Construction of Panoramic Images from Aerial Images Obtained by Aircrafts

Francisco Assis da Silva [1,3]
Almir Olivette Artero [2]
Alan Kazuo Hiraga [1]
Maria Stela Veludo de Paiva [3]

[1] University of Western São Paulo - Unoeste
Presidente Prudente - SP, Brazil
chico@unoeste.br, alanhiraga@unoeste.edu.br

[2] São Paulo State University - Unesp
Presidente Prudente - SP, Brazil
almir@fct.unesp.br

[3] São Paulo University - USP
São Carlos - SP, Brazil
mstela@sc.usp.br

Abstract. Panoramic images from aerial images captured by aircrafts are an important tool, because they increase the efficiency in obtaining results and help the decision making. The automatic panoramic images are useful to different applications, such as: water resource, environment monitoring, aerial mapping and precision farming. Panoramic images are constructed to obtain an image with a horizontal resolution much higher than an individual image, making possible to photograph a large area. The difficulty for the implementation of panoramic images lies in the way that the images need to be connected to each other, because during the capture process, the objects of the scene appear in the images in different positions, brightness and color, needing a previous rectification process, both spatial and radiometric. However, this adjustment creates distortions, which are accumulated along the process, and they needs to be repaired later, in an additional step. In this paper it is presented a new algorithm for the construction of panoramic images, which avoid the distortion that occurs by joining of several successive images, excluding the additional step to correct it. The SIFT and RANSAC algorithms are used to find overlap areas between pairs of images, as well as a Blend algorithm for smoothing the joints. As the proposed algorithm doesn't cause distortions, the subsequent correction is not necessary, contributing to a considerable performance increase. The results of experiments using commercial software and also the proposed algorithm were compared through a visual and quantitative analysis, using numeric measures calculated on a panoramic image, generated from a region mapped and georeferenced by Google Earth.

Keywords: Panoramic Images, Image Rectification, Image Adjustment, SIFT, RANSAC.

1 Introduction

Panoramic images are constructed from an image sequence in order to obtain an image with a horizontal/vertical resolution much higher than an individual image, which makes it possible to photograph a large area. They have been widely used because the visualization of a single panoramic image is much more pleasant than just viewing the individual images separately. The panoramic images can be obtained using special lenses, like fish eye (Hill, 2007), that coupled to a common photographic camera it makes possible to capture images of the environment with a larger scope. Although the technique has gained greater popularity with the advance of digital photography technology, the panoramic images have been obtained since 1839 (Newhall, 1964), when several photographs were captured using photography film, which, after they have been developed on paper, they were cut and pasted in a sequence they were taken. With the popularization of digital cameras, equipped with more and more powerful processors, the construction of panoramic images started to use a very simple strategy, that consists simply in to press and hold the camera button, and to move the camera to take the images continually. Nowadays, the construction of panoramic images uses only software and, sometimes, these softwares are embedded into the camera. When the first image of the sequence is connected with the last one, a 360 panoramic image is formed (Nayar; Karmarkar, 2000) (Bakstein; Pajdla, 2001), whose main feature is to offer a good immersion into the environment pictured, because using visualization software is possible to have the sensation of moving inside the environment. Companies like Google, with its Street View program and Microsoft, with its Streetside program, picturing streets and highways through these resources to allow users a similar sensation as if they were walking on the streets of the mapped place.

The greatest difficulty for the construction of panoramic images is due to spatial changes (translation, scaling and rotation) that occur when images are captured and also radiometric (brightness and color). During the process of joining images, they need to pass through many adjustments, so they can fit properly. However, this adjustment step creates a distortion, that is accumulated along the process (Figure 2), requiring a repair later (Brown; Lowe, 2007). Thus, in this paper is proposed a new algorithm to construct panoramic images, which avoids the distortions from the adjustments, avoiding the additional step to correct it. The remaining sections of this paper are organized in the following way: In Section 2 we presented the problem inherent to the joining images and the distortions caused by using known techniques. This section presents the SIFT (Lowe, 1999), RANSAC (Fischler; Bolles, 1981) and Blend (Uyttendaele et al., 2001) (Levin et al., 2006) algorithms used in the approach proposed in this work. In Section 3 we present a new proposal to construct panoramic images, reducing the process steps. Section 4 compares the results of experiments using two commercials software and the proposed algorithm. The results are evaluated by a quantitative analysis, using aerial images, whose coordinates are known. Finally, in Section 5 are present conclusions and suggestions to future work.

2 Related Work

Joining images is not a simple task, because during the capture of the images, spatial problems appear, due to the camera movement (translation, rotation and scale), besides the radiometric differences (color and brightness). The spatial problems are more easily solved, if the images pass through a rectification process (Wolf et al., 2000), which is described in Section 2.1. In addition, it is necessary to find the homologous points between the images to be joined. This task is done by identifying common areas in both images. In this work, it is done using SIFT (Lowe, 1999) (Lowe, 2004) and RANSAC (Fischler; Bolles, 1981) algorithms described in Section 2.2. Finally, to solve the radiometric problems we use the Blend Feathering (Levin et al., 2006) (Uyttendaele et al., 2001), as presented in Section 2.3.

2.1 Image Rectification

Image rectification (Roy et al, 1997) (Pollefeys et al., 1999) (Oram, 2001) is a widely used stage in the construction of panoramic images, because it simplifies the task of joining images. In Pollefeys et al. (1999) two ways for image rectification are discussed (planar and cylindrical rectification) to determine how the pair of images to be rectified must pass for a re-projection. The epipolar geometry (Hartley; Zisserman, 2003) is used to simplify processing for the location of homologous points, because with it, the search need not be done throughout the image, restricting only to the epipolar line.

2.2 Cylindrical Rectification

Using the epipolar geometry concept, the rectification must adjust the planar images at baseline prior to stitching them. In this way, it is ensured that the related epipolar lines will be at the same height, observing that if the epipolo of the image is too close to the image, the rectified image will be deformed. In an extreme case, if the epipolo is located inside the image, the resulted image will be stretched to infinity (Roy et al., 1997). In a cylindrical rectification, the epipolo is never located inside the same image because each pixel of the image is projected on a cylinder of radius *z* (cylindrical projection) (Roy et al., 1997), as illustrated in Figure 1.

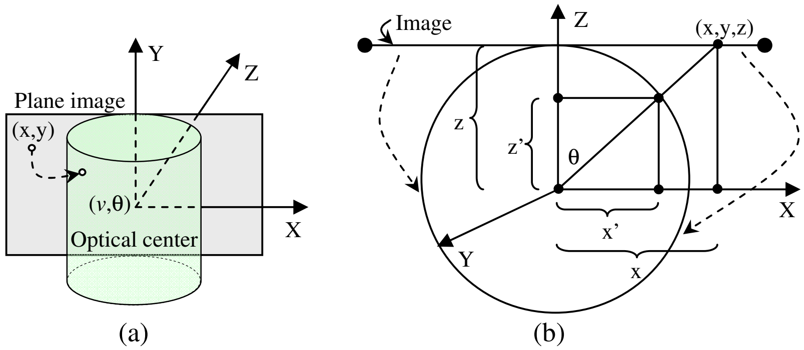


Figure 1. a) Projection of the image on cylinder; b) Top view of the cylinder.

To make the projection, it is considered the center of the image with coordinate (0,0), in the main point of the image plane, and each point of the image will have a coordinate global system given as (x,y,z). This projected position in the cylinder will have the coordinates (v,θ), where v is the height of the intersection point between the cylinder and the straight line connecting the origin to the point (x,y,z). The angle θ, between the z axis and this straight line (Szeliski; Shum, 1997) is obtained with Equation 1.

$$\theta = \tan^{-1}\left(\frac{x}{y}\right) \quad (1)$$

and v is calculated using the Equation 2.

$$v = \frac{y.z.\sin(\theta)}{x} \quad (2)$$

2.2 Identification of homologous points

In order to join two images, it is necessary to find common points between them, so that it is possible to identify the overlap region. The automatic identification of homologous points in two images is a complex task, having a lot of research for its implementation (Basri; Jacobs, 1997) (Friedman et al., 1977). The initial difficulty is to find keypoints in the first image that they can be found in second image. A possible approach (Dakun et al., 2010) (Lukashevich et al., 2011) (Fang et al., 2010) and the one used in this work, applies the SIFT algorithm to find points of interest in the two images to be joined, and then to use the RANSAC algorithm to eliminate mismatched points.

2.2.1 SIFT Algorithm

This algorithm consists in a very efficient method to identify and to describe image keypoints, which is done by performing a mapping with different views of an object or scene, resulting in a vector with 128 values, describing each image keypoint (Lowe, 1999) (Lowe, 2004), beyond other information such as positioning, gradient direction and scale.

2.2.2 RANSAC Algorithm

RANSAC (RANdom SAmple Consensus) algorithm proposed by Fischler and Bolles (1981) is a robust estimation method designed to identify the inliers[1] and outliers[2] from the set of keypoints detected by the SIFT algorithm. RANSAC is widely used for object recognition

[1] *inliers*: data points that fit a particular model within a error tolerance.
[2] *outliers*: data points that do not fit a particular model within a error tolerance.

(Okabe; Sato, 2003) (Collet et al., 2009), besides, it makes possible to find the geometrically consistent correspondences to solve the problem of joining pairs of images. RANSAC is a robust estimator, so much so that it shows fine results, even in extreme conditions, or with some kind of outlier.

2.3 Joining Images

In the construction of a panoramic image, the errors caused by the successive joining of images are propagated at the same time that the fundamental matrix is applied to each joins, causing a very strong final distortion, as illustrated in Figure 2.

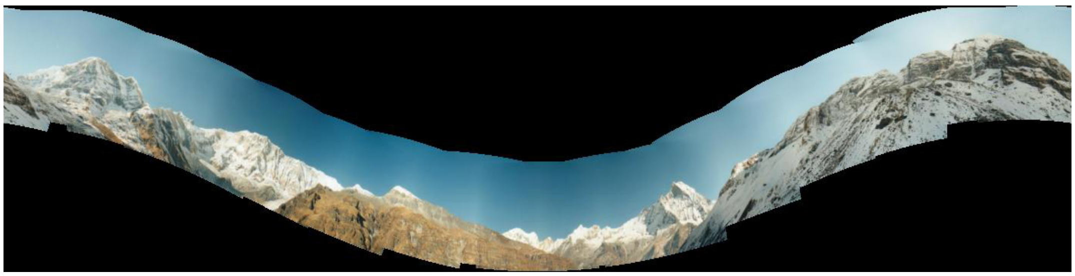


Figure 2. Panoramic image without adjustment (Brown; Lowe, 2007).

It is obvious in this figure that the error was propagated by the joins images, transforming the skyline in a sinusoidal line. Brown e Lowe (2007) suggest that this error must be corrected using an adjustment algorithm named “Bundle Adjustment”, in order to generate a panoramic image in which the horizon becomes a horizontal line again. The result of this correction is shown in Figure 3.

Figure 3. Panoramic image after “Bundle Adjustment” (Brown; Lowe, 2007).

In the case of 360 panoramic images, this adjustment is still more necessary because the first and the last image of the sequence need to be joined.

2.3.1 Blend

In the construction of panoramic images, the joined regions between two images (stitching) present a differentiation at the seams, due to factors such as small differences in the scene lighting. This makes the panoramic image present a cut aspect between the seams. The goal of a “Blending” algorithm is to camouflage the line that appears when two images with different bright are joined. There are two main approaches to image stitching in the literature, assuming that the images I_1 e I_2 have already been aligned. The first approach, optimal seam algorithms (Milgram, 1975) (Davis, 1998) (Efros and Freeman, 2001) search for a curve in the overlap region on which the differences between I_1 e I_2 are minimal. The second approach minimizes seam artifacts by smoothing the transition between the images. The Blend Feathering algorithm (Uyttendaele et al., 2001) (Levin et al., 2006), used in this work generates fine results with images that are captured in the same period of time, in spite of its simplicity. Feathering is a technique widely used in Computer Graphics to smooth the image edges and can be used to disguise the brightness and colors differences in the overlap regions between joined images. The value of each pixel in the final image I_f is determined by a color mixture of corresponding pixels in two images I_1 e I_2, given by their weighted average. The color mixture, implemented in the Blend Feathering algorithm, is given in Equation 3.

$$I_f(i,j) = (1-w).I_1(i,j) + w.I_2(i,j) \tag{3}$$

3 StitchingPHm Algorithm

In order to avoid the distortion discussed in Section 2.3 (Figure 2) and, consequently, to eliminate the subsequent correction (Figure 3), the StitchingPHm (Stitching with Partial Homographic Matrix) algorithm proposed in this paper suggests applying the homographic matrix proportionally to the proximity of the join places in the images. It has shown to be useful to prevent the errors propagation in the sequence of joined images. The proposed algorithm also showed to be quite adequate in the case of 360 panoramic images, where the first image and the last one of the sequence need to be joined too.

By the algorithm, given two images $I_1(x,y)$ and $I_2(x,y)$, to be joined, the homographic matrix (generated by RANSAC - Equation 4) is applied proportionately in I_2, only in the overlapping region with I_1. Initially it is applied 100% of the values from homographic matrix in the join points, reducing its effect gradually until it reaches 0%. In this way, the rotation and scale parameters are reduced to zero, leaving only the translation parameters in x and y.

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ h_{3,1} & h_{3,2} & 1 \end{pmatrix} \tag{4}$$

The elements of the additive matrix F, used to control the application of homographic matrix H in I_2 are obtained using the Equation 5.

$$f_{i,j} = \frac{\lambda - h_{i,j}}{jw + 0.2w} \tag{5}$$

where: $\lambda=1$ if $i=j$ and $\lambda=0$ if $i \neq j$; jw is the width of the overlap region and w is width of the image.

The pixel colors of the resulting image $I_r(x_r,y_r)$ are obtained from the image pixel colors $I_2(x,y)$, where the coordinates x_r e y_r are obtained with the Equation 6.

$$(x_r, y_r, 1) = (x, y, 1). \begin{pmatrix} f_{1,1}.x + h_{1,1} & f_{1,2}.x + h_{1,2} & h_{1,3} \\ f_{2,1}.x + h_{2,1} & f_{2,2}.x + h_{2,2} & h_{2,3} \\ f_{3,1}.x + h_{3,1} & f_{3,2}.x + h_{3,2} & 1 \end{pmatrix} \tag{6}$$

The points with coordinates not obtained with this transformation are colored by a bilinear interpolation, using the colors of the four neighboring points.

4 Experiments

This section presents an experiment in which we compared the techniques used in the process for construction of panoramic images. We compare the results using the commercials software AutoStitch (*http://www.cs.bath.ac.uk/brown/autostitch/autostitch.html*) and PTGui (*http://www.ptgui.com*) with the results obtained with the algorithm StitchingPHm, proposed for us, and implemented with C language, using the OpenCV library (Bradsky et al., 2008). An implementation of StitchingPHm can be downloaded from the site *http://fipp.unoeste.br/~chico/StitchingPHm.html*.

In this experiment, a quantitative analysis is presented in addition to a visual analysis on a set of twenty images. In the region where the experiments were done the georeferenced coordinates were obtained from Google Earth. In this comparison, we analyze the angles formed among three points *A*, *B* and *C* positioned in these regions, located in Santa Rita do Sapucaí (Brazil), as illustrated in Figure 4.

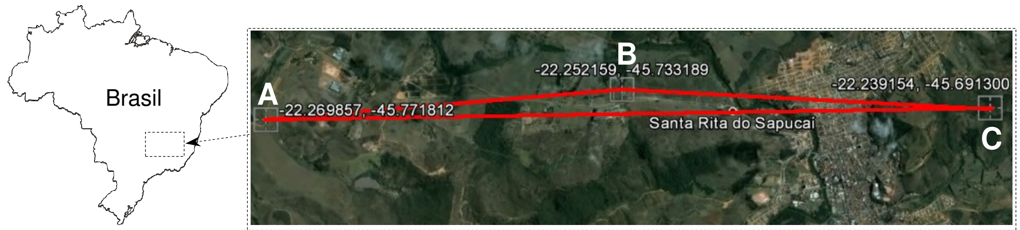


Figure 4. Region used in the experiment.

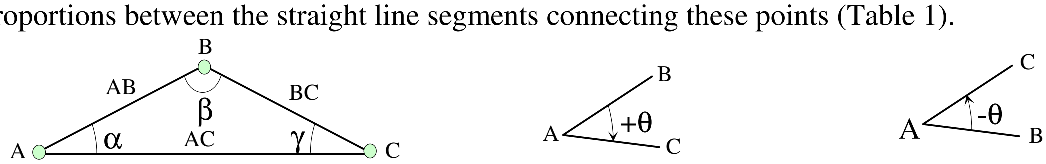
To calculate the angles among these points (Figure 5 (a)), it should be note that, given two vectors *u*, *v*, the angle between them is given by Equation 7.

$$\theta = arcos\left(\frac{u.v}{\|u\|.\|v\|}\right) \quad (7)$$

In order to obtain an accurate comparison, it must be observed the direction, clockwise (+) and counterclockwise (-), in determining the angles ∠BAC, as illustrated in Figure 5 (b). The distance *d* between *P*(*x*,*y*) and *Q*(*x*,*y*) points is calculated by Equation 8.

$$d = \sqrt{(P_x - Q_x)^2 + (P_y - Q_y)^2} \quad (8)$$

Considering that the distances among these points on Google Earth are measured in meters, and the distances among these points on the images (StitchingPHm, AutoStitch and PTGui) are measured in pixels, the way adopted to evaluate the results is to compare the proportions between the straight line segments connecting these points (Table 1).



(a) (b) (c)

Figure 5. a) Determination of the straight line segments AB, BC and AC and the angles ∠BAC (α), ∠ABC (β) and ∠ACB (γ) from tree points A, B and C; b) Determination of the angles taking into consideration the direction (clockwise and counterclockwise).

The selected Region, shown in Figure 6(a) is located among the points with the coordinates (latitude and longitude): A = -22.269857, -45.771812; B = -22.252159, -45.733189 and C = -22.239154, -45.691300.

The distances measured in these panoramic images are presented in Table 1.

Table 1. Distances measured in the region 1 on Google Earth [**G**] (georeferenced coordinates) and the pixel proportions obtained using: StitchingPHm [**S**], AutoStitch [**A**] and PTGui [**P**].

	Angles			Measures - p (pixels) - m (meters)			Proportions		
	α	β	γ	AB	AC	BC	AB/BC	AB/AC	BC/AC
G	3,941	-172,218	3,841	4435,748M	8965,388M	4550,349M	0,975	0,495	0,508
S	+2,630	-174,793	2,577	2010,647P	4058,303P	4058,303P	0,980	0,495	0,506
A	+2,436	-175,353	2,212	1754,734P	3683,925P	1932,215P	0,908	0,476	0,524
P	-3,824	+172,326	-3,850	2166,790P	4309,000P	2151,891P	1,007	0,503	0,499

(a)

(b)

(c)

(d)

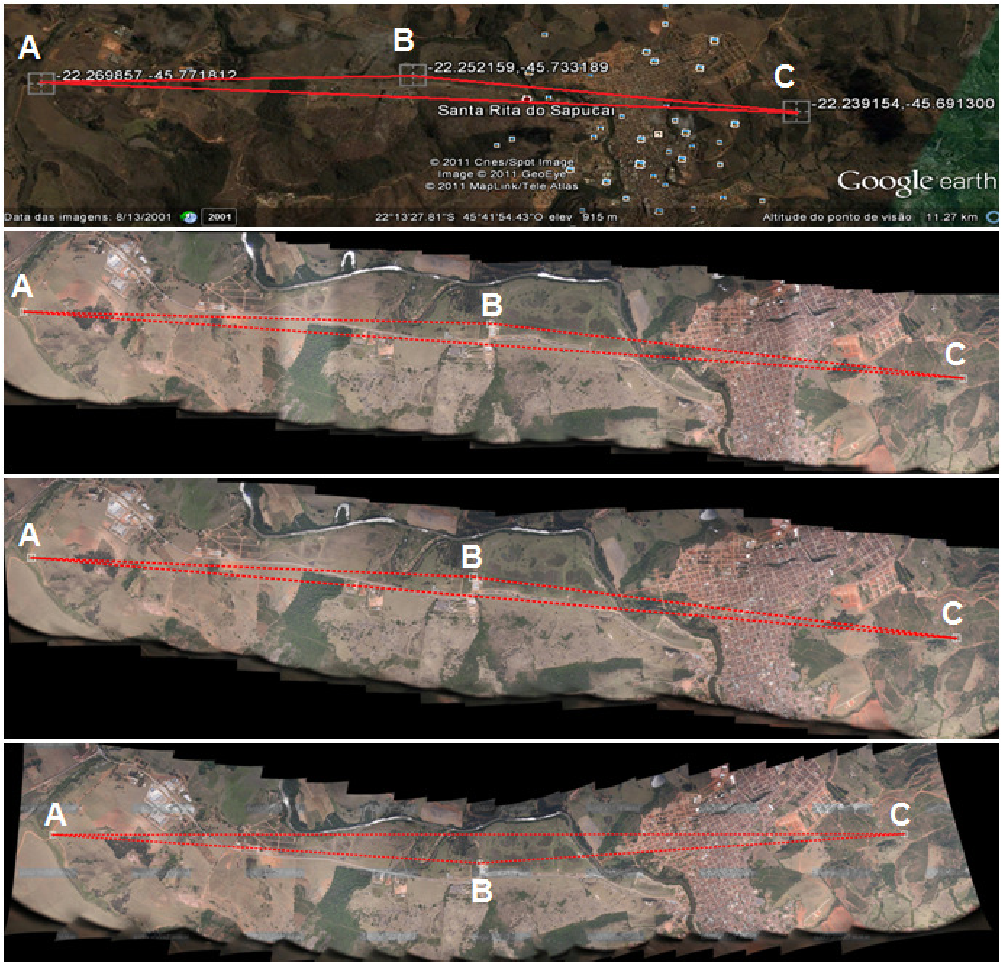


Figure 6. Selected Region; Panoramic images using: b) StitchingPHm; c) AutoStitch; d) PTGui.

The observed errors in comparison to the values obtained from *Google Earth* are presented in Table 2 and, in this case, the values obtained with the StitchingPHm are the smallest for the three angles and also for the three proportions.

Table 2. Errors between the results of Google Earth (region 1) with StitchingPHm, AutoStitch e PTGui.

	Angles			Proportions		
	α	β	γ	AB/BC	AB/AC	BC/AC
S	1,311	2,575	1,264	0,005	0,000	0,002
A	1,505	3,135	1,629	0,067	0,019	0,016
P	7,765	344,544	7.691	0,032	0,008	0,009

5 Conclusions and Future Work

The panoramic images have become very popular with the development of the electronic industry, which has incorporated this resource in the latest photographic cameras. The need to incorporate this resource in devices with modest processing power has motivated a lot of search for algorithms more and more effective for its construction. The traditional method to construct panoramic images applying the homographic matrix in the whole image produces a large deformation, requiring an additional step of adjustment. The solution proposed in this paper prevents this distortion, improving process and contributing to a considerable performance increase. The methodology used in this work was also satisfactory to the join of 360 panoramic images.

References

Bakstein, H.; Pajdla, T. 3D reconstruction from 360 x 360 mosaics, CVPR'01 conference, **Proceedings…** v. 2, p. 72–77, Loas Alamitos, CA, 2001.

Basri R.; Jacobs, D. W. Recognition using region correspondences, **International Journal of Computer Vision**, v. 25, n. 2, p. 145–166, 1997.

Bradsky, G. R.; Pisarevsky V.; Bouguet, J. **Learning OpenCV: Computer Vision with the OpenCV Library**, Springer, 2008.

Brown M.; Lowe D. G. Automatic Panoramic Image Stitching using Invariant Features, **International Journal of Computer Vision**, v. 74, n. 1, p. 59–73, 2007.

Collet, A., Berenson, D., Srinivasa, S.S., and Ferguson, D. "Object recognition and full pose registration from a single image for robotic manipulation". In: **IEEE Int Conf on Robotics and Automation**, pp. 48-55, 2009.

Dakun, Z.; Zhaoxin, L.; Guiyuan, J. A feature-based algorithm for image mosaics with moving objects, IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS), p. 26–29, 2010.

Davis, J. Mosaics of scenes with moving objects. In: **CVPR**, p. 354–360, 1998.

Efros A.; Freeman, W. Image quilting for texture synthesis and transfer. SIGGRAPH 2001 **Proceedings**, p. 341–346, 2001.

Fang, X.; Luo, B.; Zhao, H.; Tang, J.; Zhai, S. New multi-resolution image stitching with local and global alignment, **IET Computer Vision**, v. 4, n. 4, p. 231–246, 2010.

Fischler, M.A.; Bolles, R.C.; Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, **Comm. of the ACM**, v. 24, p. 381–395, 1981.

Friedman, J. H.; Bentley, J. L.; Finkel, R. A. An algorithm for finding best matches in logarithmic expected time, **ACM Transactions on Mathematical Software**, v. 3, n. 3, p. 209–226, 1977.

Hartley R.; Zisserman, **A. Multiple View Geometry in Computer Vision**, University Press, Cambridge, 2003.

Hill, R.; A lens for whole sky photographs, **Quarterly Journal of the Royal Meteorological Society**, 2007.

Levin, A.; Zomet, A.; Peleg, S.; Weiss, Y., Seamless Image Stitching in the Gradient Domain, Proceedings of 9th European Conference on Computer Vision, 2006.

Lowe, D. G.; Object recognition from local scale-invariant features, International Conference on Computer Vision, Greece, p. 1150-1157, 1999.

Lowe, D. G.; Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, v. 60, n. 2, p. 91–110, 2004.

Lukashevich, P. V.; Zalesky, B. A.; Ablameyko S. V., Medical Image Registration Based on SURF Detector, **Pattern Recognition and Image Analysis**, v. 21, n. 3, p. 519–521, 2011.

Milgram, D. Computer methods for creating photomosaics. **IEEE Trans. Computer** 23, p. 1113–1119, 1975.

Nayar S. K.; Karmarkar A., 360 x 360 mosaics. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00), Hilton Head, South Carolina, v. 2, p. 388–395, 2000.

Newhall, B. **The History of Photography, from 1839 to the Present Day**, Museum of Modern Art, NY, 1964.

Okabe T.; Sato Y., Object recognition based on photometric alignment using RANSAC, Computer Society Conference on Computeer Vision and Pattern Recognition, **Proceedings…** p. 221–228, 2003.

Oram D.; Rectification for any epipolar geometry, British Machine Vision Conference, p. 653–662, 2001.

Pollefeys M.; Koch R.; Van-Gool L. A simple and efficient rectification method for general motion, Proc. Seventh IEEE Intern. Conf. on Computer Vision, v. 1, p. 496–501, 1999.

Roy S.; Meunier J.; Cox I. Cylindrical rectification to minimize epipolar distortion, Int. Conf. on Computer Vision and Pattern Recognition, **Proceedings…** p. 393–399, 1997.

Szeliski R.; Shum H.-Y. Creating Full View Panoramic Image Mosaics and Environment Maps. Proc. 24th Int. Conf. on Computer Graphics and Interactive Techniques, ACM, p. 251–258, 1997.

Uyttendaele M.; Eden A.; Szeliski R.; Eliminating ghosting and exposure artifacts in image mosaics. Conference on Computer Vision and Pattern Recognition (CVPR 2001), **Proceedings…** 2001.

Wolf P., DeWitt B., Wilkinson B. **Elements of Photogrammetry,** Mcgraw Hill, 2000.