a difference of a single 1° may become significant), it may be sufficient to lead to an absence of a strong correlation maximum. Mutual information [6], [11], [20], [45], which had also become popular in the last years, is also not able to deal with significant rotation effects under the scope of the traditional search for a maximum on the similarity surface.

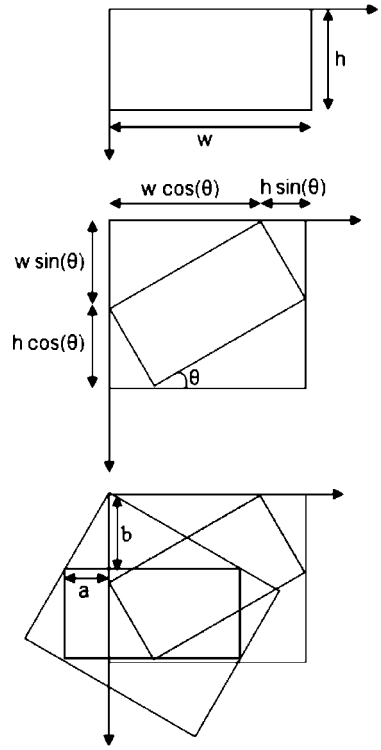

Fig. 10. Translation induced by an image backward rotation

We have compared HAIRIS with SIFT [26] and a contour-based approach [14], [23]—which are among the most popular methods for automatic image registration—in terms of accuracy and computational time. It was observed that HAIRIS generally outperformed SIFT and the countour-based approach, in particular for the remote sensing examples. The main drawback of HAIRIS is the computational time, mainly associated to the segmentation stage. However, it is expected to optimize the implementation code in the future in order to provide a faster performance.

In this work, HAIRIS was applied to single-band images at a time. However, in the future, adequate transformations (such as principal component analysis, independent component analysis, among others) of multi- (or hyper-) spectral images to single-band images will certainly lead to even better results, rather than using the information of a single spectral band.

The proposed methodology of image registration allowed for the obtention of accurate results, even in the presence of a considerable amount of noise. Furthermore, under the scope of applications with images having less evident objects, as is the case of remote sensing images, HAIRIS has shown to correctly register a pair of images at the subpixel level covering a wide range of situations (including multitemporal and multisensor).

APPENDIX

See Fig. 10.

REFERENCES

[1] A. Z. Arifin and A. Asano, "Image segmentation by histogram thresholding using hierarchical cluster analysis," *Pattern Recognit. Lett.*, vol. 27, pp. 1515–1521, 2006.

[2] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin, "An automatic image registration for applications in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2127–2137, Sep. 2005.

[3] L. G. Brown, "A survey of image registration techniques," *Comput. Surv.*, vol. 24, no. 4, pp. 325–376, 1992.

[4] J. H. Chang, K. C. Fan, and Y. L. Chang, "Multi-modal gray-level histogram modelling and decomposition," *Image Vis. Comput.*, vol. 20, pp. 203–216, 2002.

[5] C. I. Chang, Y. Du, J. Wang, S. M. Guo, and P. D. Thouin, "Survey and comparative analysis of entropy and relative entropy thresholding techniques," *IEE Proc.-Vis. Image Signal Process.*, vol. 153, no. 6, pp. 837–850, 2006.

[6] H. Chen, P. K. Varshney, and M. K. 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.

[7] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2071–2082, Dec. 2000.

[8] H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognit.*, vol. 34, pp. 2259–2281, 2001.

[9] M. Cheriet, J. N. Said, and C. Y. Suen, "A recursive thresholding technique for image segmentation," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 918–921, Jun. 1998.

[10] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proc. IEEE*, vol. 67, no. 5, pp. 773–785, May 1979.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: Wiley, 1991.

[12] P. Dare and I. Dowman, "An improved model for automatic feature-based registration of SAR and SPOT images," *J. Photogram. Remote Sens.*, vol. 56, pp. 13–28, 2001.

[13] E. Dudewicz and S. Mishra, *Modern Mathematical Statistics*. Hoboken: Wiley, 1988.

[14] D. Fedorov, L. M. G. Fonseca, C. Kenney, and B. S. Manjunath, "Automatic registration and mosaicking system for remotely sensed imagery," in *Proc. Image Signal Process. Remote Sens. VIII*, 2003, vol. 4885, pp. 444–451.

[15] M. Frigge, D. C. Hoaglin, and B. Iglewicz, "Some implementations of the boxplot," *Amer. Statist.*, vol. 43, no. 1, pp. 50–54, 1989.

[16] H. Gonçalves, J. A. Gonçalves, and L. Corte-Real, "Measures for an objective evaluation of the geometric correction process quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 292–296, Apr. 2009.

[17] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 2002.

[18] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-24, no. 3, pp. 390–399, May 1986.

[19] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 4, pp. 629–642, 1987.

[20] 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.

[21] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vis. Graph. Image Process.*, vol. 29, no. 3, pp. 273–285, 1985.

[22] G. Lazaridis and M. Petrou, "Image registration using the walsh transform," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2343–2357, Aug. 2005.

[23] H. Li, B. S. Manjunath, and S. K. Mitra, "A countour-based approach to multisensor image registration," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 320–334, Mar. 1995.

[24] Z. P. Liang, H. Pan, R. L. Magin, N. Ahuja, and T. S. Huang, "Automated image registration by maximization of a region similarity metric," *Int. J. Imag. Syst. Technol.*, vol. 8, no. 6, pp. 513–518, 1997.

[25] J. S. Lim, *Two-Dimensional Signal and Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 1990.

[26] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, 2004.

[27] B. B. Mandelbrot, "Self-affine fractals and fractal dimension," *Physica Scripta*, vol. 32, no. 4, pp. 257–260, 1985.

[28] P. Masson and W. Pieczynski, "SEM algorithm and unsupervised statistical segmentation of satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp. 618–633, May 1993.

[29] E. M. Mikhail, J. S. Bethel, and J. C. McGlone, *Introduction to Modern Photogrammetry*. Hoboken, NJ: Wiley, 2001.

[30] A. Mukherjee, M. Velez-Reyes, and B. Roysam, "Interest points for hyperspectral image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 748–760, Mar. 2009.

[31] W. Oh and W. B. Lindquist, "Image thresholding by indicator kriging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 7, pp. 590–602, Jul. 1999.

[32] P. Orbanz and J. M. Buhmann, "Nonparametric bayesian image segmentation," *Int. J. Comput. Vis.*, vol. 77, pp. 25–45, 2008.

[33] N. R. Pal and S. K. Pal, "A review on imagem segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.

[34] S. K. Pal, A. Ghosh, and B. U. Shankar, "Segmentation of remotely sensed images with fuzzy thresholding, and quantitative evaluation," *Int. J. Remote Sens.*, vol. 21, no. 11, pp. 2269–2300, 2000.

[35] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[36] P. L. Rosin, "Unimodal thresholding," *Pattern Recognit.*, vol. 34, pp. 2083–2096, 2001.

[37] P. Sahoo, C. Wilkins, and J. Yeager, "Threshold selection using renyi's entropy," *Pattern Recognit.*, vol. 30, no. 1, pp. 71–84, 1997.

[38] J. Serra, "A lattice approach to image segmentation," *J. Math. Imag. Vis.*, vol. 24, pp. 83–130, 2006.

[39] B. Sirmaçek and C. Ünsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, Apr. 2009.

[40] S. Suri and P. Reinartz, "Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 939–949, Feb. 2010.

[41] O. J. Tobias and R. Seara, "Image segmentation by histogram thresholding using fuzzy sets," *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1457–1465, Dec. 2002.

[42] J. C. Yen, F. J. Chang, and S. Chang, "A new criterion for automatic multilevel thresholding," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 370–378, Mar. 1995.

[43] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[44] X. P. Zhang and M. D. Desai, "Segmentation of bright targets using wavelets and adaptive thresholding," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1020–1030, Jul. 2001.

[45] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.

Hernâni Gonçalves was born in Porto, Portugal, in 1979. He received the Licenciatura degree in technology applied mathematics and the M.Sc. degree in computational methods in science and engineering from the Universidade do Porto, Porto, Portugal, in 2002 and 2004, respectively, and is currently pursuing the Ph.D. degree in surveying engineering from the same university.

He worked as a Researcher under the scope of a project in biomedical engineering, in the analysis of biomedical signals. Since 2006, he has been a Researcher at Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências, Universidade do Porto. His research interests include image processing and biomedical signal analysis.

José Alberto Gonçalves was born in Porto, Portugal, in 1964. He received the degree in surveying engineering from the Universidade do Porto, Portugal, in 1988, the M.Sc. degree in geographical information systems and the Ph.D. degree from the University College London, London, U.K., in 1993 and 2001, respectively.

In 1989, he joined the Universidade de Trás-os-Montes e Alto Douro. In 1997, he joined the Universidade do Porto, as a Lecturer in surveying engineering and is currently an Assistant Professor in the Departamento de Geociências, Ambiente e Ordenamento do Território, Universidade do Porto. He is also a Researcher at the Centro Interdisciplinar de Investigação Marinha e Ambiental. His research interests include photogrammetry, remote sensing and geo-information.

Luís Corte-Real (M'91) was born in Vila do Conde, Portugal, in 1958. He received the degree in electrical engineering from the Faculdade de Engenharia, Universidade do Porto, Porto, Portugal, in 1981, the M.Sc. degree in electrical and computer engineering from the Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal, in 1986, and the Ph.D. degree from the Faculdade de Engenharia, Universidade do Porto, in 1994.

In 1984, he joined the Universidade do Porto as a Lecturer of telecommunications. He is currently an Associate Professor in the Departamento de Engenharia Electrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto. Since 1985, he has been a Researcher at the Institute for Systems and Computer Engineering (INESC) Porto, Portugal. His research interests include image/video coding and processing.