Application of Stereo Vision on Determination of End-Effector Position and Orientation of Manipulators

M. H. Korayem[1], M. Irani[2], A. Hashemi[1]

[1]Robotics Research Laboratory, School of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran

[2]Department of Mechanical Engineering, Faculty of Engineering, University of Kashan, Kashan, Iran

hkorayem@iust.ac.ir

***Abstract*-the main idea of this paper is to present application of stereo vision approach to determine position and orientation of end-effector of mechanical manipulators. Stereo vision is a technique based on inferring depth of an object from two cameras. A stereo vision test set up that is made of two parallel CMOS cameras is provided and USB port is used to transmit image data to the computer. Experimental tests are exerted for 6R mechanical manipulator where the end-effector moves on circular trajectory and Scout mobile robot to move a pre-defined optimal trajectory. The image processing algorithm, matching algorithm and triangulation calculations are optimized to decrease time required for matching procedure. The results illustrate that the measurement error is less than 1 mm. Because of rapid image processing algorithm; proposed method can be used to detect end-effector in dynamic trajectories. This stereo vision set up can be applied easily as measuring unit in a closed loop control system.**

Keywords-Manipulators; End-Effector; Positioning; Orientation; Stereo Vision

I. INTRODUCTION

Determining end effector position for mechanical manipulators is an important key point in industrial applications. This problem can be considered as the process of determining the position and orientation of the end-effector of robotic manipulators with respect to a global frame. In recentyears, various methods and solutions have been proposed and applied to this problem. Earliest solution to this problem is to do relative position/orientation measurement using sensors such as odometer, or inertia measurement unit (IMU). Relative position measurement known as dead reckoning is the process of tracking the current position of end-effector based on the integration of the path that has been previously traveled by the end-effector [1]. The most common sensors are IMU that is made of accelerometers and gyroscopic rate sensors. IMU have the advantage of high frame-rate and relatively small latency, but only provide relative motion and are subject to significant drift (or significant cost). IMU sensors are often complemented by global positioning sensors (GPS) [2].

Among dead reckoning, odometry is probably the most used in mobile robots as it provides easy and cheap real time position information by the integration of incremental motion information over time. Unfortunately, this integration causes an accumulation of errors during the movement of the robot [3].

Another approach to the end-effector position measurement problem is using active beacons [1]. Measurement using active beacons has been traditionally used in the global positioning system (GPS) for the localization of ships and airplanes [1]. There are two types of active beacon systems: trilateration and triangulation [4, 5, and 6]. Trilateration is the determination of the end-effector position based on distance measurements to known beacon sources. In trilateration systems there are usually three or more transmitters mounted at known locations in the workspace and one receiver fixed onto the end-effector. An example of the active beacons localization system that makes use of trilateration is [7].

Triangulation is another way of determination of the robot end-effector position that is based on the measurements of the angle from the beacon to the end-effector or robot heading. The distance to at least one of beacons and its location must also be known. Similar to the trilateration method, three or more beacon readings must be obtained to do triangulation. Note that it is also possible to do triangulation with two beacons if the angles from these beacons to the end-effector or robot heading, the distances from these beacons to the robot, and the locations of these beacons are known [8]. An example of the triangulation with two beacons is the North Star kit [8]. Unlike dead reckoning, the errors in active beacons systems will not grow unbounded. However, the accuracy is highly dependent on the size of its random errors and precise placement of the beacons in the environment [1].

For the reason of disadvantages of relative measurement methods, sensors that provide a measure of absolute position are extremely important. These methods are named as absolute position measurements (reference-based systems). Magnetic compasses and vision systems [9, 10, and 11] are two approaches of this category. Chang [10] presented a vision-based navigation and localization system using two biologically-inspired scene understanding models which are studied from human visual capabilities: 1) Gist model which captures the holistic characteristics and layout of an image and 2) Saliency model which emulates the visual attention of primates to identify conspicuous regions in the Image.

The main idea of this paper is using stereo vision approach and proposes a simple triangulation approach for determining position and orientation using a simple and inexpensive setup. Also based on stereo vision, static and dynamic tests are implemented for 6R manipulator to show accuracy and speed of computations of proposed approach. Stereo vision represents a way for reconstructing three-dimensional information from the surrounding environment and is therefore of vital importance for a large number of vision applications [12]. Unfortunately, the key step in stereo vision is stereo matching problem, cannot be regarded as solved. Factors that complicate the matching process include image noise, untextured regions and the occlusion problem [12]. Two methods are used for matching process, local and global methods [13, 14]. In [15] an algorithm to detect depth discontinuities from a stereo pair of images is presented. The algorithm matches individual pixels in corresponding scan line pairs while allowing occluded pixels to remain unmatched, and then propagates the information between scan lines by means of a fast postprocessor. The algorithm handles large untextured regions, uses a measure of pixel dissimilarity that is insensitive to image sampling, and prunes bad search nodes to increase the speed of dynamic programming. The computation is relatively fast, taking about 1.5 microseconds per pixel per disparity on a workstation. Approximate disparity maps and precise depth discontinuities (along both horizontal and vertical boundaries) are shown for five stereo images.

II. STEREO VISION

3D measurement approach based on human vision system is named stereo vision, but in this vision approach human eyes are replaced with a pair of slightly displaced parallel cameras. 2D position information of an object is determined using one camera and third coordinate of 3D position can be computed by comparing two images. This comparison is implemented by means of technique based on stereo vision approach that is technique aimed at inferring depth of object from two cameras. Using stereo vision, with two images one can infer depth by means of triangulation method if it is available to find corresponding (homologous) points in two images [16]. According to Fig. 1, the amount to which a single pixel is displaced in the two images is called disparity, so a pixel's disparity is inversely proportional to its depth in the scene. Before computing disparity, rectification process must be applied to remove lens distortions and stereo pair to be turned in standard form.

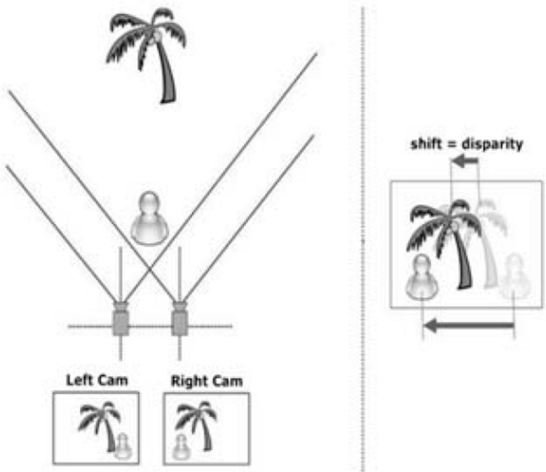


Figure 1. Relation between depth and disparity. [17]

The challenging part is to compute the disparity of each pixel and this task is known as the stereo matching problem. Two methods are used for matching process, local and global methods, local and global/semi global methods. Local algorithms use the simple WTA (Winner Takes All) disparity selection strategy but reduce ambiguity (increasing the signal to noise ratio (SNR)) by aggregating matching costs over a support window (aka kernel or correlation window).

Global (and semi-global) algorithms search for disparity assignments that minimize an energy function over the whole stereo pair using a pixel-based matching cost (sometime the matching cost is aggregated over a support) [16]. The Middlebury stereo evaluation site [18] provides a framework and a dataset for benchmarking novel algorithms. An overview of stereo vision approach is shown in Fig. 2.

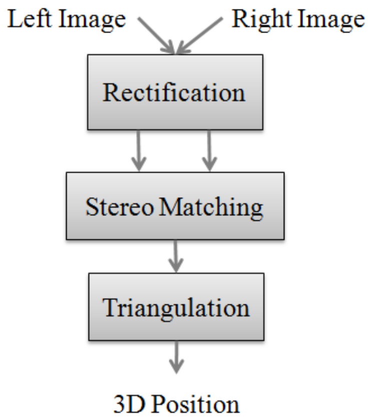


Figure 2. Overview of stereo vision approach

III. TRIANGULATION

Given the disparity of any pixel, triangulation procedure computes the position of the correspondence in the 3D space. Consider Fig. 3 that shows two images, object, image frame (local) and reference frame. Parameters that are shown in this figure are:

- OXYZ: Reference frame
- $O_1x_1y_1z_1$: Image frame of first camera
- $O_2x_2y_2z_2$: Image frame of second camera
- F_1, F_2: focal lengths
- β_1,β_2: Rotation angles of optical axis
- u_1, v_1: Pixel coordinate in first image
- u_2, v_2: Pixel coordinate in second image

Transformation between object position in reference frame and each local frame can be mentioned as Eq. 1:

$$M^T = RTP^T \quad (1)$$

P and M are local and global position of object and have following forms:

$$P = [x, y, z, 1] \quad (2)$$

$$M = [X, Y, Z, 1] \quad (3)$$

Also R and T matrixes in Eq. 1 are rotation and translation matrixes, respectively and have following forms:

$$R=\begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$T=\begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

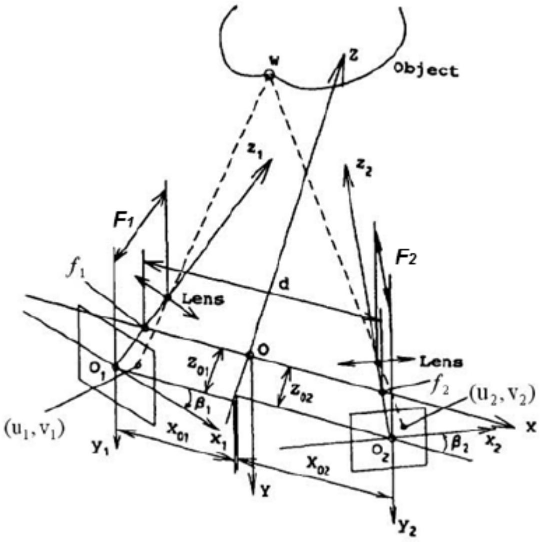


Figure 3. Coordinates in Stereo Vision [17]

In these equations β is the rotation of optical axis of considered camera and [X_0, Y_0, Z_0] is position of center of local frame in reference frame. Consider F is focal length, perspective relations for any image can be written as:

$$\frac{x}{u}=-\frac{z-F}{F} \quad ; \quad \frac{y}{v}=-\frac{z-F}{F} \quad (6)$$

Introducing Eq. 6 into Eq. 1, for two images, 3D position of object in reference frame can be computed as:

$$X=\frac{u_1F_1-m_1X_{01}-n_1\left(Z+Z_{01}\right)}{m_1}$$
$$Y=v_1-Y_{01}-\frac{v_1}{F_1}\left[\left(X-X_{01}\right)\sin\beta_1+\left(Z+Z_{01}\right)\cos\beta_1\right]$$
$$Z=\frac{m_1m_2\left(X-X_{01}\right)+m_1F_2u_2-m_2F_1u_1+m_2n_1Z_{01}-m_1n_2Z_{02}}{m_1n_2-m_2n_1} \quad (7)$$

New parameters in Eq. 7 are as follows:

$$m_1=F_1\cos\beta_1+u_1\sin\beta_1$$
$$m_2=F_2\cos\beta_2-u_2\sin\beta_2$$
$$n_1=u_1\cos\beta_1-F_1\sin\beta_1$$
$$n_2=u_2\cos\beta_2+F_2\sin\beta_2 \quad (8)$$

[u_i, v_i] is coordinate of object project into i-th image. If two cameras be parallel with $\beta=0$ then following simple relations can be considered:

$$Y_{01}=Y_{02}=0$$
$$\beta_1=\beta_2=0$$
$$F_1=F_2=F$$
$$X_{01}=X_{02}=\frac{B}{2} \quad (9)$$

Finally Eq. 7 can be rewritten as follow:

$$X=\frac{F-Z}{F}u_1-\frac{B}{2}$$
$$Y=\frac{F-Z}{F}v_1$$
$$Z=F-\frac{BF}{u_1-u_2} \quad (10)$$

u_1-u_2 in Eq. 10 is the difference between the x coordinate of two corresponding pixels that was named disparity. The disparity of an object that is in infinity is equal to zero. According to above equations parameters that affect accuracy of measurement are:

- Accuracy of stereo correspondence
- Distance between cameras and end effector
- Initial calibration

IV. END EFFECTOR ORIENTATION

According to triangulation method and using stereo vision setup one can calculate the position of three point of the robot end effector which are determine in Fig. 4 as P_1, $P_2$and P_3.

The direction vectors $\vec{x}_b$, $\vec{y}_b$ and $\vec{z}_b$ can be represented as:

$$\vec{x}_b=\frac{P_2-P_1}{\left\|P_2-P_1\right\|} \quad (11)$$

$$\vec{z}_b=\frac{\left(P_2+P_1\right)/2-P_3}{\left\|\left(P_2+P_1\right)/2-P_3\right\|} \quad (12)$$

$$\vec{y}_b=\vec{z}_b\times\vec{x}_b \quad (13)$$

Considering the rotation matrixes $R_{Z,\phi}$, $R_{Y,\theta}$ and $R_{X,\psi}$ for roll, pitch and yaw angles, rotation matrix between reference and body frames can be mentioned as below equations:

$$\left[\vec{x}_b \quad \vec{y}_b \quad \vec{z}_b\right]=\begin{bmatrix} X_x & Y_x & Z_x \\ X_y & Y_y & Z_y \\ X_z & Y_z & Z_z \end{bmatrix} \quad (14)$$

$$\begin{bmatrix} X_x & Y_x & Z_x \\ X_y & Y_y & Z_y \\ X_z & Y_z & Z_z \end{bmatrix}=\begin{bmatrix} C_\phi C_\theta & C_\phi S_\theta S_\psi-S_\phi C_\psi & C_\phi S_\theta S_\psi+S_\phi S_\psi \\ S_\phi C_\theta & S_\phi S_\theta S_\psi-C_\phi C_\psi & S_\phi S_\theta C_\psi-C_\phi S_\psi \\ -S_\theta & C_\theta S_\psi & C_\theta C_\psi \end{bmatrix} \quad (15)$$

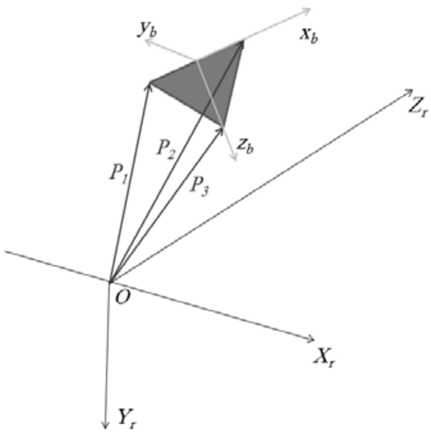


Figure 4. three point on end effector

Where $C_i = cos(i)$ and $S_i = sin(i)$

Therefore the roll, pitch and yaw angles of the robot end-effector are determined as follows:

$$\phi = A\tan 2\left(\frac{X_y}{X_x}\right)$$
$$\theta = A\tan 2\left(\frac{-X_z}{\sqrt{1-X_z^{\ 2}}}\right)$$
$$\psi = A\tan 2\left(\frac{Y_z}{Z_z}\right) \quad (16)$$

V. EXPERIMENTAL RESULTS

A. 6R manipulator

An experimental setup made of two cameras that are shown in Fig. 5 is used to implement stereo vision approach to measure end-effector position of 6R manipulator. Characteristics of cameras are mentioned in table 1 and a view of 6R manipulator is shown in Fig. 6. Also Denavit-Hartenberg parameters of this manipulator are listed in table 2.

A red LED is moved on a linear path according to (17) and measured path andx, y and z histories are displayed on Figs. 7 to 10.

$$\Delta x = 40\,mm, \Delta y = 100\,mm, \Delta z = 10\,mm.\Delta t = 6\ sec \quad (17)$$

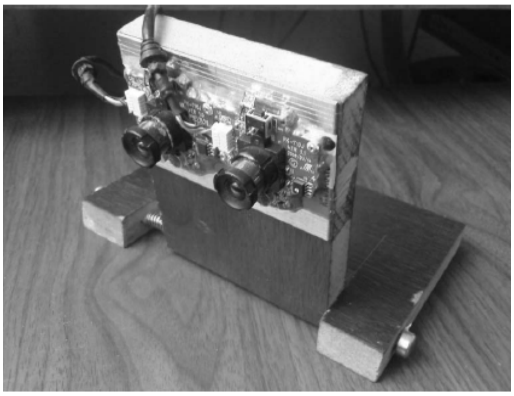


Figure 5. Stereo vision system

TABLE I. CHARACTERISTICS OF CAMERAS

ITEM	VALUE
Image sensor	¼" CMOS, 640×480 (350k pixels)
Frame rate	30 fps @ 320×240
Lens	F=2.2, f=4.6 mm
View angle	65 degree
Focus range	10 cm to infinity
Exposure control	automatic
White balance	automatic
Flicker control	50Hz, 60Hz and none
Interface	USB 2.0 port

TABLE II. DENAVIT-HARTENBERG PARAMETERS

Joint	a_i [mm]	d_i [mm]	α_i^o	Θ_i
1	36.5	438	-90	θ_1
2	251.5	0	0	θ_2
3	125	0	0	θ_3
4	92	0	90	θ_4
5	0	0	-90	θ_5
6	0	152.8	0	θ_6

$\alpha + \beta = \gamma.$ (1) (1)

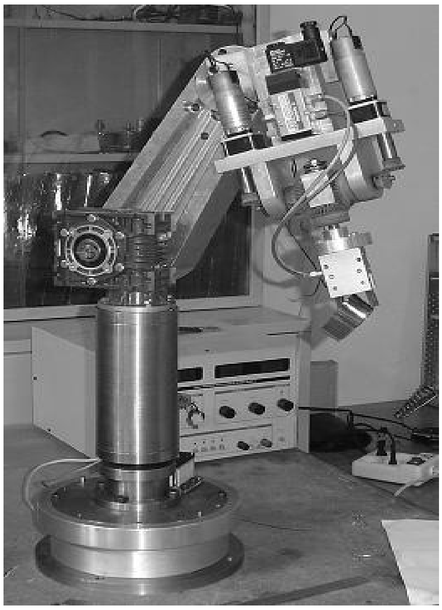


Figure 6. 6R manipulator

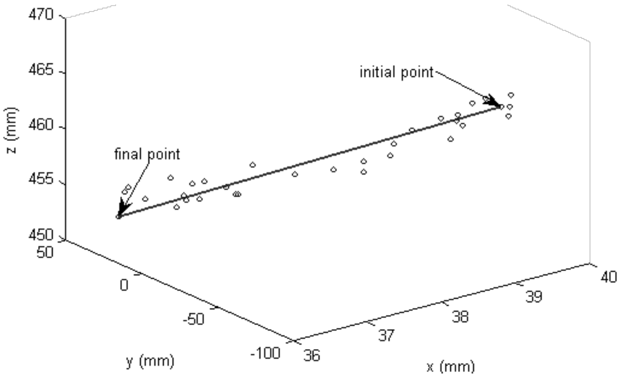


Figure 7. Measurement result for linear path

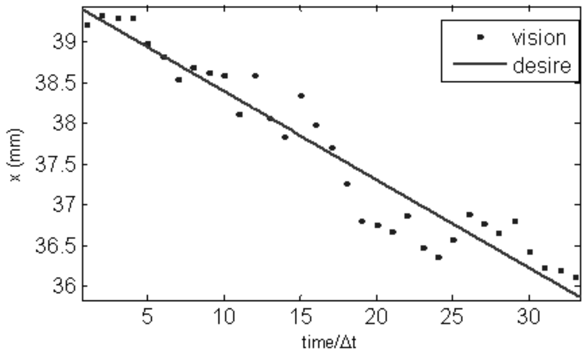


Figure 8. Measured x coordinate

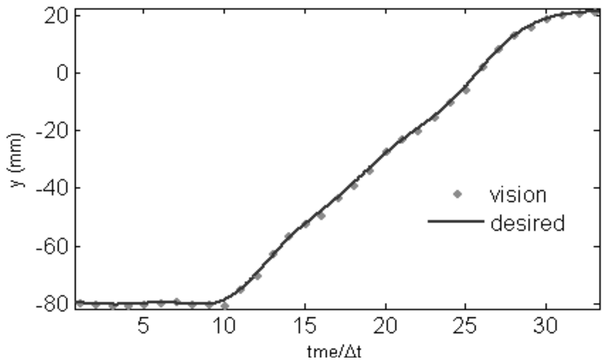


Figure 9. Measured y coordinate

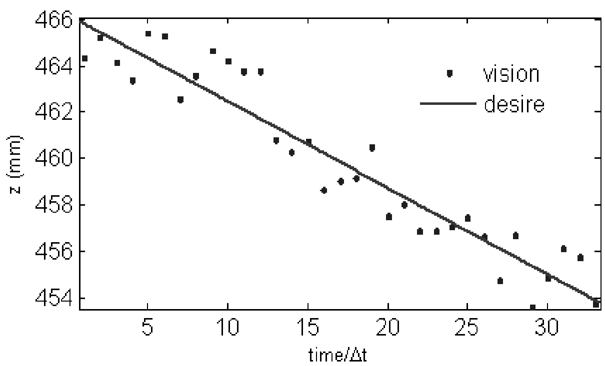


Figure 10. Measured z coordinate

As second experimental test, the end-effector of 6R manipulator is moved from initial point P_i= *(-68.5, 68, 385) mm* to final point P_f= *(-60.5, -93, 476) mm* through a circular trajectory in 6 sec. The distance between stereo vision setup and initial position of end effector is 400 mm. Resulted trajectory is shown in Fig. 11 also *x*, *y* and *z* components of end effector position are displayed in Figs. 12 to 14.

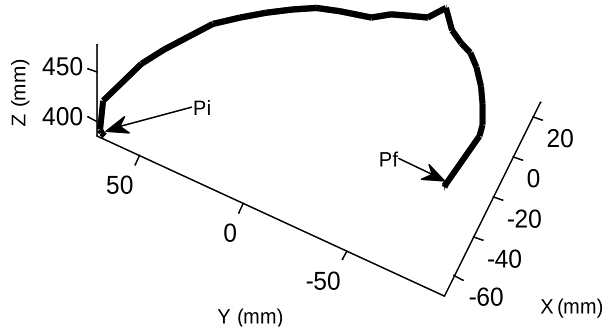


Figure 11. End effector trajectory

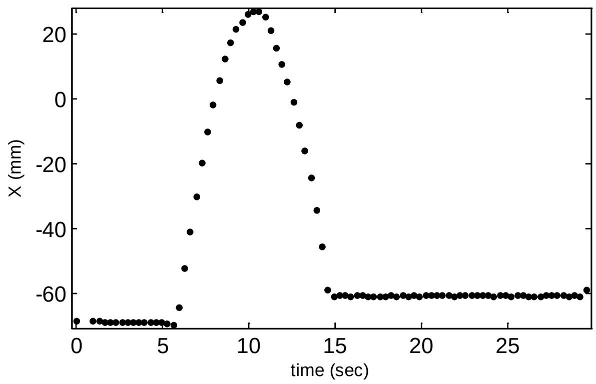


Figure 12. x coordinate

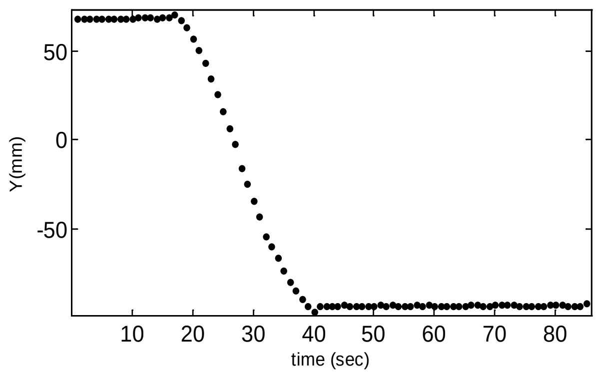


Figure 13. y coordinate

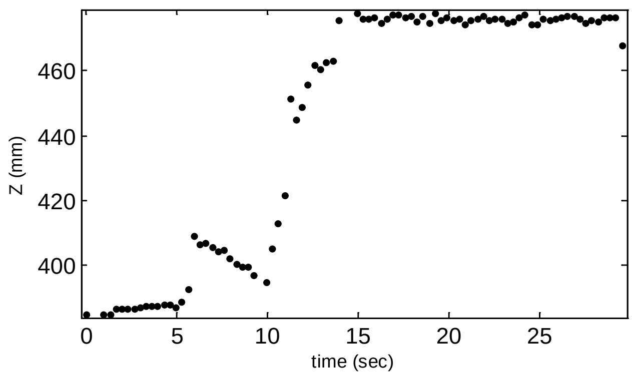


Figure 14. z coordinate

Considering figs. 12 to 14, it is shown that 3 data are computed per second and sampling frequency of this measurement unit is 3 data per second. Also by comparing measured radius of circular trajectory with actual value, it is determined that measuring error is less than 1 mm for distance about 400mm between cameras and end-effector.

B. Scout Robot (Wheeled Mobile Manipulator)

Another experimental test was done on a wheeled mobile manipulator that is shown in Fig. 15. The robot is controlled using software launched on a PC which is in contact with the robot in a wireless network. Regarding to simulation and software limitation, two joints of left arm is used for study and the others are considered to be fixed. A schematic view of this robot is shown in Fig. 16. Also parameters of robot are listed in table 3.

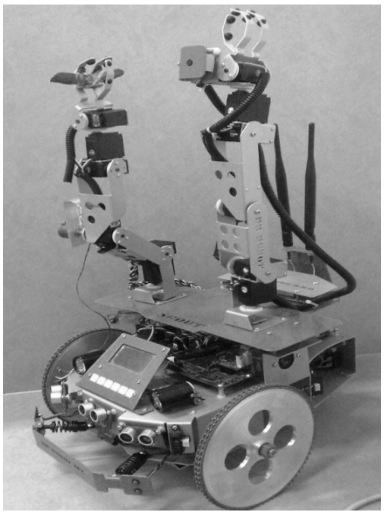


Figure 15. Scout mobile robot

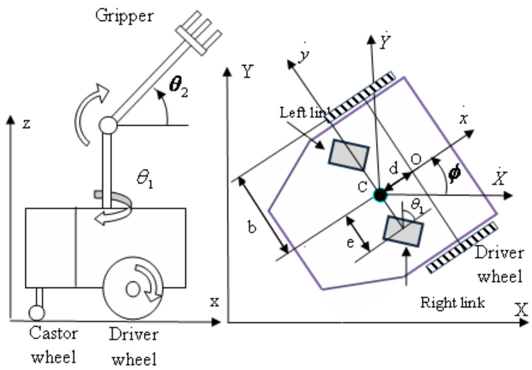


Figure 16. schematic view of Scout mobile robot

TABLE III. PARAMETERS OF SCOUT ROBOT

Parameter	Value	Unit
length of Links	L_1=0.16,L_2=0.21	m
mass of Links	m_1=0.128, m_2=0.231	Kg
moment of Inertia of link1	I_1 = 0.00005	Kg.m^2
moment of Inertia of link 2	Ix= 0.00008 Iy= 0.00091 Iz= 0.00092	Kg.m^2
mass of base and wheels	6.0, 0.32	Kg
moment of inertia of base about Z axis	0.06363	Kg.m^2
moment of inertia of wheels about rotation axis	0.0008	Kg.m^2
b,r,d,e	0.145, 0.08, 0.065, 0.08	m

Configuration of robot at initial point (P_i) and final point (P_f) are as follows:

$$P_i \begin{cases} X_c = 0m \\ Y_C = 0\,m \\ \phi = 0\,rad \\ \theta_1 = -\pi\,rad \\ \theta_2 = \dfrac{-\pi}{2}\,rad \end{cases} \tag{18}$$

$$P_f \begin{cases} X_c = 1\,m \\ Y_C = 0\,m \\ \phi = 0\,rad \\ \theta_1 = \dfrac{-\pi}{4}\,rad \\ \theta_2 = \dfrac{\pi}{4}\,rad \end{cases} \tag{19}$$

Simulation of robot motion is done for 3sec and resulted angular positions and velocities are applied to robot. Finally stereo vision set up is used to measure position of end effector that is shown in Fig. 17 where measurement process is implemented two times. For more details, a planar view of end effector path along vertical coordinate is displayed in Fig. 18.

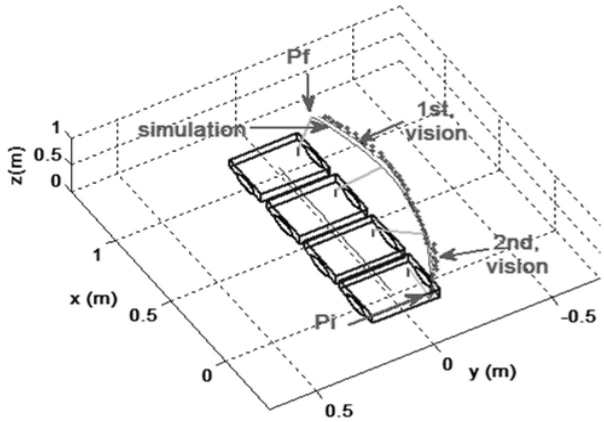


Figure 17. end effector path and robot configuration

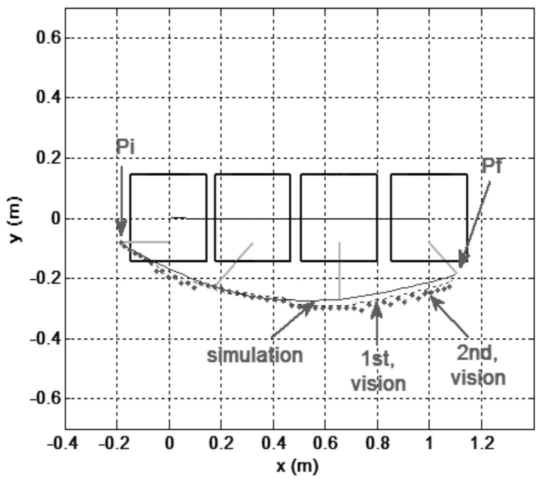


Figure 18. planar view of end effector path

Results illustrate that error between simulation and measured end-effector path is 0.7 cm that maybe as a result of flexibility of links that are not modeled also errors in encoders that cause difference between simulated and actual path.

VI. CONCLUSION

The focus of this paper is application of stereo vision approach on robotic systems to measure 3D position of end-effector. Stereo vision is a method that determines 3D position information of an object. A simple and inexpensive setup contains two cameras is made to implement stereo vision on any mechanical manipulator. Also basic formulation of triangulation process is mentioned to compute 3D position and orientation of end effector based on two resulted images. Experimental results are provided for 6R mechanical manipulator to show accuracy of proposed method for end-effector positioning. Results show positioning error less than 1 mm and good sampling frequency (5 data per sec). According to the results, following results can be mentioned:

- Performance of measurement system is related to quality of cameras and it can affect on accuracy and speed of stereo vision system.
- Other main factor is the distance between camera and end-effector. If this distance decreased then the accuracy of positioning system will be increased.
- Sampling frequency of measurement is a function of quality and speed of cameras.
- 3D position that is determined via stereo vision can be used in closed loop control system. But high sampling frequency is needed.

REFERENCES

L. G. Hee, “sensor based localization of a mobile”, phd thesis, National Univ. of Singapore 2007.

E. Jones and S. Soatto, “Visual-inertial navigation, mapping and localization: A scalable real-time causal approach”, Int. J. of Robotics Reaserch, vol. 30, pp. 407-430, 2010.

A. Gutierrez, A. Campo, F. C. Santos, C. Pinciroli and M. Dorigo, “Social Odometry in Populations of Autonomous Robots ”, Lecture Notes in Computer Science, 5217,pp. 371-378, 2008.

P. Kucsera, “Sensors for mobile robot systems”, AARMS, vol. 5(4), pp. 645-658 2006.

T. Tsukiyama, “Mobile robot localization from landmark bearings”, XIX IMEKO World Congress Fundamental and Applied Metrology, September 6-11, Lisbon, Portugal 2009.

J. Santolaria, D. Guillomía, C. Cajal, J. A. Albajez and J. J. Aguilar, “Modeling and Calibration Technique of Laser Triangulation Sensors for Integration in Robot Arms and Articulated Arm Coordinate Measuring Machines”, Sensors,vol. 9, pp. 7374-7396, 2009.

MIT Computer Science and Artifcial Intelligence Laboratory, Cambridge, MA 02139, 2004.

Evolution Robotics, Inc., 130 W., CA 91103, NorthStar Projector Kit User Guide, 1.0 edition.

A. Ramisa, A. Tapus, R. L. Mantaras, and R. Toledo. “Mobile Robot Localization using Panoramic Vision and Combinations of Feature Region Detectors”, IEEE International Conference on Robotics and Automation, Pasadena, USA, pp. 538-543, 2008.

C.K. Chang, C. Siagian, L. Itti, “Mobile Robot Vision Navigation & Localization Using Gist and Saliency”, IEEE International Conference on Intelligent Robots and Systems, Taiwan, pp. 4147-4154, 2010.

L. Spacek, C. Burbridge, “Instantaneous robot self-localization and motion estimation with omnidirectional vision”, Robotics and Autonomous Systems, vol. 55,pp. 667–674, 2007.

M. Bleyer and M. Gelautz(2008), Computer Vision Theory and Applications, VISAPP, 415-422.

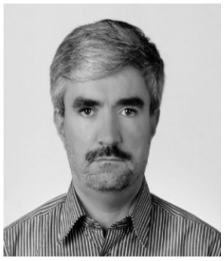
M. Humenberger, T. Engelke, W. Kubinger, “A census-based stereo vision algorithm using modified Semi-Global Matching and plane fitting to improve matching quality”, IEEE Computer Society Conference CVPRW, San Francisco, CA, pp. 77–84, 2010.

Q.Yang, “Real-time global stereo matching using hierarchical belief propagation”, British Machine Vision Conference, pp. 989–998, 2006.

B. Stan, T. Carlo, "Depth Discontinuities by Pixel-to-Pixel Stereo", International Journal of Computer Vision, vol. 35, no. 3, pp. 269–293, 1999.http://www.vision.deis.unibo.it/smatt/

A. H. Dastjerdi, “Optimizing Tracking Algorithm for Contact Probe of CMM by Non Contact Probe based On Image Processing”, MSc thesis, Amir Kabir Univ., Iran, 2010.

D. Scharstein and R. Szeliski, http://vision.middlebury.edu/stereo/eval/



M. HabibnejadKorayem was born in Tehran Iran on April 21, 1961. He received his B.Sc. (Hon) and M.Sc in Mechanical Engineering from the Amirkabir University of Technology in 1985 and 1987, respectively. He has obtained his Ph.D degree in Mechanical Engineering from the University of Wollongong, Australia, in 1994.

He is a Professor in Mechanical Engineering at the Iran University of Science and Technology. He has been involved with teaching and research activities in the robotics areas at the Iran University of Science and Technology for the last 20 years. His research interests includes dynamics of Elastic Mechanical Manipulators, Trajectory Optimization, Symbolic Modelling, Robotic Multimedia Software, Mobile Robots, Industrial Robotics Standard, Robot Vision, Soccer Robot, Nano Robotics and the Analysis of Mechanical Manipulator with Maximum Load Carrying Capacity. He has published more than 500 papers in international journal and conference in the robotic area.

M. Irani was born in Kashan, Iran on 22 July, 1980. He received his B.Sc in mechanical Engineering from the Isfahan University of Technology in 2002 and M.Sc in Aerospace Engineering from the KNTU University of Technology in 2005.
He received his Ph.D in Mechanical Engineering in the Iran University of Science and Technology, Iran in 2012. He is assistant Professor in Mechanical Engineering at the University of

Kashan.His research interests include robotic systems, nonlinear control, optimal control, mobile manipulators, Rotor Dynamicsand mechatronic systems.

A. HashemiDastjerdi was born in Tehran Iran on March 9, 1980. He received his M.Sc in Mechanical Engineering from the Amirkabir University of Technology, Tehran, in 2010 and has practiced in the Non-Contact Probe of CMM his entire career. He is an instructor in the Department of Mechanic at Iran Azad University. He is employed by aviation industry *and has*over nine years of *experience* in reverse engineering. He joined robotics researchers at the Iran University of Science and Technology and is currently working in Stereo Vision. His research interests includes Machine Vision, Stereo Vision, CMM, 3D Measurements, Image Processing, Optimizing Tracking Algorithm, Robots Vision systems, CAD/CAM.