STEREOVISION 3D TYPE WORKSPACE MAPPING SYSTEM ARCHITECTURE FOR TRANSPORT DEVICES

Janusz Szpytko, Paweł Hyla

AGH University of Science and Technology
Faculty of Mechanical Engineering and Robotics
Mickiewicza Av. 30, 30-059 Krakow, Poland
tel.: +48 12 6173103, +48 12 6173104, fax: +48 12 6173531
e-mail: szpytko@agh.edu.pl, hyla@agh.edu.pl

Abstract

One of the effects of the exponential growth universally available computing power is decrease of the hardware cost especially multimedia components. This trend combined with integration of image acquisition hardware for multimedia applications for personal computers, enable to create new kind of vision systems - stereovision. The system presented on this paper using two synchronized cameras with USB type interface.

In order to obtain an exact copy of the object in 3D type space is necessary to gather the information about projected points of the space onto the more than one image. The paper is focus on double camera stereovision system and calibration problem with image rectification. The target was achieved with use open source Computer Vision libraries and toolboxes.

Practical implementation of presented basic stereovision system (based on two cameras) is addressed for transportation devices which are operating in fixed 3D type space and include for example: device movement trajectory controlling, possible events detecting and work space of material handling devices modeling.

Paper contains the following key chapters: cameras calibration problem short discussion, single camera calibration approach, and stereovision system architecture and calibration approach.

Keywords: *stereovision, computer vision, work space mapping*

1. Introduction

The visual perception of humans and various animals was developed in course of millions years of natural evolution and deliver to the human being more than eighty % information [8] about the surrounding and environment. Near a half of our cerebral cortex (about 1.0e+10 neurons) [8] are busy with processing visual information. The architecture of our HVS (*Human Visual System*) is very complex. Each eye captures its own view, two separate captured images are sent on to the brain for processing. Left and right picture taken simultaneously are united into one picture. In the next step the human brain generated general map of differences between the combine images. It is a main sense of the three-dimensional stereo picture, because the function determining differences of the combine images are inversely proportional to the distance from the observed object. That is a main rule of binocular vision [10].

On the Human HVS base approach arising the new type of science concerned at the gain answers how to attached vision possibility to the machines. As a scientific discipline CV (*Computer Vision*) is focused for building artificial systems with the ability to obtain information from images [6, 9, 11]. The processed data in CV can take many forms: from single image [16], via views from multiple cameras [12] till to video sequence. Computer vision can be also described as a complement (but not necessarily the opposite) of biological vision.

This article focus on the implementation basic stereovision system (with two cameras) for possibility controlling trajectory [5, 13], detecting events [1] and modeling work space [17] of material handling devices.

2. Cameras calibration problem

Cameras calibration is a one of few necessary steps in computer vision in order to extract metric information from 2D images. Moreover calibration compensation among them is a form of optical aberration (distortion). In calibration dedicated for computer vision (CV) much work has been done [3, 4, 14]. Existing methods can classify into two main categories [18]: photogrammetric and auto-self calibration methods. Photogrammetric calibration mode is executed by observing a special prepare for calibration object in the 3D space. Calibration model was done with very high precision and usually consist two or three orthogonal planes. This calibration type can be done very efficiently and accurate but this method requires expensive apparatuses and an elaborate preparations.

The self-calibration method does not use any calibration object. Camera is moving on the scene and taken images. If images are taken by the same camera with fixed internal parameters, three images are sufficient to recover both the internal and external parameters and allow reconstructing 3D type scene structure. This solution is very flexible but not always can allow obtain accurate results. Other calibration method use pure rotation or vanishing points for orthogonal directions but are not popular as photogrammetric or auto-self calibration.

The search for low cost and good calibration technique generated new hybrid calibration technique which join photogrammetric accurate with auto-self calibration easiness. The calibration technique described in paper [2] only requires that calibration camera observe a planar pattern (see Fig. 1.) at a few (at least two) different orientations. Additionally this method is very flexible because the camera or the planar pattern can be moved simply by hand. The pattern translation and deviation need not be known.

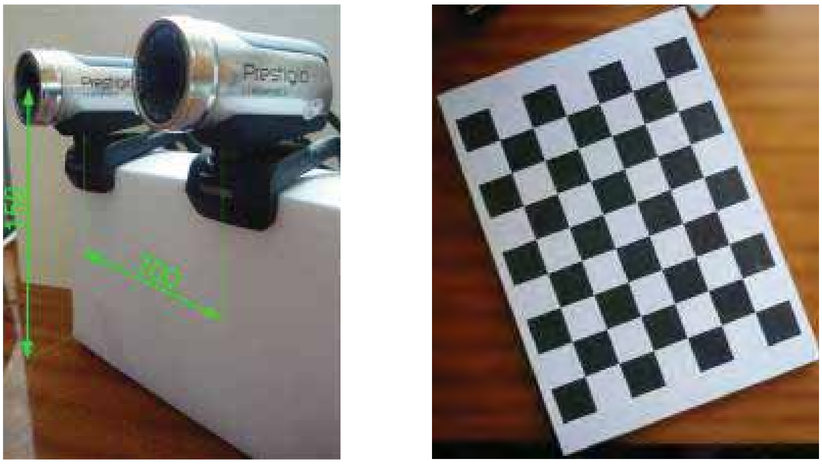

Fig. 1. Stereovision cameras position with calibration pattern board

3. Single camera calibration

The basic result of camera calibration process is a transformation maps with a 3D point $M = [X, Y, Z, 1]^T$ translated into a 2D point map $m = [u, v, 1]^T$. This transformation can be represented by a 3 x 4 projection matrix with 11 physical parameters: three angles rotations R_x, R_y, R_z, three translations t_x, t_y, t_z, the coordinates of main point (u_0, u_v), two scale factor α_u, α_v and the skewers c parameter between the image axes. Presented data not contain physical data e.g. a distortion parameter. Calibration of a standalone camera device is just enough simple. First of all we must prepare chequerboard and make with them a several photo with use the calibration device. For described calibration example was taken sixteen photos (see Fig. 2) [18].

For final calibration was use open source Computer Vision library implementing in Matlab software [2]. This calibration procedure requires few additionally steps preceding the final calibration procedure (see Fig. 3.).

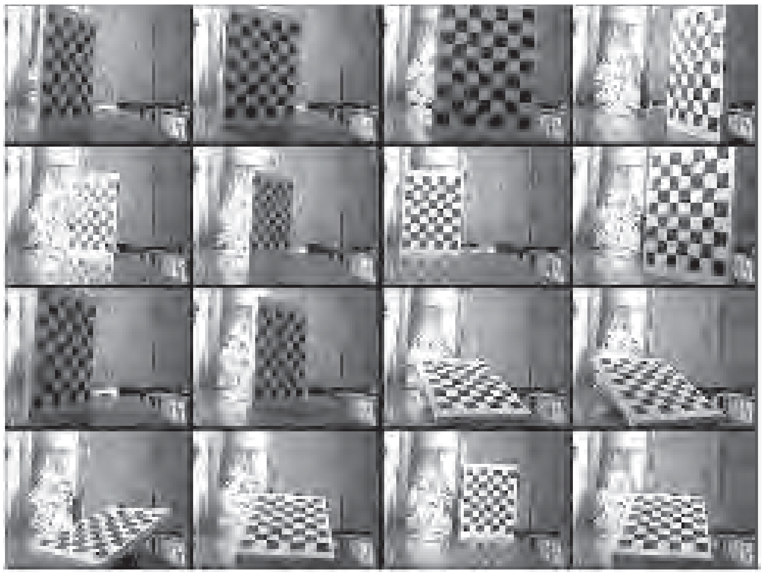

Fig. 2. Calibration images contain calibration pattern (297x210 mm) with 72 black and white squares (27x27 mm)

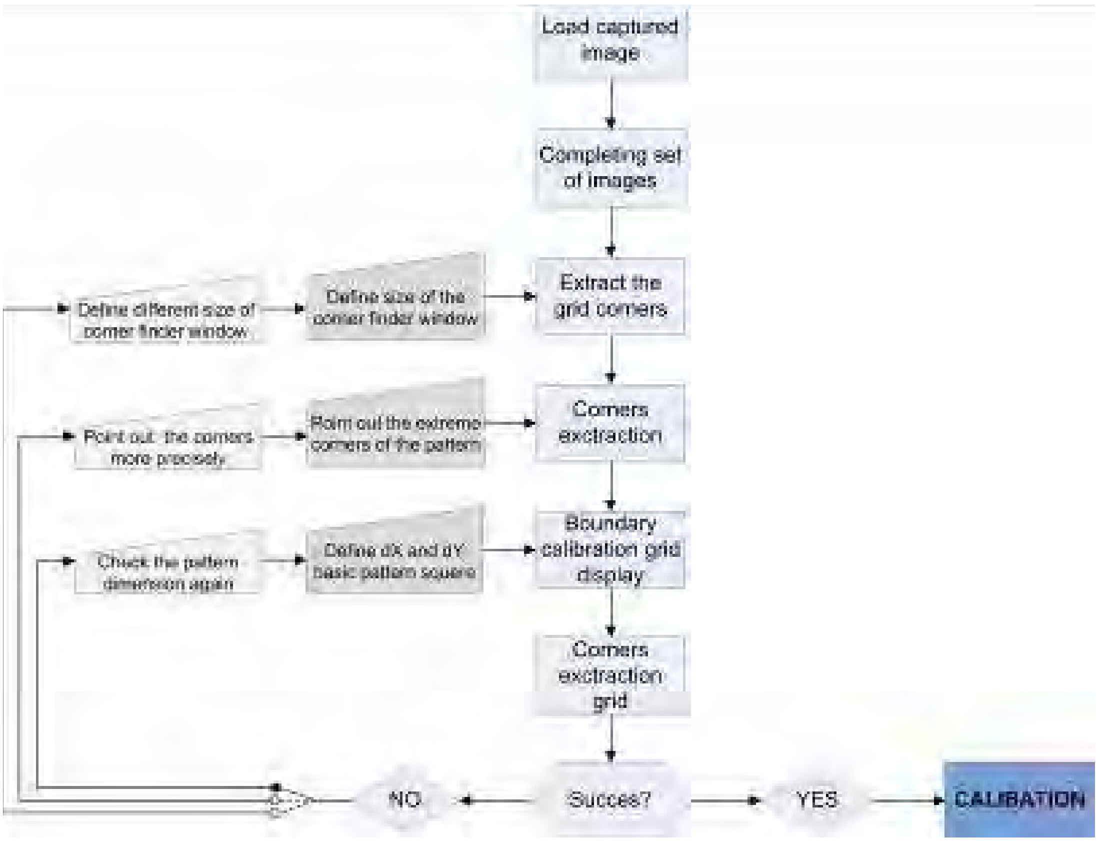

Fig. 3. Algorithm of procedure preceding calibration

After preparation and corners extraction (see Fig. 4) calibration is done in first initialization and nonlinear optimization. First initialization computes the calibration parameters not including any objective lens distortion.

The non-linear optimization minimizes the total reprojection error (see Fig. 5) over the all calibration parameters. Complete calibration was described by for calibration parameters: focal length (*fc*) give in pixels through 2x1 vector, principal point (*cc*) through 2x1 vector, skew coefficient (*alpha_c*) scalar defining the angle between the x and y pixel and distortions (*kc*) stored in the 5x1 vector.

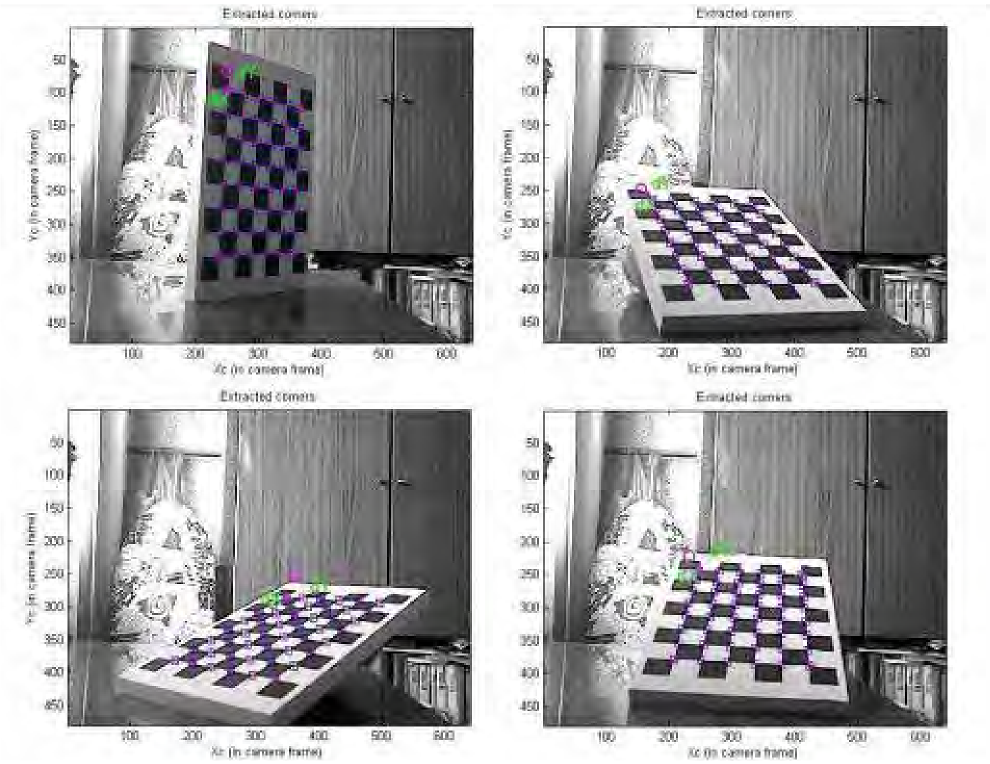

Fig. 4. Selection images (no. 10,11,13,16) with extracted corners

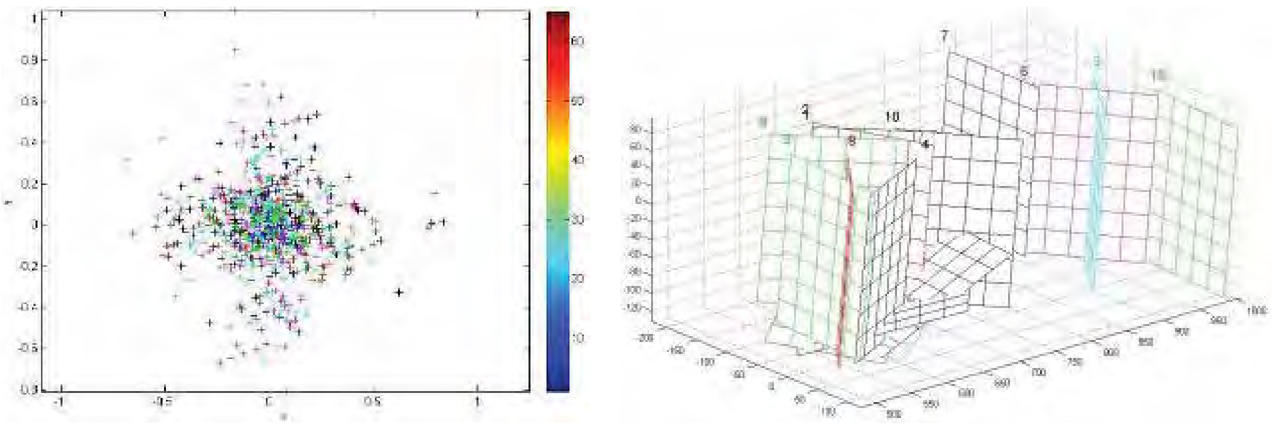

Fig. 5. Total reprojection error (in pixel) and the view of extrinsic parameters (camera-centered type)

The distortion can be corrected using Brown's distortion model (1). Brown's model [4] enables compensation for tangential distortion and radial distortion (see Fig. 6).

$$\begin{aligned} x_u &= x_d + (x_d - x_c)(K_1 r^2 + K_2 r^4 + ...) + \left\{P_1\left[r^2 + 2(x_d - x_c)^2\right] + 2P_2(x_d - x_c)(y_d - y_c)\right\}(1 + P_3 r^2 + ...) \\ y_u &= y_d + (y_d - y_c)(K_1 r^2 + K_2 r^4 + ...) + \left\{P_2\left[r^2 + 2(y_d - y_c)^2\right] + 2P_1(x_d - x_c)(y_d - y_c)\right\}(1 + P_3 r^2 + ...) \end{aligned}, \quad (1)$$

where:

(x_u,y_u) - undistorted image point,
(x_d,y_d) - distorted image point,
(x_c,y_c) - centre of distortion,
… - an infinite series,
$K_n = n^{\text{th}}$ - radial distortion coefficient,
$P_n = n^{\text{th}}$ - tangential distortion coefficient,
$r = \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2}$.

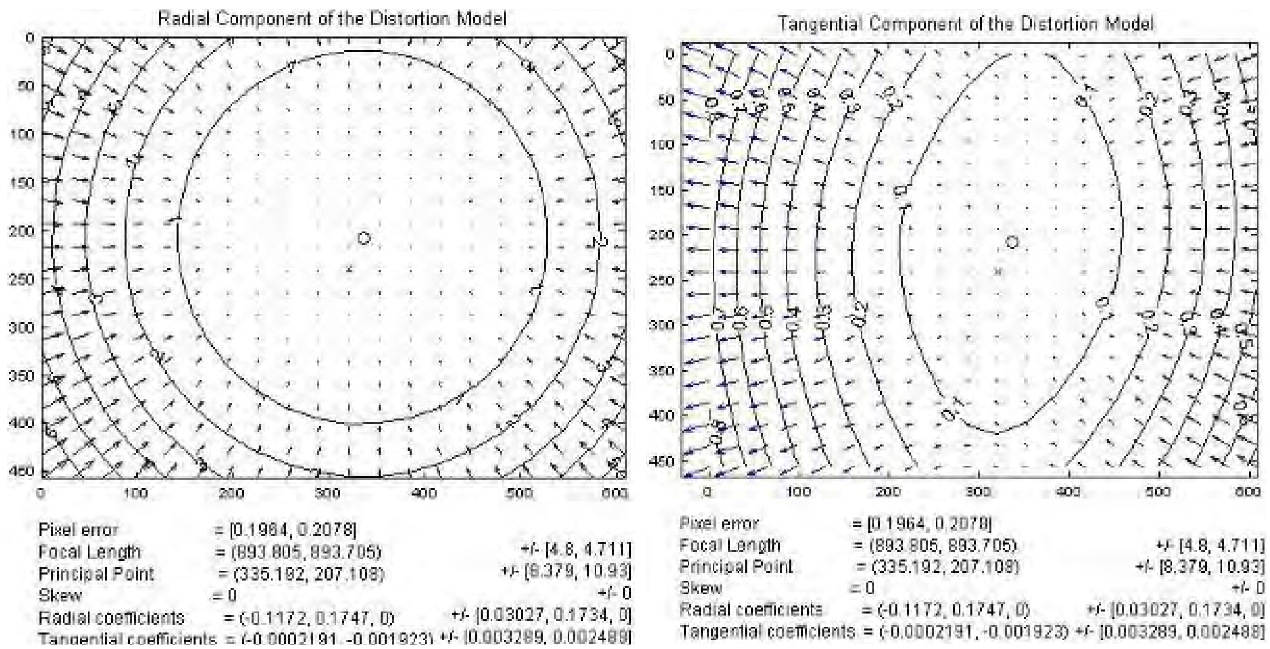

Fig. 6. Radial and tangential distortion model for Prestigio PWC2

Compared with classical techniques, the technique with use chequerboard pattern is considerably more flexible. Additionally possibility implementation the presented algorithm in numerical computing environment gives good result of calibration and data visualization. Implementing corner extraction engine includes an automatic mechanism for counting the number of squares in the grid. However to obtain good results the manual mode is better to obtain more accurate data.

4. Stereovision system architecture and calibration

One of the effects of the exponential growth universally available computing power is decrease of the hardware cost. This trend combined with integration of image acquisition hardware for multimedia applications for personal computers, enable to create new kind of vision systems - stereovision. The system presented on this paper using two synchronized cameras manufactured by Prestigio of PWC2 type, each camera acquiring the image with 1/3” CMOS sensor with 2 megapixel matrix resolution. Video stream up to 1600x1200 is transmitted via USB interface.

In order to obtain an exact copy of the object in 3D space is necessary to gather the information about projected points of the space onto the more than one image [15]. The Fig. 7 was presented three points: p_1, p_2, p_3, each of them representing the outline of the object in the stereoscopic mode. Through p_{1L}, p_{2L}, p_{3L} points in the left camera image and p_{1R}, p_{2R}, p_{3R} points in the right camera image the epipolar line was made.

In canonical arrangement, when cameras optical axis are parallel and z coordinates of image points is the same, epipolar is straight line cross pixels in the two images belonging to the same point of the scene [9].

The principal problem of reconstruction 3D image on the base stereovision images is correspondence problem [15]. For the canonical arrangement of cameras this problem is undergo a simplification, because y coordinates of images of the any point of the space are equal to oneself $y_L = y_R$. Therefore the task of searching p_L point corresponding to the p_R point is a limit to search one line with the know coordinates. The solving of the correspondence problem is appointment with the disparity images. Disparity of taken images constituting the difference of coordinates associated with images mutual moving of the space in both cameras. During disparity calculation we make establishment, that one image of the stereovision pair is a reference for another. In presented scheme right image of the stereovision system is a reference image (see Fig. 8).

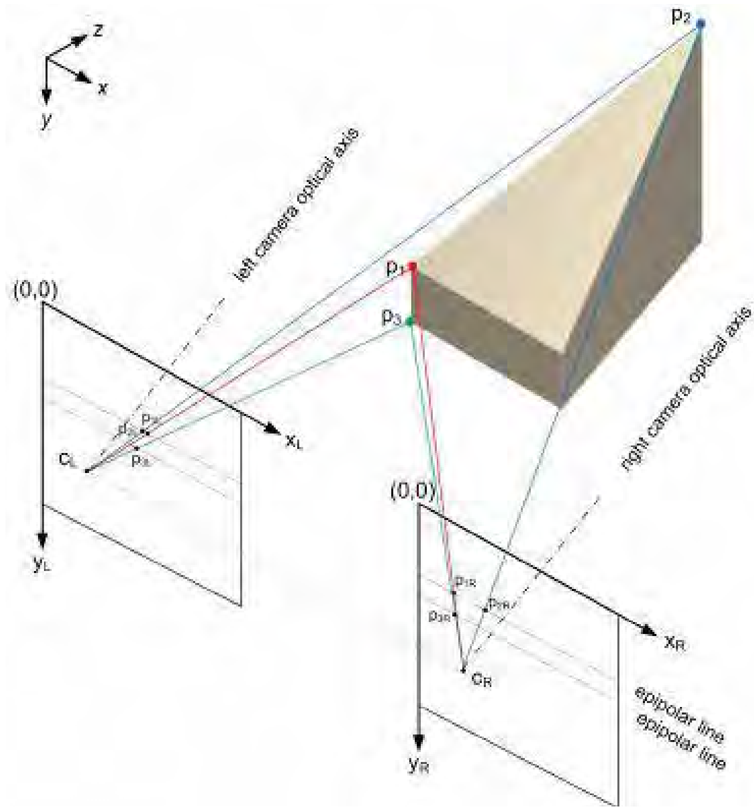

Fig. 7. The mapping scheme of the three-dimensional space

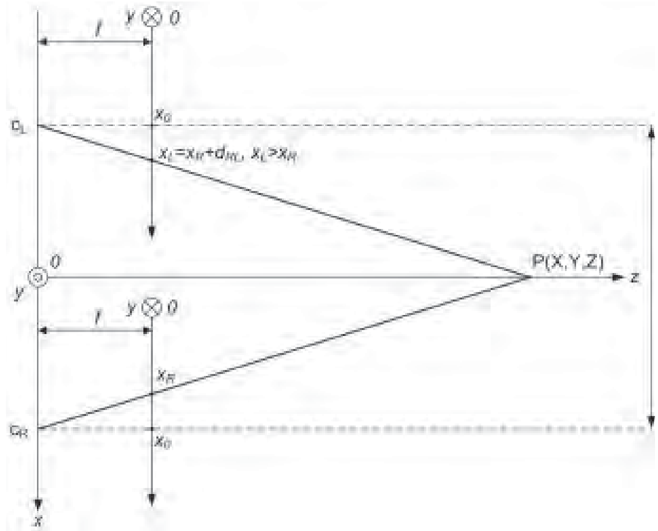

Fig. 8. Disparity calculation scheme

So, if the p_R point (reference side) with coordinates (x_R, y_R) has appropriate point in the left side p_L with (x_L, y_L), where $y_R = y_L$ is a reallocation between right and left disparity is equal (2):

$$d_{RL} = x_L - x_R \,. \tag{2}$$

Additionally an explicit relation between the pair of p_L, p_L point and the coordinates of $P(x,y,z)$ point. Coordinates of each P_x, P_y, P_z point recorded in both cameras image for canonical arrangement is possible to appoint with the equation (3), (4) and equation (5) [15]:

$$x = \frac{B(x_R + x_L - 2x_0)}{2(d_{RL})}, \tag{3}$$

$$y = \frac{B \cdot y}{d_{RL}}, \tag{4}$$

$$z = \frac{B \cdot f_p}{d_{RL}}, \tag{5}$$

where:

B - distance between cameras optical axis,
y - $yR = yL$,
d_{RL} - disparity given by equation (3),
f_p - focal length in pixel,
x_0 - coordinate of the image centre.

After successful cameras calibration is possible to do image rectification and estimate rectification error [3]. Image rectification is a specific transformation process used to project multiple images onto a common image surface. It is used to correct a distorted image into a standard coordinate system. On the Fig. 9 was presented an image from left and right after rectification.

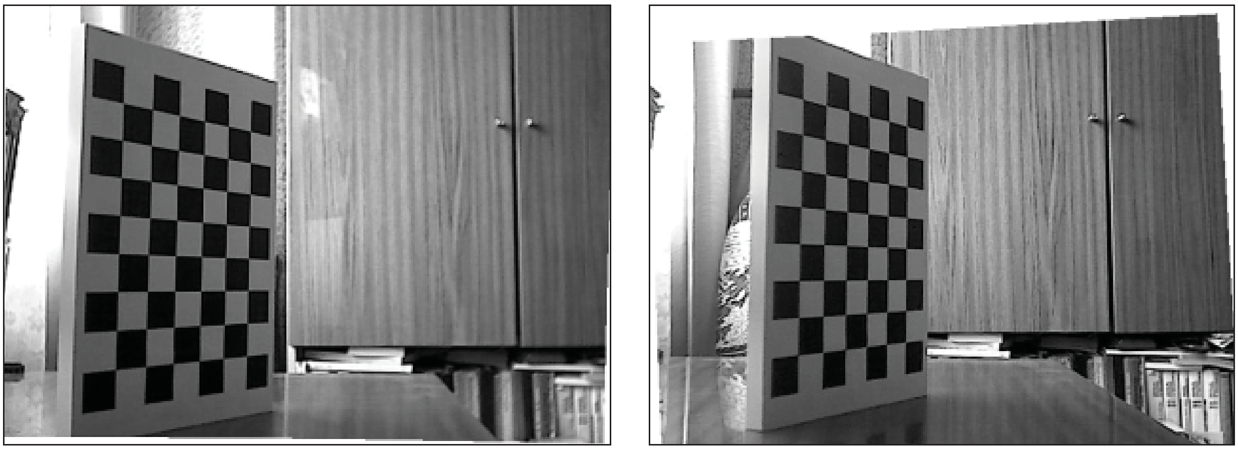

Fig. 9. Rectification process on the one of 16 calibration image (black frame presented image dimension before rectification)

Acquiring the images in stereo mode is a most important stage of action of this system type [7, 15]. The pair of images should be registered simultaneously. Preliminary processing eliminate distortions and done image rectification. All essential parameters for distortions correction for system have been presented on Fig. 10, as well as were presented in the Tab. 1.

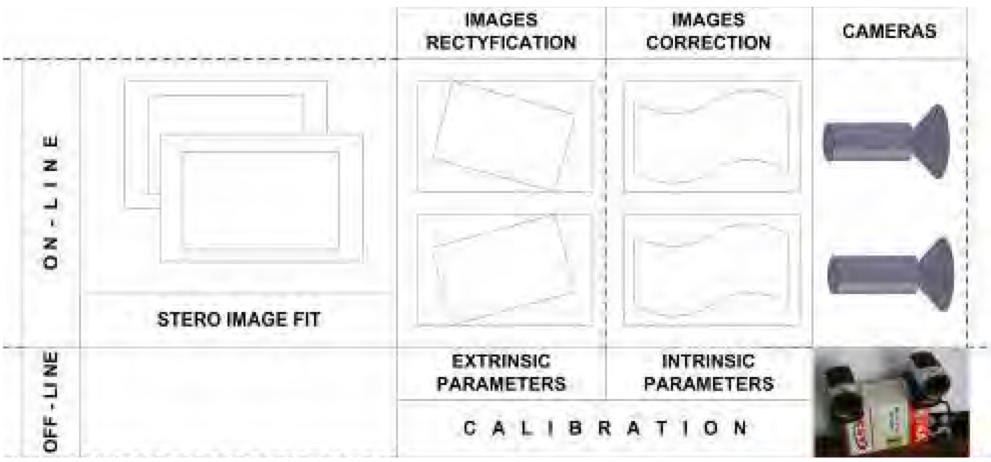

Fig. 10. The block diagram application of the stereovision

Tab. 1. Calibrated stereovision system: extrinsic and intrinsic parameters

No.	Intrinsic parameters		Left camera	Error +/–	Right camera	Error +/–
	Parameters name	Parameters label [2]				
1.	Focal Length	*fc*	893.26405 893.48998	3.06853 3.15761	897.45803 897.35879	3.04805 3.05159
2.	Principal point	*cc*	334.27900 209.55468	7.45495 7.81679	323.17409 214.07206	7.41582 7.75484
3.	Skew [deg]	*alpha_c*	0.00000	0.00000	0.00000	0.00000
4.	Distortion	*kc*	-0.11536 0.20620 -0.00027 -0.00153 0.00000	0.02880 0.17764 0.00237 0.00219 0.00000	-0.12333 0.38458 0.00008 -0.00370 0.00000	0.04929 0.60921 0.00231 0.00230 0.00000
No.	Extrinsic parameters		Position of right camera with respect to left camera			Error +/–
	Parameters name	Parameters label [2]				
5.	Rotation vector	*om*	0.05537 -0.06953 0.06751			0.00849 0.00981 0.00088
6.	Translation vector	*T*	84.77415 3.71374 2.47663			0.19504 0.14953 1.38307

If we work with one camera pair probably each of the cameras acquisition the images independently. This if fine as long as we use the stereo channels for e.g. calibration. In normal mode the cameras must works simultaneously. One possible solution is to block one of the cameras pair until the second prepares new image and informs about it. But this solution has a few flaws. When the referential camera stops working, the second video stream will be stop too. And secondary thing, if we construct a real time stereo vision the system can't be frozen and waiting for a secondary image even it was a small period of time.

5. Conclusion

The crucial element deciding about the usefulness the stereovision system is disparity. Knowing the images reallocation in both cameras images with intrinsic and extrinsic parameters (calibration process) it is possible to fix the depths of the point. When the disparity is appointed for every point of stereovision image is possible to construct a dense map of disparity which enables reconstruction the three-dimensional space.

However appointing the dense map of disparity for one pair of images require consume a lot of computational power. Every of the reference image point are connect with point correspondent point of second image of the pair. For purpose decreasing a number of connected points, the points are replaced by rectangular windows. Accuracy of 3D stereovision mapping depends on size compared window and used algorithms productivity.

Practical implementation of presented basic stereovision system (based on two cameras) is addressed for transportation devices which are operating in fixed 3D type space and include for example: device movement trajectory controlling, possible events detecting and work space of material handling devices modeling. Future research assume building a custom pan-tilt module with double cameras system for track a cargo trajectory of laboratory double girder bridge crane in stereovision mode.

Acknowledgements

The research project is financed from the Polish Science budget for the year 2008-11.

References

[1] Baha-Touzene, N., Larabi, S., *Obstacle detection with stereo vision,* IMAGE'09 Biskra, pp. 163-177, 2009.
[2] Bouguet, J. Y., *Camera Calibration Toolbox for Matlab*, Dokument elektroniczny dostępny na http://www.vision.caltech.edu/bouguetj/calib_doc/ (visited on May 30, 2010).
[3] Bradley, D., Heidrich, W., *Binocular Camera Calibration Using Rectification Error,* Proceedings of Canadian Conference on Computer and Robot Vision, pp. 1-8, Ottawa 2010.
[4] Brown, D. C., *Decentering Distortion of Lenses*, Photometric Engineering, Vol. 32, No. 3, pp. 444-462, 1966.
[5] Coifman, B., Beymer, D., McLauchlan, P., Malik, J., *A real-time computer vision system for vehicle tracking and traffic surveillance,* Transportation Research Part C 6, pp. 271-288, 1998.
[6] De Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf, O., *Computational Geometry Algorithms and Application*, Springer-Verlag, 1998.
[7] Di Stefano, L., Marchionni, M., Mattoccia, S., *A PC-based real-time stereo vision system,* Machine Graphics & Vision International Journal, Vol. 13, Iss. 3, pp. 197-220, 2004.
[8] Diamant, E., *Lecture Notes in Computer Science*, Vol. 4729, pp. 62-71, Springer, Berlin/Heidelberg 2007.
[9] Ding-Zhu, D., Hwang, F., *Computing in Euclidean Geometry, 2nd edition*, World Scientific Publishing Co., Singapore, 1998.
[10] Farag, A. A., Hassouna, M. S., El-Baz, A., Hushek, S., *Real-time vision-based approach for probe tracking in operating room.* International Congress Series, Vol. 1268, pp. 467-472, 2004.
[11] O'Rourke, J., *Computional geometry in C, 2nd edition*, Cambridge University Press, 1998.
[12] Porikli, F. M., Divakaran, A., *Multi-Camera Calibration, Object Tracking and Query Qeneration,* IEEE International Conference on Multimedia and Expo (ICME), Vol. 1, pp. 653-656, 2003.
[13] Qian, Y., Helder Arau, J., Hong, W., *A Stereovision Method for Obstacle Detection and Tracking in Non-Flat Urban Environments*, Autonomous Robots, Vol. 19, pp. 141-157, 2005.
[14] Remondino, F., Fraser, C., *Digital camera calibration methods: considerations and comparisons,* ISPRS Commission V Symposium Image Engineering and Vision Metrology, IAPRS Vol. XXXVI, Part 5, pp. 266-272, Dresden 2006.
[15] Rzeszotarski, D., Strumiłło, P., Pełczyński, P., Więcek, B., Lorenc, A., *System obrazowania stereoskopowego sekwencji scen trójwymiarowych,* Zeszyty Naukowe Elektronika, Politechnika Łódzka, Nr 10, s. 165-184, 2005.
[16] Szpytko, J., Hyla, P., *Reverse engineering vision technique applying in transport systems,* Journal of KONES Powertrain and Transport, Vol. 16, No. 3, pp. 395-400, Warsaw 2009.
[17] Szpytko, J., Hyla, P., *Work space supervising for material handling devices with machine vision assistance,* Journal of KONBiN, Warszawa 2009.
[18] Zhang, Z., *Flexible Camera Calibration by Viewing a Plane from Unknown Orientations,* Seventh International Conference on Computer Vision ICCV'99, Vol. 1, pp. 666-674, 1999.