Journal of Electronic Imaging

SPIEDigitalLibrary.org/jei

Epipolar image rectification through geometric algorithms with unknown parameters

José Herráez
José Luis Denia
Pablo Navarro
Jaime Rodríguez
M. Teresa Martín

SPIE

IS&T
imaging.org

Epipolar image rectification through geometric algorithms with unknown parameters

José Herráez
José Luis Denia
Polytechnic University of Valencia
School of Civil Engineering
Department of Cartographic Engineering, Geodesy, and Photogrammetry
Valencia 46022, Spain
E-mail: jherraez@cgf.upv.es

Pablo Navarro
Polytechnic University of Valencia
School of Architecture
Department of Graphic Expression in Architecture
Valencia 46022, Spain

Jaime Rodríguez
M. Teresa Martín
University of Santiago de Compostela
Polytechnic College of Lugo
Lugo 27002, Spain

Abstract. *Image processing in photogrammetry is commonly used for scene reconstruction. Although two-dimensional applications can be solved using isolated images, reconstruction of three-dimensional scenes usually requires the use of multiple images simultaneously. Epipolar image rectification is a common technique for this purpose. It typically requires internal orientation parameters and, therefore, knowledge of camera calibration and relative orientation parameters between images. A reparameterization of the fundamental matrix through a completely geometric algorithm of seven parameters that enables the epipolar image rectification of a photogrammetric stereo pair without introducing any orientation parameters and without premarking ground control points is presented. The algorithm enables the generation of different stereoscopic models with a single photogrammetric pair from unknown cameras, scanned from a book, or frames from video sequences. Stereoscopic models with no parallaxes have been obtained with a standard deviation of* <0.5 pixels. *© 2013 SPIE and IS&T [DOI: 10.1117/1.JEI.22.4.043021]*

1 Introduction

The use of photogrammetry has become increasingly common for a variety of technical applications, such as civil engineering, architecture, and robotics. The reason for this increase in popularity is that photographic images provide complete information about the object being studied, thereby complementing its metric capability. From this point of view, two main routes have been developed for its use: single images and sets of images. Isolated images are ideally suited for two-dimensional (2-D) applications, whereas multiple series of images are necessary for the reconstruction of three-dimensional (3-D) images and scenes.

The accurate reconstruction of the 3-D scene structure from two different projections and the estimation of the camera scene geometry is of paramount importance for many computer vision tasks.[1] A stereo pair consists of two images of the same rigid scene with purely horizontal parallax taken from two different viewpoints, which are related by the so-called epipolar transformation, also called epipolar geometry.[2] Thus, epipolar geometry from stereo images is commonly used for multiple-image 3-D scene depth reconstruction[3] in which all of the available image frames are considered simultaneously.[4] The epipolar geometry is completely characterized by the fundamental matrix,[5] where most of the information about the camera-scene geometry is encapsulated. Estimating the fundamental matrix has been an objective of research for many years and continues to be a challenging task in current computer vision systems. Hartley[6] explores the relationship between two images of a stereoscopic scene through projective geometry using the fundamental matrix by minimizing the deformation of the original images. Mallon and Whelan[7] used the Jacobian transformation of the fundamental matrix to study the minimum deformation of the original images to be rectified. Zhang[8] presented an introduction to epipolar geometry, a complete review of the current techniques for estimating the fundamental matrix and its uncertainty, and a new specification for the fundamental matrix that is valid for both the perspective and affine cameras based on a measurement that is obtained through sampling the whole visible 3-D space.

Paper 13097 received Mar. 1, 2013; revised manuscript received Oct. 2, 2013; accepted for publication Oct. 22, 2013; published online Dec. 2, 2013.

Zhang and Loop[9] analyzed all possible cases, where epipolar image rectification could fail due to the epipoles located at the infinity. Seitz[10] presents a theory of stereo image formation that enables complete classification and modeling of all possible stereo views through the concept of the quadratic view. Toward this goal, the notion of epipolar geometry has been generalized to apply to multiperspective images and, in recent years, to linear pushbroom images that result from different processes with perspective images[11] or spaceborne linear array scanner scenes, which are used increasingly often.[12]

The basic problem of recovering the 3-D structure of a scene from a set of images is the correspondence search.[13] For a single point in one of the images, its correspondences in the other images must be detected for the 3-D reconstruction. Depending on the algorithm, two or more point correspondences, as well as the camera geometry, are used to estimate the depth of that point.[14] Several approaches use the epipolar constraint to establish correspondences between two perspective images,[15] where the cost for a single match is computed from a large number of pixels rather than from only two pixels, as in traditional stereo pairs.[16] An epipolar plane image (EPI) is a 2-D spatiotemporal image usually obtained from a dense image sequence that is rectified so that each scene point is projected to the same row for all of the frames. Matousek and Hlavác[17] presented the idea of how to use the set of EPIs for seeking correspondences. The well-organized structure of EPIs enables all of the image data from a sequence to be employed directly in computation. This approach works well when the data are processed in accordance with the strict assumptions of Lambertian surfaces and accurately rectified sequences. Criminisi et al.[18] described an approach for automatically recovering 3-D layers from extended multiview sequences by analyzing the data in the entire EPI volume. The approach is based on decomposing the EPI volume into a set of EPI tubes, each of which represents a coherent subvolume that corresponds to a coherent portion of 3-D space. The EPI tubes are the basis for a complete 3-D layered sprite representation and for novel techniques to separate diffuse and specular components in static scenes.

Many of these stereo techniques require the knowledge of internal orientation (IO) parameters (camera calibration). For uncalibrated cases, the epipolar geometry is the only information that can be derived from point matches. Some authors consider epipolar geometry to be a type of stereo calibration, called weak calibration.[19,20] Hartley et al.[21] present an approach that uses the essential matrix defined by Longuet-Higgins for stereo reconstruction and avoids the need for camera calibration. That general method relies strongly on techniques of the projective geometry, in which a configuration of points may be subject to projective transformations in both 2-D image space and 3-D object space without changing the projective configuration of the points. Hartley and Gupta[22] further develop projective geometry applications, applying calibration-free methods to the stereo problem, using the essential matrix and ground-control points or some other constraints to reconstruct the absolute configuration of the scene. A quasi-Euclidean methodology developed by Fusiello and Irsara[23] can be used when the calibration information is unknown and exhibits sufficient accuracy for most applications. Wexler et al.[24] determined the epipolar geometry of a photogrammetric stereo pair using only measures, without known parameters for the camera, by analyzing pixel intensity values for pairs of images obtained with stereo cameras with fixed configurations and aggregating information from multiple pairs to create a dense map of epipolar curves. This approach does not model the 3-D geometric configuration of the cameras, but learns the shape of the epipolar curves by accumulating matching evidence over multiple image pairs. Therefore, any optical configuration can be used, but the technique requires tens to hundreds of images to obtain a representative calibration.

This manuscript shows a reparameterization of the fundamental matrix through a completely geometric algorithm based on homology that requires no knowledge of any orientation parameter or the use of multiple images (sequence) or their radiometric properties. IO parameters (focal length and camera center) will not be needed (radial distortion will be removed before the process by applying the method of straight lines[25] although other well known methods, such as plumb lines,[26] direct linear transformation,[27] and others,[28] could also be applied). Likewise, relative orientation parameters (rotations of each image and vector components between the optical centers) and external orientation data (ground control points) are not necessary. Stereoscopic models can be obtained using only a stereo pair obtained from a camera with unknown parameters or through scanned images obtained with an unknown camera from a textbook or through video frames. Only homologous marking points in the two images are used through a geometric algorithm based on the determination of seven parameters (while the fundamental matrix has nine parameters and requires two restrictions) derived from the homology between the epipolar lines in the two images.

2 Epipolar Image Rectification without Orientation Parameters

2.1 *Algorithm of Epipolar Image Rectification*

The epipolar geometry between images of a stereo pair is essentially the geometry of the intersection of the image planes with the "pencil of epipolar planes" (planes containing the baseline).[29] The baseline (vector defined by the line joining the two optical centers, O_1 and O_2) intersects each image plane at the left epipole (C_1) and the right epipole (C_2), respectively [so, an epipole is the image in one view of the camera center of the other view, and it is also the vanishing point of the baseline (translation) direction].

Any point $P(X, Y, Z)$ in the object space will be impressed on both images [passing through the optical centers leading to a point $p_1\ (x_1, y_1)$ in the left image and a point $p_2\ (x_2, y_2)$ in the right image] and will define an epipolar plane that intersects the image planes in corresponding epipolar lines (e_1 and e_2), defining a correspondence between them. The correspondence established between epipolar lines (not for image planes) consists of a homology[30] or central collineation.[31] As the position of P varies, a pencil of epipolar planes, called an epipolar pencil, is generated. Each pair of epipolar lines generated by an epipolar pencil intersects at a point I on the axis of homology (line intersection defined by extending the image planes) (Fig. 1).

According to the above, the homologous epipolar line of any epipolar line on an image is constructed by joining its

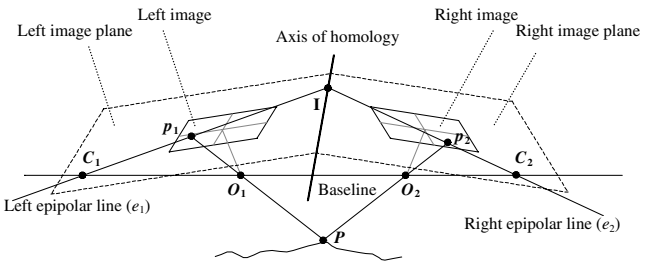

Fig. 1 Definition of homology in a photogrammetric stereo pair.

intersection point I on the axis of homology with the epipole of the other image (C_j), such that a point p_j on an epipolar line has its homolog on the homologous epipolar line.

In either of the two image coordinate systems and maintaining the same value of the t parameter for the same point, the axis of homology can be parameterized according to Eq. (1):

$$x_j^{(I)} = a_j t + b_j, \quad y_j^{(I)} = a'_j t + b'_j, \tag{1}$$

where $[x_j^{(I)}, y_j^{(I)}]$ are the coordinates of any point I on the axis of homology projected onto one of the two images, (a_j, b_j) are the coefficients of the line in parametric form for the x-coordinate, (a'_j, b'_j) are the coefficients of the line in parametric form for the y coordinate, and t is the parameter defining the position of a point on the axis of a pair of epipolar lines.

Equation (1) may be formulated in parametric form for both images, joining the point I with each epipole C_j, which can be described by the following expression:

$$\frac{x_j^{(I)} - x_j}{y_j^{(I)} - y_j} = \frac{x_j - x_{C_j}}{y_j - y_{C_j}}, \tag{2}$$

where (x_j, y_j) are the coordinates of any point on one epipolar line (in left or right image) and (x_{C_j}, y_{C_j}) are the coordinates of the corresponding epipole (in left or right image).

Substituting Eq. (1) into Eq. (2), we find that the value of t, the ratio of the differences between abscissas and ordinates, is constant and results in a line

$$t = \frac{(b'_j - y_{C_j})x_j + (x_{C_j} - b_j)y_j + (b_j y_{C_j} - b'_j x_{C_j})}{-a'_j x_j + a_j y_j + (a'_j x_{C_j} - a_j y_{C_j})}. \tag{3}$$

This equation relates the value of t directly with the coordinates of each epipole C_j, with the coefficients of the equation of the axis of homology (a_j, b_j, a'_j, b'_j) and with the coordinates of any point on the epipolar line $p_j(x_j, y_j)$ that intersects the axis of homology at the point defined by such a value of t.

By choosing a t value, we are choosing a single point on the image plane (located on the axis of homology). Therefore, choosing any point on one image plane, we are subsequently choosing a single t value (corresponding to the intersection point of the line defined by this point and the corresponding epipole with the axis of homology). Applying this fact to the stereo pair, taking any point $p_1(x_1, y_1)$ from the left image and $p_2(x_2, y_2)$ from the right image and taking into account that both equations have the same value of t, we obtain an equation that contains 12 unknowns and relates the coordinates of a point from the left image with a right image line and vice versa:

$$\frac{A_1 x_1 + B_1 y_1 + C_1}{A'_1 x_1 + B'_1 y_1 + C'_1} = t = \frac{A_2 x_2 + B_2 y_2 + C_2}{A'_2 x_2 + B'_2 y_2 + C'_2}, \tag{4}$$

where, $A_j = (b'_j - y_{Cj})$, $B_j = (x_{Cj} - b_j)$, $C_j = (b_j y_{Cj} - b'_j x_{cj})$, $A'_j = -a'_j$, $B'_j = a_j$, and $C'_j = (a'_j x_{Cj} - a_j y_{Cj})$.

Specifying $j = 1$ for the left image and $j = 2$ for the right image, we find that $A_1, B_1, C_1, A'_1, B'_1, C'_1$ are the unknowns that result from the development of Eq. (3) for the left image plane, and $A_2, B_2, C_2, A'_2, B'_2, C'_2$ are the unknowns resulting from the development of Eq. (3) for the right image.

Equation (4) is expressed for a value of t that contains 12 unknowns but only 7 are independent, so it must be simplified. The left term of Eq. (4), which corresponds to the left image, can be greatly simplified by subtracting B_1/B'_1 to eliminate the term B_1 (coefficient of y_1), inverting the resulting equation and subtracting $A'_1/[A_1 - (B_1/B'_1)A'_1]$ to eliminate the term A'_1 and then dividing the numerator by B'_1 and the denominator by $[C_1 - (B_1/B'_1)C'_1]$. In addition, applying these reductions to the right term of the equation (which corresponds to the right image) and reducing to unity the independent term of its denominator, we finally obtain:

$$\frac{y_1 + C_{R1}}{A'_{R1} x_1 + 1} = t' = \frac{A_{R2} x_2 + B_{R2} y_2 + C_{R2}}{A'_{R2} x_2 + B'_{R2} y_2 + 1}, \tag{5}$$

where $A'_{R1} = (A'_1 B_1 - A_1 B'_1)/(C'_1 B_1 - C_1 B'_1)$, $C_{R1} = (A'_1 C_1 - A_1 C'_1)/(A'_1 B_1 - A_1 B'_1)$, $A_{R2} = C_{R1} A'_{R2} - (C_1 A'_2 - C'_1 A_2)/(B_1 C'_2 - B'_1 C_2)$, $B_{R2} = C_{R1} B'_{R2} - (C_1 B'_2 - C'_1 B_2)/(B_1 C'_2 - B'_1 C_2)$, $C_{R2} = C_{R1}[1 - (C_1 C'_2 - C'_1 C_2)/(B_1 C'_2 - B'_1 C_2)]$, $A'_{R2} = (B_1 A'_2 - B'_1 A_2)/(B_1 C'_2 - B'_1 C_2)$, $B'_{R2} = (B_1 B'_2 - B'_1 B_2)/(B_1 C'_2 - B'_1 C_2)$, $A_{R1} = B'_{R1} = 0$, and $B_{R1} = C'_{R1} = C'_{R2} = 1$.

Equation (5) establishes the same correspondence as the fundamental matrix[29] does through $x'^T F x = 0$. According to the described algorithm, the reparameterization of the fundamental matrix particularized only for seven coefficients results in Eq. (6). As can be seen, it is not necessary to solve it as the third row results from linear combination of the first and the second one:

$$\begin{pmatrix} x_2 & y_2 & 1 \end{pmatrix} \cdot \begin{pmatrix} A'_{R1}A_{R2} & -A'_{R2} & -A'_{R2}C_{R1} + A_{R2} \\ A'_{R1}B_{R2} & -B'_{R2} & B_{R2} - B'_{R2}C_{R1} \\ A'_{R1}C_{R2} & -1 & C_{R2} - C_{R1} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = 0. \tag{6}$$

The system has a unique solution for the seven coefficients through seven pairs of homologous points. However, if we have a larger number of measured points, then the calculation is performed by least squares adjustment (establishing a system of linear equations that facilitates analyzing the accuracy and locating unsuitable points by reliability tests).

Epipolar image rectification implies that epipolar pairs of lines obtain the same y coordinate. To accomplish this goal, we will apply a projective transformation to the original

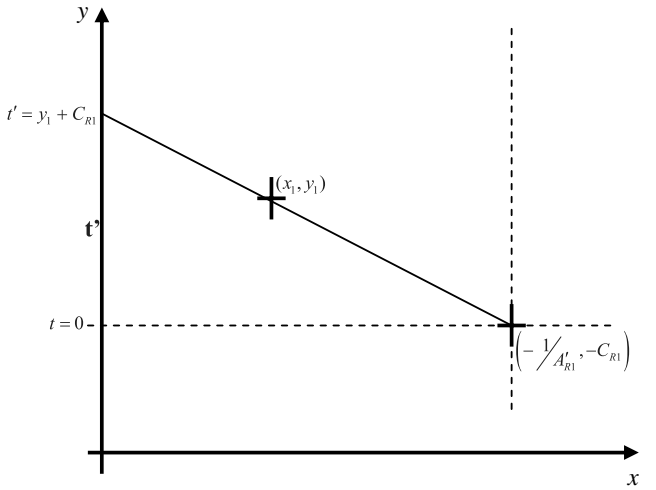

Fig. 2 Geometric meaning of the coefficients of the left image.

images so that each point in the rectified image has the value of t' as ordinate, maintaining the abscissa value to be compatible with its projective transformation. Thus, epipolar lines in both images have necessarily the same height. The calculation of the rectified images by the developed transformation is performed by calculating each pixel of the image, first generating the environment in which the image is generated by calculating the coordinates of its corners (to establish the framework beyond which there should be no image). The corner coordinates in the left rectified image [Eq. (7)] are obtained by replacing the coordinates of the original corners of the left image in Eq. (5), resulting in (x^i, y^i). Similarly, the right corner coordinates [Eq. (8)] are obtained by substituting the coordinates of the original corners of the right image in Eq. (5), resulting in (x^d, y^d) using the corresponding term in each case.

$$x^i = \frac{x_1}{A'_{R1}x_1 + 1} \quad y^i = \frac{y_1 + C_{R1}}{A'_{R1}x_1 + 1}, \tag{7}$$

$$x^d = \frac{x_2}{A'_{R2}x_2 + B'_{R2}y_2 + 1} \quad y^d = \frac{A_{R2}x_2 + B_{R2}y_2 + C_{R2}}{A'_{R2}x_2 + B'_{R2}y_2 + 1}. \tag{8}$$

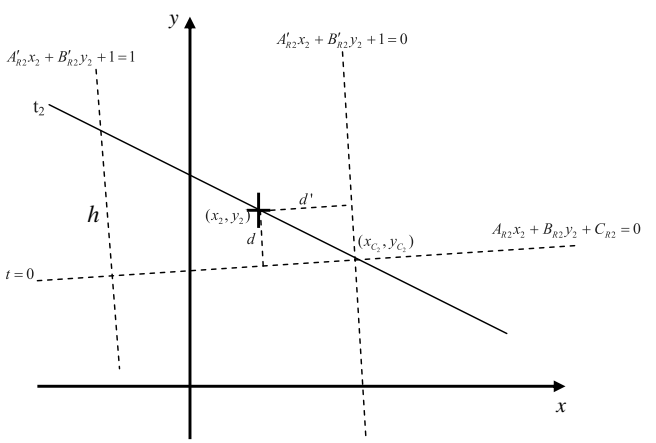

Fig. 3 Geometric meaning of the coefficients of the right image.

For both cases, the abscissa must keep the denominator of the ordinate to maintain the projectivity.

Proposed homological Eqs. (7) and (8) applied to original images results in Eqs. (9) and (10) for the left and the right ones:

$$\begin{pmatrix} x^i \\ y^i \\ 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & C_{R1} \\ A'_{R1} & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}, \tag{9}$$

$$\begin{pmatrix} x^d \\ y^d \\ 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 0 & 0 \\ A_{R2} & B_{R2} & C_{R2} \\ A'_{R2} & B'_{R2} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}. \tag{10}$$

So, the condition for the epipolar rectification $\boldsymbol{F} = \boldsymbol{H}_L^T \boldsymbol{F}_R \boldsymbol{H}_R$ is fulfilled.[25]

2.2 *Calculation of the Epipoles and the Axis of Homology*

The homologic transformation results are well defined by the epipole of each image plane and the axis of homology. The coordinates of each epipole can be obtained easily by developing Eq. (4) in its classical form:

$$y_i = x_i \frac{tA'_i - A_i}{B_i - tB'_i} + \frac{tC'_i - C_i}{B_i - tB'_i} = m(x_i - x_{C_i}) + y_{C_i}, \tag{11}$$

where

$$m = \frac{tA'_i - A_i}{B_i - tB'_i}$$

$$x_{C_i} = \frac{B'_i C_i - B_i C'_i}{B_i A'_i - B'_i A_i}$$

$$y_{C_i} = \frac{A'_i C_i - A_i C'_i}{A_i B'_i - A'_i B_i}.$$

The coordinates of the epipoles do not depend on t, which implies that all epipolar lines of each image plane pass through each one of them. If we consider that all the epipolar lines of Eq. (4) must comply with the foregoing, we have the equations of the epipole in terms of the parameters such that all epipolar lines have the same epipole and are distinguished by their different slopes, which depend on t.

With respect to the axis of homology, it cannot be determined unequivocally. In accordance with Eq. (8), the equation of an epipolar line for the right image, obtained by applying any rotation ϑ_2 results in Eq. (12):

$$\begin{aligned}&[(A'_{R2}(\lambda_2 x_2 + \mu_2 y_2) + B'_{R2}(-\mu_2 x_2 + \lambda_2 y_2) + 1]t' \\ &\quad = A_{R2}(\lambda_2 x_2 + \mu_2 y_2) + B_{R2}(-\mu_2 x_2 + \lambda_2 y_2) + C_{R2},\end{aligned} \tag{12}$$

where $x'_2 = \lambda_2 x_2 + \mu_2 y_2$ is the x coordinate and $y'_2 = \lambda_2 y_2 - \mu_2 x_2$ is the y coordinate, for any point on the epipolar line rotated ϑ_2, with $[\lambda_2^2 + \mu_2^2]^{1/2} = 1$ and $\tan \vartheta_2 = \lambda_2/\mu_2$.

By normalizing Eq. (12) for $x_2 = 0$ and reducing it to its canonical form $y = (R/t' + S) + Q$, we obtain the value of

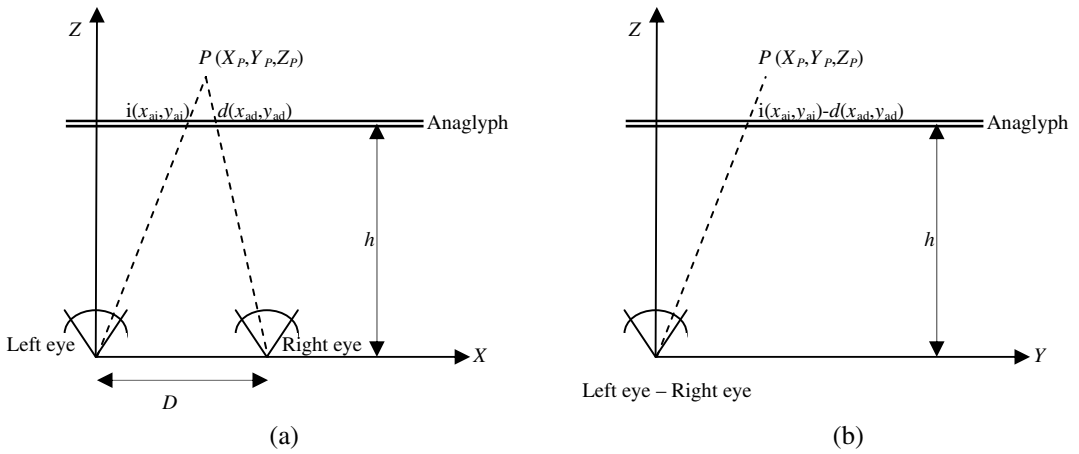

Fig. 4 Properties of stereoscopic vision. (a) XZ vision plane. (b) YZ vision plane.

S for the right image according to the following Eq. (13):

$$S^d = -\frac{\mu_2 A_{R2} + \lambda_2 B_{R2}}{\mu_2 A'_{R2} + \lambda_2 B'_{R2}}. \tag{13}$$

Using the same procedure, we obtain the value of S for the left image by applying a rotation ϑ_1 to the left epipolar line [Eq. (7)] as a function of λ_1, μ_1, according to the following Eq. (14):

$$S^i = -\frac{\lambda_1}{\mu_1 A'_{R1}}. \tag{14}$$

The lengths in both images are preserved if the value of S is the same. Thus, by equating Eqs. (13) and (14), we obtain the value of ϑ_1 and ϑ_2 according to the following Eq. (15):

$$\begin{aligned} \tan\vartheta_1 &= A'_{R1}\frac{A_{R2} + B_{R2}\tan\vartheta_2}{A'_{R2} + B'_{R2}\tan\vartheta_2} \\ \tan\vartheta_2 &= \frac{A'_{R2}\tan\vartheta_1 - A_{R2}A'_{R1}}{A'_{R1}B_{R2} - B'_{R2}\tan\vartheta_1}. \end{aligned} \tag{15}$$

This result shows that it is not possible to determine the real axis because it is always mathematically possible to calculate an axis for the left image for which the rotation (ϑ_1) corresponds to another chosen rotation for the right image (ϑ_2) and vice versa.

2.3 *Meaning of the Coefficients*

According to Eq. (5), both the left and right image planes can be studied separately and the two sides of the equation can be equated to determine the parameter $t\prime$. For the left image plane, the numerator becomes null for $y_1 = -C_{R1}$ (corresponding to a line parallel to the x axis), whereas the denominator becomes null for $x_1 = -1/A'_{R1}$ (corresponding to a line parallel to the y axis). Thus, the point of intersection of both lines is the only point where $t\prime$ is undefined. Taking into account that each value of $t\prime$ corresponds to an epipolar line and that they all intersect at the epipole (where the infinite epipolar lines pass by), the point $(-1/A'_{R1}, -C_{R1})$ must be the left epipole (Fig. 2).

The intersection of any epipolar line passing through (x_1, y_1) with the y axis results in $t' = y_1 + C_{R1}$ because the denominator of Eq. (5) becomes unitary. Taking into account that $t\prime$ becomes null for $y_1 = -C_{R1}$, the value of $t\prime$ for the epipolar line corresponds to the segment defined by both points in the y axis.

Applying the same procedure on the right side of Eq. (5), we obtain the coefficients of the right image plane. The numerator becomes null for $A_{R2}x_2 + B_{R2}y_2 + C_{R2} = 0$, whereas the denominator becomes null for $A'_{R2}x_2 + B'_{R2}y_2 + 1 = 0$. Thus, at the intersection of both lines, we will locate the right epipole (Fig. 3).

According to the expression that determines the distance from one point to a line, it is easy to observe that Eq. (5)

Table 1 Photo-coordinates of the homologous points measured in the original images to calculate the epipolar image rectification.

	Left image		Right image	
Point	x_1 **(pixel)**	y_1 **(pixel)**	x_2 **(pixel)**	y_2 **(pixel)**
1	143	1517	182	1492
2	658	1533	702	1525
3	930	1434	980	1436
4	1493	1247	1565	1267
5	104	923	161	898
6	822	906	873	903
7	97	464	168	440
8	875	362	940	361
9	1068	470	1096	474
10	1348	466	1378	478
11	1389	1143	1400	1158
12	1046	204	1075	207

Fig. 5 Original images with epipolar lines of the measured points.

corresponds to a proportionality of distances between the point and the defined lines. Therefore, the value of t' must be the ratio of these distances. Just as in the left image plane, the magnitude t' was defined geometrically on the y axis; in the right image plane, it will be defined by a perpendicular line that cancels the numerator so that the height h is the value of t'. The position of that perpendicular line obviously corresponds to $A'_{R2}x_2 + B'_{R2}y_2 + 1 = 1$.

Finally, there are two situations where improper positions of the epipoles can be obtained. The first case will occur when the value of the x coordinate or y coordinate of any of the epipoles is null. In this case, the epipole will be exactly on the edge of the image (which is not possible in normal shooting). The second case will occur when the value of the x coordinate or y coordinate of any of the epipoles is infinite. The solution is immediately solved through a single rotation, so that the image is necessarily parallel to the baseline.

Table 2 Coefficients of transformation.

Coefficient	Value
A_{R2}	-1.903772×10^{-3}
B_{R2}	1.012251
C_{R2}	5548.018985
A'_{R2}	3.177330×10^{-6}
B'_{R2}	1.260175×10^{-6}
A'_{R1}	-1.392718×10^{-6}
C'_{R2}	5522

2.4 *Properties of Stereoscopic Vision*

It is interesting to analyze the possible transformations that can be applied to the stereoscopic view of the model constructed by the coefficients of the generated algorithm. Such a model must be analyzed in a coordinate system with the origin placed at the left eye, with the X axis passing through both eyes, the Z axis perpendicular to the direction of the stereoscopic model, and the Y axis in the direction of the direct trihedral (Fig. 4). By studying the similar triangles generated by a point on the stereoscopic model P [with projections on the anaglyph that result in $i(x_{ai}, y_a)$ and $d(x_{ad}, y_a)$] in Fig. 4(a), we obtain the coordinates (X_P, Z_P). By performing the same procedure on the point P in Fig. 4(b) (both points are superimposed because they have the same Y coordinate), we obtain the coordinate Y_P. The spatial position of point P of the stereoscopic model is given by

$$X_P = \frac{x_{ai}D}{D - (x_{ad} - x_{ai})}$$
$$Y_P = \frac{y_a D}{D - (x_{ad} - x_{ai})}$$
$$Z_P = \frac{hD}{D - (x_{ad} - x_{ai})}, \tag{16}$$

where $i(x_{ai}, y_a)$ are coordinates of P on the left image of the anaglyph, $d(x_{ad}, y_a)$ are coordinates of P on the right image of the anaglyph, $P(X_P, Y_P, Z_P)$ is spatial coordinate

Table 3 Coordinates of the epipoles.

	Left image		Right image	
Point	x_{C1} (pixel)	y_{C1} (pixel)	x_{C2} (pixel)	y_{C2} (pixel)
Epipole	718020	−5522	6140772	6068

Table 4 Photo-coordinates of the homologous points in the rectified images. Residual parallax is generated in the stereoscopic model.

	Epipolar rectification (left image)			Epipolar rectification (right image)			Stereoscopic model
Point	x_1 **(pixel)**	y_1 **(pixel)**	t'	x_2 **(pixel)**	y_2 **(pixel)**	t'	**Residual parallax**
1	143.5	1510.9	7041.0	185.8	1511.1	7041.2	−0.2
2	658.9	1532.2	7062.4	707.0	1531.7	7061.9	+0.5
3	931.2	1435.8	6965.8	981.8	1436.1	6966.1	−0.3
4	1496.6	1253.3	6783.1	1559.1	1253.4	6783.2	−0.1
5	104.6	916.9	6446.3	147.3	917.1	6446.5	−0.1
6	823.4	906.3	6435.7	859.4	906.1	6435.5	+0.2
7	97.7	458.0	5986.9	141.2	457.8	5986.6	+0.3
8	876.5	363.0	5891.8	911.2	363.2	5892.0	−0.2
9	1070.0	472.3	6001.2	1070.0	472.7	6001.6	−0.4
10	1351.4	471.2	6000.1	1351.4	470.7	5999.6	+0.5
11	1392.0	1148.7	6678.4	1392.0	1148.8	6678.5	−0.1
12	1047.9	205.8	5734.4	1041.6	205.9	5734.5	−0.1

Table 5 Photo-coordinates of the corners of rectified images.

	Left image		Right image	
Corner	x_1 **(pixel)**	y_1 **(pixel)**	x_2 **(pixel)**	y_2 **(pixel)**
Upper left	0.0	2352.8	28.1	2384.2
Upper right	1656.6	2371.2	1675.9	2340.4
Lower left	0.0	−6.1	−40.8	19.9
Lower right	1656.6	6.1	1611.2	−11.8

of P on the stereoscopic model, and D is the interpupillary distance.

The spatial coordinates of any point are affected by the same homothetic factor with respect to the left image of the anaglyph as a function of the interpupillary distance, the distance from which the anaglyph is observed and the overlapping of the images that generate it.

The proposed Eqs. (7) and (8) provide the coordinates of the images that will be used for the construction of the anaglyph (x_i, y_i) and (x_d, y_d). If we change the anaglyph using a linear transformation for one or both images, we obtain a new anaglyph with a different stereoscopic "appearance." Let us modify the X coordinate and/or the Y coordinate of the points on the right image of the anaglyph.

The original Eq. (4) is reduced to Eq. (5) by reducing from 12 coefficients to 7. A linear transformation for Y in both images of type $Y = (at + b)/(ct + a)$ makes possible to obtain reduced coefficients to get the original Eq. (4). Thus, if $b = c = 0$, then $Y = t$ as in Eqs. (7) and (8), but if only $c = 0$ then a homothetic and translation factor is obtained for the Y coordinate. Finally if c is not null, we obtain a perspective effect for the Y coordinate. The transformation for the X coordinate in both images will be compatible in any case, maintaining the denominator as $X = (mx + ny + s)/(ct + 1)$, respectively.

In this case, if $c = 0$, then there is no perspective effect on the image. It is worth noting that the modification of the X coordinate of the points of either the right or the left image would require the same process, which is indicated by $X = (1 + m)x_1 + (1 + n)y_1 + s$. The linear transformation on X causes three effects by varying the value of m, n, or s. A variation of s results in the translation of the X coordinate of the right frame, which directly increases the value of $(x_{ad} - x_{ai})$ by that amount and introduces a homothetic factor $f_{sx} = (1 + s)/[D - (x_{ad} - x_{ai})]$ for each of the coordinates of P in Eq. (16). This homothetic factor scales and enlarges the stereoscopic model. A variation of m results in the translation of the X coordinate, which linearly increases the value of $(x_{ad} - x_{ai})$ and introduces a variable homothetic factor $f_{mx} = (1 + mx)/[D - (x_{ad} - x_{ai})]$ for each of the coordinates of P in Eq. (16). The homothetic factor scales the stereoscopic model in the direction of the Z axis linearly with X. This fact implies a rotation in the X coordinate by a value of $\alpha = \arctan\{(m)/[D - (x_{ad} - x_{ai})]\}$. A variation of n results in the translation of the X coordinate, which also linearly increases the value of $(x_{ad} - x_{ai})$, thereby introducing a variable homothetic factor $f_{nx} = (1 + ny)/[D - (x_{ad} - x_{ai})]$ to each of the coordinates of P in Eq. (16). This change maintains the Y coordinate of the stereoscopic model linearly in the direction of the Z axis.

Fig. 6 Left and right epipolar image rectification.

In this case, there is a rotation by a value of $\beta = \arctan\{(n)/[D - (x_{ad} - x_{ai})]\}$. Thus, by choosing the appropriate values of m, n, and s, we can zoom in or out on the displayed stereoscopic model and rotate it in both longitudinal and transverse directions.

However, it is not possible to apply a transformation of this type to the Y coordinate of the right frame [Eq. (8)] because if the X coordinate is modified for various values of m such that $m \neq 0$, then we get different values for Y that depend on the effect of cancellation of the parallax and hence stereoscopic vision. The linear transformation that is applicable is therefore of the type $Y = (1 + n)y + s$, if applied to both images simultaneously. A variation of s will translate the observed stereoscopic model in the direction of the Y axis, and a variation of n will scale the model by introducing a homothetic factor with a value that is given by $f_{ny} = (1 + n)$.

Finally, it is possible to apply a projective transformation making c not null, obtaining a perspective effect in the Y coordinate. By applying transformations on the X coordinate, we get homothetic and translation factors and rotation effects. By applying transformations on the Y coordinate, we get homothetic and translation factors and a perspective effect.

Fig. 7 : Stereoscopic model.

3 Results

To verify the validity of the algorithm, several images have been taken with a semimetric camera Rolleiflex 6008 Integral 2 with Phase One db20p 16 mp digital back (its parameters are not shown as they will not be used). The original images have a size of 1653×2362 pixels. The measurement of photo-coordinates has been performed using commercial software for image processing (with the origin of the coordinates at the lower left corner). The measurement of 12 homologous points (Table 1) in two images (Fig. 5) has been made, generating a system of 12 equations with seven unknowns according to Eq. (5) and then calculating the corners of the rectified images according to Eqs. (7) and (8).

The final results for the seven coefficients have a standard deviation of 0.45 pixels (Table 2). The coordinates of the respective epipoles are shown in Table 3.

The rectified coordinates of the 12 homologous points with their corresponding values of t' show residual parallaxes better than 0.5 pixels (Table 4), which demonstrates that the generated stereoscopic model is correct and free from parallaxes.

The corners of the rectified images calculated by Eqs. (7) and (8) with the obtained coefficient values (Table 2) are shown in Table 5 and correspond to Fig. 6, which displays the rectified images.

Fig. 8 Stereoscopic models obtained by varying properties of stereoscopic vision from a single stereo pair.

For stereoscopic model generation, digital photogrammetric stations base the epipolar image rectification on a calculation system through coplanarity algorithms, using camera calibration parameters (focal length, camera center,...) and the vector components between the projection centers and the image rotations. For the proposed system [Eq. (5)], it is only necessary to enter the photo-coordinates of at least seven homologous points, providing the seven transformation coefficients. So that the epipolar image rectification is calculated using the corners of both images through Eqs. (7) and (8) and finally obtaining the stereoscopic model (Fig. 7) by placing each rectified image (filtered with complementary colors) according to calculated coordinates (Table 5).

As demonstrated in Sec. 2.4, it is possible to stretch or compress the image horizontally by applying linear transformations to one or both images through homothetic factors maintaining the correct formation of the stereoscopic model or changing the stereoscopic perspective by rotating the model because the epipolar lines are not altered. However, any modification of the images in the vertical direction will lead to the loss of epipolarization properties by moving epipoles, which causes loss of stereoscopy.

Accordingly, it is possible to select points so that the values of m, n, and s cause an estimated effect on the perspective. This procedure facilitates the possibility of rectifying the images that give rise to the anaglyph on a defined plane of interest. By varying the properties of the stereoscopic vision of the stereoscopic model, we can generate different stereoscopic models based on the same photogrammetric stereo pair (Fig. 8).

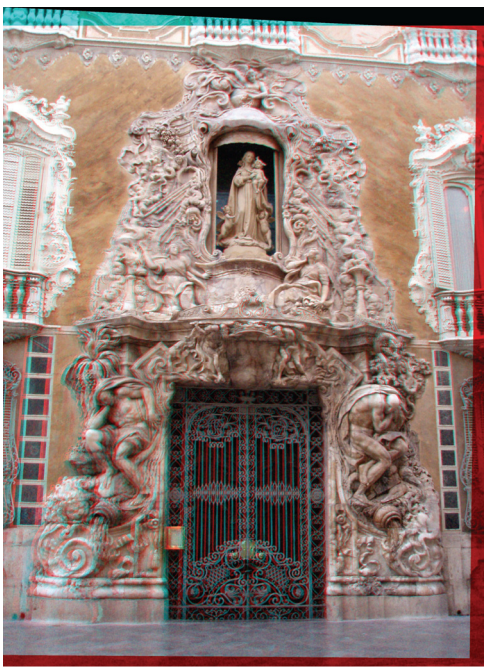

Fig. 9 Close-range photogrammetry.

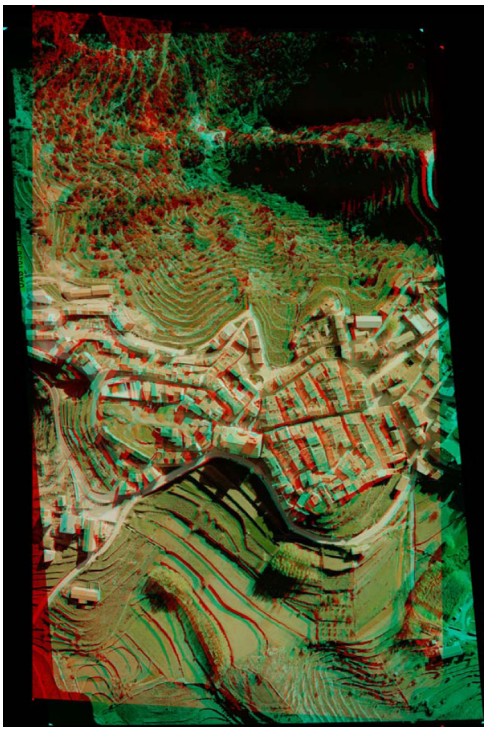

Fig. 10 Aerial photogrammetry.

As can be seen in Fig. 8, the first model obtained provides opposing parallaxes at the top and bottom of the palm tree, which implies a shift forward in the top of the model to the X axis ($m = 0.08033$, $n = -1.03919$, and $s = -67.42$). In the second model, the monument and palm tree remain parallel vertical to the plane of the anaglyph though at different depths ($m = 0.00569$, $n = -0.96911$, and $s = -83.73$). The third model reduces the relative depths of the palm tree and the monument, causing a substantial shift parallel to the Y axis ($m = -0.00483$, $n = -0.97040$, and $s = -41.37$). The parameters m, n, and s indicated in each case result from eliminating the parallax for the three points that define the front plane of the observer (points 3, 7, and 11 for the first model, points 1, 3, and 7 for the second, and points 9, 10, and 11 for the third).

Other examples applied to close-range photogrammetry (Fig. 9) or aerial photogrammetry (Fig. 10) can be seen.

4 Conclusions

Coplanarity algorithms, on which digital photogrammetric stations base stereoscopic model generation (based on the use of calibrated cameras with known IO parameters), involve the calculation of epipolar planes that intersect both images in epipolar lines. Therefore, it is obvious that the use of epipolar lines as a mathematical model should serve the same purpose.

The algorithm we developed for the epipolar image rectification of a photogrammetric pair can be considered as a reparameterization of the fundamental matrix. It seems to be optimal because it provides the same results as the application of collinearity and coplanarity without using any calibration or orientation parameters and without premarking ground control points. This process enables the generation of stereoscopic models even from scanned images or video frames with only seven homologous points (although it is advisable to measure a greater number of points for reliability tests). This article shows this process manually in order to verify the algorithm validity. However, algorithms work with measurements obtained from matching processes as well.

The epipolar image rectification is quickly and easily calculated through just seven independent coefficients. The geometric algorithm is able to calculate the transformation even in cases where the axis or epipoles turn out to be improper (the denominators of the algorithm tend to zero or where the y coordinate of epipoles is infinite). However, this problem can be solved by applying rotations to the images in the iterative process (getting the epipoles at infinity in x axis).

At the same time, it is undesirable to select a point that is near the line determined by any of the two denominators. In any case, through reliability tests, the least square process is capable of locating optimal equations as a function of the distribution of the homologous points, meaning that these situations are avoidable.

Finally, the algorithm displays great versatility in the interpretation and transformation of its coefficients. For instance, stereoscopic images can even be generated with just five coefficients that are scaled by different homothetic factors. Thus, several perspectives of anaglyphs can be generated by applying transformations to the coefficients such that the images can be rectified with respect to a selected plane within the scene.

References

1. J. X. Chai and H. Y. Shum, "Parallel projections for stereo reconstruction," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 493–500, Hilton Head Island, South Carolina (2000).
2. S. Maybank, *Theory of Reconstruction from Image Motion*, p. 261, Springer-Verlag, New York (1992).
3. A. Shashua, "Projective structure from two uncalibrated images: structure from motion and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(8), 778–790 (1994).
4. R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar image analysis: an approach to determine structure form motion," *Int. J. Comput. Vis.* **1**(1), 7–55 (1987).
5. H. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature* **293**(5828), 133–135 (1981).
6. R. Hartley, "Theory and practice of projective rectification," *Int. J. Comput. Vis.* **35**(2), 115–127 (1999).
7. J. Mallon and P. F. Whelan, "Projective rectification from the fundamental matrix," *Image Vis. Comput.* **23**(7), 643–650 (2005).
8. Z. Zhang, "Determining the epipolar geometry and its uncertainty: a review," *Int. J. Comput. Vis.* **27**(2), 161–195 (1998).
9. Z. Zhang and C. Loop, "Estimating the fundamental matrix by transforming image points in projective space," *Comput. Vis. Image Understanding* **82**(2), 174–180 (2001).
10. S. M. Seitz, "The space of all stereo images," in *Proc. 8th IEEE Int. Conf. Computer Vision*, Vol. 1, pp. 26–33, Vancouver, Canada (2001).
11. T. Kim, "A study on the epipolarity of linear pushbroom images," *Photogramm. Eng. Remote Sens.* **66**(8), 961–966 (2000).
12. M. Morgan et al., "Epipolar resampling of space-borne linear array scanner scenes using parallel projection," *Photogramm. Eng. Remote Sens.* **72**(11), 1255–1263 (2006).
13. J. P. Mellor, S. Teller, and T. Lozano-Perez, "Dense depth maps from epipolar images," Technical Report, AIM-1593, MIT (1996).
14. P. Beardsley, O. Torr, and A. Zisserman, "3D model acquisition from extended image sequences," in *Proc. European Conf. Computer Vision*, Vol. 2, pp. 683–695, Springer-Verlag, Cambridge, UK (1996).
15. Z. Zhang et al., "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry," *Artif. Intell. J.* **78**(1–2), 87–119 (1995).
16. O. Faugueras, *Three-dimensional Computer Vision: A Geometric Viewpoint*, p. 695, MIT Press, Cambridge, Massachusetts (1993).
17. M. Matousek and V. Hlavác, "Correspondences from epipolar plane images, experimental evaluation," in *Proc. Computer Vision*, Vol. 1, pp. 11–18, Winter Workshop, Bad Aussee, Austria (2002).
18. A. Criminisi et al., "Extracting layers and analyzing their specular properties using epipolar plane image analysis," Technical Report, MSR-TR-2002-19, Microsoft Research (2002).
19. L. Robert and O. Faugueras, "Relative 3D positioning and 3D convex hull computation from a weakly calibrated stereo pair," in *Proc. 4th Int. Conf. on Computer Vision*, pp. 540–544, IEEE Computer Society Press, Berlin, Germany (1993).
20. I. A. Lourakis and R. Deriche, "Camera self-calibration using the singular value decomposition of the fundamental matrix: from point correspondences to 3D measurements," Technical Report, No. 3748, INRIA, Sophia Antipolis, France (1999).
21. R. Hartley, R. Gupta, and T. Chang, "Stereo from uncalibrated cameras," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 761–764, Urbana Champaign, Illinois (1992).
22. R. Hartley and R. Gupta, "Computing matched-epipolar projections," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 549–555, New York (1993).
23. A. Fusiello and L. Irsara, "Quasi-Euclidean uncalibrated epipolar rectification," in *19th Int. Conf. Pattern Recognition*, pp. 1490–1493, Tampa, Florida (2008).
24. Y. Wexler, A. Fitzgibbon, and A. Zisserman, "Learning epipolar geometry from image sequences," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Madison, Wisconsin, Vol. 2, pp. 209–216 (2003).
25. R. Swaminathan and S. K. Nayar, "Nonmetric calibration of wide-angle lenses and polycameras," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(10), 1172–1178 (2000).
26. D. C. Brown, "Close range camera calibration," *Photogramm. Eng.* **37**(8), 855–866 (1971).
27. Y. I. Abdel-Aziz and H.M. Karara, "Direct linear transformation into object space coordinates in close-range photogrammetry," in *Proc. Symp. Close-Range Photogrammetry*, pp. 1–18 (1971).
28. F. Remondino and C. Fraser, "Digital camera calibration methods: considerations and comparisons," *Proc. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **36**(5), 266–272 (2006).

29. R. J. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, p. 672, Cambridge University Press, Cambridge, UK (2003).
30. R. Hartshorne, *Foundations of Projective Geometry*, p. 167, Benjamin, New York (1967).
31. H. S. M. Coxeter, *Projective Geometry*, p. 162, Blaisdell, New York (1964).

José Herráez received his MS and PhD in civil engineering from Polytechnic University of Valencia, Spain, where he is in the Cartographic, Geodesic, and Photogrammetric Engineering Area. He is also with the Cultural Heritage Institute of Valencia, and his research interests are centered on close-range photogrammetry and 3-D laser scanning.

José Luis Denia received his MS in civil engineering from Polytechnic University of Valencia, Spain, and PhD from the University of Santiago de Compostela, Spain. He is with Polytechnic University of Valencia, Spain, in the Cartographic, Geodesic and Photogrammetric Engineering Area. He is also with the Cultural Heritage Institute of Valencia, and his research interests are centered on close-range photogrammetry and 3-D laser scanning.

Pablo Navarro received his MS and PhD in architecture from Polytechnic University of Valencia, Spain, where he is in the Graphic Expression in Architecture Area. He is also with the Cultural Heritage Institute of Valencia, and his research interests are centered on close-range photogrammetry and 3-D laser scanning.

Jaime Rodríguez received his MS in geodesy and cartography from Polytechnic University of Valencia, Spain, and PhD from the University of Vigo, Spain. He is with University of Santiago de Compostela, Spain, in the Cartographic, Geodesic and Photogrammetric Engineering Area. His research interests are centered on close-range photogrammetry and 3-D laser scanning.

M. Teresa Martín received her MS in geodesy and cartography from Polytechnic University of Valencia, Spain, and PhD from the University of Vigo, Spain. She is with University of Santiago de Compostela, Spain, in the Cartographic, Geodesic and Photogrammetric Engineering Area. Her research interests are centered on cartographic generalization and close range photogrammetry.