ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2009. No. 3(91)
ELEKTRONIKA IR ELEKTROTECHNIKA

T125 AUTOMATION, ROBOTICS
AUTOMATIZAVIMAS, ROBOTECHNIKA

A Stereovision System for 3-D Perception

A. Lipnickas

The Mechatronics Centre for Studies, Information and Research, Kaunas University of Technology, Studentu str. 48-109, LT-51367 Kaunas, Lithuania, email: arunas.lipnickas@ktu.lt

A. Knyš

Department of Control Technology, Kaunas University of Technology, Studentu str. 48-325a, LT-51367 Kaunas, Lithuania, e-mail: andrius.knys@ktu.lt

Introduction

3-D depth measurement and perception is useful in many applications, such as robot navigation, teleoperation or modelling in virtual/augmented reality. Recent advantages in computer stereo vision and new low cost high speed synchronized cameras have simplified many application tasks.

The stereo problem is based on the physical phenomena, that a 3-D object has multiple projections depending on the point of view. Therefore, it is possible to reconstruct the 3-D scene from at least two images of the same object taken from two distinct points of view. The corresponding pixels from one image are found in the other one and according to this information the disparity map is build. Later, the 3-D scene is reconstructed according to the disparity map. The main problem in stereo computation lies in design of corresponding points searching algorithms. According to [1] there are two classes of stereo matching algorithms: *global* (with off-line exhausting computations) and *local* – less accurate, but appropriate in many real-time applications.

When the information about the 3-D scene is presented, then the specific algorithms for 3-D objects shape extraction and object recognition might be applied [2].

This paper deals with the design of stereo vision system for 3-D perception, starts with the brief description of stereo geometry and errors in depth estimation, follows by the systems' hardware design, and then 3-D measurements.

Basis of stereo vision

A number of textbooks give adequate descriptions of the theory and practice of stereo reconstruction and 3-D world perception [3, 4]. Any stereo system must solve two basic problems: correspondence (i.e. which pixel in the left image corresponds to which pixel in the right image?) and reconstruction (what is the real world location of the corresponding point of interest?).

Stereo application takes stereo pairs of 2-D images as input and produces the reconstructed 3-D data by finding the corresponding points. The corresponding problem requires a specific searching and matching technique, the robustness of witch determines quality and precision of reconstructed 3-D data [1].

Most stereo reconstruction methods are based on the usage of the pinhole camera model and parallel geometry [3, 4]. Therefore, given two points coordinates (one from the left and the second from the right images) identified by any stereo matching process, the depth is estimated from the points disparity, or the separation of the two points in image pixels coordinates [3, 4]. Given pixel coordinate (x_L,y_L) in the left and (x_R,y_R) right images, then the 3D world coordinates (X,Y,Z) are computed as:

$$X = \frac{x_L b}{d},\ Y = \frac{y_L b}{d},\ Z = \frac{fb}{d}, \quad (1)$$

where disparity $d = x_L - x_R$ [in pixels], b – the parallax or inter-ocular separation [in mm] of cameras, f – the focal length of camera lenses, measured in pixels (see Fig. 1).

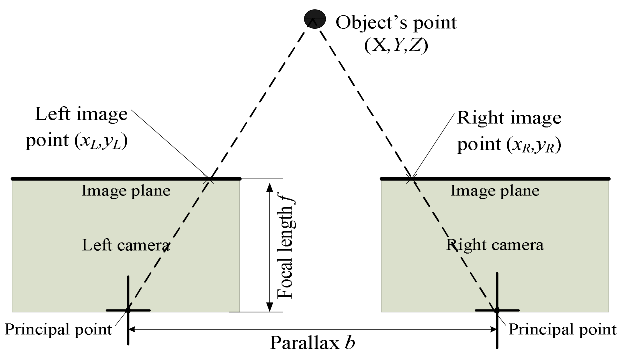

Fig. 1. Illustration of the stereo geometry

The accuracy of 3-D data reconstruction depends on the accuracy of the disparities, stereo system calibration, images rectification and overall stereo system construction.

Error in depth estimation

For the depth error definition, it is necessary to assume that the stereo system is calibrated and the given images are rectified. Therefore the correspondence problem reduces to the single line search problem and the scene depth is estimated according to (1). As it is shown in [5, 6] the error in depth estimation *Z* is sensitive to errors in disparity *d:*

$$\Delta Z = \frac{Z^2 \Delta d}{fb}, \qquad (2)$$

where Δd comes from the estimate errors in x_L and x_R. It means that the error in range estimation is proportional to the square of the range and independent of the object location along x-axis. Therefore the depth perception is less reliable for the distant objects and this error might be reduced by increasing the parallax of the stereo system or by active zooming (increasing the focal length of the lenses).

Systems' hardware setup

The designed system consist of two digital cameras (type FFMV-03MTC, mfg. Point Grey, Canada), mounted in parallel. The reason to choose these cameras was their technical parameters:

- small-sized (25 x 40 mm);
- IEEE-1394a 400Mbps digital interface ("firewire");
- Video resolution up to 752(640) x 480 pixels @60/30/15/7.5fps;
- colour or b/w picture;
- possibility for operator to change essential parameters: brightness, exposure, shutter interval, gain, etc.
- internal (over IEEE-1394a) or external (hardware) trigger for shutter synchronisation.

The most important requirement for the system is to insure that each frame of both cameras is captured at the same time with same brightness, exposure, shutter time and gain parameters. Otherwise, if the cameras are mounted on a moving object (robot) and there is no synchronisation between them, the determined depth map will be wrong. If the picture is taken with one camera and after a moment the second camera takes a picture, they will not correspond and the computed depth map will be incorrect. The simplest way to avoid this phenomenon is to ensure that the capturing is done synchronously at the same moment. Fig. 2 shows how the cameras are synchronized by rising front of triggering signal to capture a frame.

Modern cameras provide multiple ways of frames synchronisation, like through an external triggering signal or by using software drivers to issue synchronized triggers to multiple cameras over IEEE-1394 bus. Manufacturer of FFMV cameras also offer drivers for essential parameters setup in manual or automatic mode. In our system setup, one camera is used for automatic parameters setup for brightness, exposure and shutter times; and each frame the same setup parameters are broadcasted to the second camera for better correspondence of images. Fig. 3 shows the implemented cameras synchronization and essential parameters cloning over the IEEE-1394 bus. The advantage is that this solution has a minimal phase shift and meets the requirements for the stereo systems to capture the frames at the same moment with identical cameras setup parameters.

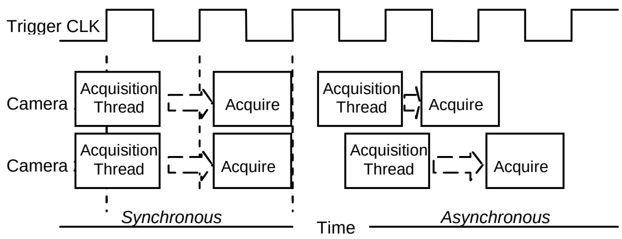

Fig. 2. Synchronous/asynchronous acquisition of images

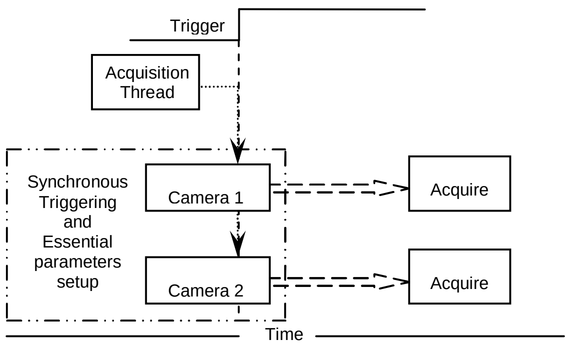

Fig. 3. Implemented triggering technique

Stereo Cameras Calibration

Several multiple-camera calibration methods exist that may simultaneously solve cameras' calibration for intrinsic and extrinsic parameters [7, 8]. In our work we have applied Jean-Yves Bouguet's well known and widely-used Camera Calibration Toolbox for MATLAB to calibrate the stereo systems' parameters. The toolbox uses a pinhole camera model with nonlinear radial and tangential distortion compensation which was inspired by Heikkila [8].

Image rectification

Rectification methods are very well known and have been widely studied for years. The aim of these techniques is to adjust captured images to simplify stereo correspondence problem. When the image is captured with an optical device the resulting image differ from the real world geometry due to the optics. There are basically two factors that have to be adjusted in stereo applications: the image distortion and the image epipolar geometry (see Fig. 4). This process is called rectification. When the stereo pair of images is rectified, then the stereo correspondence problem simplifies and reduces from 2D search of order N^2 (for an image of $N \times N$ pixels) to a 1D search of order N for each corresponding pair of points on the same epipolar line (i.e. on the same height of image matrixes). The Camera Calibration Toolbox for MATLAB offers stereo image rectification procedure after the stereo system calibration.

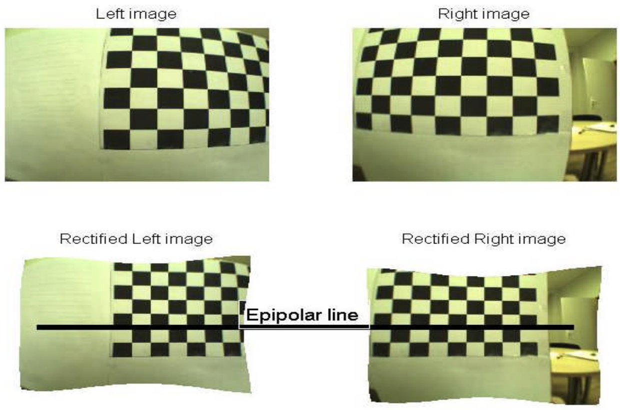

Fig. 4. Pair of images: distorted and after rectification

Parallax of cameras

For parallel stereo cameras' the epipolar lines lie parallel to the image rasters, thereof from (1) the estimated depth, *Z*, is inversely proportional to the cameras separation (parallax), *b*, therefore the errors in the depth estimation are inversely proportional to *b*. Large parallax gives smaller depth errors, but unfortunately also less scene overlap for depth computation [5]. This is the trade-off between stereo system errors and visible stereo scene. In our case we chose the cameras parallax (see Fig. 5) close to 130 mm.

Lenses of cameras

FireFly MV03MTC cameras from factory are supplied with a 36° angle of view and 6 mm focal length lenses, which is obvious not enough for close environment perception. Therefore the original ones have been replaced by lenses from Edmund optics stock No. # 57-681 and increased the field of view up to 90° (see Fig. 5).

Experimental investigation

All experiments were performed with calibrated stereo system and rectified images at resolution 640 x 480 pixels. To evaluate perception capabilities of the designed system, the two types of lenses have been applied. As it was mentioned earlier, the lenses with focal length at 6 mm and 2.5 mm are used; unfortunately, this length should be converted into pixels. When cameras were fixed on the metal support at ~130 mm distance one from the other; the performed calibration has indicated that the exact parallax *b* is 128 mm.

To obtain the focal length of system in pixels, for the depth error estimation, the target point was manually measured at distances from 100 mm to 2000 mm. Fig. 6 shows the measured data and extrapolated curves according to equation (1). The measured data (distance *Z* and disparity *d*) were fitted by the least square regression method. As was mentioned, the parallax *b* of the system was found during the calibration procedure and was kept as *b*=128 mm. Therefore the focal lengths for 6 mm and 2.5 mm were found to be 877.68 pixels and 408.26 pixels correspondingly. The lines in Fig. 6 show the functional dependency with obtained parameter *f* in (1).

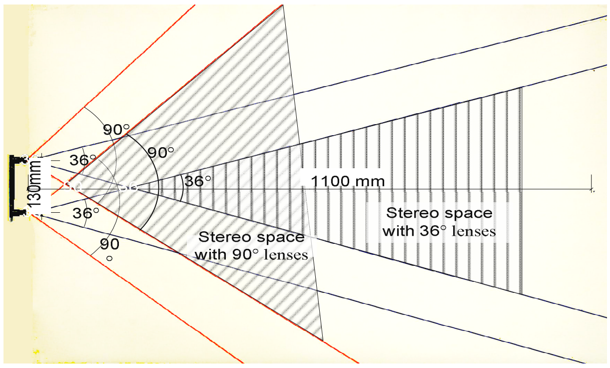

Fig. 5. Stereo couple with view angles 36° and 90° at 130 mm of stereo couple parallax

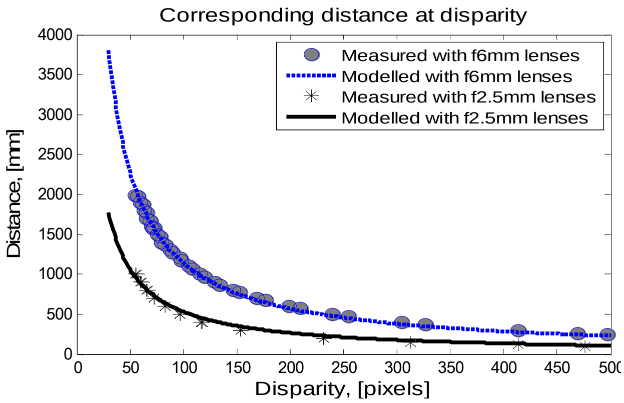

Fig. 6. Measured and extrapolated distances at disparities with 6mm lens (doted line) and 2.5mm lens (solid line)

With the received parameters of the designed system, the depth estimation error was calculated according to eq. (2) by assuming that $\Delta d = 1$ (i.e. point correspondence is mismatched only by one pixel). The functional dependency of measured error vs. measured range is modelled and shown in Fig. 7. As it is seen from Fig. 6 and Fig. 7, the system with short focal length is most suitable to perceive objects with low error at close distance, vice versa the designed system with 6 mm focal length lenses is more precise in measuring distant objects; but Fig. 5 shows that this precision at big range is obtained by small visible angle or small field view of the stereo system.

Conclusions

In this paper we presented the design of 3-D perception system based on the binocular stereovision. In order to provide good quality images for the depth estimation, the software triggering and cameras setup parameters cloning technique have been implemented. The system has been calibrated and images rectified according to the techniques described.

Manually collected data of object distance and disparity allowed to calculate the focal distance of used lens, in pixels. Afterwards, the functional dependency of measuring error and measured distance were evaluated.

The experiment allows concluding that the selection of lenses and parallax of stereo system is application dependent. We have shown that with large field of view lenses (f=2.5) it is possible to obtain better perception of close objects, and, vice versa, lenses with longer focal length are more suitable for precise long range perception

The presented stereo system is to be applied on robot for the autonomous navigation and obstacles avoidance purposes.

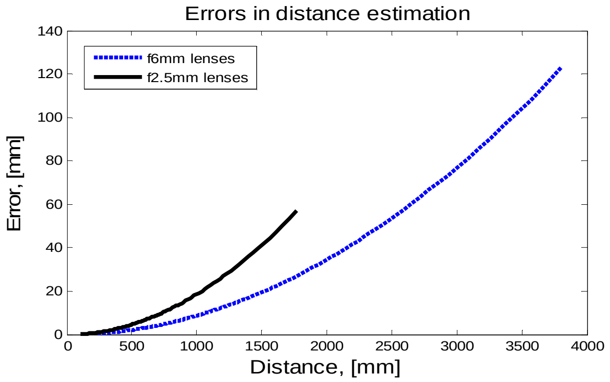

Fig. 7. Error of distance measurement distribution according to the measured range

Acknowledgments

This work has been support by the Lithuanian state science and studies foundation.

References

1. **Scharstein D., Szeliski R.** A taxonomy and evaluation of dense two-frame stereo correspondence algorithms // Int. J. of Computer Vision.- 2002.- 47(1/2/3).- P.7-42.
2. **Lipnickas A., Raudonis V., Jankauskas K.** Extraction of constructive primitives based on clustering of normals and curvature // Electrical and Control Technologies – 2008, selected papers of the 3rd international conference.- Kaunas, 2008.- P.51-56.
3. **Serafinavičius P., Sajauskas S., Daunys G.,** Evaluation of Hand Pointing System Based on 3-D Computer Vision // Electronics and Electrical Engineering. –Kaunas: Technologija, 2008. – No. 8(88). P. 95-98.
4. **Faugeras O.** Three-Dimensional Computer Vision - A Geometric Viewpoint. MIT Press, Cambridge, MA, 1993.
5. **Buonaccorsi G. A., Lacey A. J., Thacker N. A.** Characterisation of a Stereo Matching and Object Location System. http://www.tina-vision.net/docs/ memos/2003-006.pdf. Accessed 2008-11-24.
6. **Sridhar B., Suorsa R.** Comparison of motion and stereo methods in passive ranging systems // IEEE Trans. on Aerospace and Electronic Systems.- 1991.- Vol. 27, Issue: 4.- P.741-746.
7. **Serafinavičius P.** Investigation of Technical Equipments in Computer Stereo Vision: Camera Calibration Techniques // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005. – No. 3(59). – P. 24–27.
8. **Heikkila J., Silven O.** A four-step camera calibration procedure with implicit image correction // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).- 1997.- P.1106–1112.

Received 2008 11 25

A. Lipnickas, A. Knyš. A Stereovision System for 3-D Perception // Electronics and Electrical Engineering. – Kaunas: Technologija, 2009. – No. 3(91). – P. 99–102.

The results of investigation on a stereovision system for 3-D perception are presented. A 3-D space data can be reconstructed from a pair of 2-D images, the basis of stereo vision and computation equations are given. Also error of depth estimation and ways to minimise it are surveyed. The systems' hardware design and the ways to synchronise cameras to avoid incorrect map reconstruction due to asynchronous frame capturing are presented. The results of experimental investigation: measured and extrapolated distances at certain disparities using 6 mm and 2.5 mm lenses, error of distance measurement distribution are presented. The results show that usage of longer focus lenses is more useful at long range perception. Also the results show that the system surveyed is suitable for autonomous robot navigation system. Ill. 7, bibl. 8 (in English; summaries in English, Russian and Lithuanian).

А. Липницкас, А. Кныш. Стереосистема для регистрации трёхмерного пространства // Электроника и электротехника. – Каунас: Технология, 2009. – № 3(91). – С. 99–102.

Представлены результаты анализа стерео системы для регистрации трёхмерного пространства. Как известно, данные о трёхмерном пространстве могут быть восстановлены из пары двухмерных картинок. Для этого представлены основы трёхмерного видения, а также методика вычислений. Анализируются причины возникновения погрешностей определения расстояний и методы их минимизации. Представлена структура системы, использованная аппаратура, возможные методы синхронизации камер с целью избежания неправильного построения карты из-за асинхронной работы указанных камер. Представлены экспериментные результаты, показывающие погрешность при использовании камер с расстоянием фокуса 6 мм и 2.5 мм. Полученные результаты позволяют делать выводы о том, что линзы с большим расстоянием фокуса более подходят для измерения больших расстояний, а сама система может быть использована для реализации автономной навигационной системы робота. Ил. 7, библ. 8 (на английском языке; рефераты на английском, русском и литовском яз.).

A. Lipnickas, A. Knyš. Stereosistema trimatei erdvei registruoti // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2009. – Nr. 3(91). – P. 99–102.

Pateikiami stereosistemos, skirtos trimatei erdvei registruoti, tyrimo rezultatai. Trimatės erdvės duomenys gali būti atkurti iš dviejų dvimačių paveikslų, todėl pateikiami stereo regos pagrindai ir skaičiavimo metodika. Taip pat analizuojami nuotolio įvertinimo paklaidų šaltiniai bei šių paklaidų minimizavimo būdai. Pateikiama sistemos struktūra ir naudojama aparatūra, galimi naudoti vaizdo kamerų sinchronizavimo būdai, siekiant išvengti erdvės žemėlapio sudarymo klaidų dėl nesinchronizuoto vaizdų fiksavimo ir perdavimo. Pateikti eksperimentinio tyrimo rezultatai, atspindintys išmatuotų ir ekstrapoliuotų atstumų paklaidas, esant fiksuotiems vaizdų poslinkiams, naudojant 6 mm ir 2,5 mm fokuso atstumo lęšius. Gauti rezultatai leidžia teigti, kad tolimesnį fokuso atstumo turintį lęšį geriau naudoti tolimesniam nuotoliui iki objekto nustatyti, o nagrinėjamoji sistema tinka autonominei roboto navigacijos sistemai sukurti. Il. 7, bibl. 8 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).