

STEREOSCOPIC ANALYSIS AND DEPTH MAP CREATION

Ondřej Švehla¹

¹Mendel University in Brno, Czech Republic



EUROPEAN JOURNAL
OF BUSINESS SCIENCE
AND TECHNOLOGY

Volume 2 Issue 2
ISSN 2336-6494
www.ejobsat.com

ABSTRACT

This contribution is focused on the use of the stereoscopic image for the purpose of depth map creation. Further, methods for calibration of the camera(s) are discussed. A stereoscopic head was constructed for the purpose of creating a stereoscopic image. Two Basler acA1600-20uc industrial cameras with Computar M2514-MP2 lenses were used for constructing this head. Furthermore, the algorithm for obtaining the depth map is described. The programming language C# and EmguCV library were used for the implementation of the algorithm. The algorithm consists of 4 parts. The calibration of the camera(s) and image acquisition is solved as first. Calibration of the camera(s) is solved by detection of intersections on the chessboard. Further, methods for the purpose of obtaining the depth map are described. Finally the implemented algorithm is tested.

KEY WORDS

stereoscopic image, depth map, EmguCV, Basler

JEL CODES

L86

1 INTRODUCTION

Today, robotic assistants are at the forefront of scientific research. Robotic assistants are not only used in the industrial field, but are increasingly used in the home. Therefore, these assistants must be able to manipulate with objects of everyday use. For these purposes, robotic assistants must be able to detect dif-

ferent objects in the scene and subsequently manipulate with them.

Currently there exist many approaches to the detection of an in an image and subsequent manipulation. These approaches differ especially in the technology used for image acquisition and in the data storage format. Approaches based

upon a stereoscopic image or storing data in RGB-D format are used especially in practice. Each approach brings various problems, e.g. the image quality and the sensor calibration. These problems have to be addressed before creating the stereoscopic image. The depth map can be used for the detection of objects, manipulation sequence and obstacles in the scene, which is necessary for manipulation with objects.

The image quality and calibration of the camera(s) are decidedly possible problems when a stereoscopic image is used. Light conditions affect image quality in most cases. Digital noise is produced especially in low light conditions and extreme oversaturation can be caused by direct sunlight. Therefore, it is necessary to

use a polarizing filter or to use subtle image smoothing algorithms. The calibration of the camera(s) is the most important step when working with a stereoscopic image. A perfect parallel type of stereoscopic head is almost impossible to achieve. Therefore, calibration of the camera(s) is necessary. Deformation of the depth map is the next possible problem. This deformation can be detected when the depth map is obtained.

The main goal of this contribution is the creation of a configurable stereoscopic head, the subsequent acquisition of the stereoscopic image, obtaining the depth map and finding the right values of the key parameters of the depth map.

2 THEORETICAL FRAMEWORK/CALCULATION

Two basic approaches exist for the purpose of obtaining the stereoscopic image. These approaches differ especially in the orientation of the cameras. These approaches are shown in Fig. 1. The ideal approach (Fig. 1a) is almost unattainable in practice. Cameras are mounted perfectly parallel and the image is not distorted in this approach. Therefore, a conventional approach (Fig. 1b) can be often found in practice. The image is distorted by using some lens and cameras are not mounted parallel in this approach. The image must be calibrated for the purpose of firstly compensating the distortion. The OpenCV library or Matlab has an integrated calibration tool. Many experts are currently focused on the process of calibration of the camera(s). A 3D image can be created after the calibration process. (National Instruments, 2016)

Sun et al. (2011) deal with calibration of the camera(s). These authors are focused on the calibration of the camera(s) with a large field of view. The conventional model of cameras was used to obtain of the stereoscopic image. Each camera is modelled as a pinhole. Radial distortion can be seen on the lenses of the cameras which are used in this model. All calculations are performed by using rotation or transformation matrices. It is necessary to

firstly obtain the distortion. Further, a linear model must be obtained. This model is obtained by using a non-linear optimization method. This method minimizes linearity between key points in the image. Further, the image is converted by using vanishing points to a 1D space (three key points in 1D space are obtained from one key point in 2D space). The intrinsic and extrinsic parameters of the cameras are detected by this conversion. The obtained image in 1D space is affected by digital noise. Therefore, this image is smoothed by using a Lavenberg-Marquart algorithm. When the intrinsic and extrinsic parameters are detected, the rotation and transformation matrices are then calculated. The distortion of the cameras is removed by using these matrices. Cao and Foroosh (2006) also use calibration of the camera(s) in 1D space.

An alternative approach to calibration of the camera(s) uses a pinhole as the model of the cameras. The rotation and transformation matrices are used to the conversion of 3D coordinates from the coordinates system of the cameras to the general coordinates system. The Lavenberg-Marquardt, Extend Kalman Filter and Bard-Deming algorithms were combined with the above matrices and were then used to obtain the camera parameters. The Bard-

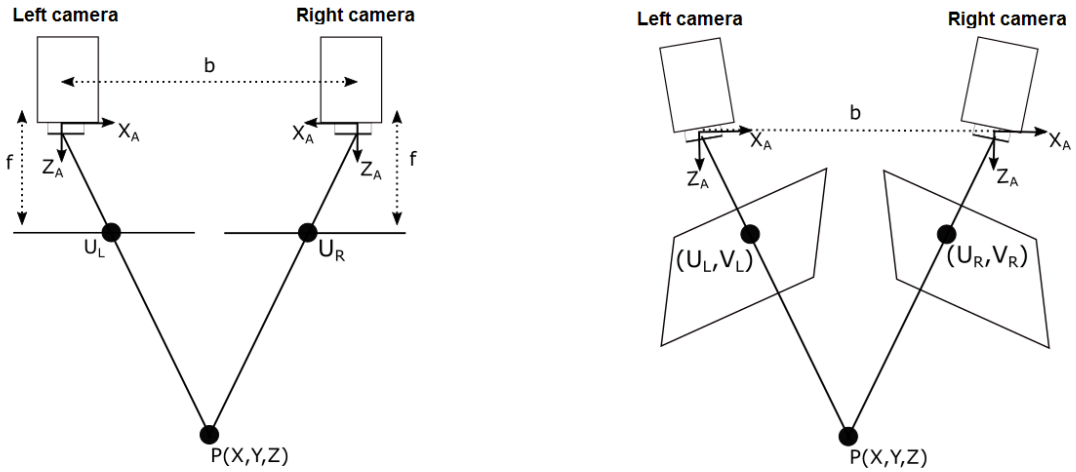


Fig. 1: The ideal (left) a conventional (right) approaches for the purpose of obtaining the stereoscopic image (National Instruments, 2016)

Deming method achieved the best result. However, this method uses a lot of memory and is time-consuming. The method of global nonlinear minimization was performed in a multi-view context. The Sorensen method was used for the first estimation of the intrinsic parameters. Sub-pixels were extracted from the image. An interpolation of these sub-pixels is used for the optimization of the calibration accuracy. A grid is used for the calibration. Intersections are detected on the grid. The calibration of the pictures is performed in greyscale (Devynck et al., 1997). Chen et al. (2012a) solve this problematic similarly to these authors.

Lindner et al. (2010) develop an alternative approach, which uses detection of the chessboard for the purpose of camera(s) calibration. Chessboard detection is also used by Chen et al. (2012b), de la Escalera and Armingol (2010), Fathi and Brilakis (2016), Prokoshin et al. (2012), Bennett and Lasenby (2014). Placht et al. (2014) use the OpenCV library for the purpose of image processing. Chu et al. (2013) also deal with camera(s) calibration by using a chessboard pattern and he uses Matlab for the purpose of image processing. Laureano et al. (2015) are also focused on camera(s) calibration by using a chessboard pattern and these authors use the Matlab image database for the testing process.

Kolomazník et al. (2013) deal with the approach towards obtaining a depth map. This approach uses the BoofCV library for the purpose of image processing. The methods SURF fast Hessian and Canny edge are used in this approach. Kamencay et al. (2012) develop an alternative approach, which uses a modified Sum of Absolute Difference (SAD) algorithm to obtain the depth map. George and George (2014) and Gu et al. (2014) use the OpenCV library for camera(s) calibration and obtaining the depth map. Wang et al. (2017) developed a similar approach and this author uses the Matlab toolbox for camera(s) calibration and the OpenCv library for obtaining the depth map. The use of the OpenCV library for the purpose of calibration and obtaining the depth map is described in Dröppelmann et al. (2010). Revuelta Sanz et al. (2011) use PCI (pseudo-color image) for matching process and segmentation of objects in the scene.

If cameras are properly calibrated, it is possible to obtain the stereoscopic image or the depth map. There exist many toolboxes for obtaining the stereoscopic image, e.g. OpenCV, BoofCV, EmguCV or MatlabCV. These toolboxes are related to the programming language of the user's choice. The toolbox OpenCV is related to the programming language C++. The toolbox BoofCV is related to the programming

language Java. The toolbox EmguCV is related to the programming language C#. Various methods can be used to obtain the depth map. These methods differ according to the type of algorithm which is used to compare of the image. These methods are the Block-Matching (BM) method, the Graph-Cut (GC) method and the Semi-Global Block-Matching (SGBM) method. The SGBM method uses the semi-global block-matching algorithm for computing stereo correspondence. The BM method uses the block-matching algorithm for computing stereo correspondence. The GC method uses the graph-cut algorithm for computing stereo correspondence. When a method is used it is then necessary to set up its key parameters. These parameters affect the quality of the depth map. The key parameters are published in Rambhia (2013) and these are the key parameters:

- **minDisparity** – minimum possible disparity value;
- **numDisparities** – maximum disparity minus minimum disparity (this parameter must be divisible by 16);
- **SADWindowSize** – matched block size (it must be an odd number ≥ 1);
- **disp12MaxDiff** – maximum allowed difference (in integer pixel units) in the left-right disparity check;
- **preFilterCap** – truncation value for the pre-filtered image pixels;
- **uniquenessRatio** – margin in percentage by which the best (minimum) computed cost function value should “win” the second best value to consider the found match correct (normally, a value within the 5–15 range is good enough);
- **speckleWindowSize** – maximum size of smooth disparity regions to consider their noise speckles and invalidate;
- **speckleRange** – maximum disparity variation within each connected component.

3 METHODOLOGY AND DATA

A stereoscopic head was constructed for the purpose of creating a stereoscopic image as the first step. This stereoscopic head was constructed by using two Basler acA1600-20uc industrial cameras with Computar M2514-MP2 lenses and is shown in Fig. 2. A frame rate of 20 fps, resolution 1624×1234 and chip size $1/1.8''$ are the key parameters of these cameras. A focal length 25 mm and optical size $2/3$ are the key parameters of these lenses.

The programming language C# was used for developing this solution. The EmguCV library was used for the purpose of image processing. The designed algorithm is shown in Fig. 3. This algorithm consists of 4 parts. The first part is the acquisition of the image. The key camera properties are set up in the process of initialization of the camera. The digital noise, brightness and shutter mode are these properties. These properties are set up for the purpose of obtaining a noiseless image. Only one image is acquired from both cameras, which means that the cameras do not operate as

a stream. The obtained data is converted to PNG format in the process of image acquisition. Furthermore, images are slightly smoothed by using the bilateral filter. Further, images are converted to the image in greyscale.

Camera calibration is next part of this solution. A chessboard detection in the acquired image is used for purpose of camera calibration. The detection is performed 10 times at different positions of the chessboard. Intersections are detected on the chessboard and the resulting lines are subsequently drawn. The detection of lines is performed by using EmguCV functions. The functions **FindChessboardCorners** and **FindCornerSubPix** were used for this purpose.

The detection of intrinsic and extrinsic camera parameters is performed in another part. The detection is performed again by using EmguCV functions. The functions **StereoCalibrate** and **cvStereoRectify** were used for this purpose.

The depth map is obtained in the last part. The key properties of the depth map, which



Fig. 2: The stereoscopic head

are discussed in Chapter 2, are firstly set up. Further, the depth map is obtained by using the SGBM method. This method uses a modified

Hirschmuller algorithm HH08. The depth map is reprojected to the 3D space in the final step.

4 RESULTS

The process of implementation has been examined in the previous chapter. Several problems have been found in the implementation of the depth map. The first problem was the format of data acquired by the Basler camera. The camera captures the scene as raw data. Therefore, the conversion to PNG format has been made using a storage sequence. The resolution of the image was another problem. Calculation of the depth map insists that the resolution of 1624×1234 is too long.

Therefore, the obtained image was scaled to the resolution of 800×600 . A GUI was developed for testing purposes. This GUI allows rendering the obtained image and the depth map. Furthermore, it allows setting parameters of the depth map, which were discussed in Chapter 2. Objects of various shapes, e.g. cups or cones, were used for the purpose of testing the process. The testing process was

realized in several lighting conditions. Methods for obtaining the depth map were tested in the first step. The Block-Matching method (Fig. 4a) and the Semi-Global Block-Matching method (Fig. 4b) were tested in this step. The SGBM method was evaluated as the best. The quality of the depth map was tested in the next step. The results of the testing process are shown in Fig. 5. The quality of the depth map was tested in a special cage with LED lighting, which is placed in the laboratory of intelligent systems, as the first step. The testing process in the special cage with LED was performed with a special photo background (Fig. 5a–5c) and without it (Fig. 5d–5f). The quality of the depth map was also tested in normal lighting conditions and this result is shown in Fig. 5g–5i. The quality of the depth map was evaluated as acceptable.

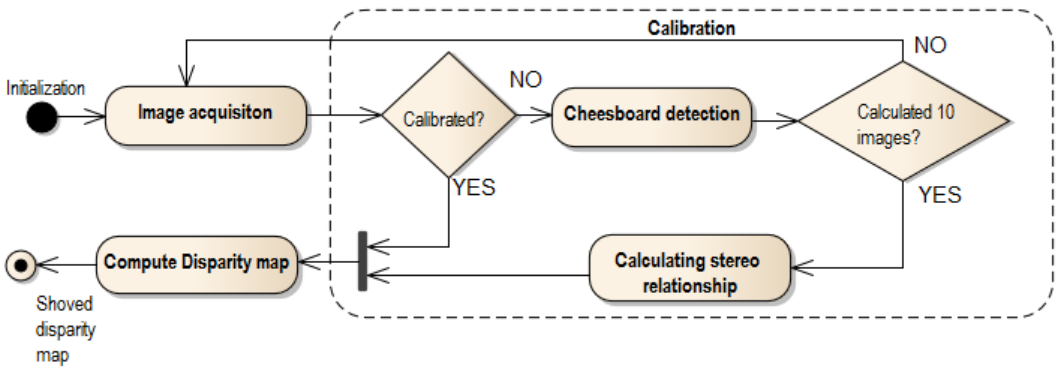


Fig. 3: Outline of the designed algorithm for image processing

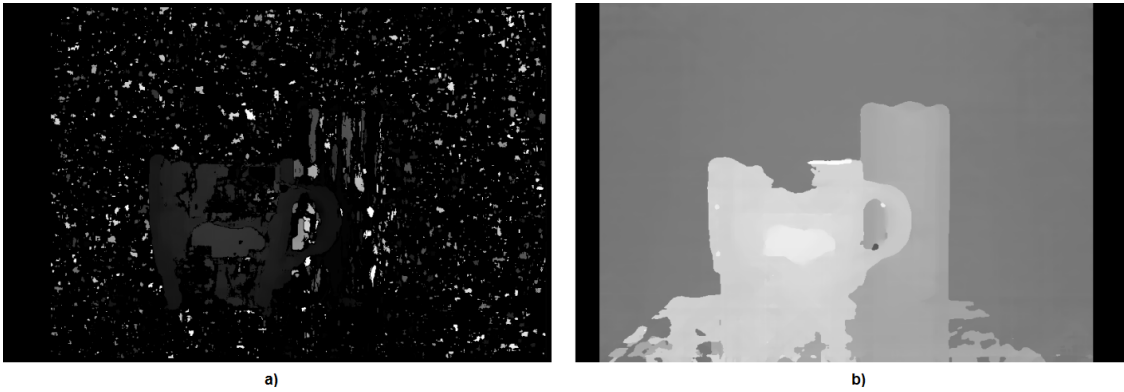


Fig. 4: The result of methods for obtaining the depth map

Some shortcomings were revealed in the testing process. The deformation of the depth map is the biggest shortcoming of this solution. The lighting conditions probably affect deformations of the depth map, because it may cause oversaturation of some parts in the image or digital noise may be produced. The deformation of edges and digital noise in the depth map are these kinds of deformations. Therefore, further

research will be focused on removing these deformations. Furthermore, further problems are included with the lighting conditions. The global setup of key parameters does not exist for different lighting conditions. Therefore, the key parameters must be reset at the change of lighting conditions. The final depth map is shown in Fig. 6 as the result of author's work.

5 DISCUSSION AND CONCLUSIONS

This article deals with the use of stereoscopic images for the purpose of creating a depth map and with the associated problems. The image quality and camera calibration need to be addressed when working with stereoscopic images. These problems are gradually examined in Section 1. The stereoscopic head was constructed

for the purpose of obtaining of the stereoscopic image. Furthermore, the algorithm for camera calibration and the creation of the depth map was implemented. The programming language C# was used for developing this algorithm. The use of the programming language C# allows control of the Katana 300s robotic arm. The

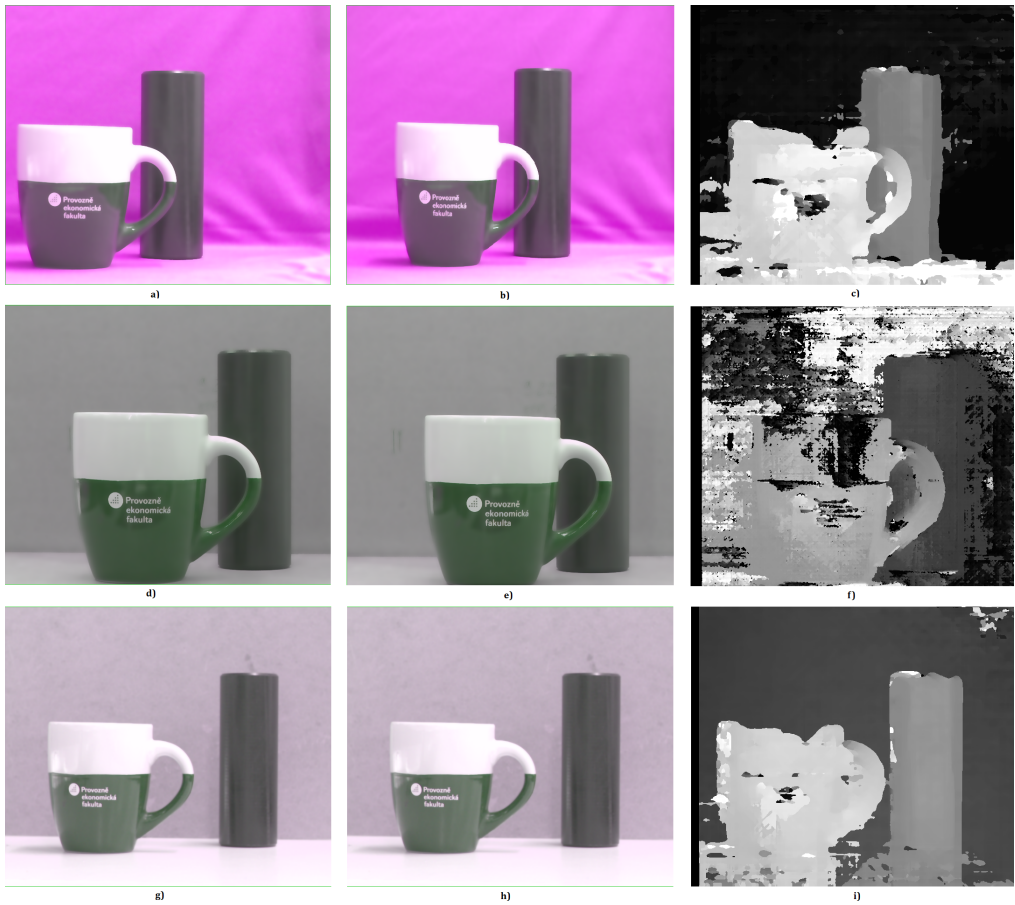


Fig. 5: Testing of the depth map quality



Fig. 6: Final depth map

problem with the image quality was solved by using a bilateral filter and setting key properties for the camera. Furthermore, the implemented solution was tested.

The result of the developed solution was also compared with another solution in the EmguCV tutorial (Emgu CV, 2012). A human head is included in the depth map in the EmguCV tutorial and some everyday objects are included in author's depth map. The comparison of these solutions evaluates author's solution as more

usable in normal conditions. The depth map of the EmguCV tutorial is better in an ideal setting and conditions.

This solution includes some shortcomings, which are discussed above in the Results chapter. These shortcomings will be removed in the future. The use of the depth map includes some benefits. The detection of obstacles in a scene is the first benefit. The detection of the object manipulation sequence or detection of distance of each object are further benefits.

6 ACKNOWLEDGEMENTS

Published results were acquired using the subsidization of the Ministry of Education, Youth and Sports of the Czech Republic, research

plan IGA MENDELU MP PEF_DP_2016002 "Stereoscopic analysis and objects recognition".

7 REFERENCES

- BENNETT, S. and LASENBY, J. 2014. ChESS – Quick and Robust Detection of Chess-Board Features. *Computer Vision and Image Understanding*, 118, 197–210.
- CAO, X. and FOROOSH, H. 2006. Camera Calibration Using Symmetric Objects. *IEEE Transactions on Image Processing*, 15 (11), 3614–3619.
- CHEN, G., GUO, Y., WANG, H., YE, D. and GU, Y. 2012a. Stereo Vision Sensor Calibration Based on Random Spatial Points Given by CMM. *Optik – International Journal for Light and Electron Optics*, 123 (8), 731–734.
- CHEN, J., BENZEROUAL, K. and ALLISON, R. S. 2012b. Calibration for High-Definition Camera Rigs with Marker Chessboard. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 29–36.
- CHU, J., GUOLU, A. and WANG, L. 2013. Chessboard Corner Detection under Image Physical Coordinate. *Optics & Laser Technology*, 48, 599–605.
- DE LA ESCALERA, A. and ARMINGOL, J. M. 2010. Automatic Chessboard Detection for Intrinsic and Extrinsic Camera Parameter Calibration. *Sensors*, 10 (3), 2027–2044.
- DEVY, M., GARRIC, V. and ORTEU, J. J. 1997. Camera Calibration from Multiple Views of a 2D Object, Using a Global Nonlinear Minimization Method. In: *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems*, pp. 1583–1589.
- DRÖPPELMANN, S., HUETING, M., LATOUR, S. and VAN DER VEEN, M. 2010. Stereo Vision using the OpenCV Library. [online]. Available at: <https://pdfs.semanticscholar.org/8149/5e6c1e6c3460b14f4be8e7876f1d6659f5c1.pdf>. [Accessed 2016, December 20].
- Emgu CV. 2012. *Stereo Imaging*. [online]. Available at: http://www.emgu.com/wiki/index.php/Stereo_Imaging. [Accessed 2016, December 8].
- FATHI, H. and BRILAKIS, I. 2016. Multistep Explicit Stereo Camera Calibration Approach to Improve Euclidean Accuracy of Large-Scale 3D Reconstruction. *Journal of Computing in Civil Engineering*, 30 (1).
- GEORGE, M. A. and GEORGE, A. M. 2014. Stereovision for 3D Information. In: BABU, B. V. et al. (eds.). *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012)*, pp. 1595–1602.
- GU, W., YIN, J., YANG, X. F. and LIU, P. 2014. Disparity Map Acquisition Based on Matlab Calibration Toolbox and OpenCV Stereo Matching Algorithm. *Advanced Materials Research*, 926–930, 3030–3033.
- KAMENCAY, P., BREŽŇAN, M., JARINA, R., LUKÁČ, P. and ZACHARIÁŠOVÁ, M. 2012. Improved Depth Map Estimation from Stereo Images based on Hybrid Method. *Radioengineering*, 21 (1), 70–78.

- KOLOMAZNÍK, J., ONDROUŠEK, V. and VYTEČKA, M. 2013. Stereoscopic Analysis of the Technological Scene. In: *19th International Conference on Soft Computing MENDEL 2013*. Brno: Vysoké učení technické v Brně, pp. 353–356. ISBN 978-80-214-4755-4.
- LAUREANO, G. T., DE PAIVA, M. S. V., DA SILVA SOARES, A. and COELHO, C. J. 2015. A Topological Approach for Detection of Chessboard Patterns for Camera Calibration. *Emerging Trends in Image Processing, Computer Vision and Pattern Recognition*, chapter 34, pp. 517–531.
- LINDNER, M., SCHILLER, I., KOLB, A. and KOCH, R. 2010. Time-of-Flight Sensor Calibration for Accurate Range Sensing. *Computer Vision and Image Understanding*, 114 (12), 1318–1328.
- National Instruments. 2016. *3D Imaging with NI LabVIEW*. [online]. Available at: <http://www.ni.com/white-paper/14103/en/>. [Accessed 2016, April 10].
- PLACHT, S., FÜRSATTEL, P., MENGUE, E. A., HOFMANN, H., SCHALLER, C., BALDA, M. and ANGELOPOULOU, E. 2014. ROCHADE: Robust Checkerboard Advanced Detection for Camera Calibration. In: FLEET, D. et al. (eds.). *ECCV 2014, Part IV*, pp. 766–779.
- PROKOS, A., KALISPERAKIS, I., PETSA, E. and KARRAS, G. 2012. Automatic Calibration of Stereo-Cameras Using Ordinary Chess-Board Patterns. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B5, 45–49.
- RAMBHIA, J. 2013. *Disparity Map*. [online]. Available at: <http://www.jayrambhia.com/blog/disparity-mpas>. [Accessed 2016, October 31].
- REVUELTA SANZ, P., RUIZ MEZCUA, B., SÁNCHEZ PENA, J. M. and THIRAN, J.-P. 2011. Stereo Vision Matching over Single-Channel Color-Based Segmentation. In: *Proceedings of the International Conference on Signal Processing and Multimedia Applications (SIGMAP)*, pp. 126–130.
- SUN, J., LIU, Q., LIU, Z. and ZHANG, G. 2011. A Calibration Method for Stereo Vision Sensor with Large FOV based on 1D Targets. *Optics and Lasers in Engineering*, 49 (11), 1245–1250.
- WANG, F., JIA, K. and FENG, J. 2017. The Real-Time Depth Map Obtainment Based on Stereo Matching. In: PAN, J. et al. (eds.). *Intelligent Data Analysis and Applications*, pp. 138–144.

AUTHOR'S ADDRESS

Ondřej Švehla, Department of Informatics, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: ondrej.svehla@mendelu.cz