Journal of KONES Powertrain and Transport, Vol. 19, No. 3 2012

DISPARITY COMPUTE METHODS IN THREE-DIMENSIONAL SCENE RECONSTRUCTION FOR OVERHEAD TRAVELLING CRANE WORK SPACE VISUALIZATION

Janusz Szpytko, Paweł Hyla

AGH University of Science and Technology
Faculty of Mechanical Engineering
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

The problem of ensuring the safe and efficient cranes operations in automated manufacturing processes involves the needs of the operating workspace identification, especially in the automation mode. However, this paper describes the problem of overhead travelling crane's workspace identification with the stereovision. Main authors' attention was focused on the stereo pictures matching problem through optimization exist disparity computation methods with algorithms optimization by completion with implement correlation matching algorithms. The stereo system is based on the single camera localized under the crane trolley that allows obtaining the sequences of stereo snapshots of crane workspace during the crane or trolley movement. Typically, disparity was produce by calculating the displacement of each pixel of the stereo image along an epipolar line. The disparity software optimization was based on matching algorithms architecture. In this article, the disparity optimization procedure achieved by produce disparity on the ROI (Region of Interest) method is described. Disparity was obtained with the help of sets homologous pixels collections. These sets were determinate with variety kind of similarity measures, which was implemented to disparity search algorithm. Each disparity map (based on the separate similarity measure) was tested under outward appearance and computing time criteria.

Keywords: *material handling devices, stereovision, disparity map, work space modelling*

1. Introduction

The process of manufacturing of both tools and machines evolved from creation of devices that were supposed to make human's life easier up to those that thanks to fully automatic performance started replacing human activity in a more and more prominent way. Nowadays the role of machine operator has been shifted from direct supervision over the whole process to maintenance of the system responsible for introducing the boundary conditions directly connected to surroundings as well as the character of the task itself. The autonomy of contemporary machines and appliances requires implementation of mechanisms allowing not only to acquire information concerning current localization of the machine [18, 19], but also additional data specifying the type of environment it is functioning in.

The scope of activity of the self-moving machines, apart from being focused on highly specialized tasks, is strictly connected to environment in which human vision developed. Therefore, equipping such devices with stereoscopic vision inspired by optical system of perspective evaluation among all organisms, especially on Human Binocular Vision possibility, seems to be a natural phenomenon [18, 20]. Enabling the machines to analyze the reality by means of information derived from visual stimuli brought to achieving a higher level of interaction between the device and its operator with both of them working on the same plane [11, 22].

The computing power of modern computers and digital processing systems [5, 9] has increased dramatically in comparison to that of machines from the 90s of the last century. Although an effective real-time navigation (based on variety kind of image processing technique) has not been fully applied [14], nonetheless contemporary computers enable to implement algorithms that proved useful in deriving information about the surrounding world [2, 8, 15].

The article was focused on the stereo pictures matching problem through optimization exist disparity computation methods with algorithms optimization by completion with implement correlation matching algorithms. Moreover the test of different similarity methods useful in disparity estimation procedure [6, 12, 14, 21] was done. The article includes also evaluation of the algorithms usefulness in disparity calculation or disparity map [10, 23] produce.

2. Disparity in three-dimensional scene reconstruction

Image representation of three-dimensional scene is only, disregarding the method and equipment applied, created as a result of projection of three-dimensional space on a limited surface (determined by measurements of the light-sensitive converter, the so-called sensor matrix). Parameters of the geometric conversion in question depend first and foremost on intrinsic parameters of optical system in the image capture device. The final image is determined by two parameters: scale and perspective. Projection of the same object placed in a proper distance from the image-recording device is also connected to the value of focal length. The key lies in obtaining data related to transformation parameters of the optical system through its calibration. Calibration procedure depends on taking photographs with object with strictly defined geometry [3, 18, 19]. After certain computations is possible to obtain a piece of information concerning the real value of focal length of the optical system as well as generated deformations (tangential and radial distortions). Thanks to data collected in conditions of real representation, it is also possible to determine a distance between the image plane connected to the recording device and the object plane referring to the represented thing.

A specific case including calibration is the analysis of the images recorded on two cameras simultaneously whose optical axes were shifted towards each other by some strictly specified value. In this case additional calibration parameter is a distance between optical axes of the recording devices [13, 14]. It enables to estimate the so-called disparity effect [2, 9, 10, 23], which means the level of position incompatibility between corresponding points in the right and the left image being a part of stereo image [7].

2.1. Scene reconstruction from stereo image

In order to gain a unique representation of a single point in three-dimensional space, information of projection of that single point in more than just one image is required. In a stereo optic system, the crucial task consists in images rectification process (Fig. 1). To obtain useful and high reliable disparity map each acquired snapshot of the stereo pair must be correlated through so-called epipolar line, additionally must be known location of all corresponding points between each photos. Furthermore, if the left and right stereo pictures are aligned to be a coplanar, the disparity search problem is simplified, because the stereo matching is limited only to the horizontal line search.

The coordinates of a point W on showed on the Fig. 1 in three-dimensional workspace are mapped based on the coordinates of conjugated pairs of m_L and m_R points on image plane and rectifying retinal plane. The perspective projection matrixes (PPMs) of the left image plane before (focal plane) and after rectification (retinal plane), denoted $\widetilde{\mathbf{P}}_{Lo}$ and $\widetilde{\mathbf{P}}_{Ln}$ respectively, in the homogenous coordinates can be written as the composition of the rotation matrixes $\widetilde{\mathbf{Q}}_{Lo}$ and $\widetilde{\mathbf{Q}}_{Ln}$, and the translation vectors $\widetilde{\mathbf{q}}_{Lo}$ and $\widetilde{\mathbf{q}}_{Ln}$:

$$\begin{aligned} \widetilde{\mathbf{P}}_{Lo} &= \left[\widetilde{\mathbf{Q}}_{Lo}, \widetilde{\mathbf{q}}_{Lo}\right]. \\ \widetilde{\mathbf{P}}_{Ln} &= \left[\widetilde{\mathbf{Q}}_{Ln}, \widetilde{\mathbf{q}}_{Ln}\right] \end{aligned} \tag{1}$$

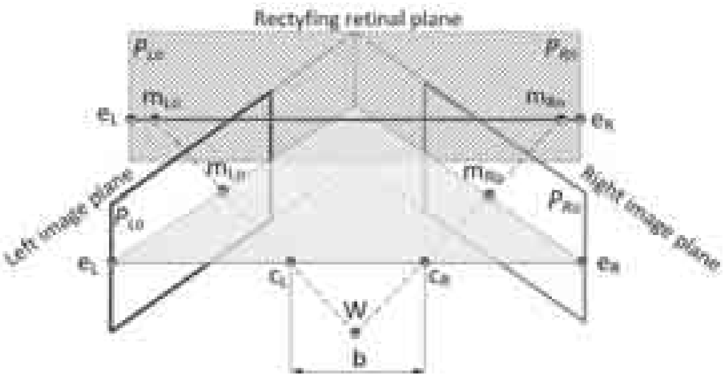

Fig. 1. Epipolar geometry describing rectification procedure

The **w** vector of coordinates of the given point *W* obtained through perspective projection can by derived from:

$$\begin{cases} \mathbf{w} = c_L + \lambda_o \cdot Q_{Lo}^{-1} \cdot \tilde{\mathbf{m}}_{Lo}, \\ \mathbf{w} = c_L + \lambda_n \cdot Q_{Ln}^{-1} \cdot \tilde{\mathbf{m}}_{Ln}, \end{cases} \tag{2}$$

where:
c_L – the centre point of the left focal plane,
$\lambda_o \in \Re, \lambda_n \in \Re$.

Thus, the homogenous coordinates of any point m_L on the retinal plane can be calculated as the transformation of coordinates of a corresponding point on the original image plane, performed according the formula:

$$\tilde{\mathbf{m}}_{Ln} = \frac{\lambda_o \tilde{\mathbf{Q}}_{Ln}}{\lambda_n \tilde{\mathbf{Q}}_{Lo}} \cdot \tilde{\mathbf{m}}_{Lo} \cdot \tag{3}$$

3. Disparity estimation algorithm

This chapter includes presentation an algorithms determining the disparity map on the base of correlation coefficient and comparison of particular regions on rectified image pairs. In the assumed system of a left and right image projection, tables *A* and *B* might be distinguished where *A* is a reference image while *B* is used to look for points or regions that would correspond to points or regions in the image *A*, still bearing in mind the shift of optical axis of the recording device. In the process of disparity map creation, digital photograph is changed into a table of numbers consisting of columns and lines directly corresponding to the width and the height of the photograph dimensions in pixels. Additionally each pixel has its own address to which a vector of colour intensity in the set of *I (m,n) = (R,G,B)* is ascribed. The vector is expressed in the value of threshold of red, green and blue from *I (m,n) = (0-255, 0-255, 0-255) value*. Therefore, if there is a particular point in the image *A*, described with the coordinates *(m,n)* what we are looking for is a corresponding point out of the set of points included in matrix ***B*****,** ascribed to it the address *I (m,n) = (0,0)*. In the a.m. transformation, the set of coordinates may be expressed through correlated pair of indexes *(x, y)* for the point in the reference image *A* and (μ, ν) for the correlated point in the image *B*.

3.1. Digital image correlation for disparity estimation

Digital image correlation methods might be understood as “finding an equivalent” or matching (*Digital Image Matching*) particular regions of images. In fact, it involves finding strictly determined equivalent in other digital image. DIM methods most often embrace four basic stages:

a) Defining the matching elements (point, line or region),

b) Finding equivalents to define elements in the second image (in case of a stereo pair),
c) Calculation of a spatial localization of the matched elements against the reference image (disparity),
d) Estimation of the matching level.

Algorithms determining the correlation coefficient through comparing regions that contain a dozen or even a couple of dozens of points have to embrace the necessity to establish the measurements of region comparisons were made to. In the describing procedure the mini matrix dimensions was determining to the searching area were expressed by the constant *U* factor with the x-index for the width of the region and *V* for its height. Chosen similarity measure useful in disparity estimation algorithms are presented below [17]:

- Locally scaled Sum of Absolute Differences (LSAD),
- Locally scaled Sum of Squared Differences (LSSD),
- Normalized Cross-Correlation (NCC),
- Sum of Absolute Differences (SAD),
- Sum of Squared Differences (SSD),
- Zero mean Normalized Cross-Correlation (ZNCC),
- Zero mean Sum of Absolute Differences (ZSAD),
- Zero mean Sum of Squared Differences (ZSSD).

Obtaining a reliable disparity map [2, 5, 9, 23] reflecting the distance between the image and object planes has a substantial meaning especially in applications created for the needs of autonomic systems enabling creation of digital map of the work space for dedicated transport devices. Stereo vision is not the only method providing this kind of information, but due to application of more and more efficient computational units, it's possible to apply algorithms serving possibility of the disparity estimation. On the Fig. 2. was present an algorithm to three-dimensional scene reconstruction on the disparity base. However, digital presentation of obstacles was additionally enriched by produce an anaglyph [4] image (Fig. 5). These images contain two differently filtered coloured images, cyan and red combined together. With the special anaglyph glasses, an overhead travelling crane operator can see three-dimensional view of obstacles spaced in workspace.

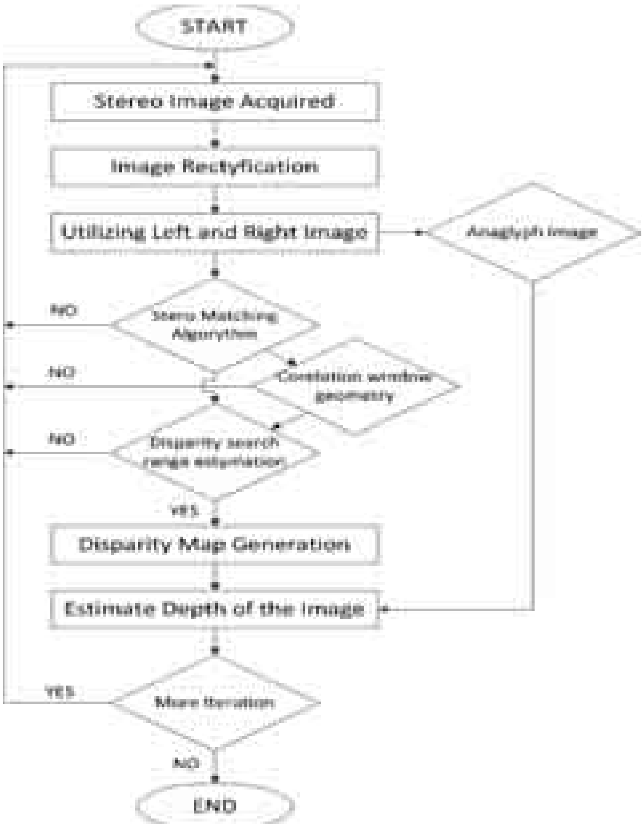

Fig. 2. An algorithm of the three-dimensional scene reconstruction on the disparity base

In order to evaluate the usefulness of algorithms in disparity estimation, the crucial is proper selection of the stereo matching algorithm (includes similarity measure, correlation window size) and disparity search range. The key to evaluate the usefulness of such algorithm lies in reliable reflection of demand on computing power and the calculation time presented in the form of disparity search function. An example disparity calculation time from the stereo picture of the basic dimension 322x244 (after rectification) was present on the Fig. 3.

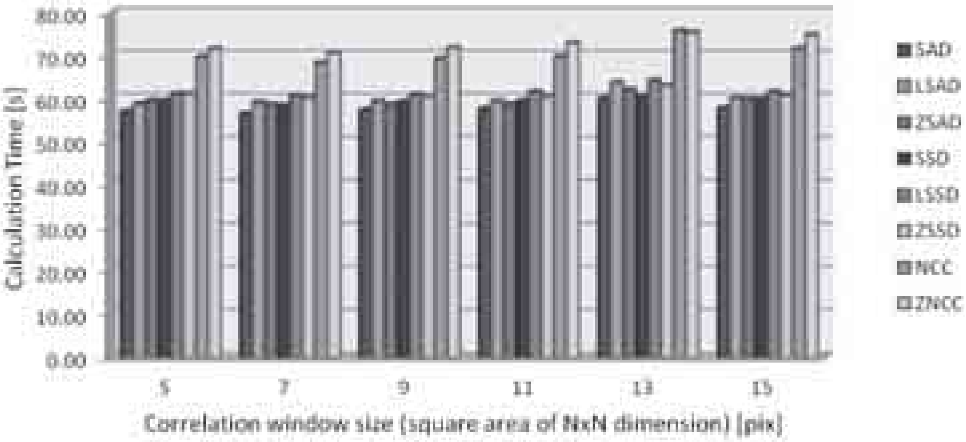

Fig. 3. Disparity calculation time in correlation windows size and use similarity measure function

Fig. 4. Left and the right rectification picture of the laboratory crane workspace (camera suspend under crane trolley)

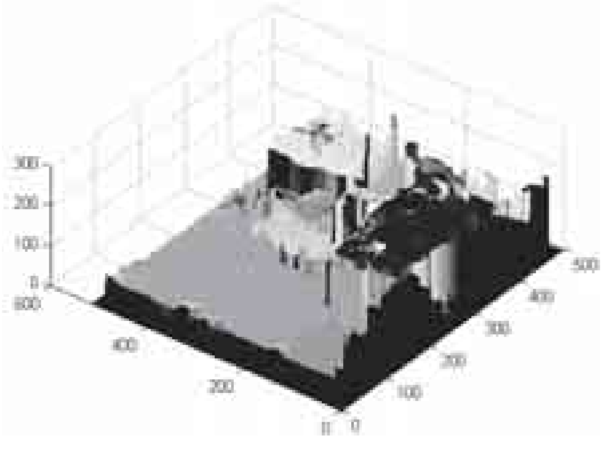

Fig. 5. An anaglyph image: left and right rectification pictures connected together (with marked disparity range)

Fig. 6. Dense disparity map

The described stereo system (Fig. 8) is based on the single camera localized under the crane trolley that allows obtaining the sequences of quasi-stereo snapshots of crane workspace during the crane or trolley movement. The baseline shift between cameras was known through data collected from group encoder recorded bridge and trolley movement.

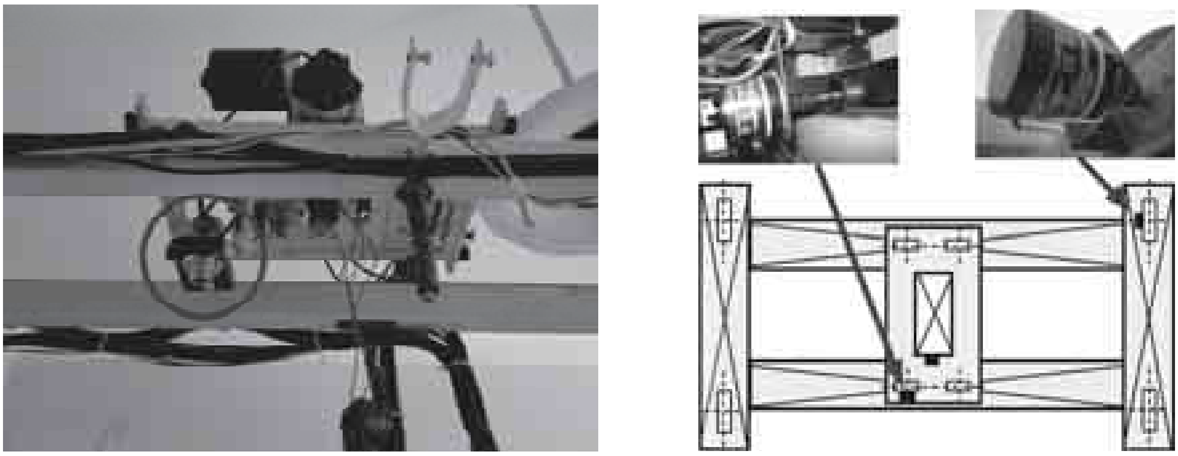

Fig. 7. Mono camera stereovision system with the sets of encoders mounted on the crane mechanism

4. Conclusion

This paper show a challenge of disparity estimation being a key element in reconstruction of 3D scene with the help of stereo vision especially to poses an information to obtain reliable map of material handling devices workspace.

The problem of safe and efficient crane operations is solved in the paper through unconventional stereovision system based on the single camera installed under the trolley of a crane that allows acquiring the stereo pairs of snapshots of crane workspace during the crane\trolley motion. This type of the hardware architecture in the future can be improved. The possibility of use one camera make possible exclude necessity of stereo pair rectification before disparity calculation, this step will be very useful because enable to apply the parallel programming in disparity calculation.

At present time a great challenge are research a new methods (especially algorithms) that enable to obtain an reliable dense disparity map with simultaneously relatively low demand on computing power (shorter calculation time). In this domain, authors rest their hopes in a new computing method called multi-GPU, which seems to offer full-quality stereo vision image together with possibility to analyze it in the real-time mode. This technique offers all calculations being made not by CPU (*Central Processor Unit*), but by stream processors of the graphic card. CUDA technology (*Compute Unified Device Architecture*) using the stream processors built in the graphic systems for the needs of general numerical problems (including disparity estimation) constitutes an interesting alternative for the traditional computing based on commonly used sequence processors.

Acknowledgement

The research project is financed by the Polish National Science Centre.

References

[1] Azzari, P., *Reconstruction from image correspondences,* Universit´a di Bologna, PhD. Thesis, December 2008.

[2] Brewer, N., Liu, N., Wang, L., *Stereo disparity calculation in real-world scenes with Informative Image Partitioning*, 25th International Conference of Image and Vision Computing New Zealand (IVCNZ), pp. 1-8, 2010.

[3] Cramer, M., *Digital camera calibration,* European Spatial Data Research, Dublin 2009.

[4] Dubois, E., *A projection method to generate anaglyph stereo images*, IEEE Proceedings of International Conference on Acoustics, Speech and Signal Processing ICASSP'01, Vol. 3, pp. 1661-1664, 2001.

[5] Georgoulas, Ch., Andreadis, I., *FPGA based disparity map computation with vergence control,* Microprocessors and Microsystems, Article in Press, 15, 2010.

[6] Karge, A., *Implementation and test of area based correlation detection algorithms,* Department of Informatics and Engineering Science, University of Applied Science Cologne, pp. 1-29, Cologne 2003.

[7] Kytö, M., Nuutinen, M., Oittinen, P., *Method for measuring stereo camera depth accuracy based on stereoscopic vision,* Proc. SPIE 7864, pp. 1-9, 2011.

[8] Lowe, D. G., *Distinctive image features from scale-invariant keypoints,* International Journal of Computer Vision, pp. 1-28, 2004.

[9] Mandai, S., Shi, B., Dudek, P., *Binocular disparity calculation on a massively-parallel analog vision processor*, 12th International Workshop on Cellular Nanoscale Networks and Their Applications CNNA, pp. 1-5, 2010.

[10] Mühlmann, K., Maier, D., Hesser, J., Männer, R., *Calculating dense disparity maps from color stereo images, an efficient implementation,* International Journal of Computer Vision, Vol. 47(1/2/3), pp. 79–88, 2002.

[11] Ogale, A. S., Aloimonos, Y., *A Roadmap to the integration of early visual modules,* Int. J. Comput. Vision, Vol. 72, Is. 1, pp. 9-25, 2007.

[12] Rieck, K., Laskov, P., Sonnenburg, S., *Computation of similarity measures for sequential data using generalized suffix trees,* In Proceedings of NIPS, pp. 1177-1184, 2006.

[13] Rzeszotarski, D, Sawicki, P., Ostrowski B., *Prototypowy system stereowizyjny typu machine vision – Działanie i dokładność,* Archiwum Fotogrametrii, Kartografii i Teledetekcji, Vol. 17b, s. 717-727, 2007.

[14] Rzeszotarski, D., Strumiłło, P., Pełczynski, P., Wiecek, B., Lorenc, A., *System obrazowania stereoskopowego sekwencji scen trójwymiarowych,* Elektronika, prace naukowe, Nr 10, s. 165-184, 2005.

[15] Storch, B., Wierucka, I., *Optyczne pomiary zarysów powtarzalnych z wykorzystaniem technik przetwarzania obrazu,* Acta Mechanica et Automatica, Vol. 1, No. 2, s. 59-62, 2007.

[16] Szpytko, J., Hyla, P., Gbyl, M., *Sterowanie dźwignicą przy pomocy gestów z użyciem hybrydowego urządzenia wejścia z trójwymiarowym układem odniesienia,* Logistyka, Nr 6, s. 3657-3669, 2011.

[17] Szpytko, J., Hyla, P., *Miary podobieństwa w wyznaczaniu dysparycji obrazu stereowizyjnego,* Logistyka, Nr 3, s. 2217-2226, 2012.

[18] Szpytko, J., Hyla, P., *Odwzorowanie przestrzeni roboczej środka transportu technologicznego z użyciem układu wizyjnego,* Czasopismo Techniczne, T. II, Z. 7, s. 505-514, Wydawnictwo Politechniki Krakowskiej, Kraków 2011.

[19] Szpytko, J., Hyla, P., *Stereovision 3D type workspace mapping system architecture for transport devices,* Journal of KONES Powertrain and Transport, Vol. 17, No. 4, pp. 495-504, Warsaw 2010.

[20] Szpytko, J., Hyla, P., *Work space supervising for material handling devices with machine vision assistance,* Journal of KONBiN, Safety and Reliability Systems, No. 3-4, pp. 7-16, Warsaw 2009.

[21] Watman, C. Austin, D. Barnes, N. Overett, G., Thompson, S., *Fast sum of absolute differences visual landmark detector*, Proceedings of International Conference on Robotics and Automation, ICRA '04, Vol. 5, pp. 4827-4832, April-May 2004.

[22] Yang, Q., Wang, L., Yang, R., Stewenius, H., Nister, D., *Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 3, pp. 1-13, 2009.

[23] Yong, Y. S., Hon, H. W., *Disparity estimation for objects of interest,* World Academy of Science, Engineering and Technology, Vol. 43, pp. 536-539, 2008.