

3D scene stereo reconstruction with the use of epipolar restrictions

Fursov V.A., Goshin Y.V.

Samara State Aerospace University,
Image Processing Systems Institute, Russian Academy of Sciences

Abstract. In the present paper a new approach to scene digital model reconstruction from pair of stereo images is considered. We propose to perform image matching with use of weight coefficient as a penalty function for the distance from the point to the correspondent epipolar line. Technology implementation for unknown camera parameters is also considered. In this case, identification of the fundamental matrix from corresponding points is performed on the initial stage. The main advantage of this technology is the absence of the image rectification stage which causes additional distortions due to image interpolation. There is an example of scene reconstruction from pair of test images and a digital model reconstruction from satellite images.

Keywords: stereo image processing, image matching, 3D reconstruction, projective geometry, epipolar geometry

Citation: Fursov V.A., Goshin Y.V. 3D scene stereo reconstruction with the use of epipolar restrictions. Proceedings of Information Technology and Nanotechnology (ITNT-2015), CEUR Workshop Proceedings, 2015; 1490: 268-276. DOI: 10.18287/1613-0073-2015-1490-268-276

1. Introduction

The main problem with 3D-scene reconstruction is image matching. Solving this problem presents some difficulties because projective distortions on stereo pairs are usually significantly different. To overcome these difficulties, an image preprocessing rectification technique is used. Rectification of stereo images is a transformation, in which corresponding points on stereo images are arranged in the same rows.

In [1], the theoretical basis for the method of projective rectification is provided, and the basic algorithm for rectification is introduced. The algorithm involves calculation of the fundamental matrix and projective transformation. The method of projective rectification is also covered in [2]. The idea of this method is to decompose the matrix of the projective transformation into several matrices. But a common disadvantage of projective rectification is its inapplicability in case when the epipoles are located on the images as it results in infinite resolution images. In addition, images can become very large, for example, when the epipole is close to the image.

In [3], polar rectification is introduced as an alternative method of rectification. The method consists in polar scanning of images around the epipoles. Operation of the proposed approach is illustrated on real pairs of images. This method has two main advantages: an opportunity to operate with epipoles on images and guaranteed minimum size of images without losing pixels.

Though it solves many problems of projective rectification mentioned above, polar rectification has a number of disadvantages too. For instance, it does not operate correctly when an epipole is located at infinity. In particular, there may be cases where one epipole is located at infinity, and another one is located on the image. In papers [4] and [5], the idea of polar rectification is extended for these cases.

The above methods of rectification do not take into account possible differences in the internal parameters of the camera, such as focal lengths. This problem is described in [6], in which an algorithm for rectification of heterogeneous and uncalibrated image pairs is proposed, in particular, for the pairs of images obtained from static and dynamic cameras of different focal lengths and/or different resolutions. Rectification is performed in two steps. The first step is the correction of heterogeneity (different scale) by the expansion, compression or shift of images. The second step is a standard rectification of images. This approach avoids image distortions associated with differences in scale (compression or tension as a result of rectification).

At the same time, necessity for conversions directly on an image is a common disadvantage of the rectification approach. The images and objects on them are distorted considerably, owing to polar transformation. Feature points detection is performed on the new interpolated image which causes additional errors. Although polar rectification is more universal than the projective one, it still impairs some problems connected with distortions of images during conversions.

An alternative technique of matching is an approach of scene elements tracing, in particular, using a method of optical flow [7], [8]. The method of an optical flow operates efficiently with a sequence of images, for example, with sequences of video frames. However, when the images are obtained from cameras which are located relatively far apart, losses of objects during the tracing are possible.

Earlier authors proposed a technology [9] in which a projective transform is constructed from reliable corresponding points in an informative areas of images. And then this transform is used to determine and adjust the corresponding points in uninformative areas. Number of correspondence points on images given by this technology is low and resulting 3D model is too rough.

The technology of 3D scene reconstruction proposed in this paper to a considerable degree allows to avoid the above mentioned disadvantages. The basic idea is to take into account epipolar restrictions in the course of points matching.

We consider the approach to the epipolar restrictions implementation in which the most similar image fragments are selected in the neighborhood of the epipolar line. Then, according to the given criterion, the best points with real coordinates belonging to the epipolar line are selected.

2. Problem definition

In order to solve the problem of image matching, the camera obscura model is used [9]. Let us consider the case when the parameters of the cameras, as well as their position and orientation are known. To characterize them, camera parameters matrices are defined:

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{K}' = \begin{bmatrix} f' & 0 & u'_0 \\ 0 & f' & v'_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where f and f' is focal length of the camera, (u_0, v_0) , (u'_0, v'_0) are cameras' principal points location in the coordinate system associated with these cameras [10].

Let us introduce a global coordinate system and the coordinate systems of first and second cameras with their centers at points \mathbf{c} , \mathbf{c}' in the global coordinate system. Neither of these two points, in general, coincides with the origin of the global coordinate system.

Suppose \mathbf{M} is a coordinate vector of some point in the global coordinate system. Coordinate vectors of this point in the coordinate systems of the first and second cameras \mathbf{m} and \mathbf{m}' are defined as

$$\mathbf{m} = \mathbf{K}[\mathbf{R}; \mathbf{t}]\mathbf{M}, \quad (2)$$

$$\mathbf{m}' = \mathbf{K}'[\mathbf{R}'; \mathbf{t}']\mathbf{M}, \quad (3)$$

\mathbf{R} , \mathbf{R}' are the matrices of 3×3 -dimension, describing the rotation of the coordinate systems of the first and second cameras on the global coordinate system, and $\mathbf{t} = [t_x, t_y, t_z]^T$, $\mathbf{t}' = [t'_x, t'_y, t'_z]^T$ are coordinates of the origin of the global coordinate system in the coordinate systems of the first and second cameras respectively, defined as

$$\mathbf{t} = -\mathbf{R}\mathbf{c},$$

$$\mathbf{t}' = -\mathbf{R}'\mathbf{c}',$$

If projections \mathbf{m} and \mathbf{m}' (2), (3) on the first and second camera images are known, point \mathbf{M} coordinates in three-dimensional space can be calculated as the intersection of rays (\mathbf{c}, \mathbf{m}) and $(\mathbf{c}', \mathbf{m}')$. Due to errors in the course of determining of the corresponding points coordinates, these rays will not probably cross. The errors occur in the internal and external calibration matrices and image distortions due to rectification. Therefore, some inaccurate estimates of the point coordinates are usually computed.

A detailed review of methods for image matching is given in [11]. It presents a classification of methods for stereo matching by the cost function, aggregation area, and a disparity map construction approach. The paper [12] describes stereo matching by use of so-called super pixels. Superpixel is a set of pixels that are homogeneous in their brightness and texture. In [13], the method of matching in a rectangular area, based on cross-correlation and the maximum-surface method is proposed. In [14] an intensity-based method that takes into account discontinues and occlusions are presented.

In our technology, rectified images are not generated. Matching points are searched directly on the epipolar lines belonging to the same plane. Then, for each pair of corresponding points, spatial coordinates of the scene are computed.

3. Points matching using weighting coefficients

We assume that the fundamental matrix is calculated with use of camera parameters or estimated on the set of corresponding points (for example, with use of the 8-th point algorithm [15]). We will designate the points coordinates on the first image of stereo images as (u, v) , and the corresponding points coordinates on the second stereo images as $(u + \Delta u, v + \Delta v)$, where Δu and Δv are relative shifts of the coordinates respectively. Let $I(u, v)$ and $I'(u + \Delta u, v + \Delta v)$ are values of the brightness distribution functions of these images. We will use Euclidean norm as a measure of proximity between values of brightness for point (u, v) and the corresponding point $(u + \Delta u, v + \Delta v)$:

$$e(u, v, \Delta u, \Delta v) = \|I(u, v) - I'(u + \Delta u, v + \Delta v)\|$$

The problem of the determination of shifts values is formulated as a problem of the following criterion minimization:

$$E(u_0, v_0, \Delta u, \Delta v) = \sum_{(u, v) \in D(u_0, v_0)} a(u, v) e(u, v, \Delta u, \Delta v)$$

where $D(u_0, v_0)$ is the area around point (u_0, v_0) , and $a(u, v)$ is the weight function defined as multiplication of three weighting coefficients:

$$a(u, v) = w_c \cdot w_d \cdot w_f$$

where w_c , w_d are the coefficients which reduce projective distortions, and w_f is the coefficient providing the location of the point on the epipolar line. These coefficients depend on point coordinates (u, v) in area $D(u_0, v_0)$.

Coefficients w_d , w_c are determined similar to algorithm SimpleFlow [16] by the following equations:

$$w_d = \exp\left\{-\|(u_0, v_0) - (u, v)\|^2\right\}, \quad (u, v) \in D(u_0, v_0),$$

$$w_c = \exp\left\{-\|I(u_0, v_0) - I'(u, v)\|^2\right\}, \quad (u, v) \in D$$

Coefficient w_d increases the weight of the points values depending on the proximity to the center of area $D(u_0, v_0)$. It reduces distortions influence, as the distance from the center to the edges of area increases.

Weighting coefficient w_c performs the same function; however, in this case, the values of the brightness distribution functions are used. This coefficient extends the effective area of the fragments comparison in case of sufficiently different intensity values in area $D(u_0, v_0)$.

Along with the weight coefficients w_d , w_c , which were considered in paper [16], we introduce weight coefficient w_f , which characterizes the distance from a point to an epipolar line. For each point (u_0, v_0) on the first image there is a corresponding epipolar line $l' : au' + bv' + c = 0$ on the second image:

$$(u, v) \rightarrow \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$a = u_0 F_{11} + u_0 F_{12} + F_{13},$$

$$b = u_0 F_{21} + u_0 F_{22} + F_{23},$$

$$c = u_0 F_{31} + u_0 F_{32} + F_{33}.$$

Coefficient w_f is calculated as

$$w_f = \exp\{-d((u', v'), l')\}$$

where

$$d((u', v'), l') = \left| \frac{au' + bv' + c}{\sqrt{a^2 + b^2}} \right|$$

is a distance from point $(u + \Delta u, v + \Delta v)$ on the second image to the epipolar line l' . This coefficient is used as penalty function for the distance from the point to the correspondent epipolar line.

4. Examples

4.1. Test scene reconstruction

We compared the algorithm of the corresponding points determination using the epipolar restrictions with similar SimpleFlow algorithm. For this experiment we use a set of test stereo images "Tsukuba" (Figure 1) in the daylight and in the flashlight. Figure 2 shows the results of the disparity map reconstruction using proposed method and SimpleFlow. For quantitative assessment of quality we use the number of the disparity map pixels, which are not occlusions and differ more than by 10% from their exact values. The number of such pixels is given in Table 1. This table contains the results of processing by Simple Flow algorithm as well. The first number is the absolute number of pixels. The relative number of pixels (to total number of pixels) is given in brackets.

4.2. Digital terrain model reconstruction

Stereo images obtained from a spacecraft and shown in Figures 2a and 2b were used in the test. As parameters of the spacecraft at the moment of recording by an optical device are unknown, the problem of the fundamental matrix determination using the corresponding points of the stereo images shown in Figure 3 should be solved. Let us refer to this stage as preliminary comparison of the stereo images.

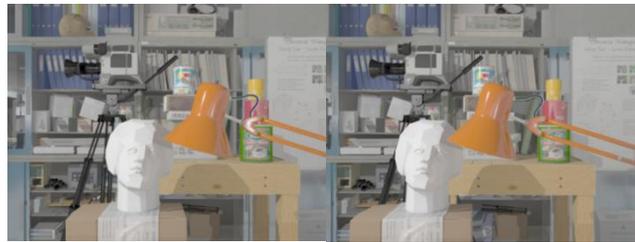


Fig. 1. – Scene stereo images: a) left view; b) right view

Table 1. Number of wrong correspondences

Images registration conditions	Proposed method	SimpleFlow
Daylight	19 191 (~8,3%)	21 187 (~9,1%)
Direct light	15 645 (~6,7%)	21 088 (~9,1%)

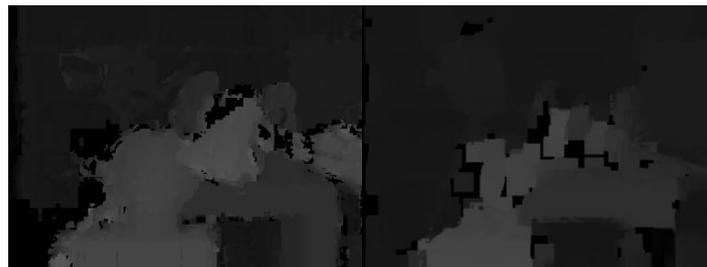


Fig. 2. – Disparity map: a) proposed method; b) SimpleFlow



Fig. 3. – Initial pair of stereoimages

Figure 4 (a) shows the fragment of one (left) image of a stereo pair with white lines illustrating the sizes and the directions of shifts between images (optical flow). Let us note that the relative number of chaotically directed white lines is rather great. When using a small number of the corresponding points these errors can lead to rough

errors when calculating a fundamental matrix. Therefore we form a system of linear equations and use conforming estimate method as it is described in [18], [19]. Thus, the fundamental matrix is obtained:

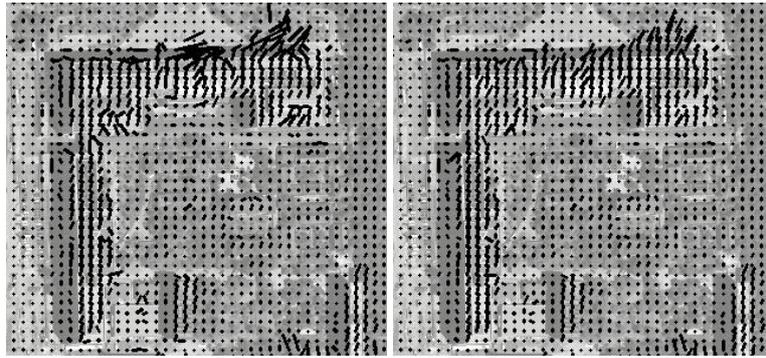


Fig. 4. – Images with the results of comparison of points: a) at a preliminary stage; b) taking into account epipolar restrictions

$$F = \begin{pmatrix} -6.73 \cdot 10^{-5} & -6.56 \cdot 10^{-5} & -5.94 \\ 1,22 \cdot 10^{-4} & 6.25 \cdot 10^{-6} & -0.65 \\ 5.98 & 0.62 & 1 \end{pmatrix}$$

After the fundamental matrix is determined, it is possible to find the corresponding points with using the epipolar restrictions. Figure 4 (b) gives a fragment of the same image with the black lines showing sizes and directions of shifts errors under the epipolar restrictions. It is easy to notice that the number of the chaotic black lines characterizing errors in this case is significantly lower.

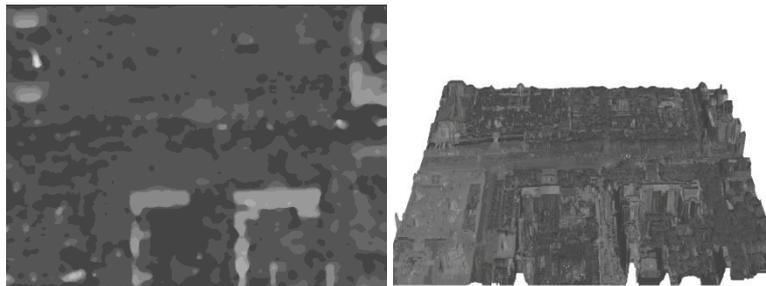


Fig. 5. – a) the disparity map, b) restored DTM

In Figures 5 the disparity map (a) and the restored digital terrain model (b) of the image segment, constructed with use of the corresponding points, are shown.

5. Conclusion

When the fundamental matrix is not known precisely, high quality of the reconstruction can be reached by the use of epipolar restrictions in the form of weight coefficient in the penalty function. These coefficients correspond to the distance from the point to the epipolar line. The results of the comparison with the most popular SimpleFlow algorithm demonstrate the prospects of the proposed approach.

Acknowledgements

Study was funded by RFBR according to the research project #13-07-12030 ofi_m, 13-07-97000-r_povolzhye_a.

References

1. **Hartley RI.** Theory and Practice of Projective Rectification. *International Journal of Computer Vision*, 1999; 35: 115-127.
2. **Monasse P, Morel JM., Tang Z.** Three-step image rectification. *BMVC 2010-British Machine Vision Conference*, 2010; 89.1-89.10.
3. **Pollefeys M.** A simple and efficient rectification method for general motion. *The Proceedings of the 7th IEEE International Conference on Computer Vision*, 1999; 1: 496-501.
4. **Häming K, Peters G.** Extension of the generalized image rectification. *Catching the infinity cases. Proceedings 4th International Conference on Informatics in Control, Automation, and Robotics (ICINCO 2007)*, 2007; 2: 275-279.
5. **Oram D.** Rectification for any epipolar geometry. *British Machine Vision Conference*, 2001; 653-662.
6. **Kumar S, Micheloni C, Piciarelli C, Foresti GL.** Stereo rectification of uncalibrated and heterogeneous images. *Pattern Recognition Letters*, 2010; 31(11): 1445-1452.
7. **Fleet D, Weiss Y.** Optical flow estimation. *Handbook of Mathematical Models in Computer Vision*, 2006; 237-257.
8. **Barron JL, Fleet DJ, Beauchemin SS.** Performance of optical flow techniques. *International Journal of Computer Vision*, 1994; 12(1): 43-77.
9. **Goshin YeV, Kotov AP, Fursov VA.** Two-stage formation of a spatial transformation for image matching. *Computer Optics*, 2014; 38(4): 886-891. [in Russian]
10. **Forsyth D, Ponce J.** *Computer Vision: A Modern Approach*. Moscow: "Williams" Publisher, 2004. 928 p. [in Russian]
11. **Hartley R, Zisserman A.** *Multiple view geometry in computer vision*. Cambridge university press, 2003.
12. **Scharstein D, Szeliski R.** A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 2002; 47(1-3): 7-42.
13. **Mičušik B, Košecká J.** Multi-view superpixel stereo in urban environments. *International journal of computer vision*, 2010; 89(1): 106-119.
14. **Sun C.** Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques. *International Journal of Computer Vision*, 2002; 47(1-3): 99-117.
15. **Luo A, Burkhardt H.** An intensity-based cooperative bidirectional stereo matching with simultaneous detection of discontinuities and occlusions. *International Journal of Computer Vision*, 1995; 15(3): 171-188.
16. **Hartley R.** In defense of the eight-point algorithm. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1997; 19(6): 580-593.

17. **Tao M, Bai J, Kohli P, Paris S.** SimpleFlow: A Non-iterative, Sublinear Optical Flow Algorithm. *Computer Graphics Forum*, 2012; 31(2): 345-353.
18. **Fursov V, Goshin Ye.** Conformed Identification of the Fundamental Matrix in the Problem of a Scene Reconstruction, using Stereo Images Image Mining. *Theory and Applications. Proceedings of IMTA-4*, 2013; 29-37.