Jakub KORTA, Piotr KOHUT, Tadeusz UHL
AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF MECHANICAL ENGINEERING AND ROBOTICS,
DEPARTMENT OF ROBOTICS AND MECHATRONICS, AL. A. MICKIEWICZA 30, 30-059 KRAKOW

OpenCV based vision system for industrial robot-based assembly station: calibration and testing

M.Sc. Jakub KORTA

The author graduated from AGH University of Science and Technology in 2010, obtaining M.Sc. (Eng) title. Since then he has been continuing scientific work towards Ph.D degree. His interests in the field of computer vision systems include motion tracking algorithms, stereovision and applications of camera based production control systems working in conjunction with industrial manipulators.

e-mail: korta@agh.edu.pl

Ph.D. Piotr KOHUT

Piotr Kohut is an adjunct professor at the Department of Robotics and Mechatronics of AGH University of Science and Technology in Krakow. His scientific interests focus on mechatronics, vision systems, methods of image processing and analysis as well as 3D measurement techniques.

e-mail: pko@agh.edu.pl

Prof. Tadeusz UHL

Prof. Tadeusz Uhl is the head of the Department of Robotics and Mechatronics of AGH University of Science and Technology in Krakow. His scientific interests are connected with construction diagnostics and Structural Health Monitoring, dynamics, experimental modal analysis and control systems used in mechatronics. He participated in many national and international projects finalized by valuable monographs and papers.

e-mail: tuhl@agh.edu.pl

Abstract

The paper presents the vision system implemented on a laboratory test station, simulating industrial robot-operated electronic circuits assembly line. Verification of the developed algorithm, which is based on the Open Source Computer Vision library, has also been supplemented by measurements of the system speed and uncertainties. By that means, it has been shown that expensive off-the-shelf systems can be replaced by the elaborated one, without compromising production cycle time and repeatability.

Keywords: vision system, OpenCV, robot, assembly, image processing.

System wizyjny zrobotyzowanego stanowiska produkcyjnego oparty na bibliotekach OpenCV: kalibracja i testy

Streszczenie

W artykule zaprezentowano system wizyjny zaimplementowany na laboratoryjnym stanowisku pomiarowym, symulującym działanie przemysłowej linii produkcyjnej wytwarzającej układy scalone. Testowane algorytmy powstały w oparciu o otwarte biblioteki Open Source Computer Vision (OpenCV) co sprawia, że ich użycie pozwala na obniżenie kosztów wytwarzania jakie ponosi przedsiębiorstwo. Jest to szczególnie istotne w przypadku produkcji masowej, w której nawet najmniejsza zmiana czasu pojedynczego cyklu pracy przekłada się na znaczącą zmianę ogólnej opłacalności produkcji. Wydajność opracowanego systemu została potwierdzona poprzez testy wykonane w sprzężeniu z komercyjnym odpowiednikiem, czym udowodniono, że kosztowne rozwiązania mogą z powodzeniem zostać zastąpione przez zbudowany system. Cel ten można osiągnąć bez strat na dokładności i czasie cyklu produkcyjnego. Zastosowanie opisanego rozwiązania nie powoduje również obniżenia jakości produkowanych elementów, nie generując tym samym dodatkowych kosztów związanych z ich kontrolą. W pierwszej części tekstu zamieszczono opis metod wykorzystanych do realizacji zamierzonego celu, jak również objaśnienia zastosowanych technik przetwarzania obrazu aby ułatwić zrozumienie zagadnienia czytelnikowi niezaznajomionemu z tą dziedziną robotyki. W dalszych paragrafach przedstawiono analizę dokładności i szybkości działania systemu, porównując te parametry z odpowiednikami uzyskanymi przy użyciu profesjonalnego systemu.

Słowa kluczowe: system wizyjny, zrobotyzowane stanowisko produkcyjne, przetwarzanie obrazów.

1. Introduction

Camera-based systems were first introduced in the production site in the automotive industry in the 20th century as applications integrated with programmable manipulators. They proved their usefulness, providing an important capability of real-time control, what before had been impossible. Ever since then, vision systems have entered the production plants gradually, becoming a common part of modern automated production lines. Due to significant repeatability, accuracy and efficiency, application of this type in modern high-productive industry is inevitable. Similar to human cognitive capabilities based on the sense of sight, vision systems acquire visual data and process them in order to extract the data of interest.

Nowadays, advancements in electronics with simultaneous improvements in optical hardware, provide a possibility of processing large amount of data in a relatively short time. Specialized algorithms are capable of analyzing signals from single and multiple sensors [1, 2], and provide information regarding quantity, spatial orientation and quality of the observed objects. These features are especially helpful in large-scale production sites, where improvements in production pace are desired along with elimination of any inaccuracies.

An example of a computer vision based control application guiding a manipulator is described in [3]. The authors faced the problem of collecting information about randomly positioned objects in order to estimate parameters regarding their position and orientation, which were used by a trajectory computing algorithm. In [4] an approach for tracking and classifying objects on a conveyor belt is presented. By eliminating redundant areas of the acquired image, the described system that was implemented on an average PC hardware could determine the desired parameters in real-time. In [5] the authors describe the application of feedback control based on a vision sensor mounted at the robot's end-effector. A multidisciplinary vision application controlling sheet metal formation working in conjunction with force-acoustic sensors is presented in [6]. A comprehensive guide of the control hardware implementation is enclosed in [7].

A problem regarding the estimation of 3D position and orientation of an obstacle by vision systems mounted on mobile robots is described in [8, 9]. Although it covers different fields of application, it can be helpful in understanding the principles of basic image processing and geometric transformation procedures.

The system elaborated in the project and presented in this paper was based on the Open Source Computer Vision programming library, that allows a real-time image processing. A thorough text on this subject can be found in [10, 11, 12]. Moreover, the authors

of [13] provided comprehensive OpenCV function descriptions supplemented with pertinent examples.

The following paragraphs describe an attempt of implementing a vision system based on non-commercial OpenCV libraries that could cooperate with their commercial counterpart, in order to control a robot-based industrial assembly line. The constructed test rig was composed of two industrial manipulators operating in conjunction with separate conveyor belts, high-end video cameras and auxiliary systems. The main goal was to demonstrate that it was possible to integrate a low-cost system with off-the-shelf expensive high-efficient solutions to provide an opportunity of significant cost reduction of production-oriented enterprises.

The test rig emulated electronic circuit assembly lines. The components were delivered to the working station in a random order and positioned to simulate real conditions. They were subsequently recognized, picked up from the conveyor, classified and assembled into the final product.

2. Methodology

The majority of vision systems are based on some fundamental, commonly used image transformations, which are the starting point for more sophisticated processing algorithms. Examples of this type of architecture can be found in [11-15]. The process of extracting data from an image can be divided into the following steps:

1) Acquisition and storage of a vision signal – capture of the image from an input device and possible storage in memory,
2) Image pre-processing techniques– choosing RoI (Region of Interest), thresholding, filtration and morphological operations conducted in order to eliminate noise,
3) Image analysis:
   a. segmentation – extraction of regions representing objects,
   b. extraction of object features – estimation of quantities of interest (e.g. shape coefficients, area, color, etc.),
   c. object localization – detection of features like CoG (Center of Gravity) co-ordinates or an orientation.
4) Image recognition (object identification) – classification of the objects found on an image in the preceding steps. This information can be treated as an input for more advanced processing techniques.

In addition, an action called camera calibration that precedes the signal acquisition has to be undertaken in order to eliminate geometrical image distortions and therefore enhance accuracy of the system. A detailed description of this operation can be found in [13-16, 18].

The described system took advantage of the methodology presented above. Due to memory saving requirements, an acquired image was not stored but analyzed in real-time. During the image pre-processing step, median filtering preceded threshold operation in order to reject undesired pixels from the video signal. Hence, a grayscale input was transformed into its binary representation, on which morphological operations for noise elimination and image refinement were carried out. Further action included establishing a RoI on a chosen area, which decreased the computational time needed for subsequent operations. The main purpose of the image pre-processing stage was to estimate CoG co-ordinates and orientation angle describing the spatial position of a tracked item. This was done by means of moment method analysis. The last step was a transformation of the obtained values between two co-ordinate systems: the global one and the one attached to the robot's end-effector. The information was an input for the manipulator control unit calculating trajectories for the gripper. A simplified block diagram of the described algorithm is presented in Fig. 1.

The camera calibration process was conducted to eliminate image distortions. Although many factors can affect the quality of a vision signal, the most influential are radial and tangential. These are caused by production errors and by design simplifications (i.e., lens shape). The undesirable deflection of light on the edge of the lens causes an effect known as a radial distortion. Because of camera assembly inaccuracies, perfect alignment between a sensor and lenses is unattainable, therefore tangential distortions arise. Imprecise imager attachment also results in offset between its center point and lens optical axis.

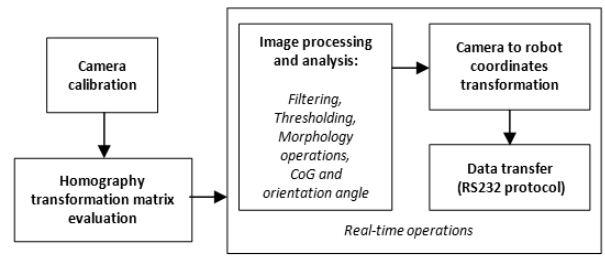


Fig. 1. Block diagram of the built algorithm
Rys. 1. Schemat blokowy zbudowanego algorytmu

In measurement applications, it is of paramount importance to compensate for the mentioned imperfections, thus a camera calibration process must be conducted. The main goal is to retrieve intrinsic camera parameters of the device along with a set of distortion coefficients.

In the implemented system, the distortion vector elements are estimated on the basis of multiple images of a calibration pattern – black and white squares "chessboard". By examination of the edge distortions, the algorithm computes all the correction coefficients. The second step before running the system is to find the transformation matrix which allows for expressing the object position in the manipulator's co-ordinate system used by the control unit to calculate the trajectory.

In order to conduct the mentioned transformations in an unambiguous manner, a homography projection needs to be performed. Because parallelism between an imager and working plane is impossible to obtain, perspective distortions must be taken into consideration as well. The mentioned projection is described by Eq. 1, in which the image coordinate vector **q** is related to the object's coordinates **Q** on the observed plane by means of the homography matrix **H**. It consists of rotation $\mathbf{R_{2x2}}$ and translation $\mathbf{T_{2x1}}$ matrixes that describe mutual position of the sensor and the observed plane. Additionally, matrix $\mathbf{K_{1x2}}$ containing perspective correction factors (equal to zero if the transformation is affine), a scale factor in homogenous coordinates S and an arbitrary image scale factor s are employed. Co-ordinates of points in the robot space can be expressed using vector **q** and inverted matrix **H**:

$$\mathbf{q} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{HQ} = s \begin{bmatrix} \mathbf{R}_{2x2} & \mathbf{T}_{2x1} \\ \mathbf{K}_{1x2} & S \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{1}$$

In order to find the homography matrix, co-ordinates of four pairs of corresponding points were extracted simultaneously from an image and manipulator working spaces. The former was done automatically by the dedicated algorithm, while the latter was done manually by moving a mechanical arm equipped with a special pointer into the predefined positions.

3. Image processing and analysis

As shown in the block diagram in Fig. 1, every frame of the captured video signal is processed in the following order: filtration, grayscale thresholding to binary image transformation, morphology operations and object position and orientation estimation. This comprehensive approach ensures robust performance and short response time, sufficient for systems operating in real-time.

In order to reduce the level of signal noise on the analyzed image, a median filter was used. The main principle of this technique is to replace the analyzed pixel by a median value

chosen from among the adjacent points. A more detailed description of this method can be found in [10]. The series of tests conducted confirmed that median filtering can lead to a significant improvement in the quality of an acquired image mainly by reducing the impulsive noise.

The operation known as thresholding is a method of band-pass filtering, where a single image point is compared to a certain cut-off level. The simplest example, binary threshold, assigns a maximum allowable value to pixels that are above a specified limit. Other variations of the method are known as well, and are described in detail in [13]. Histogram representation of an image is helpful and hence often used in defining the adequate cut-off level value. An image after binarization with the corresponding histogram chart is presented in Fig. 2.

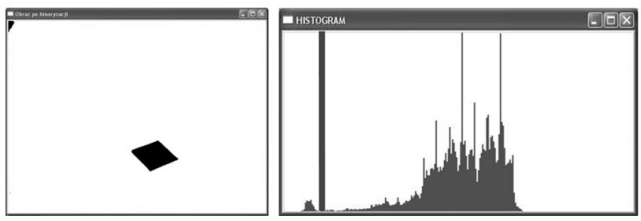


Fig. 2. Image after thresholding (left) and a corresponding histogram chart (right)
Rys. 2. Przykład obrazu po operacji progowania (po lewej) oraz jego histogram (po prawej)

The problem that arises after the operation described above is the presence of residual undesired areas. This is due to similar brightness values of pixels representing analyzed objects and other areas. An example can be seen in the upper left corner of the image in Fig. 2. Another possible error is discrimination of regions of different pixel values that belong to the analyzed object. In order to omit significant inaccuracies during subsequent analyses, morphological operations are often performed.

To execute this kind of operation, a proper structuring element must be defined. The most popular kernel shape is a unit circle, which in the raster representation is a 3-by-3 square mask with a central point distinguished. The structuring element is then moved along rows and columns of an image to analyze each pixel and apply some predefined operations. The two fundamental functions of morphological processing are erosion and dilation, that are the basis for widely used opening and closing techniques, implemented also in the described system.

When applied, dilation substitutes the central kernel point with the maximum value enclosed by the structuring element. In effect, a removal of grain noise from binary signals is achieved. The converse result is obtained by means of an erosion process. The difference is that the kernel anchor point is replaced by a minimum value from the pixels enclosed by the mask. This tool is useful for closing undesired tunnels and filling up missing pixels in a binary representation of the observed object [17].

A combination of these two operations results in the most widely used noise removing techniques: closing and opening. The former is a dilation followed by an erosion, while the latter consists of the same basic operations applied in the reversed order.

Evaluation of a center of gravity co-ordinates and an orientation angle of the object is done by means of the image spatial moments analysis, as described in [13].

Spatial moment of zero order described by Eq. 2 evaluates the area of the observed object:

$$m_{0,0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b(x,y)dxdy, \tag{2}$$

where *b(x,y)* is a binary image function.

In order to calculate CoG co-ordinates, spatial moments of the first order (Eq. 3) must be employed:

$$m_{1,0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xb(x,y)dxdy, \quad m_{0,1} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yb(x,y)dxdy, \tag{3}$$

To express the position of an object, the operations described by Eq. 4 must be carried out:

$$x_c = \frac{m_{1,0}}{m_{0,0}}, \quad y_c = \frac{m_{0,1}}{m_{0,0}}, \tag{4}$$

where x_c and y_c are the co-ordinates of a CoG.

The orientation is computed by means of the definition of the central spatial moment, which is expressed by Eq. 5:

$$\mu_{p,q} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-x_c)^p (y-y_c)^q b(x,y)dxdy. \tag{5}$$

It is noticeable that:

$$\mu_{0,0} = m_{0,0}, \quad \mu_{0,1} = 0, \quad \mu_{1,0} = 0. \tag{6}$$

To evaluate the angle θ between the X axis of an image co-ordinate system and an axis of a minimum object inertia moment, Eq. 7 is applied:

$$\theta = \frac{1}{2} arctg \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}. \tag{7}$$

4. The experimental set-up

As mentioned in the introduction, the experimental set-up shown in Fig. 3 consisted of two vision systems. The first, described in detail in the previous paragraphs, was implemented using the OpenCV libraries, while the second was the commercial AdeptSight®.

Electronic components were supplied in a random order and orientation on the conveyor belt operating in conjunction with the Adept Viper s650 robot and the AdeptSight® vision system, equipped with the Basler A601f camera. Objects of interest were tracked, recognized, picked up and palletized. Segregated components were then transported to the second robot (Mitsubishi Melfa RV-2AJ), which collected them from the conveyor and, being guided by the developed vision system, assembled the final product. The system used the EverFocus EQ 150/C camera and a backlight type lighting system. The process did not require any external actions, making the experimental production line fully autonomous.

Working cycle time is a figure of merit describing every vision system. It should be understood as the time needed to acquire a frame from a camera, process and analyze it, to extract data of interest. In the discussed example, the commercial system AdeptSight® working cycle time oscillated between 61ms and 79ms. These were the values provided automatically by the software regarding the time needed for recognition of the object on the conveyor and calculation of all the data sent to the robot controller.

In the case of the elaborated system, this time was measured on the basis of the required CPU working cycles of the PC which it was implemented on. Measurements were started at the moment the request for a frame was sent and continued until the co-ordinates for the manipulator controller were calculated. The range of acquired values was from 67ms to 72ms.

The presented results were obtained for both systems on the basis of 30 consecutive measurements. Table 1 presents the uncertainty of the single working cycle time measurement, expressed by the standard deviation of the mean value. By definition, for a set of n values $\{x_1, x_2, \ldots x_n\}$, the measurement uncertainty is expressed by Eq. 8:

$$u(X) = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n(n-1)}}, \tag{8}$$

where $\overline{X}$ is an arithmetic mean of the set.

Tab. 1. Working cycle time measurements
Tab. 1. Długość zmierzonych cykli pracy

	AdeptSight®	OpenCV based system
Mean value [ms]	72.40	69.27
Standard deviation [ms]	6.05	1.66

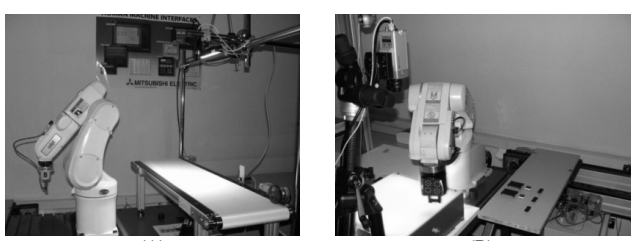
Tab. 2. The camera calibration parameters for the elaborated vision system
Tab. 2. Parametry kalibracyjne kamery prototypowego system pomiarowego

Focal lengths	$f_x = 708.9 \pm 4.5$pixel $f_y = 708.5 \pm 4.9$pixel
Principal point co-ordinates	$c_x = 346.1 \pm 3.6$pixel $c_y = 229.0 \pm 3.0$pixel
Distortion coefficients	$k_1 = 0.0660 \pm 0.0059$ $k_2 = -0.307 \pm 0.0330$ $p_1 = 0.0018 \pm 0.0014$ $p_2 = -0.0042 \pm 0.0017$

Tab. 3. Spatial resolution of the vision systems
Tab. 3. Rozdzielczość przestrzenna systemów wizyjnych

AdeptSight® [mm/pix]	OpenCV based system (full image error) [mm/pix]	OpenCV based system (RoI error) [mm/pix]
$\delta_w = 0.401878 \pm 1.3 \cdot 10^{-5}$ $\delta_l = 0.401709 \pm 1.1 \cdot 10^{-5}$	$\delta_w = 0.69579 \pm 0.0433$ $\delta_l = 0.69619 \pm 0.0307$	$\delta_w = 0.69579 \pm 0.02603$ $\delta_l = 0.69619 \pm 0.02136$

In order to examine the performance of the systems, a set of calibration parameters was investigated. In reality, the overall accuracy of the test station depends also on the mechanical characteristic of the manipulators. Therefore, mechanical measurements could not express the accuracy of the image processing algorithms. Nevertheless, it is worth noting that in the case of the system operating in conjunction with the assembling robot, visual measurements were precise enough to undertake the process of mounting integrated circuits on PCB boards without errors.



(A) (B)

Fig. 3. (A) Robot Adept Viper650s, the conveyor belt and AdeptSight® system (B) The Mitsubishi Melfa RV 2-AJ robot, the EverFocus EQ 150/C camera and a backlight illuminator
Rys. 3. (A) Robot Adept Viper650s, podajnik taśmowy oraz systemem AdeptSight® (B) Robot Mitsubishi Melfa RV 2-AJ, kamera EverFocus EQ 150/C oraz oświetlacz typu „backlight"

In the case of the AdeptSight® system, the pixel width δ_w and length δ_l were automatically measured by the system during the calibration process. The OpenCV libraries give access to more sophisticated parameters used in the process of eliminating distortions.

Table 2 presents the camera calibration results obtained for the prototypic vision system. Using Eq. 2 and Eq. 10, the parameters describing both cameras are compared in Table 3. The presented values were calculated for all of the pixels on the sensor and for the chosen RoI.

$$x_{screen} = f_x\left(\frac{X}{Z}\right) + c_x, \quad y_{screen} = f_y\left(\frac{Y}{Z}\right) + c_y, \tag{10}$$

where x_{screen}, y_{screen}, X, Y describe the position of the point on the screen and on the working plane respectively [13]. The obtained resolution level was close to the limits that ensured proper assembly.

5. Conclusions

The prototypic system described in the paper was based on the classical image processing techniques including median filtering, thresholding and morphological transformations, which preceded further shape analysis of the extracted objects. The purpose of the performed calculations was to determine the CoG position and orientation angle of an item.

The most crucial parameters describing the tested system were the cycle time and spatial resolution. After detailed investigations, it was shown that those figures of merit of the elaborated system were comparable with the equivalents obtained by commercial products. Therefore, an application based on OpenCV libraries can operate in conjunction with specialized commercial systems, providing a possibility of easy modifications for further production enhancements. As a result, the system implementation and operational costs can be considerably decreased.

6. References

[1] Rosselot D., Hall E. L.: Processing real-time stereo video for an autonomous robot using disparity maps and sensor fusion, Proceedings of SPIE Intelligent Robots and Computer Vision XXII: Algorithms, Techniques and Active Vision, Vol. 5608, pp.70-78, 2004.
[2] Rosati G., Boschetti G., Biondi A., Rossi A.: On-line dimensional measurement of small components on the eyeglasses assembly line, Optics and Lasers in Engineering, 47, pp. 320 – 328, 2009.
[3] Nagchaudhuri A., Kuruganty S., Shakur A.: Introduction of mechatronics concepts in a robotics course using an industrial SCARA robot equipped with a vision sensor, Mechatronics 12, pp. 183-193, 2002.
[4] Bozma H. I., Yalçin H.: Visual processing and classification of items on a moving conveyor: a selective perception approach, Robotics and Computer-Integrated Manufacturing 18, pp. 125 – 133, 2002.
[5] Kosmopoulos D.I., Varvarigou T.A., Emiris D.M., Kostas A.A., MD-SIR: A methodology for developing sensor – guided industry robots, Robotics and Computer Integrated Manufacturing, 18, pp. 403-419, 2002.
[6] Fillatreau P., Bernard F.X., Aztiria A., Saénz de Argandoña E., García C., Arana N., Izaguirre A.: Sheet metal forming global control system based on artificial vision system and force – acoustic sensors, Robotics and Computer Integrated Manufacturing, 24, pp. 780 – 787, 2008.
[7] Veiga G., Pires J.N., Nilsson K.: Experiments with service-oriented architectures for industrial robotic cells programming, Robotics and Computer-Integrated Manufacturing, 25, pp. 746-755, 2009.
[8] Hallmann I.: Mobile robot position estimation due the camera's image, Pomiary Automatyka Robotyka, R. 6, nr 1, pp. 20-24,2002.
[9] Rekleitis I., Meger D., Dudek G.: Simultaneous planning, localization, and mapping in camera sensor network, Robotics and Autonomous Systems, 54, pp. 921-932, 2006.
[10] Tadeusiewicz R., Korohoda P.: Digital image analysis and processing (Komputerowa analiza i przetwarzanie obrazów), FTP, Kraków, 1997.
[11] Jahne B.: Digital image processing: concepts, algorithms, and scientific application, Springer-Verlag, Berlin, 1995.
[12] Weeks A.: Fundamentals of electronic image processing, IEEE, Inc., New York, 1996.
[13] Bradski G., Kaehler A.: Learning OpenCV , O'Reilly Media, Sebastopol, 2008.
[14] Uhl T., Kohut P., Holak K.: Zastosowanie metod wizyjnych do monitorowania stanu konstrukcji, Pomiary Automatyka Kontrola, vol. 56, nr. 6, pp. 569-572, 2010.
[15] Kohut P.: Rapid prototyping of visual servoing systems, Pomiary Automatyka Kontrola, R. 48, nr 12, pp. 25 – 28, 2002.
[16] Kohut P.: Vision methods in robotics. Pt. 1, Przegląd Spawalnictwa, R. 80, nr 12, str. 21 - 25, 2008.
[17] Szpytko J., Tekielak M., Kohut P.: Optoelektroniczna metoda analizy przemieszczeń obiektów, Pomiary Automatyka Kontrola, pp. 5-7, 11, 2003
[18] Ma L., Chen Y., and Moore K. L.: Rational radial distortion models of camera lenses with analytical solution for distortion correction, International Journal of Information Acquisition, vol. 1, no. 2, pp. 135-147, 2004.

otrzymano / received: 26.03.2013
przyjęto do druku / accepted: 02.12.2013 *artykuł recenzowany / revised paper*