Choice of Window Size in Calibrating the Geometry of Manipulators Based on the Regions Correlation

Petar Maric and Velibor Djalic

***Abstract*—To achieve the full flexibility of industrial robots, in addition to the mechanical flexibility, it is necessary to achieve the flexibility in control. Precise automatic calibration of these robots manipulators is essential precondition for achieving this goal. Possibilities and limitations of stereo vision system implementation for automatic calibration of manipulators are presented in this paper. Furthermore, a practical solution for the problem of corresponding points using an area based algorithm is given. Analysis has been conducted of how the marker positioning on the manipulator end-effector influences the shape of the cost function. Choice of window size relative to marker size which provides the best reliability for corresponding area determination is proposed. Results of practical realization, which confirms conducted analysis and reliability of proposed calibration procedure, are presented.**

***Index Terms*—Camera calibration, Computer vision, Modular reconfigurable robot, Robot calibration.**

I. Introduction

A high level of positioning accuracy is an essential requirement in a wide range of industrial robots applications. This accuracy is affected by geometric factors (geometrical parameters accuracy) and non-geometric factors (gear backlashes, encoder resolution, flexibility of links, thermal effects, etc.).

The error due to geometric factors accounted for 90% of the total error. A common approach is to calibrate the current geometric parameters and treat the non-geometric factors as a randomly distributed error. The calibration procedure is very important for robot programming using CAD systems where the simulated robot must reflect accuracy the real robot. During a manipulator control system design, and periodically in the course of task performing, manipulator geometry calibration is required [1].

Vision systems have developed significantly over the last ten years and now have become standard automation components. They represent qualitative bounce in the area of metrology and sensing because they provide us with a remarkable amount of information about our surroundings, without direct physical contact [2]. At the same time, vision systems are the most complex sensors [3].

P. Maric and V. Djalic are with the University of Banja Luka, Faculty of Electrical Engineering, Banja Luka, Bosnia and Herzegovina (e-mail: {petar.maric, velibor.djalic}@etfbl.net).

Calibration of cameras is necessary first step in vision system using. Camera calibration is the process of determining the internal camera (geometric and optical) characteristics and the 3D position and orientation of the camera frame relative to a world coordinate system [2], [4]. There are a number of techniques which only requires the camera to observe a planar pattern(s) shown at a few different orientations. Precise set points are placed on the calibration plain. Calibration process is automatically performed based on the correspondence between the positions of calibration points and the positions of their images [5], [6].

If the camera calibration is performed then for every scene point in a world coordinate system it is possible to determine the position of its image point in image plain. This transformation is called perspective transformation. Inverse perspective transformation is very important for computer vision application in industrial automation. This transformation examines the problem of how to identify the point position in a world coordinate system based on the position of its image point for different camera positions or for several cameras at the same time.

Inverse perspective transformation and advances in processing and image analysis have a wide range of applications in industrial automation, and allow companies to achieve previously impossible levels of efficiency and productivity [7]. Effective co-operation in the computer aided manufacturing depends on the recognition and perception of typical production environments as well as on the understanding of tasks in their context. Vision systems are the basis for scene analysis and interpretation, both in time and 3D space. Measurement of the dimensions of objects and parts in many industrial fields is very important, as the quality of the product depends especially on the reliability and precision of each object part. The use of vision-based metrology allows calculating a set of 3D point coordinates and/or estimate the dimension and pose of a known object. They enable industrial robots to perform different and complex tasks reliably and accurately [8].

II. Robot Calibration

A. Manipulator Geometry Modeling

The first step of manipulator calibration is concerned with a mathematical formulation that results in model which gives

relation between the geometric parameters, the joint variables and end-effector position. Many researchers have been looking for the suitable kinematic models for robot calibration, since Richard Paul's book [9]. The most popular among them is the Denavit-Hartenberg (D-H) method. For this reason we will use this notation.

Prior description kinematic model let us define the basic coordinate systems as follows (Fig. 1.):

$O_BX_BY_BZ_B$ – base coordinate system of the manipulator

$O_EX_EY_EZ_E$ – end-effector (tool) coordinate system of the manipulator (we denote the origin O_E as the endpoint of the robot)

$O_iX_iY_iZ_i$ (i=1, n) – coordinate system fixed to the i[th] link ($O_nX_nY_nZ_n$ – coordinate system fixed to the terminal link) of the manipulator.

The original D-H representation of a rigid link depends on geometric parameters. Four parameters a,d,α and θ denote manipulator link length, link offset, joint twist and joint angle, respectively. Composite 4x4 homogenous transformation matrix $A_{i-1,i}$ known as the D-H transformation matrix for adjacent coordinate system i and $i-1$, is:

$$A_{i-1,i}=\begin{bmatrix}\cos\theta_i & -\cos\alpha_i\sin\theta_i & \sin\alpha_i\sin\theta_i & a_i\cos\theta_i\\ \sin\theta_i & \cos\alpha_i\cos\theta_i & -\sin\alpha_i\cos\theta_i & a_i\sin\theta_i\\ 0 & \sin\alpha_i & \cos\alpha_i & d_i\\ 0 & 0 & 0 & 1\end{bmatrix} \quad (1)$$

The homogenous matrix $A_{B,i}$ which specifies the location of the i[th] coordinate system with respect to the base coordinate system is the chain product of successive coordinate transformation matrices $A_{i-1,i}$, and expressed as:

$$A_{B,i}=A_{B,1}A_{1,2}\ldots A_{i-1,i}. \quad (2)$$

Particularly, for $i=n$ we have $A_{B,n}$ matrix which specifies the position and orientation of the end-effector of the manipulator with respect to the base coordinate system. Matrix $A_{B,n}$ is a function of the $4n$ geometrical parameters which are constant for constant robot geometry, and n joint coordinates that change their value when manipulator moves.

Moreover, a robot is not intended to perform a single operation at the workcell, it has interchangeable different tools. In order to facilitate the programming of the task, it is more practical to have transformation matrix defining the tool coordinate system with respect to the terminal link coordinate system $A_{n,E}$.

Thus, the transformation matrix $A_{w,E}$ can be written as:

$$A_{w,E}=A_{w,B}A_{B,n}A_{n,E}. \quad (3)$$

Since the world coordinate system can be chosen arbitrarily by the user, six parameters are needed to locate the robot base relative to the world coordinate system. From independence to some manipulator parameters it follows that consecutive coordinate systems are represented at most by four independent parameters.

Since the end-effector coordinate system can be defined arbitrarily with respect to the terminal link coordinate system ($O_nX_nY_nZ_n$), six parameters are needed to define the matrix $A_{n,E}$. If we extend the robot notation to the definition of the end-effector coordinate system, it follows that the end-effector coordinate system introduces four independent parameters. For more details the reader can refer to [1].

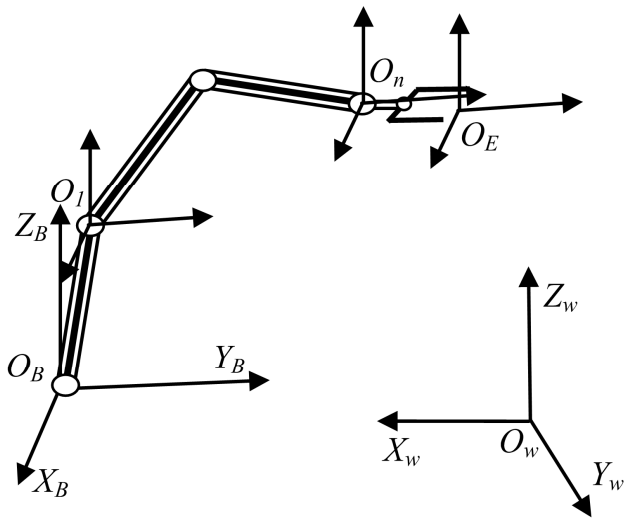

Fig. 1. Coordinate systems assignment for robot modeling.

With above mentioned equations (1), (2), (3) dependence between joint coordinates and geometrical parameters, and endpoint location of the tool can be written as:

$$x=f(q,g^0) \quad (4)$$

where x,q,g^0 denotes end-effector position vector expressed in the world coordinate system, vector of the joint variables, and vector of the geometric parameters, respectively. Dimension of the vector $\boldsymbol{x}$ is 6 if measurement can be made on the location and orientation of the end-effector. However, the most frequently only a location of the endpoint is measured, and therefore dimension of a vector $\boldsymbol{x}$ is 3. Dimension of the vector q is equivalent to the number of DOF for manipulator. Dimension of the vector g^0 is at most $4n+6$.

B. Geometric Parameters Estimation Based on the Differential Model

The calibration of the geometric parameters is based on estimating the parameters minimizing the difference between a function of the real robot variables and corresponding mathematical model. Many authors [10]–[13] presented open-loop methods that estimate the kinematic parameters of manipulators performing on the basis of joint coordinates and the Cartesian coordinates of the end-effector measurements. The joint encoders outputs readings are joint coordinates. It is assumed that there is a measuring device that can sense the position (sometime orientation) of an end-effector Cartesian coordinates.

A mobile closed kinematic chain method has been proposed that obviates the need for pose measurement by forming a manipulator into a mobile closed kinematic chain [14], [15]. Self motion of the mobile closed chain places manipulator in a number of configurations and the kinematic parameters are determined from the joint position readings alone.

The calibration using the end-efector coordinates (open-loop method) is the most popular one. The model represented by equation (4) is nonlinear in g^0, and we must linearize it in order to apply linear estimators. The differential model provides the differential variation of the location of the end-effector as a function of the differential variation of the

geometric parameters. The function to be minimized is the difference between the measured ($\boldsymbol{x}$) and calculated end-effector location (x^m). Let $\Delta x=x-x^m$, and $\Delta g= g^0$-g be the pose error vector of end-effector and geometric parameter error vector, respectively (g – vector of geometric parameters estimation). From equation (4), the calibration model can be represented by the linear differential equation

$$\Delta x=J_g\,\Delta g = x-x^m\,, \tag{5}$$

where:

g is the (p x 1) vector of geometric parameters estimation

$\Delta x=x-x^m$ is the (r x 1) pose error vector of end-effector

$\Delta g= g^0$-g is the geometric parameter error vector

J_g is the (r x p) sensitivity matrix relating the variation of the endpoint position with respect to the geometric parameters variation (calibration Jacobian matrix) [10], [13].

To estimate Δg we apply equation (5) for a number of manipulator configurations. It gives the system of equations:

$$\Delta X=\Phi\,\Delta g +\mathrm{E} \tag{6}$$

where is:

$$\Delta X=\begin{bmatrix}\Delta x^1\\ \Delta x^2\\ \vdots\\ \Delta x^k\end{bmatrix},\ \Phi=\begin{bmatrix}J_g^1(q^1,g)\\ J_g^2(q^2,g)\\ \vdots\\ J_g^k(q^k,g)\end{bmatrix}, \tag{7}$$

and E is the error vector which includes the effect of unmodeled non-geometric parameters:

$$E=\begin{bmatrix}e^1\\ e^2\\ \vdots\\ e^k\end{bmatrix}. \tag{8}$$

Equation (6) can be used to estimate iteratively the geometric parameters. This equation is solved to get the least-squares error solution to the current parameters estimate. The least-squares solution can be obtained from:

$$\Delta g= (\Phi^T\Phi)^{-1}\,\Phi^T\Delta X. \tag{9}$$

At the each iteration, geometric parameters are updated by adding Δg to the current value of g:

$$g=g+\Delta g. \tag{10}$$

By solving equations (9) and (10) alternately, the procedure is iterated until the Δg approaches zero.

Calibration a manipulator is an identification process, and hence, one should take a careful look at the identifability of the model parameters [10], [14]. A general method to determine these parameters have been proposed in [14]. Determination of the identifiable (base) geometric parameters is based on the rank of the matrix Φ. Some parameters of manipulator related to the locked passive joints may become unidentifiable in the calibration algorithm due to the mobility constraints. It reduces number of identifiable parameters in general for the closed-loop kinematic chain approach, compared with open-loop case.

As the measurement process is generally time consuming, the goal is to use set of manipulator configurations that use limited number of optimum points on the parameters estimation. Furthermore, goal is to minimize the effect of noise on the parameters estimation. The condition number of the matrix Φ gives a good estimate of the persistent excitation [1]. Therefore, much work was led on finding the so-called optimal excitation. The task of selecting the optimum manipulator configurations to be used during the calibration is discussed and solutions are proposed in [14]–[16]. It is worth noting that most of geometric calibration methods give an acceptable condition number using random configurations. The paper [17] presents an updating algorithm to reduce the complexity of computing and observability index for kinematic calibration of robots. An active calibration algorithm is developed to include an updating algorithm in the pose selection process.

III. Computer Vision

A. Camera Model

This section describes the camera model. Fig. 2. illustrates the basic geometry of the camera model. The camera performs transformation from the 3D projective space to the 2D projective space. The projection is carried by an optical ray originating (or reflected) from a scene point P. The optical ray passes through the optical center O_c and hits the image plane at the point p.

Prior describing the perspective transformation, and camera model, let us define the basic coordinate systems. The coordinate frames are defined as follows:

$O_wX_wY_wZ_w$ - world coordinate system (fixed reference system), where O_w represents the principal point. The world coordinate system is assigned in any convenient location.

$O_cX_cY_cZ_c$ - camera centered coordinate system, where O_c represents the principal point on the optical center of the camera. The camera coordinate system is the reference system used for camera calibration, with the Z_c axis the same as the optical axis.

$O_iX_iY_iZ_i$ - image coordinate system, where O_i represents the intersection of the image plane with the optical axis. X_iY_i plane is parallel to X_cY_c plane.

Let (x_w, y_w, z_w) are the 3D coordinates of the object point P in the 3D world coordinate system, and (u,v) position of the corresponding pixel in the digitized image. A projection of the point P to the image point p may be represented by a 3x4

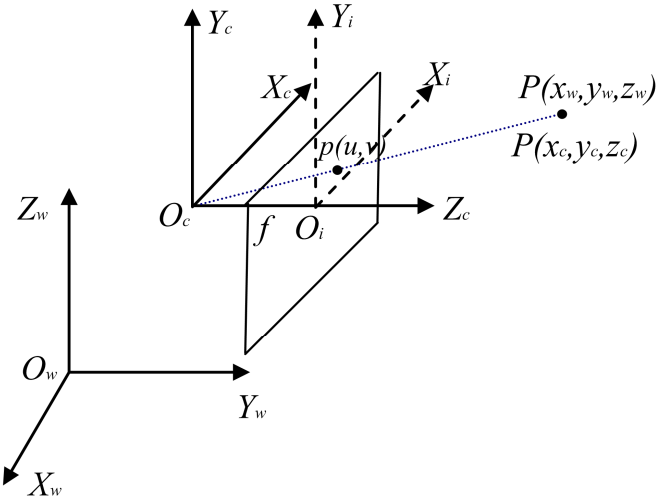

Fig. 2. The basic geometry of the camera model.

projection matrix (or camera matrix) M [2], [5]:

$$p = K\begin{bmatrix} R & T \end{bmatrix} P = MP \ . \tag{11}$$

Matrix:

$$K = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

is called the internal (intrinsic) camera transformation matrix. Parameters α, β, u_0 and v_0 are so called internal distortion-free camera parameters.

R and T, a 3x3 orthogonal matrix representing the camera's orientation and a translation vector representing its position, are given by:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}, \tag{13}$$

respectively. The parameters r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}, t_x, t_y, t_z are external (extrinsic) parameters and represent the camera's position refereed to the word coordinate system.

Projection in an ideal imagining system is governed by the pin-hole model. Real optical system suffers from a number types of distortion. The first one is caused by real lens spherical surfaces and manifests itself by radial position error. Radial distortion causes an inward or outward displacement of a given image point from its ideal (distortion free) location. This type of distortion is mainly caused by flawed radial curvature curve of the les elements. A negative radial displacement (a point is imaged at a distance from the principle point that is smaller than predicted by the distortion free model) of the image point is referred to as barrel distortion. A positive radial displacement (a point is imaged at a distance from point that is larger than the predicted by the distortion free model) of the image point is referred to as pin-cushion distortion. The displacement is increasing with distance from the optical axis. This type of distortion is strictly symmetric about the optical axis. Fig. 3. illustrates the effect of radial distortion.

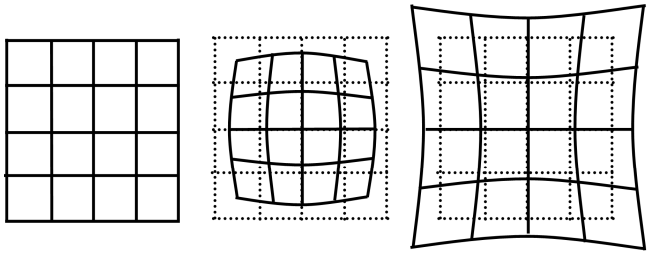

Fig. 3. Effect of radial distortion illustrated on a grid.

The radial distortion of a perfectly centered lens is usually modeled using the equations:

$$\Delta x_r = x_i(k_1 r^2 + k_2 r^4 + ...) , \tag{14}$$

$$\Delta y_r = y_i(k_1 r^2 + k_2 r^4 + ...) , \tag{15}$$

where r is the radial distance from the principal point of the image plane, and $k_1,k_2,...$ are coefficients of radial distortion. Only even powers of the distance r from the principal point occur, and typically only the first, or the first and the second terms in the power series are retained.

The real imagining systems also suffer from tangential distortion, which is at right angle to the vector from the center of the image. That type of distortion is generally caused by improper lens and camera assembly. Like radial distortion, tangential distortion grows with distance from the center of distortion and can be represented by equations:

$$\Delta x_t = -y_i(l_1 r^2 + l_2 r^4 + ...) , \tag{16}$$

$$\Delta y_t = x_i(l_1 r^2 + l_2 r^4 + ...) . \tag{17}$$

Fig.4. illustrates the effect of tangential distortion.

The reader is referred to [2]–[4] for more elaborated and more complicated lens models.

Note that one can express the distorted image coordinates as a power series using undistorted image coordinates as variables, or one can express undistorted image coordinates as a power series in the distorted image coordinates. The r in the above equations can be either based on actual image coordinates or distortion-free coordinates.

Bearing in mind the radial and tangential distortion, correspondence between distortion-free and distorted pixels image coordinates can be expressed by:

$$x_d = x_i + \Delta x_r + \Delta x_t , \tag{18}$$

$$y_d = y_i + \Delta y_r + \Delta y_t . \tag{19}$$

The parameters representing distortion of an image are: $k_1,k_2,...,\ l_1,l_2,...$ The distortion tends to be more noticeable with wide-angle lenses than telephoto lenses. Electro-optical systems typically have larger distortions than optical systems made of glass.

B. Camera Calibration

Camera calibration is considered as an important issue in computer vision applications. With the increasing need for higher accuracy measurement in computer vision, if has also attracted research effort in this subject. Task of camera calibration is to compute the camera projection matrix M from a set of image-scene point correspondences. By correspondences it means a set $\{(p_i, P_i)\}_{i=1}^{m}$ where p_i is a homogeneous vector representing image point and P_i is a homogeneous vector representing scene point, at the i^{th} step. Equation (11) gives an important result: the projection of a point P to an image point p by a camera is given by a linear mapping (in homogeneous coordinates):

$$p \ = MP \ . \tag{20}$$

The matrix M is non-square and thus the mapping is many-to-one. All scene points on a ray project to a single image point.

To compute M, it have to be solved the system of homogeneous linear equations

$$s_i p_i = MP_i \ , \tag{21}$$

where s_i are scale factors.

Camera calibration is performed by observing a calibration object whose geometry in 3D space is known with very good precision. The calibration object usually consists of two or three planes orthogonal to each other. These approaches

require an expensive calibration apparatus. Accurate planar targets are easier to make and maintain than three-dimensional targets. There is a number of techniques which only requires the camera to observe a planar pattern(s) shown at a few different orientation (Fig. 5). The calibration points are created by impressing a template of black squares (usually chess-board pattern) or dots on top of white planar surface (steel or even a hard book cover [5]). The corners of the squares are treated as a calibration points. Because the corners are always rounded, it is recommended to measure the coordinate of a number of points along the edges of the square away from the corners, and then extrapolate the edges to obtain position of the corners which lie on the intersection of adjacent edges.

Due to the high accuracy performance requirement for

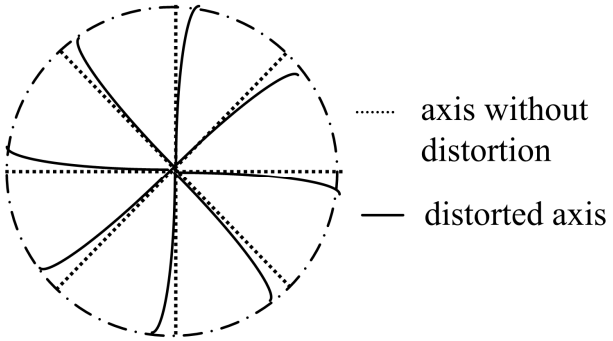

Fig. 4. Effect of tangential distortion.

camera calibration, a sub-pixel estimator is desirable. It is a procedure that attempts to estimate the value of an attribute in the image to greater precision than that normally considered attainable within restrictions of the discretization. Since the CCD camera has relatively low resolutions, interest in a sub-pixel method arises when one applies CCD-based image systems to the computer integrated manufacturing [6].

Camera calibration entails solving for a large number of calibration parameters, resulting in the large scale nonlinear search. The efficient way of avoiding this large scale nonlinear search is to use two-stage technique, described in [2]. The methods of this type in the first stage use a closed-form solution for most of the calibration parameters, and in the second stage iterative solution for the other parameters.

In [4] a two-stage approach was adopted with some

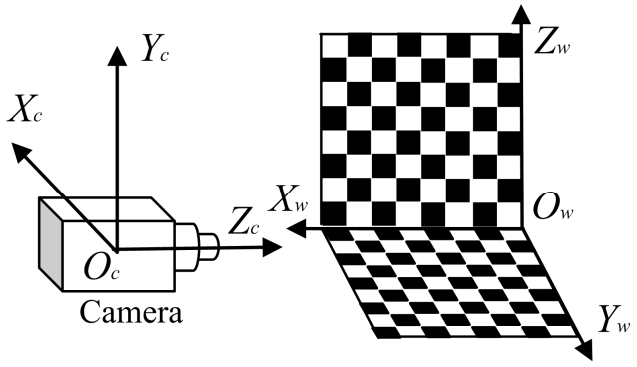

Fig. 5. Illustration of experimental setup for camera calibration using coplanar set of points.

modification. In the first step, the calibration parameters are estimated using a closed-form solution based on a distortion-free camera model. In the second step, the parameters estimated in the first step are improved iteratively through a nonlinear optimization, taking into account camera distortion. Since the algorithm that computes a closed-form solution is no iterative, it is fast, and solution is generally guaranteed. In the first step, only points near the optical axis are used. Consequently, the closed-form solution isn't affected very much by distortion and is good enough to be used as an initial guess for further optimization. If an approximate solution is given as an initial guess, the number of iterations can be significantly reduced, and the globally optimal solution can be reliably reached.

C. Stereo Vision

Calibration of one camera and knowledge of the coordinates of one image point allows us to determine a ray in space uniquely (back-projection of point). Given a homogeneous image point p, we want to find its original point P from the working space. This original point P is not given uniquely, but all points on a scene ray from image point p. Here, we will consider how to compute 3D scene point P from projections p_i in the several cameras, or projections p_i in one camera at different positions (different images are denoted by superscript i). Assume that m views are available, so that we have to solve linear system

$$s_i p_i = M_i P \quad , i=1,\ldots,m. \tag{22}$$

This approach is known as triangulation (it can be interpreted in terms of similar triangles). Geometrically, it is a process of finding the common intersection of m rays given by back-projection of the image points by the cameras. In the reality, image points p_i are corrupted by noise, and the rays will not intersect and the system would have no solution. We might compute P as the scene point closest to all of the skew rays.

If two calibrated cameras observe the same scene point P, its 3D coordinates can be computed as the intersection of two of such rays. The epipolar geometry is a basis of a system with two cameras (principle of stereo vision). It is illustrated on Fig. 6.

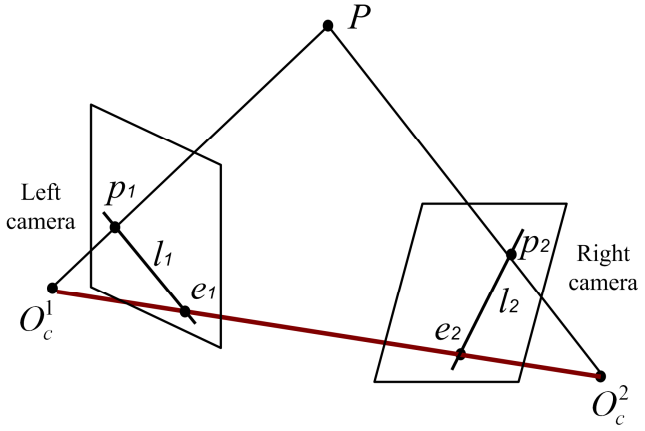

Fig. 6. The epipolar geometry.

Let O_c^1, O_c^2 represents the optical centers of the first and second camera, respectively. The same consideration holds if one camera takes two images from two different locations. In that case O_c^1 represents optical center of the camera when the first image is obtained, and O_c^2 represents the optical center for the second image.p_1 and p_2 denote the images of the 3D point P. The base line is the line joining the camera centers O_c^1 and O_c^2. The baseline intersects the image planes in the epipoles e_1 and e_2. Alternatively, an epipole is the image of the optical center of one camera in the other camera. Any scene point P and the two corresponding rays from optical centers O_c^1 and O_c^2 define an epipolar plane. This plane intersects the image plane in the epipolar line. It means, an epipolar line is the projection of the ray in one camera into the other camera. Obviously, the ray $O_c^1 P$ represents all possible positions of P for the first image and is seen as the epipolar line l_2 in the second image. The point p_2 in the second image that corresponds to p_1 must thus lie on the epipolar line in the second image l_2, and reverse. The fact that the positions of two corresponding image points are not arbitrary is known as the epipolar constraint. This is a very important statement for the stereo vision. The epipolar constraint reduces the dimensionality of the search space for a correspondence between p_1 and p_2 in the second image from 2D to 1D.

A special relative position of the stereo cameras is called rectified configuration. In this case image planes coincide and

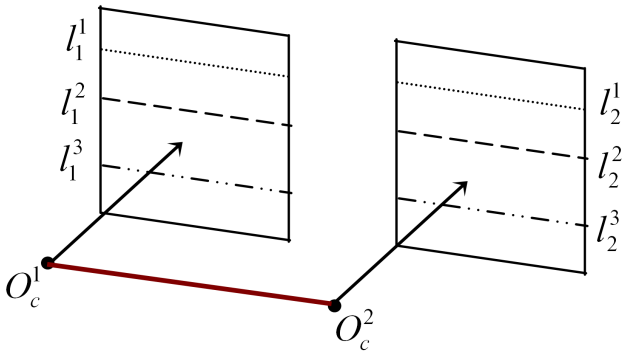

Fig. 7. The rectified configuration of two cameras.

line $O_c^1 O_c^2$ is parallel to them, as shown in Fig. 7.

The epipoles e_1 and e_2 go to infinity, and epipolar lines coincide with image rows, as a consequence. For the rectified configuration, if the internal calibration parameters of both cameras are equal, it implies that corresponding points can be sought in 1D space along image rows (epipolar lines).

The optical axes are parallel, which leads to the notion of disparity that is often used in stereo vision literature. Top view of two cameras stereo configuration with parallel optical axes is shown in Fig. 8. World coordinate system is parallel to cameras coordinate systems. The principal point O_w of the world coordinate system is assigned on the midway on the baseline. The coordinate z_w of point P represents its distance from the cameras (z_w =0), and can be calculated from the disparity $d=u_1- u_2$. Values $u_1- u_2$ are measured at the same height (same rows of images). Noting that:

$$\frac{u_1}{f}=\frac{x_w+\frac{B}{2}}{z_w},\ \frac{u_2}{f}=\frac{x_w-\frac{B}{2}}{z_w}, \tag{23}$$

we have:

$$z_w=\frac{Bf}{d}. \tag{24}$$

The remaining two coordinates of the 3D point P can be calculated from equations:

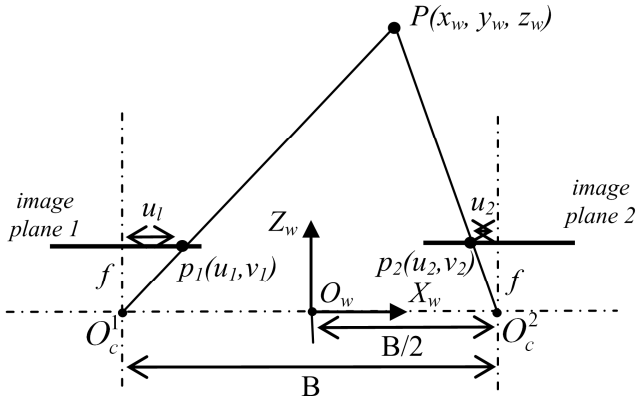

Fig. 8. Top view of two cameras with parallel optical axes rectified configuration.

$$x_w=\frac{-B(u_1+u_2)}{2d},\ y=\frac{Bv_1}{d}. \tag{25}$$

The position of the point P in the 3D scene can be calculated from the disparity d. It is a question, how the same point can be found in two images if the same scene is observed from two different viewpoints. The solution of this correspondence problem is a key step in any stereo vision. Automatic solution of the correspondence problem is under extensive exploration. Until now there is not solution in general case. The inherent ambiguity of the correspondence problem can in practical cases be reduced using several constrains. A vast list of references about this task can be found in the [3].

The geometric transformation that changes a general cameras configuration with non-parallel epipolar lines to the parallel ones is called image rectification. More deep explanation about computing the image rectification can be found out in [3].

IV. Robot Calibration Using Computer Vision

Measurement of robot manipulator end-effector pose (i.e., position and orientation) in the reference coordinate system is unquestionably the most critical step towards a successful open-loop robot calibration. A variety of measurement techniques ranging from coordinate measuring machines, proximity measuring systems, theodolites, and laser tracking interferometer systems to inexpensive customized fixtures have been employed for calibration tasks. These systems are very expensive, tedious to use or with low working volume [12], [18], [19]. In general, the measurement system should be

accurate, inexpensive and should be operated automatically. The goal is to minimize the calibration time and the robot unavailability.

To overcome the above limitations, advances in robot calibration allow the start using a computer vision to calibrate a robot. Compared to those mechanical measuring devices, the camera system is low cost, fast, automated, user-friendly, non-invasive and can provide high accuracy [20].

There are two types of setups for vision-based robot pose measurement. The first one is to fix cameras in the robot environment so that the cameras can see a calibration fixture mounted on the robot ende-ffector while the robot changes its configuration. The second typical setup is to mount a camera or a pair of cameras on the end-effector of the robot manipulator.

The stationary camera configuration requires the use of stereo system placed at fixed location. It is not possible compute 3D scene point *P* position from only one projection *p,* on the camera plane. The stereo system has to be placed in location that maintains necessary field-of-view overlap. The proper camera position needs to be selected empirically. The stereo system must be calibrated before manipulator calibration. The manipulator is placed in a number of configurations. From pair of images the location (position and orientation) of the calibration board is computed for every configuration (Fig. 9.). At the each configuration, geometric parameters are updated by adding Δg (calculated in accordance with equation (9)) to the current value of g [21].

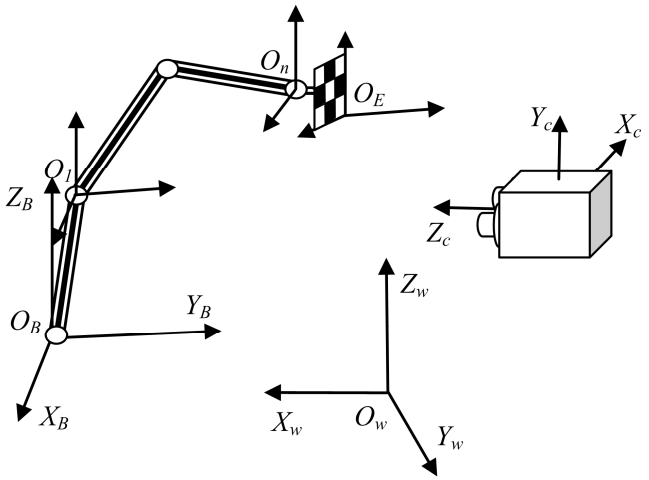

Fig. 9. A manipulator calibration using stationary camera configuration.

If it is enough to measure only the end-effector pose (usually tool's tip) for robot calibration, then it is not necessary to use a calibration plate. In that case it is enough to place a marker on the end-effector of the manipulator. Tests were conducted with square markers. In this way the calibration of the manipulators, from the viewpoint of practical implementation, is simplified. Also, this is an important prerequisite for increasing the flexibility of the manipulators, because in the case of automatically interchangeable tools, a marker can be placed on the end-effector of the manipulator. Then, for any change of tools, by tracking the position of marker it is possible to accomplish the recalibration of the manipulator.

The position of a marker (end-effector) is determined on the basis of pairs of images from stereo camera system [22]. Any point of the marker can be assumed as referent point of end-effector of the manipulator. For this reason, hereinafter a term „marker" is considered as a whole, rather than a one specified point. In this case the main problem is the automatic detection of corresponding points. The corresponding points are represented by a set of marker points on both images.

This paper uses algorithm based on the most similar intensity area correlation. The algorithm assumes that more pixels have similar intensity (color). Therefore, correlation of two pixels does not provide sufficient information because of the existence of more similar candidates. Thus, correlation of more adjacent pixels which are forming the window of *hxw* pixels is determined. When stereo system with parallel optical axes is used, the epipolar lines of both cameras lie on the same height on both images, as shown on Fig. 10.

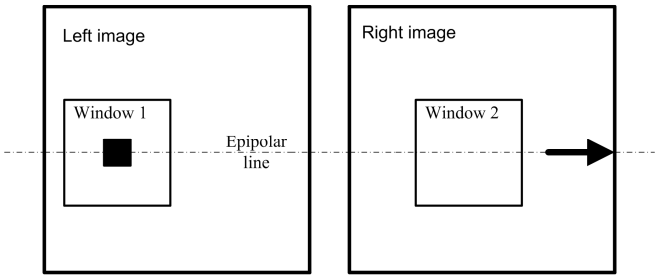

Fig. 10. Windows position of two corresponding points.

The algorithm principle is as follows. Window of hxw pixels is formed. The window central pixel represents the referent point on one of two images from stereo system (Ex. left image). This window is used as referent area to be searched on the second image (i.e. right image). On the second image the same size window is observed on the same height as on the first image. By changing window disparity d the second window is sliding along *u* axe. Measure of two windows intensity likelihood, i.e. cost function, is calculated as Sum of Squared Differences of all pixels intensities in both windows.

$$c(u,v,d)=\sum_{(h,w)}\left[\mathrm{Im}_L(u+h,v+w)-\mathrm{Im}_R(u+h+d,v+w)\right]^2 \quad (26)$$

The value of disparity *d*, for which is obtained minimal value of cost function, gives the position of window which is the best correlated with the reference window. Therefore, the corresponding windows are on the same height on both images, but shifted along *u* axe for:

$$disparity(u,v)=\min c(u,v,d) \quad (27)$$

For the purpose of calibration, stereo system with two standard web color cameras was used. The algorithm was tested with color marker placed on the end-effector of a modular Robix manipulator. Adopted window size is 5x5 pixels. Fig. 11. and Fig. 12. present images from left and right cameras, respectively. A detail of marker found on the second image is shown on Fig. 12. Fig. 13. shows graphical representation of cost function for disparity change along epipolar line, from minimum to maximum value. It is obvious, as it shown on Fig. 13, that a reliable method of determining the corresponding points is obtained by using marker and

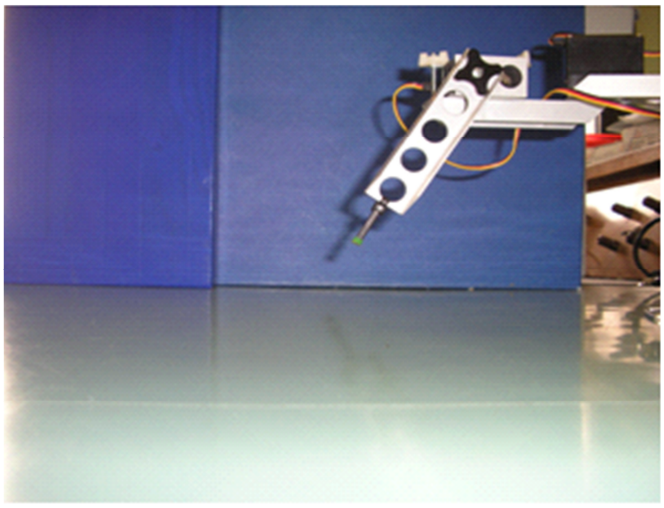

Fig. 11. Image from left camera.

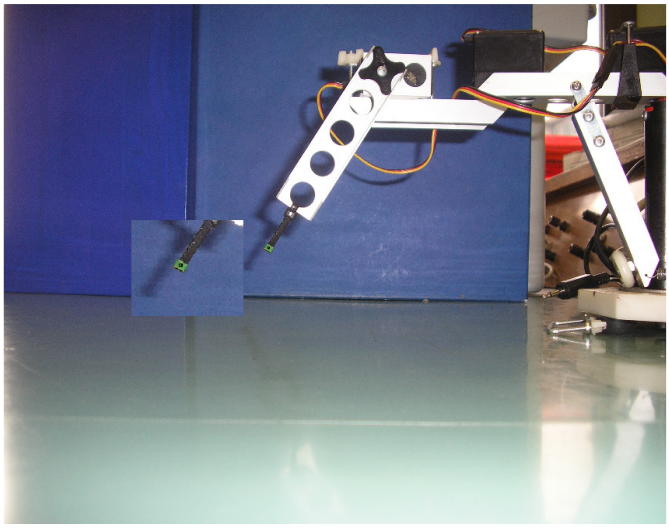

Fig. 12. Image from right camera.

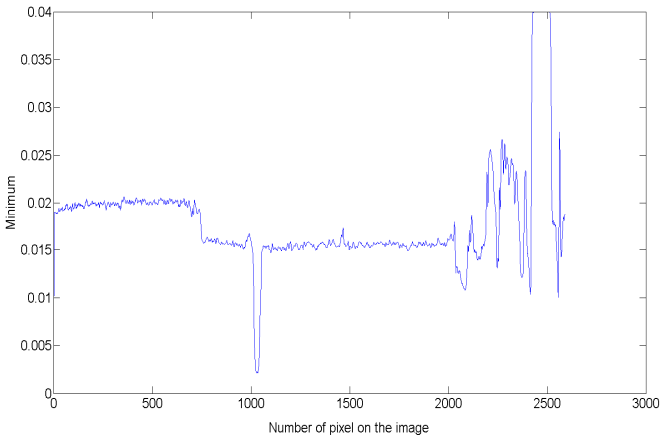

Fig. 13. Graphical representation of cost function for disparity change along epipolar line.

selected cost function. Selected cost function has a pronounced global minimum.

For successfully finding of corresponding points, choice of window size is crucial. In the classical problem of correspondence, if the window size is too small, it increases probability of occurrence of a large number of candidates for correspondence.

This increases probability of wrong selection of corresponding points. On the other hand, if the window size is too large, there is a possibility for error because of a constant value of disparity within the window. Therefore, there is no single recommendation for the best window size. In special cases, even an adaptive window size is suggested, but such algorithms are generally very complex, compute demanding and not widely accepted in practice.

In accordance with previous demonstration, windows size will depend on the size of the marker when it is necessary to determine markers correspondence on two images. Marker is an area with nearly constant intensity (color). By adopting a window smaller than the marker, all the windows, which are contained in the marker, will be detected as candidates. Thus, there will be no single solution for the problem of correspondence. In another case, by adopting a window size larger than the size of marker, by increasing a size of window the value of cost function will not depend on the compliance of marker points. Value of the cost function will be influenced by compliance of region points out of marker. Therefore, it is reasonable assumption that the cost function will have less pronounced global minimum. Further augmentation of window size (relative to marker size) may produce error in correspondence estimation. Points surrounding marker will have significant influence on window correlation compared to marker points. Based on this analysis, it is concluded that the best results are achieved by adopting window size approximately equal to marker size.

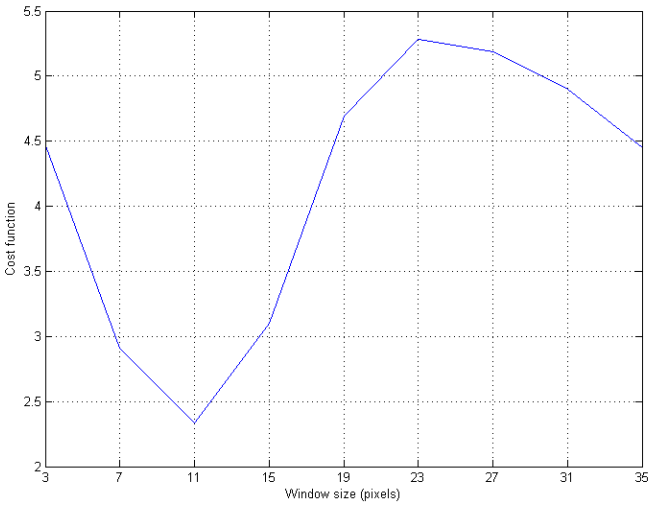

Fig. 14a. Graphical representation of cost function for disparity change along epipolar line.

Assumptions about the window size effects (relative to marker size) on the reliability of the correspondence procedure have been tested in several cases. Window size has been altered for different marker sizes. Diagrams of minimum values change of cost functions with change of window size, for three different sizes of markers, are illustrated on Fig. 14. In order to compare results, values of cost functions, shown on diagrams, are divided with number of pixels that belongs to

window. In this way, the cost functions represents the average inconsistency for every pixel of two windows. The first marker (Fig.14.a.) is the size of 11x11 pixels, the second marker (Fig.14.b) is the size of 21x21 pixels and a third marker (Fig.14.c) is the size of 37x37 pixels. The illustrations confirm that the best results are achieved by adopting that window size is equal to marker size.

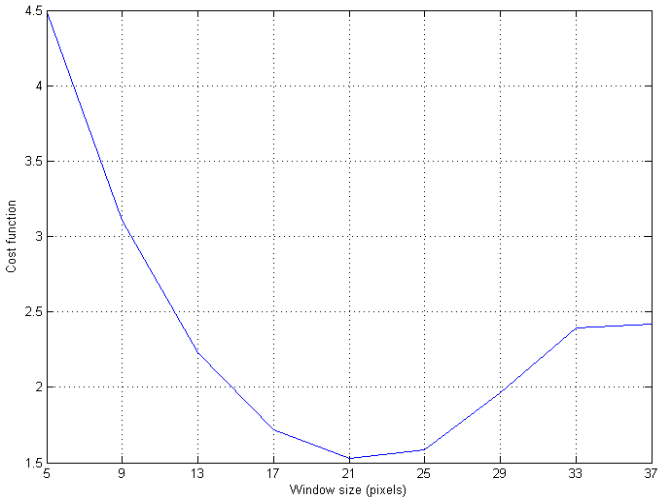

Fig. 14b. Graphical representation of cost function for disparity change along epipolar line.

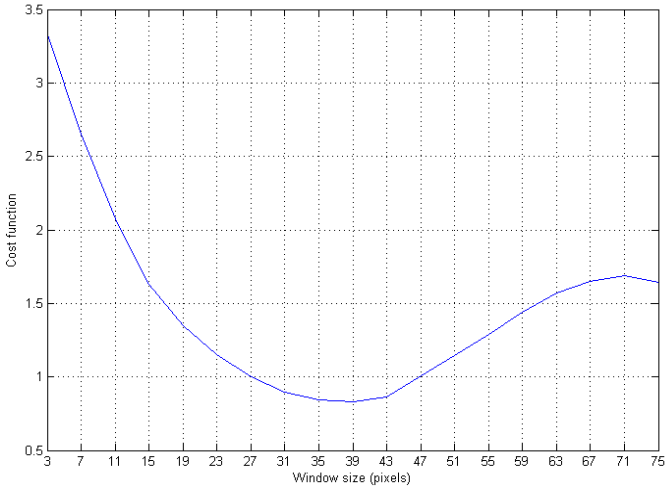

Fig. 14c. Graphical representation of cost function for disparity change along epipolar line.

V. CONCLUSION

An overview of new features for improving the calibration of industrial manipulators using visual systems is presented in this paper. The focus is on the stereo system with parallel optical axes. The practical aspects of using the algorithm based on area correlation are particulaty analyzed. In order to increase the reliability of the corresponding areas determination the next is proposed: set the marker on end-effector of manipulator, choose the window size equal to marker size and scale cost function with total number of pixels for each window. Conducted experiments and shown illustrations confim that presented calibration method allows eficient and reliable manipulator calibration in standard conditions.

REFERENCES

[1] W. Khalil and E. Dombre, *Modeling, Identification and Control of Robots*, Kogan Page Science, 2004.

[2] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses", IEEE J. Robotics and Automation, Vol. RA-3, No. 4, pp. 323-344, 1987.

[3] M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis, and Machine Vision*, Thomson, 2008.

[4] J. Weng, P. Cohen and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 10, pp. 965-981, 1992.

[5] Z. Zhuang, "A Flexible New Technique for Camera Calibration", Technical Report MSR-TR-98-71, 2008.

[6] D. J. Kang, J. E. Ha, and M. H. Jeong , "Detection of Calibration Patterns for Camera Calibration with Irregular Lighting and Complicated Backgrounds", Int. J. of Control, Automation, and Systems, Vol. 6, No. 5, pp. 746-754, 2008.

[7] P. Maric "Computer Vision Systems for the Enhancement of Industrial Robots Flexibility" (Invited Paper), Proc. X Triennial International SAUM Conference on Systems, Automatic Control and Measurements, pp. 80-88, 2010.

[8] X. Tian, X. Zhang, K. Yamazaki and A. Hansel, "A study on three-dimensional vision system for machining setup verification", Int. J. Robotics and Computer-Integrated Manufacturing, No. 26, pp. 46–55, 2010.

[9] B. Paul, *Robot manipulatosr: mathematics, programming, and control*, Cambridge, MIT Press, 1981.

[10] W. Khalil, M. Gautier and Ch. Enguehard, "Identifiable parameters and optimum configurations for robots calibration", Robotica, Vol. 9, pp. 63-70, 1991.

[11] E. Jackson, Z. Lin and D. Eddy, "A global formulation of robot manipulator kinematic calibration based on statistical considerations", Proc. IEEE Conf. on Systems Man and Cybernetics, pp. 3328-3333, 1995.

[12] J. Renders, E. Rossignol, M. Besquetand and R. Hanus, "Kinematic calibration and geometrical parameter identification for robot", IEEE Trans. on Robotics and Automation, Vol. 7, No. 6, pp. 721–732, 1991.

[13] P. Maric and V. Potkonjak, "Geometrical Parameters Estimation for Industrial Manipulators Using Two-step Estimation Schemes", J. Of Intelligent and Robotic Systems, vol. 24, pp. 89-97, 1999.

[14] D. J. Bennett, J. M. Hollerbach, "Autonomous calibration of single-loop closed kinematic chains formed by manipulators with passive endpoint constraints", IEEE Trans. on Robotics and Automation, Vol. 7, No. 5, pp. 597–606, 1991.

[15] W. Khalil, G. Garcia and J. Delagarde, "Calibration of geometrical parameters of robots without external sensors", Proc. IEEE Int. Conf. On Robotics and Automation, Vol. 3, pp.3039-3044, 1995.

[16] S. Bay, "Autonomous parameter identification by optimal learning control", IEEE Control Systems, pp. 56-61, 1993.

[17] Y. Sun and J. Hollerbach, "Active Robot Calibration Algorithm", Proc. ICRA, pp. 1276-1281, 2008.

[18] M. Vincze, J. Prenninger and H. Gander, "A laser tracking system to measure position and orientation of robot end effectors under motion", Int. J. Robotics Research, Vol.13, No. 4, pp. 305-314, 1994.

[19] M. Driels, "Automated partial pose measurement system for manipulator calibration experiments", IEEE Trans. on Robotics and Automation, Vol. 10, No. 4, pp. 430–440, 1994.

[20] H. Zhuang and Z. *S.* Roth, "On Vision-Based Robot Calibration", Proc. SOUTHCON 94, pp. 104-109, 1994.

[21] D. Kosić, V. Đalić, P. Marić, "Robot Geometry Calibration in an Open Kinematic Chain Using Stereo Vision", Proc. International Scientific Conference UNITECH 2010, Gabrovo, Bulgaria, vol. 1, pp. 528-531, 2010.

[22] N. Sebe, M. S. Lew, Robust Computer Vision – Theory and Applications, Kluwer Academic Publishers, volume 26, 2003.