Przemysław SZEWCZYK
TECHNICAL UNIVERSITY OF LODZ, INSTITUTE OF AUTOMATIC CONTROL,
ul. Stefanowskiego 18/22, 90-924 Łódź, Poland

Stereovision system with active vergence and gaze control mechanism

MSEE Przemysław SZEWCZYK

Przemysław Szewczyk graduated from the Faculty of Electrical, Electronic, Computer and Control Engineering of the Technical University of Lodz. Currently he is a PhD student at the Institute of Automatic Control Technical University of Lodz. He conducts research work on the usage of stereovision systems with variable geometry of cameras to detect and track objects of robotic scene. He has been employed in the company Teleca Poland Sp. z o. o. as a Software Engineer since 2008.

e-mail: przemyslaw.szewczyk@o2.pl

Abstract

The paper describes a prototypical stereovision system with active vergence and gaze control mechanism. The model and mechanical construction enable the accurate positioning of cameras. Increase in the measurement resolution of the binocular vergence angle in the selected range of depth is specified. The vergence control mechanism uses a kinematic chain with 3-DOF joints. In order to obtain a linear characteristic of the measurement system, there was performed the optimisation process of the kinematic chain parameters.

Keywords: stereovision system, vergence and gaze control mechanism, binocular vision.

System stereowizyjny z aktywnym mechanizmem zmiany kąta widzenia oraz zbieżności osi optycznych kamer

Streszczenie

W artykule przedstawiony jest opis prototypowego układu stereowizyjnego z aktywnym systemem zmiany kąta widzenia oraz zbieżności osi optycznych kamer. Zmiana kąta zbieżności osi optycznych kamer wraz z obrotem głowicy umożliwia koncentrację systemu na pewnym, dowolnie wybranym obiekcie w rzeczywistej przestrzeni trójwymiarowej. Prowadzi do nałożenia pól widzenia obu kamer i niewielkich wartości dysparycji bliskich zeru dla wybranego, śledzonego obiektu umożliwiając szybką segmentację obrazu. Zaprezentowano model oraz konstrukcję mechaniczną umożliwiającą dokładne pozycjonowanie kamer, a także zwiększenie rozdzielczości pomiarowej kąta zbieżności w wybranym przedziale mierzonej głębi. W mechanizmie kontroli zbieżności osi optycznych kamer zastosowano łańcuch kinematyczny wykorzystujący złącza o trzech stopniach swobody, a następnie dokonano optymalizacji parametrów tego łańcucha tak, aby zapewnić możliwie linową charakterystykę przetwarzania układu pomiarowego.

Słowa kluczowe: system stereowizyjny, widzenie dwuoczne.

1. Introduction

The most natural method for measuring depth and determining 3-dimensional structure of a scene is computer stereovision, which uses two or more digital images taken from different viewpoints [1, 2]. By comparing information about a scene from several camera perspectives in the scene, limited 3D information can be extracted by examination of the relative perspectives. This process is similar to the biological process – stereopsis, which is one aspect of a more general problem connected with visual perception and binocular vision.

Binocular systems, unlike canonical (standard) stereovision systems with coplanar, row-aligned imaging planes and parallel optical axes [3], have additional utility – vergence, the process of adjusting the angle between optical axes of the eyes (or cameras) so that both eyes are directed at the same world point (fixation point). In the human vision, the ability to change the vergence angle allows registering an object on the fovea of each eye – the central, high-resolution region of the retina, so the greatest possible amount of information about the object can be extracted [4]. The fovea comprises less than 1% of retinal size but takes up over 50% of the visual cortex in the brain.

The vergence is mainly obvious for foveate systems such as the human visual system and follows from the non-uniform spatial resolution of the photoreceptors array in the retinas. However, most robot vision systems use digital cameras with CMOS or CCD image sensors, which do not have foveas, so the most obvious motivation for vergence control in humans does not apply to them [5]. Nevertheless, vergence has many advantages and it is a useful strategy even for a non-foveate binocular vision system.

In many cases, giving a vision system the ability to change camera geometric configuration simplifies scene analysis and completes visual information limited by occlusion or even light reflections.

Fixation of an object of interest puts points on the object near the optical axis in both cameras that permits use of the mathematical simplification of the camera model by replacing perspective projections with ortography. This significantly reduces computational complexity of the geometrical transformation and makes analysis easier [6].

Gaze and vergence movement allows tracking an object of interest in real 3-dimensional scene by both cameras and setting fixation point close to that object, which in turn increases the binocular visual field. By definition, the fixation point has a stereoscopic disparity of zero, and points nearby tend to have small disparities. This feature permits use of fast and amenable to hardware implementation stereo algorithms that accept only a limited range of disparities and can be applied to disparity-based segmentation process. As a consequence, disparity may be used to filter objects that are not currently of interest out of the scene (ignore features at non-zero disparities) [5, 7].

The all described above features of gaze and vergence control in stereovision system facilitate stereo fusion and 3D reconstruction in comparison with a canonical system [8, 9], but an appropriate positioning mechanism is required. The model and mechanical construction enabling the accurate positioning of cameras, as well as increasing the measuring resolution of the vergence angle of a binocular system are specified in the next sections.

2. Model of vergence control mechanism

The function of a camera positioning mechanism is to control the distance from the cameras to the fixation point along some specified gaze direction. In many cases the motivation for vergence control is to set and keep the fixation point near an object of interest. Since the target vergence angle is directly related to the distance between the cameras system and target object, any sensory cue to depth or depth changes may be useful to the camera positioning system. Obviously, the most natural cue is binocular disparity, but also the focus error can be used in the case of using zoom lenses in the visual system, or any other cues connected with depth information, which can be extracted from object motion, texture, or shading analysis [5].

Vergence control is a part of the larger problem of visual object tracking, which involves also gaze direction and focal distance control. Ability to change the gaze direction and adjust the angle between optical axes of the cameras is related to the degrees of freedom offered by the positioning mechanism. Evidently, it is possible to build the hardware, which allows each camera to rotate about the axis perpendicular to the camera optical axis independently, but in this type of a system vergence and gaze are dependent. The vergence angle is equal to the difference between two camera angles, and the gaze angle is a function of both camera angles [5, 7]. Thus, in this paper the positioning mechanism that allows independent control of vergence and gaze is presented.

Fig. 1 shows a prototypical model of the vergence control mechanism in the form of a kinematic chain which consists of one active joint, two 1-DOF joints and four ball joints (3-DOF). The use of such a kinematic chain allows a simultaneous change in the vergence angle with identical $\Delta\alpha$ for both cameras (Fig. 1).

According to Fig. 1, the distance L as a function of the angle α can be determined from the equation:

$$L = \frac{a}{2}\operatorname{ctg}(\alpha) \quad (1)$$

and a suitable changes of this value can be determined from the differential equation:

$$dL = -\frac{a}{2}\frac{1}{\sin^2(\alpha)}d\alpha \quad (2)$$

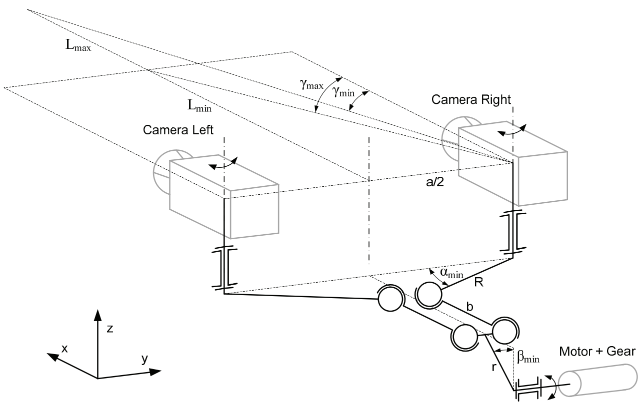

Fig. 1. Model of vergence control mechanism
Rys. 1. Model mechanizmu sterowania kątem zbieżności osi optycznych kamer

Obviously, in all measurement systems the differential linearity of a transducer is desirable – constant relation between the change at the output and input. According to equation (1), the relation between distance L and convergence angle α is extremely non-linear. However, it is still possible to change the conversion characteristic by using the kinematic chain (Fig. 1) and optimising its parameters, so that the function $L(\beta)$ can be as linear as possible (ΔL is small and constant over the whole range of the angle β).

The kinematic chain parameters are denoted as follows:

- R, b, r - length of subsequent links,
- $\alpha_{\min}, \beta_{\min}$ - extreme values of joint variables,
- $\hat{\alpha}, \hat{\beta}$ - ranges of joint variables.

In view of the fact that the system shown in Fig. 1 is symmetric, only the selected half of it can be considered in the further analysis (Fig. 2).

Assuming that the coordinates of the point s_0 lie at the origin of the basic Cartesian coordinate frame and that there are known coordinates of s_3, it is possible to determine the coordinates of other points s_1, s_2, and join them as follows:

$$\begin{cases} x_1 = r\sin(\beta_{min} + \beta) \\ y_1 = 0 \\ z_1 = r\cos(\beta_{min} + \beta) \end{cases}$$

$$\begin{cases} x_2 = x_3 - R\sin(\alpha_{min} + \alpha) \\ y_2 = y_3 - R\cos(\alpha_{min} + \alpha) \\ z_2 = z_3 \end{cases}$$

$$[x_3 - R\sin(\alpha_{\min} + \alpha) - r\sin(\beta_{\min} + \beta)]^2 + [y_3 - R\cos(\alpha_{\min} + \alpha)]^2 + [z_3 - r\cos(\beta_{\min} + \beta)]^2 = b^2 \quad (3)$$

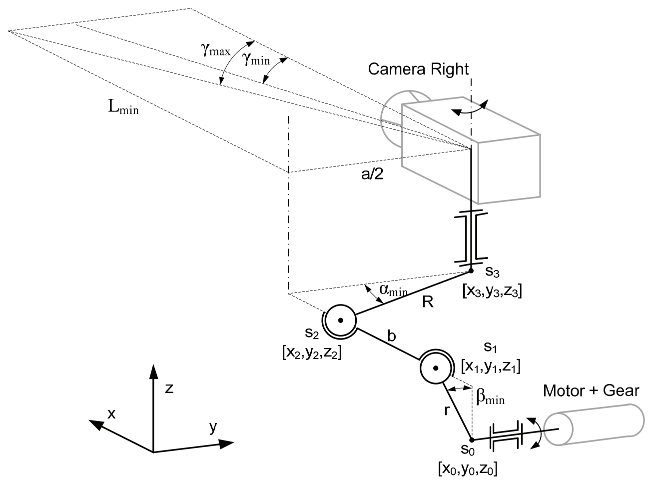

Fig. 2. Model of vergence control mechanism.
Rys. 2. Model mechanizmu sterowania kątem zbieżności osi optycznych kamer

Taking into account equations (3) and (1), it is possible to find a solution:

$$L = f(R, b, r, \alpha_{\min}, \beta_{\min}, \beta) \quad (4)$$

L is a nonlinear function of β. Using an optimization procedure and the constraints resulting from the mechanical design limitations, it is possible to determine the parameters of the kinematic chain R, b, r that allows treating the function as very close to a linear one within the regions of $\alpha_{\min}$, $\beta_{\min}$.

3. Mechanical construction of vergence and gaze control mechanism

The mechanical construction of the active stereovision system was made of aluminium and plastic materials to ensure light weight and good dynamic properties of the system (Fig. 5). The vergence control mechanism consists of rotating aluminium handles with camera remote heads, the kinematic chain using ball joints, and the integrated miniature drive system (Faulhaber) containing a 24-step stepper motor and precision zero backlash spur gearbox with the reduction ratio 1670:1. The drive of the same parameters was also used in the gaze control mechanism. The control system of the motor is based on Trinamic Motion Control integrated circuits, which allows parallel operation of motors in microstep mode (up to 64 microstep) and sensorless motor stall detection.

Due to the complexity of equation (3), Matlab Optimization Toolbox enabling optimization of a multivariable nonlinear function with constraints was used to determine the parameters of the kinematic chain.

Assuming the following parameters of the mechanism: (Fig. 2): x_3 = 85 mm, $y_3 = a/2$, z_3 = 20 mm, where the value a corresponds to the length of the stereo baseline and is equal to 200 mm, the kinematic chain parameters are constrained as follows:

$$\begin{cases} d_{\min} \le r \le \frac{a}{2} \\ d_{\min} \le R \le \frac{a}{2} \\ d_{\min} \le b \le a \\ -\frac{\pi}{4} \le \alpha_{\min} \le \frac{\pi}{4}, \\ -\frac{\pi}{4} \le \beta_{\min} \le \frac{\pi}{4} \\ 0 \le \alpha \le \hat{\alpha} \end{cases} \tag{5}$$

where a range of angle $\hat{a}$ is equal to the required range of the vergence angle and depends on the extreme values of depth which the tracked object of interest may have (Fig. 1):

$$\hat{\alpha} = \Delta\gamma = arcctg\frac{2L_{\min}}{a} - arcctg\frac{2L_{\max}}{a}. \tag{6}$$

For the required distance $L_{\min}$ = 400 mm ($L_{\min}$ is related to the construction of the lens used, and possible to achieve depth of field) and $L_{\max}$ = 10 m, the vergence angle $\hat{a}$ range is approximately 13.5 deg.

The value $d_{\min}$ depends on physical dimensions of the ball joints used. The value 50 mm was assumed for further analysis.

As a result of optimization, the configuration of the kinematic chain was obtained as shown in Fig. 3(b). Furthermore, Fig. 3 shows the configuration before starting the optimization process (a) and the configuration acquired by decreasing the first module length constraint from the maximum value $a/2$ to a (b).

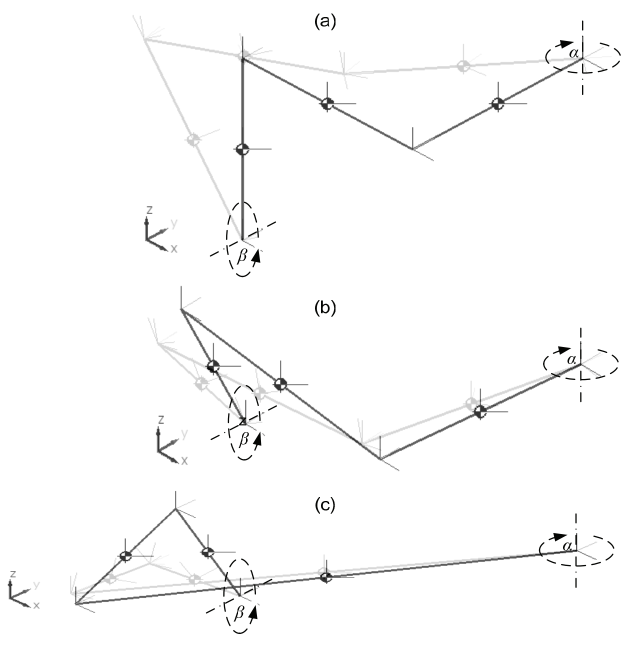

Fig. 3. Model of the kinematic chain of the vergence control mechanism (dark shape - a basic configuration, the maximum depth($L_{\max}$), light shape - a configuration, corresponding to a minimum depth ($L_{\min}$)). a) - configuration before the optimization process, b) - configuration after optimization with constraints, c) - configuration after optimization with constraints (the first module length constraint was reduced)

Rys. 3. Model łańcucha kinematycznego mechanizmu zmiany zbieżności osi optycznych kamer (kolor ciemny – konfiguracja podstawowa, kolor jasny – konfiguracja odpowiadająca minimalnej głębi ($L_{\min}$)). a) – konfiguracja początkowa przed procesem optymalizacji, b) – konfiguracja po optymalizacji z ograniczeniami na parametry układu, c) – konfiguracja po optymalizacji z ograniczeniami na parametry układu (zmniejszone ograniczenie na długość pierwszego modułu)

Each model of the kinematic chain is shown in the basic configuration, corresponding to the maximum depth ($L_{\max}$), and the configuration for the minimum depth ($L_{\min}$).

Assuming that the theoretical angular stepper motor resolution including the gear reduction ratio is $r_{\deg} = 360/(1670\cdot24) \approx 0.009$ deg (not including microsteps), it is possible to calculate the relative error of depth estimate $|\Delta L|/L$, based on equations (3) and (2). Tab. 1 presents a comparison between kinematic chain configurations (b, c) and configuration with a direct drive ($\beta=\alpha$). The table contains the extreme values of depth estimation relative error $|\Delta L|/L$, the minimum distance L_{min}, for which the system can adjust the convergence angle, and the range of convergence angle $\Delta\alpha$ (depth changes from $L_{\max}$ to $L_{\min}$).

Fig. 4 shows the depth estimation L and absolute measurement error ΔL as functions of the normalized angle $\beta/\Delta\beta$.

As shown in Tab. 1 and Fig. 4, the depth measurement resolution can be significantly increased in the selected range by using an appropriate configuration of the kinematic chain. However, increase in the resolution of depth measurement is encumbered with the range of vergence angle reduction, which in turn, reduces the range of depth possible to measure by the system. Nevertheless, it is known that the depth measuring of a scene based on disparity is accurate for objects of interest that are very close the visual system. Therefore, despite the fact that the convergence angle is limited to the value L_{min} greater than the required, it is still possible to accurately measure the depth below this value based only on the disparity value.

Tab. 1. Comparison between kinematic chain configurations (b, c) and configuration with direct drive ($\beta=\alpha$)

Tab. 1. Porównanie konfiguracji łańcuchów kinematycznych (b, c) z układem z napędem bezpośrednim ($\beta=\alpha$)

	Direct drive	**Configuration (b)**	**Configuration (c)**
$\min\frac{\lvert\Delta L\rvert}{L}$	$6.66\times10^{-4}(L_{\min})$	$5.09\times10^{-4}(1.68\text{ m})$	$1.22\times10^{-6}(L_{\max})$
$\max\frac{\lvert\Delta L\rvert}{L}$	$0.015(L_{\max})$	$0.0035(L_{\max})$	$0.0012(7.7m)$
$L_{\min}$	0,4 m	1,68 m	1.84 m
$\Delta\alpha$	$\hat{\alpha}$	2.8 deg	2.53 deg

Another important aspect of the mechanical construction is the location of the axes of rotation with respect to the nodal points of the cameras. The nodal point is one of the two points in an optical system, located in such way that a light ray directed through the first point will leave the system through the second one, parallel to its original direction. The location of nodal points relative to the axes of rotation is not negligible for the purpose of gaze and vergence control, and it is desirable to mount the cameras so that their axes of rotation pass through the nodal points (the front nodal points for an ideal thick lens system) [5]. Rotation of the optical system around the nodal point results in a change in the projected image that depends only on rotation without a translation component. Designing a system, that rotates the cameras about their nodal points, is a difficult task, because the nodal points are often far from the cameras centres of gravity and they also depend on lens constructions and change with changes in optical parameters (adjusting focus or zoom). The errors resulting from inaccurate mounting the camera can be ignored only if the targets are close to the camera system. Therefore, the presented stereovision system uses lightweight cameras with separate heads containing CCD sensors and microlenses, as well as mounting brackets, which allow fine adjustment of head positions in 3 directions.

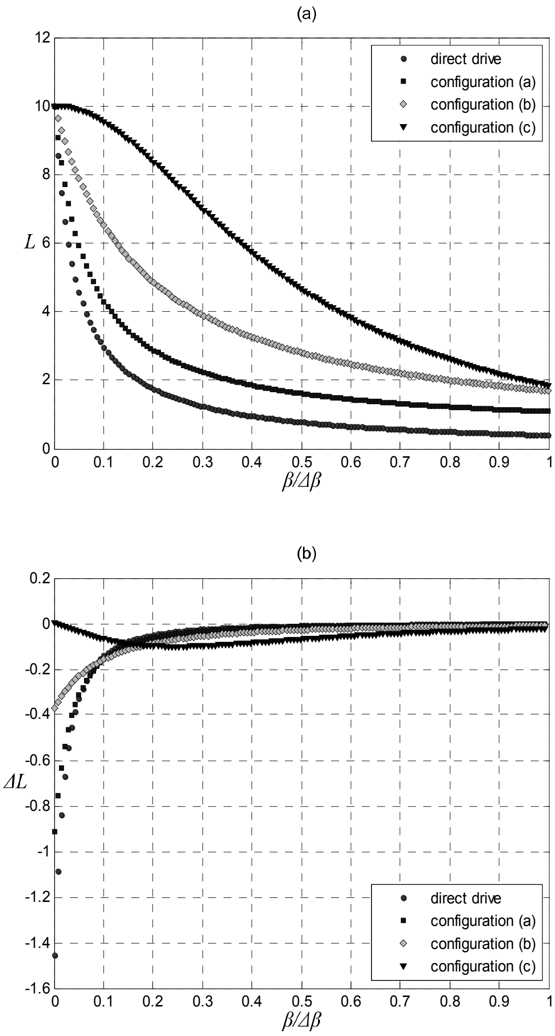

Fig. 4. Comparison between kinematic chain configurations (a, b, c) and configuration with direct drive (the angular resolution $r_{deg}=0.1deg$ was assumed to increase the clarity of charts). a) the depth estimation L as function of the normalized angle $\beta/\Delta\beta$, b) the absolute measurement error ΔL as function of $\beta/\Delta\beta$

Rys. 4. Porównanie konfiguracji łańcuchów kinematycznych (a, b, c) z układem z napędem bezpośrednim (ze względu na przejrzystość wykresów założono, że rozdzielczość kątowa wynosi $r_{deg}=0.1deg$). a) zależność pomiaru głębi L w funkcji kąta $\beta/\Delta\beta$, b) – bezwzględny błąd pomiaru ΔL w funkcji kąta $\beta/\Delta\beta$

Choice of the rotation axis that passes through the center of the stereo baseline is quite obvious for the gaze control mechanism. Similar to the previous vergence control system, the gaze control mechanism uses ball joints, but in this case, they form only a simple chain drive.

4. Conclusions

The kinematic chain configuration shown in Fig. 3 (b) was used to build the active stereovision system (Fig. 5). The length of the first module in the model of a kinematic chain (c) has a value close to the length of the stereo baseline, which in turn causes kinematic chain components crossing, after applying this model to both cameras. It is possible to modify the kinematic chain to eliminate overlapping workspaces and to avoid collisions, but it increases the mechanism complexity, provides additional backlash and can be the reason for stiffness reduction. In view of the above limitations, the kinematic chain configuration shown in Fig. 3(c) was excluded from further consideration.

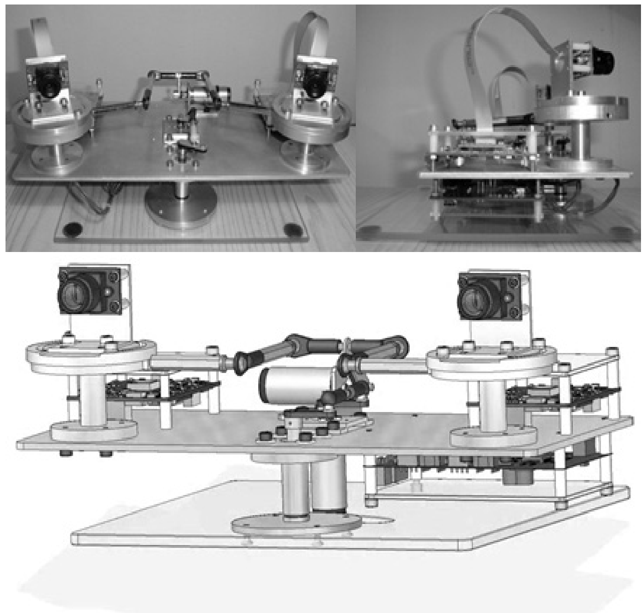

Fig. 5. Stereovision system with active vergence and gaze control mechanism

Rys. 5. System stereowizyjny z aktywnym mechanizmem pozycjonowania kamer

The construction of the active stereovision system allowed verifying the proposed solution of the gaze and vergence control, which is an important element of the visual tracking systems. Currently, the work on a visual feedback control of the active vision system is being continued.

Dr Krzysztof Mianowski for valuable discussion about zero-backlash transmission systems.

The project was done under supervision of Prof. Edward Jezierski.

5. References

[1] Brown M. Z., Burschka D., Hager G. D.: Advances in computational stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(8). Aug 2003.

[2] Trucco E., Verri A.: Introductory Techniques for 3-D Computer Vision. Prentice Hall, 1998.

[3] Cyganek B.: Komputerowe przetwarzanie obrazów trójwymiarowych. Akademicka Oficyna Wydawnicza EXIT, 2002.

[4] Bochenek A., Reicher M.: Anatomia człowieka, Vol. V. Wydawnictwo Lekarskie PZWL, Warsaw 2004.

[5] Olson T., Coombs D.: Real-time vergence control for binocular robots. International Journal of Computer Vision, Vol. 7, 1991.

[6] Ballard A., Ozcandarli D.H.: Eye fixation and early vision: Kinetic depth. Second Int. Conference on Computer Vision. Dec 1988.

[7] Gu Y., Sato M., Zhang X.: Active depth estimation with gaze and vergence control using gabor filters. 13th Int. Conference on Pattern Recognition. Aug 1996.

[8] Bradski G., Kaehler A.: Learning OpenCV - Computer Vision with the OpenCV Library. O'Reilly, 2008.

[9] Hansen M., Sommer G.: Active depth estimation with gaze and vergence control using gabor filters. 13th International Conference on Pattern Recognition. Aug 1996.

otrzymano / received: 01.04.2011
przyjęto do druku / accepted: 03.10.2011 *artykuł recenzowany*