Determination of the Position and Orientation of Rigid Bodies by Using Single Camera Images

Varlık Kılıç and Bülent E. Platin

Abstract - **This study aims to present a new reconstruction method which enables reconstruction of 3D configuration of an object using single camera images. A secondary planar target which is a white circle with two internal black spots, one is located at the center, is used. The proposed reconstruction method is monocular and non-iterative. The elliptical contour and spot center locations of the target in an image is used to determine the 6-DOF configuration parameters of an object, on which the secondary target is rigidly attached. These six parameters represent the three basic translations and three basic rotations of the object with respect to camera coordinate system. The reconstruction algorithm is implemented and tested using an experimental setup composed of a digital imaging system and 6-DOF positioning unit. 512x512 pixels grayscale images are used to determine position and orientation of the secondary target with respect to the camera, which are controlled through the positioning unit. Theoretical accuracy limits of the reconstruction algorithm are evaluated and presented.**

Index Terms— **monocular vision, 3D reconstruction, camera calibration.**

I. INTRODUCTION

In an automated environment when a physical task has to be accomplished, the relative position and orientation of the tool and the workpiece should be known. The need for gathering real-time information about the environment is the natural consequence of making intelligent actions. Depending on the nature of the required information, a variety of tactile, sonar, visual, infrared and proximity sensors are available. Among them the most powerful information can be obtained using image sensors. This is the reason why so many efforts have been spent in domain of machine vision.

This study is a part of the project conducted in the Control Systems Laboratory of Mechanical Engineering Department at METU. The aim of the project is to develop a vision based non-iterative, accurate and robust sensing system. In the previous studies of the project, a reconstruction algorithm was developed and its theoretical limits were tested by Kılınç [1]. In the study conducted by Acar [2], this reconstruction algorithm was implemented using a digital imaging system and the validity of the theoretical limits of the algorithm was examined through a set of experiments. Digital image processing techniques such as automatic thresholding and ellipse detection were implemented by Özkılıç [3, 4] to avoid the human guidance. Finally, the proposed methods by the previous studies are complied and the image processing techniques are optimized or replaced by Kılıç [5] to increase overall speed and robustness of the reconstruction algorithm. An overall algorithm, Figure 1, composed of three main processing blocks, which are target detection, reconstruction and target tracking, is proposed [5], Figure 1.

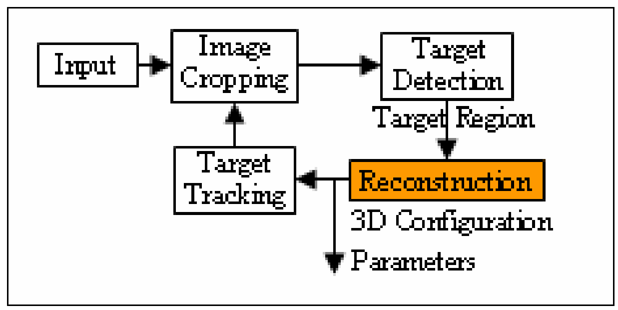

Figure 1: Flowchart of the Overall Algorithm [5]

The idea of utilizing a secondary passive planar target to determine the position and orientation of an object using a single image goes beyond the referred studies [1-5]. Olgac, Gan and Platin [6] used a circular secondary target and orthographic projection model but also made the assumption that the target center lies on the optical axis of the camera. However, the use of a full circular target made the determination of the in-plane rotation inherently impossible. Platin et al. [7] and Olgac et al. [8] applied the same method by replacing the orthographic projection model with a perspective projection model. The major disadvantages of these algorithms were the constraints on the target configuration and the requirements of a priori knowledge of some configuration variables. This method was further improved by Kılınç [1] by adding two internal spots to the circular secondary target, Figure 1, such that all six configuration parameters can be reconstructed. The method developed by Kılınç [1] was implemented on an experimental set-up and tested by Acar [2] and satisfactory results for the reconstruction of all 6 external parameters were obtained.

Manuscript received March 6, 2007.
Varlık KILIÇ is with Ford OTOSAN, Kocaeli, TURKEY phone: 90-262-3156762; e-mail: vkilic@ford.com.tr.
Prof. Dr. Bülent E. Platin is with Middle East Technical University, Mechanical Engineering Department, Ankara, TURKEY, e-mail: platin@metu.edu.tr.

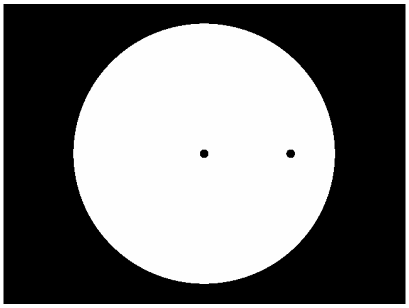

Figure 2: Secondary Planar Target.

This paper is organized to present the reconstruction method proposed by Kılınç [1] and the experimental results achieved by Acar [2]. In the following sections, the secondary target, used coordinate systems, image formation, reconstruction algorithm and the experimental results are examined. And the last section concludes the preceding ones.

II. Secondary Target

Kılınç [1] and Acar [2] used a circular planar target of radius R with two internal spots as a passive secondary target (Figure 3). One of these spots was located at the center of the circle and this spot was used to determine the location of true center of the circle in the image plane. The other spot was placed (at a known distance) r_f away from the center. Both spots were identical circles with radii of r_0. While the boundary of the elliptical image of the projected circular target was used to determine the two rotation parameters and the depth information, the image of the center spot was used for two translation parameters, and the image of the second spot was used for the remaining rotation parameter.

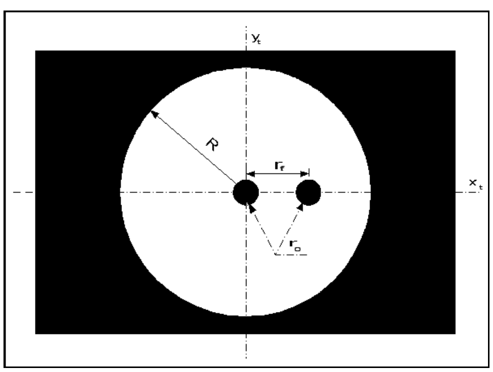

Figure 3: Secondary Planar Target.

III. 3D Configuration Parameters

The purpose of the reconstruction algorithm is to compute the 3D configuration of the target coordinate system $(O_t x_t y_t z_t)$ with respect to camera coordinate system $(O_i x_i y_i z_i)$ as presented in Figure 4. Therefore, the reconstruction parameters which determine the target plane configuration with respect to the camera can be defined as follows:

ϕ : Tilt angle of the camera
θ : Pan angle of the camera
d : Distance of the target plane along the optical center
α : Tilt angle of the target
β : Yaw angle of the target
γ : In-plane rotation of the target

These six parameters are to be reconstructed using the image points of the elliptical contour and spot centers of the secondary target formed on the image plane.

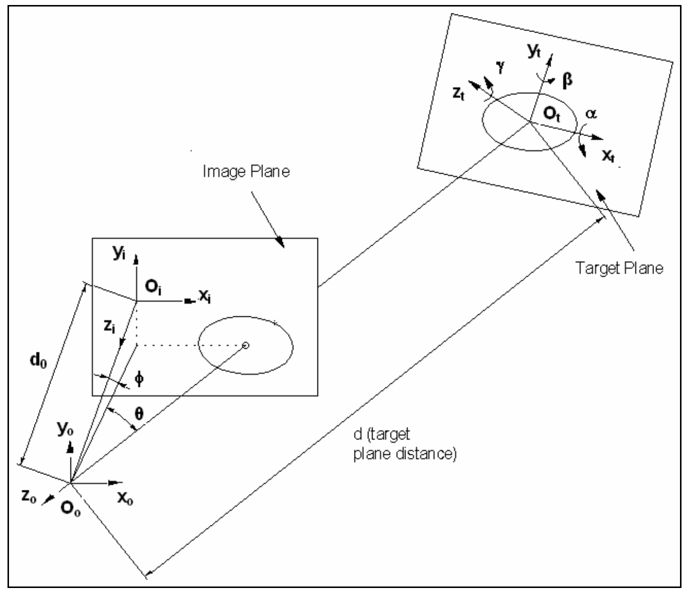

Figure 4: 3D configuration parameters

IV. Image Formation And Processing

Although the relation between the features of the target's image and the target configuration parameters is complex to realize due to nature of the perspective projection, it is possible to identify the relation between the reconstruction parameters and image features. The position of the center spot with respect to image center is related with the tilt and pan motion of the camera and the aspect ratio, alignment and the position of the geometric center of the elliptical contour with respect to the center spot are related with pan and tilt motion of the target plane. The distance of the target plane to the image plane contributes to the area of the elliptical region. And position of the outer spot with respect to the center spot is obviously related with the in-plane rotation of the target.

The digital image of the target obtained using the experimental setup constructed by Acar [2] is presented in Figure 5. The setup is composed of a frame grabber, a CID camera and a 6-DOF positioning unit. 512x512 pixel grayscale images of the scenes including the secondary target are used for the experiments.

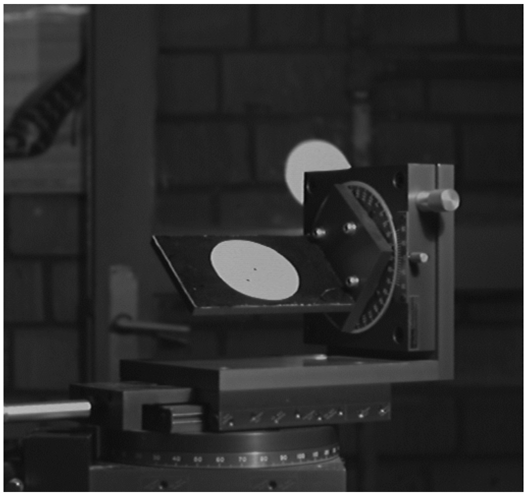

Figure 5: Image of the target and the experimental setup

A real-time target detection algorithm is proposed by Kılınç [5] to identify the feature points of the target, namely the elliptical contour and the spot center positions. And these features, Figure 6, are utilized by the reconstruction algorithm which is discussed below in detail.

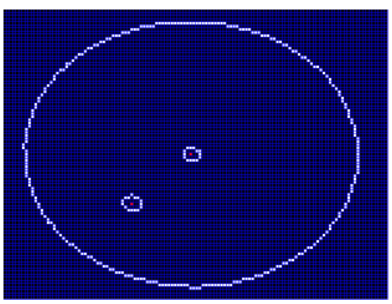

Figure 6: Feature points detected in the target image

V. RECONSTRUCTION ALGORITHM

The reconstruction method proposed by Kılınç and presented in this study is aimed to reconstruct the parameters, ϕ, θ, d, α, β and γ seen in Figure 4 using the image features presented in Figure 7. These parameters define the transformation matrix between the image and target coordinate systems, $H_{it}^{(i,t)}$, can be decomposed to a series of basic rotation and translation matrices as follows:

$$H_{it}^{(i,t)} = T_z(d_0)R_x(-\phi)R_y(-\theta)T_z(-d)R_x(\alpha)R_y(\beta)R_z(\gamma)$$

where

$T_z(d_0)$: Translation d_0 along z_i axis.

$R_x(-\phi)$: Rotation $-\phi$ about x_i axis.

$R_y(-\theta)$: Rotation $-\theta$ about y_i axis.

$R_x(\alpha)$: Rotation α about x_t axis.

$R_y(\beta)$: Rotation β about y_t axis.

$R_z(\gamma)$: Rotation γ about z_t axis.

$T_z(d)$: Translation d along z_t axis.

According to the developed algorithm by Kılınç [1]:

- First, ϕ and θ angles are determined from the centroid of the center spot.
- Then, d, α and β parameters are determined from the contour information for the image ellipse.
- Finally, the in-plane rotation γ of the target is determined from the centroid of the outer spot.

A. *Determination of ϕ and θ*

The image of the central spot of the target represents the image of the true center of the circle on the image plane and its image coordinates are determined only by the $R_x(-\phi)$ and $R_y(-\theta)$ rotations. The other rotations and translations will not affect the position of center spot on the image plane. Since these coordinates (x_0, y_0) are known for a given image, ϕ and θ can be obtained using the following equations:

$$\phi = \tan^{-1}\left(-y_0/d_0\right)$$
$$\theta = \tan^{-1}\left(-x_0 \cos\phi/d_0\right)$$

Once ϕ and θ are determined, then the direction in which the target center is translated can be found. It is possible to define a virtual image plane of a fictitious camera whose optical axis (z_{v1} axis) goes through the target center (Figure 8). A coordinate system is defined in this virtual image plane and is denoted by $\left(O_{v1}x_{v1}y_{v1}z_{v1}\right)$.

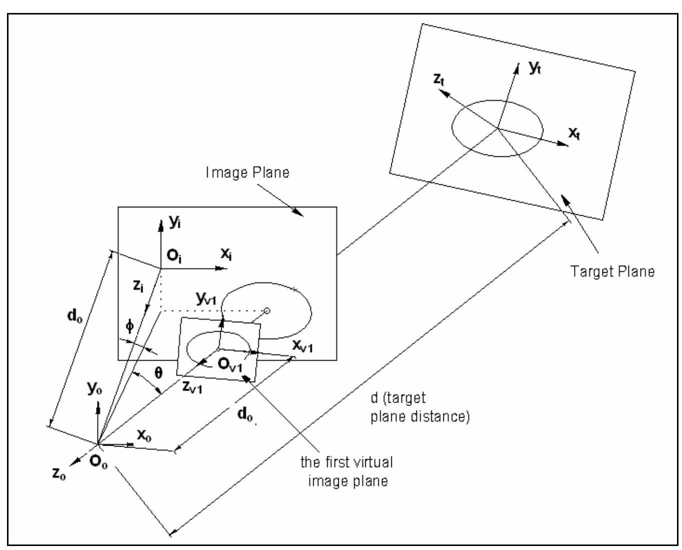

Figure 8: First Virtual Image Plane

If one finds the coordinates of the image points in the first virtual image plane, then it will be possible to apply the target centered solution, will be discussed in the following section, to find d , α and β .

Since ϕ and θ are known, it is possible to write the transformation matrix $H_{v1i}^{(v1,i)}$ from the virtual image plane to the actual image plane as follows

$$H_{v1i}^{(v1,i)} = T_z(d_0)R_y(\phi)R_z(\theta)T_z(-d_0)$$

Then the coordinates of any point lying on the image plane can be expressed in the virtual image plane coordinate system by the following homogeneous transformation as

$$\begin{bmatrix} X_{V1} & Y_{V1} & Z_{V1} & 1 \end{bmatrix}^T = H_{v1i}^{(v1,i)} \begin{bmatrix} x_i & y_i & 0 & 1 \end{bmatrix}^T$$

where X_{V1}, Y_{V1} and Z_{V1} are the coordinates of the image point lying on the actual image plane expressed in the first virtual image coordinate system whereas x_i and y_i are the coordinates of this image point in the actual image coordinate system. On the other hand, the projection of a point in the actual image plane to the first virtual image plane can be achieved using the following perspective transformation:

$$\begin{bmatrix} x_V \\ y_V \end{bmatrix} = \begin{bmatrix} (X_{V1}d_0/(d_0 - Z_{V1}) \\ (Y_{V1}d_0/(d_0 - Z_{V1}) \end{bmatrix}$$

where x_V and y_V are the coordinates of the projected point on the first virtual plane coordinate system. All of the feature points on the actual image plane, elliptical contour and spot centers are projected to first virtual image plane and reconstruction problem is reduced to target centered case, for which only d , α and β are unknown.

B. Determination of α , β and d

After projecting all of the feature points on the actual image plane to first virtual image plane the solution for d , α and β are reduced to target centered solution for which ϕ, θ and γ are zero.

The transformation matrix between the first virtual image plane coordinate system and the target plane coordinate system, $H_{v1t}^{(v1,t)}$, can be decomposed in to a set of basic translations and rotations as follows:

$$H_{vt}^{(v,t)} = T_z(d_0 - d)R_x(\alpha)R_y(\beta)$$

And the coordinates of any point on the target plane expressed in virtual image plane coordinate system

$$\begin{bmatrix} X_{V1} & Y_{V1} & Z_{V1} & 1 \end{bmatrix}^T = H_{v1t}^{(v1,t)} \begin{bmatrix} x_t & y_t & 0 & 1 \end{bmatrix}^T$$

And the projection of a point on the target plane to the virtual image plane can be obtained as follows

$$\begin{bmatrix} x_V \\ y_V \end{bmatrix} = \begin{bmatrix} (X_{V1}d_0/(d_0 - Z_{V1}) \\ (Y_{V1}d_0/(d_0 - Z_{V1}) \end{bmatrix}$$

Inverting the above set of equations on can write the coordinates of a point on target plane in terms of its image points on the virtual image plane as expressed in the following equations:

$$x_t = \frac{(d\cos\alpha)x_v}{(-\sin\beta)x_v + (\sin\alpha\cos\beta)y_v + (\cos\alpha\cos\beta)d_0}$$

$$y_t = \frac{d\left[(-\sin\alpha\sin\beta)x_v + (\cos\beta)y_v\right]}{(-\sin\beta)x_v + (\sin\alpha\cos\beta)y_v + (\cos\alpha\cos\beta)d_0}$$

If this inversion is applied to the feature points corresponding to the elliptical contour projected to the virtual image plane, the following circle equation can definitely be written, since these feature points representing the circular boundary of the secondary target.

$$x_t^2 + y_t^2 = R^2$$

If the previously obtained expressions for x_t and y_t in terms of virtual image coordinates are substituted, the above equation turned out to be a conic equation, representing an ellipse if $B^2 - 4AC < 0$ holds, which can be expressed as follows

$$Ax_v^2 + Bx_v y_v + Cy_v^2 + Dx_v + Ey_v + 1 = 0$$

where

$$A = \frac{-\left[d^2 + (d^2 - R^2)\tan^2\beta/\cos^2\alpha\right]}{(Rd_0)^2}$$

$$B = \frac{2(d^2 - R^2)\tan\alpha\tan\beta}{(Rd_0)^2\cos\alpha}$$

$$C = \frac{-\left[d^2 + (d^2 - R^2)\tan^2\alpha\right]}{(Rd_0)^2}$$

$$D = \frac{-2\tan\beta}{d_0\cos\alpha} \quad \text{and} \quad E = \frac{2\tan\alpha}{d_0}$$

The parameters A, B, C, D and E can be obtained by least squares ellipse fitting using the image coordinates of projected points of the elliptical contour on actual image plane to the virtual image plane. And the parameters, d α and β can be obtained by inversions of the above equations. However, due to the fact that D and E values come out to be very small in magnitude in case of α and β angles are close to $0°$, an intermediate coordinate transformation is used to obtain a canonical representation of the ellipse. Later the parameters d, α and β can be obtained using the following derived equations:

$$d = Rd_0\sqrt{-\frac{A+C}{2}-\frac{\sqrt{(A-C)^2+B^2}}{2}}$$

$$\alpha = \operatorname{sgn}[E]\tan^{-1}\left[\sqrt{\frac{-(CR^2d_0^2+d^2}{d^2-R^2}}\right]$$

$$\beta = \operatorname{sgn}[D]\tan^{-1}\left[\cos\alpha\sqrt{\frac{-(AR^2d_0^2+d^2}{d^2-R^2}}\right]$$

C. Determination of γ

In order to get in plane rotation of the target, a new virtual image plane is defined such that it is parallel to the target plane and its origin lying on the line O_oO_t. Also to ease the solution procedure, one more constraint to the location of the virtual image plane is added. It is positioned such that the center of the outer spot lies on the intersection line between the actual image plane and the virtual image plane, Figure 9. The second virtual image plane will be parallel to the target plane and as a consequence z_t will be also parallel to z_{v2}.

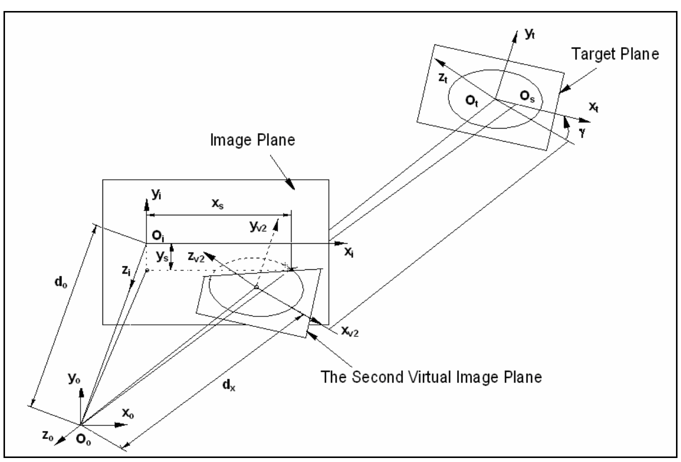

Figure 9: The virtual image plane for the calculation of in-plane rotation

It is possible to write down the transformation matrix from this second virtual image plane to the actual image plane as:

$$H_{v2i}^{(v2,i)} = R_y(-\beta)R_x(-\alpha)T_z(d_x)R_y(\theta)R_x(\phi)T_z(-d_0)$$

where the parameter d_x is used to locate the center of the outer spot to lie on the intersection of the second virtual image plane and the actual image plane. And, for the image position (x_s, y_s) it is possible to write the following equation since this point is on the intersection line of second virtual image plane and the actual image plane.

$$\begin{bmatrix} x_s^{\iota} & y_s^{\iota} & 0 & 1 \end{bmatrix}^T = H_{v2i}^{(v2,i)}\begin{bmatrix} x_s & y_s & 0 & 1 \end{bmatrix}^T$$

The unknowns, d_x, y_s^{ι} and y_s^{ι} can easily be obtained through the above equation. As the second virtual image plane and the target plane are parallel, γ can be calculated using a double argument arctangent function as follows:

$$\gamma = \tan_2^{-1}\left(\frac{y_s^{\iota}}{x_s^{\iota}}\right)$$

VI. Experimental Results

The 3D reconstruction method was tested using a positioning unit and associated absolute reconstruction errors were evaluated. The reconstruction parameters ϕ, θ, d, α, β and γ obtained using the new algorithm are converted to T_x, T_y, T_z, R_x, R_y and R_z. T_x, T_y and T_z represents the three translations along the x, y and z axes of the camera coordinate system respectively. R_x, R_y and R_z represents the three rotation angles about x, y and z axes for rotated frame based 213 sequence. Mainly two sets of experiments have been conducted using the full range of the experimental setup.

- Target located at the maximum possible T_z
- Target located at the minimum possible T_z

Sets of experiments were conducted mainly varying one parameter at a time while keeping the others constant. The occurred reconstruction errors for each parameter and for each experiment set were tabulated and averaged. The theoretical limits of the algorithm were verified as follows:

$$10° < |\alpha| < 80°$$
$$10° < |\beta| < 80°$$
$$500mm < d < 1300mm$$

ISBN:978-988-98671-5-7

Results of the experiments can be summarized as follows

- T_x and T_y are reconstructed successfully in the whole range of rotations, because they are reconstructed using central spot.
- Reconstruction errors for T_z are low but as the target moves away from the camera, error increases.
- R_x and R_y are determined successfully but the errors increase when the rotations get closer to the ultimate limits.
- R_z is reconstructed successfully since the location of outer spot is used for reconstruction.

The average reconstruction errors obtained in the experiments are presented in Table 1.

Table 1: Average Reconstruction Errors

Parameter	Average Reconstruction Error
T_x	0.5mm
T_y	0.5mm
T_z	1.5mm
R_x	0.4°
R_y	0.4°
R_z	0.5°

VII. Conclusion

An alternative approach to the stereo vision has been presented to reconstruct the configuration of rigid bodies by means of the image formation of a passive secondary target. The secondary target is planar and rigidly attached to the object of interest. The system requires a single camera that supplies gray scale images and an image processing sequence to detect the target and its features. The new algorithm enables the reconstruction of 6 parameters that define the 3D configuration of the target using the features of the target detected in single image by target detection process.

The main activity in this study was to develop a reconstruction algorithm and to test its accuracy limits using a digital imaging system. It has been proved that the method works properly within the determined limits. These limits correspond to highly rotated and unrotated configurations. When the rotations are close to $80°$, a very thin ellipse has been obtained. In the opposite case when the target rotations are close to $10°$, the α and β parameters cannot be successfully reconstructed. However satisfactory results are obtained in the reconstruction of the other parameters. The target has been tested in the available range of experimental setup. And in these regions, the algorithm provides satisfactory results and can be used to determine the configuration parameters of a rigid object in an automated environment.

Finally, the main advantages of the new reconstruction method can be summarized as follows:

- Monocular vision
- Simple non-iterative solution
- Capability of reconstructing 6 configuration parameters
- Uniqueness of solution

Acknowledgment

The authors would like to express their appreciation to Mr. İ. Alpar Kılınç, Dr. Ercan U. Acar, and Ms. Sibel Özkılıç. Without their contributions, this study would not have been what it is now.

References

[1] Kılınç, İ. A., (1994), "Determination of the Position and Orientation of Rigid Bodies by Using Single Camera Images", M.Sc. Thesis, Mechanical Engineering Department, Middle East Technical University, Ankara, Turkey.
[2] Acar, E. U., (1995), "Experimental Investigation and Implementation of a 3D Configuration Reconstruction Algorithm for an Object Using a Single Camera Image", M.Sc. Thesis, Mechanical Engineering Department, Middle East Technical University, Ankara, Turkey.
[3] Özkılıç, S. and Platin B. E., (2002), "Elliptical Target Recognition for Automation of a 3-D Configuration Reconstruction Algorithm of an Object Using a Single Camera Image", Proceedings of the 6th International Conference on Mechatronic Design and Modeling, September 2002, Cappadocia, Turkey.
[4] Özkılıç, S., (2003), "Performance Improvement of a 3D Configuration Reconstruction Algorithm for an Object Using a Single Camera Image", M.Sc. Thesis, Mechanical Engineering Department, Middle East Technical University, Ankara, Turkey.
[5] Kılıç, V., (2005), "Performance Improvement of a 3D Reconstruction Algorithm Using a Single Camera Image", M.Sc. Thesis, Mechanical Engineering Department, Middle East Technical University, Ankara, Turkey.
[6] Olgaç, N., Gan, Z. and Platin, B. E., (1989), "3D Reconstruction of Object Configurations by Hybrid Projection Analysis Using a Single Camera Image", Proceedings of the 1st National Conference on Applied Mechanisms and Robotics, Vol. 1, Paper No. 89AMR-4C-6, November, Cincinnati, USA.
[7] Platin, B. E., Gan, Z. and Olgaç, N., (1990), "3D Object Configuration Sensor Utilizing Single Camera Images", ASME Paper No. 90-WA/DSC22, presented at the Winter Annual Meeting, November 25-30, Dallas, USA.
[8] Olgaç, N., Craig, P. D. and Platin, B. E., (1991), "Advancements in Utilizing Monocular Object Configuration Sensor, Part II", Advances in Instrumentation, DSC-Vol. 30, ASME Book No. H00691, pp. 21-24, also presented at the Winter Annual Meeting, December 1-6, Atlanta, USA.