Integrating 3D Vision Measurements into Industrial Robot Applications

by

Frank S. Cheng
cheng1fs@cmich.edu
Engineering and Technology
Central Michigan University

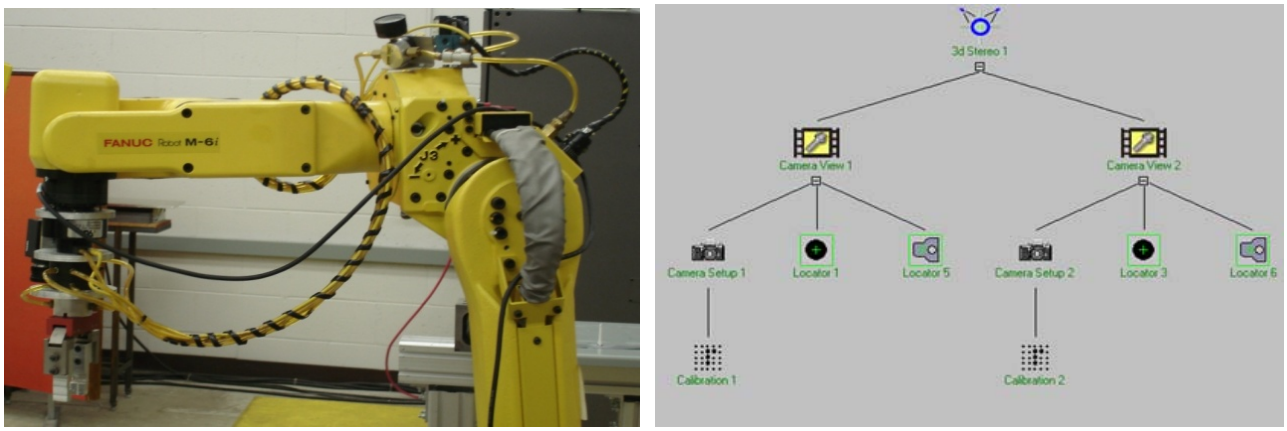
Xiaoting Chen
Graduate Student
Engineering and Technology
Central Michigan University

Abstract: Vision systems are used more and more in recent industrial robot applications for enhancing robot flexibility and intelligence. This paper presents the integration of a three-dimensional (3D) binocular stereo vision system into an industrial robot system. Examples are used to illustrate the developed method and its implementation including vision system setup and calibration, robot hand-eye integration, and vision measurement analysis. The results show that a 3D binocular stereo vision system is able to aid an industrial robot in performing robot motions to a given 3D part through accurate vision measurements and integrated robot programming.

I. Introduction

Industrial robots are often equipped with vision systems for recognizing objects in a robot work environment. Among various vision technologies, a three-dimensional (3D) binocular stereo vision system utilizes two cameras to obtain the positioning information (x, y, z, roll) of a 3D object. This technology is used in industrial robot applications due to the simple setup of lighting conditions and the efficient image processing of the vision system. Fig. 1 shows an industrial robot workcell where the FANUC M6i robot is assisted by the FANUC VisLOC 3D vision system in performing robot motions to a given 3D part on the table. However, in order to make the 3D vision application successful, the robot programmer must correctly set up the 3D measuring functions of the vision system and integrate the precise vision measurements into the robot application programs. Industrial practice shows that vision-guided industrial robot applications not only use the technology of vision cameras, image processing, system calibration, and communication [1] [2] [3] [4][5] but also require the robot programmer to have a solid understanding of other issues including camera positions and views, vision measurements and accuracy, reference frames and frame transformations, and robot operations and programming.

This paper presents the method for integrating the measuring functions of a 3D binocular stereo vision system into an industrial robot system. The study deals with the design techniques and procedures related to vision system setup, robot hand-eye integration, and vision measurement analysis. The



(a) (b)

Figure 1: (a) FANUC M6i robot workcell and (b) FANUC VisLOC 3D vision process

discussion includes the concepts of Epipolar geometry and perspective calibration used by 3D binocular vision systems, the procedures for unifying the coordinates of vision measurements and robot points, and the variations of vision depth measurements due to pattern distortion. With the developed method and solution, the robot programmer is able to better apply the 3D binocular stereo vision technology for enhancing flexibility and intelligence of industrial robot operations.

II. 3D Binocular Stereo Vision and Calibration

The geometry of 3D binocular stereo vision is referred to as Epipolar geometry. When two cameras view a 3D scene from two distinct positions as shown in Fig. 2, there are a number of geometric relationships between a 3D point and its projection in a 2D image as shown in Fig. 3. The concept of 3D binocular stereo vision is illustrated in Fig. 4 where two cameras look at point P and the projection problem is simplified by placing a virtual image plane in front of the focal point of each camera to produce an unrotated image. Specifically, O_L and O_R represent the focal points of the two cameras. Points p_L and p_R are the projections of point P on the left and right image planes, respectively. Each camera captures a 2D image of the 3D world. This conversion from 3D to 2D is referred to as a perspective projection. Since the two cameras are at distinct locations, the focal point of one camera projects a distinct point on the image plane of the other camera. These two projected image points are called epipoles and denoted as E_L and E_R, respectively. It is clear that epipoles, E_L and E_R, and focal points, O_L and O_R, lie on a single line. Although the left camera sees line PO_L as a point because it is directly aligned with its focal point, the right camera sees the same line as epipolar line p_RE_R on its image plane. Symmetrically, the left camera sees line PO_R as epipolar line p_LE_L on its own image plane. Alternatively, each epipolar line is actually the intersection of the corresponding camera image plane and the epipolar plane formed by points P, O_L and O_R. Thus, all epipolar lines intersect the corresponding epipoles regardless of P point locations. If projection point p_L is known, then corresponding epipolar line p_RE_R is known and projection point p_R must lie on this particular epipolar line. This epipolar constraint is applied to all corresponding image points. This means that if image

points p_L and p_R are known, their projection lines are also known and must intersect precisely at point P. Then, the position of point P can be calculated from the coordinates of the two image points.

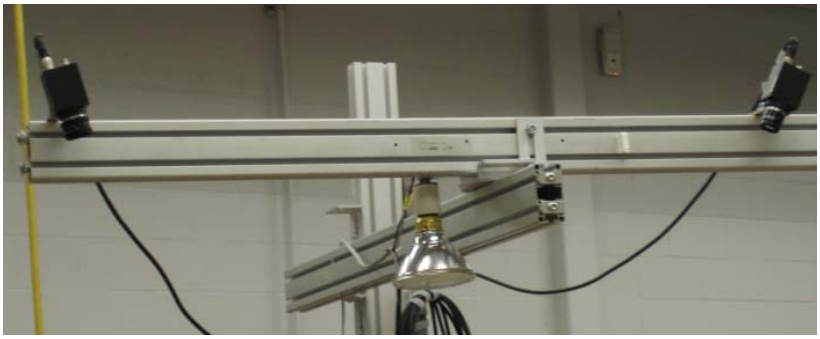


Figure 2: Camera setup for 3D binocular stereo vision

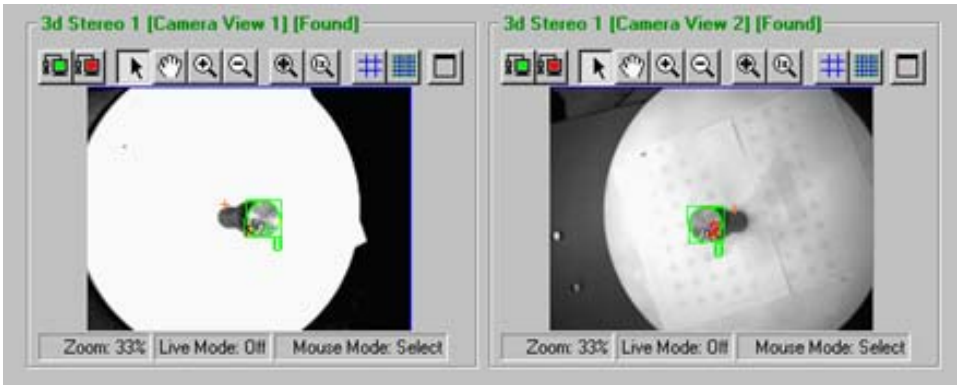


Figure 3: Target views from both cameras in 3D binocular stereo vision

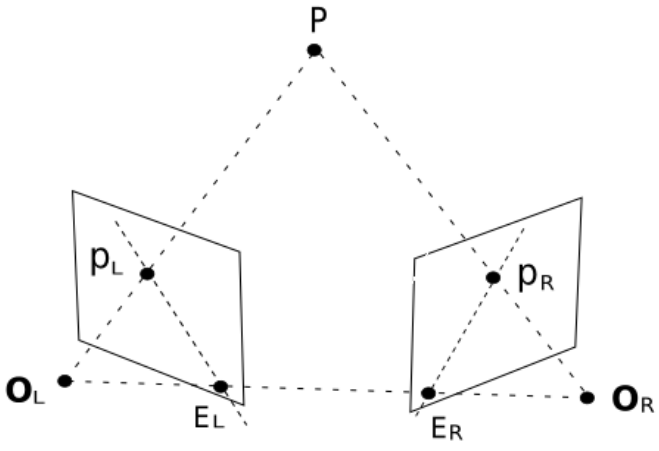
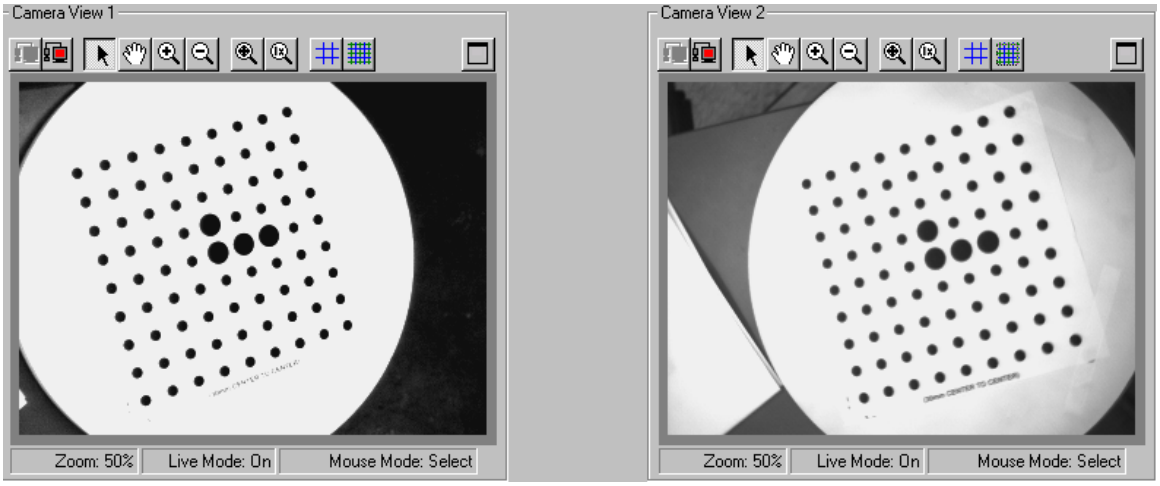
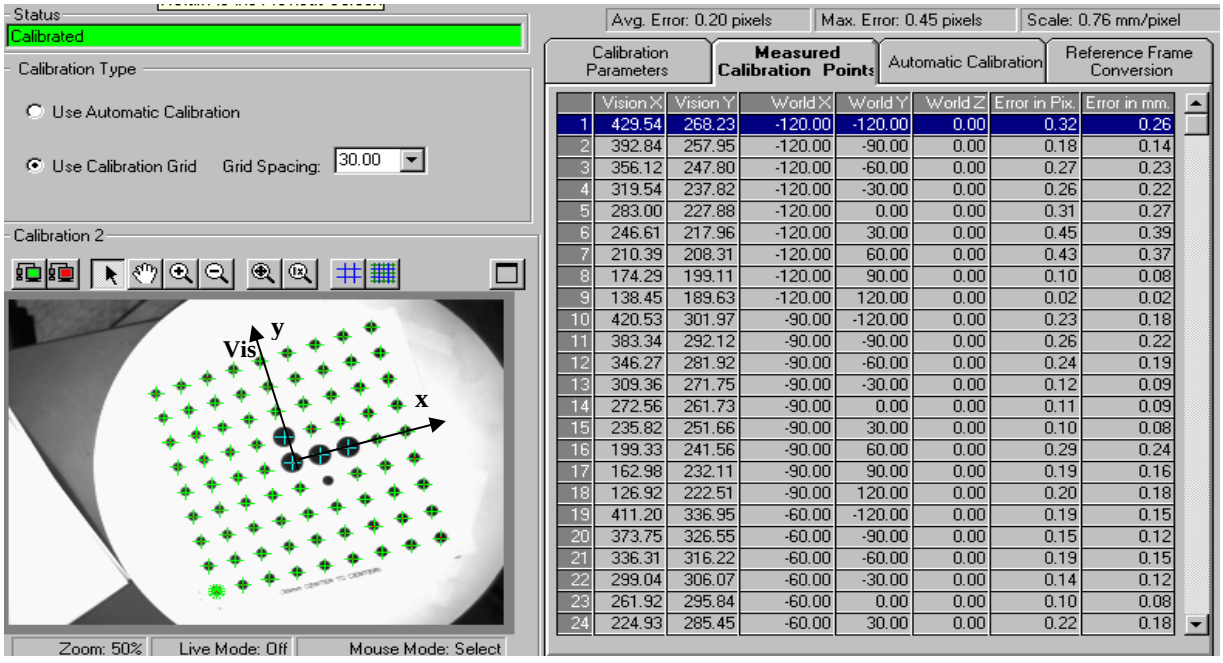


Figure 4: Epipolar line of binocular stereo vision

Vision camera calibration is a process that establishes the vision frame, Vis (x, y, z), and the pixel value for a vision system. A 3D binocular stereo vision system conducts the perspective calibration through two 2D camera view pictures of a common calibration grid sheet as shown in Fig. 5, where the large circles define the x- and y-axes of vision frame Vis and the small circles calibrate the camera pixel value (mm) in Vis.



(a) Calibration view planes from both cameras



	Vision X	Vision Y	World X	World Y	World Z	Error in Pix	Error in mm
1	429.54	268.23	-120.00	-120.00	0.00	0.32	0.26
2	392.84	257.95	-120.00	-90.00	0.00	0.18	0.14
3	356.12	247.80	-120.00	-60.00	0.00	0.27	0.23
4	319.54	237.82	-120.00	-30.00	0.00	0.26	0.22
5	283.00	227.88	-120.00	0.00	0.00	0.31	0.27
6	246.61	217.96	-120.00	30.00	0.00	0.45	0.39
7	210.39	208.31	-120.00	60.00	0.00	0.43	0.37
8	174.29	199.11	-120.00	90.00	0.00	0.10	0.08
9	138.45	189.63	-120.00	120.00	0.00	0.02	0.02
10	420.53	301.97	-90.00	-120.00	0.00	0.23	0.18
11	383.34	292.12	-90.00	-90.00	0.00	0.26	0.22
12	346.27	281.92	-90.00	-60.00	0.00	0.24	0.19
13	309.36	271.75	-90.00	-30.00	0.00	0.12	0.09
14	272.56	261.73	-90.00	0.00	0.00	0.11	0.09
15	235.82	251.66	-90.00	30.00	0.00	0.10	0.08
16	199.33	241.56	-90.00	60.00	0.00	0.29	0.24
17	162.98	232.11	-90.00	90.00	0.00	0.19	0.16
18	126.92	222.51	-90.00	120.00	0.00	0.20	0.18
19	411.20	336.95	-60.00	-120.00	0.00	0.19	0.15
20	373.75	326.55	-60.00	-90.00	0.00	0.15	0.12
21	336.31	316.22	-60.00	-60.00	0.00	0.19	0.15
22	299.04	306.07	-60.00	-30.00	0.00	0.14	0.12
23	261.92	295.84	-60.00	0.00	0.00	0.10	0.08
24	224.93	285.45	-60.00	30.00	0.00	0.22	0.18

(b) Calibration Accuracy

Figure 5: 3D binocular stereo vision calibration

A 3D binocular stereo vision system also utilizes the user-trained pattern of an actual object to identify similar objects from the camera view. Fig. 6 shows a real block and the corresponding trained patterns for the two cameras. The robot programmer must define the object frame, OBJ (x, y), for each object pattern by choosing a point that represents a common point on the real object. Through pattern comparison and matching, the vision software finds the similar object from the camera view and measures the corresponding OBJ frame location OBJ[n] as frame transformation ${}^{Vis}T_{OBJ[n]}$ as shown in Fig. 7.

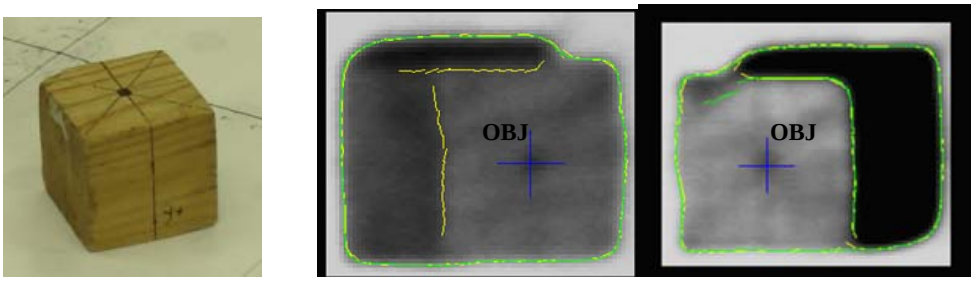


Figure 6: The trained patterns of an object in two camera views

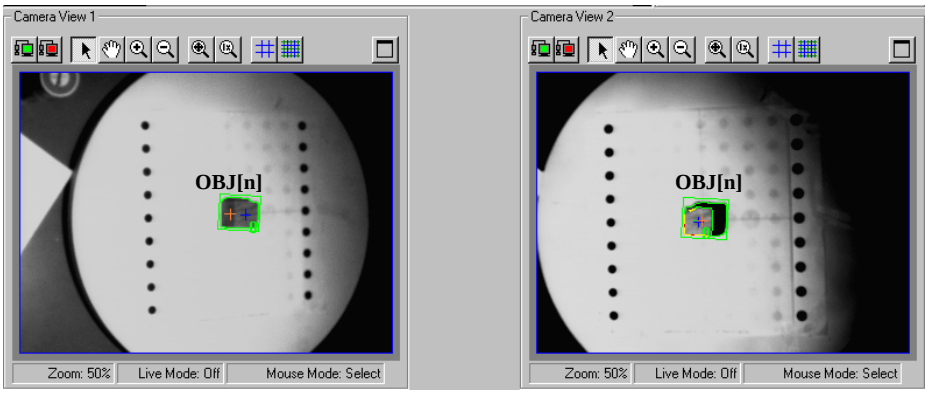


Figure 7: Vision-found object and its OBJ frame location

III. Robot Hand-Eye Integration

An industrial robot system employs Cartesian reference frames and frame transformations to represent the position and orientation of the robot hand as shown in Fig. 8 [6]. Among them, the robot base frame R(x, y, z) is a fixed one and the default tool-center-point frame Def_TCP (n, o, a) is located at the center of the robot's wrist faceplate. The robot point $P[n]^{R}_{Def\_TCP}$ is the frame transformation ${}^{R}T_{Def\_TCP}$ that represents the position and orientation of frame Def_TCP relative to frame R. The robot programmer may also define the robot user frame UF[i] (for i = 1, 2, …, n) relative to frame R to represent the workpiece location, and the user tool frame UT[k] (for k = 1, 2, .., n) relative to frame Def_TCP to represent the tool-tip of the robot end-effector. After the robot programmer teaches a robot point with the selected UT[k] and UF[i], the robot point $P[n]^{UF[i]}_{UT[k]}$ is defined as frame transformation ${}^{UF[i]}T_{UT[k]}$ in the robot system.

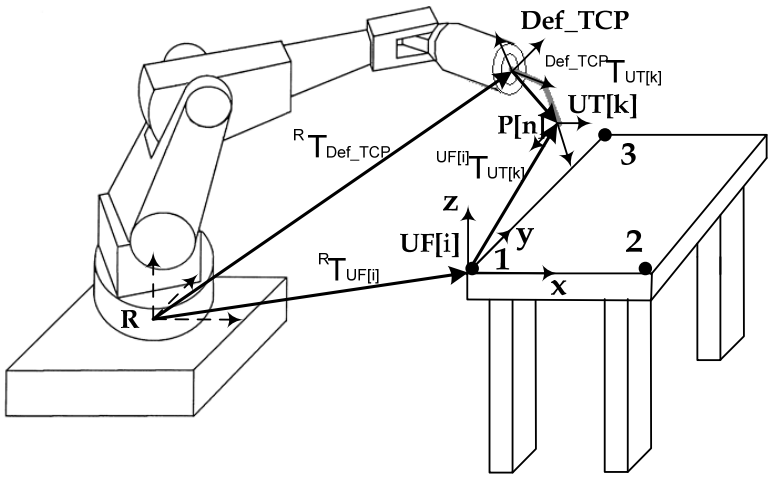


Figure 8: Robot reference frames and frame transformations

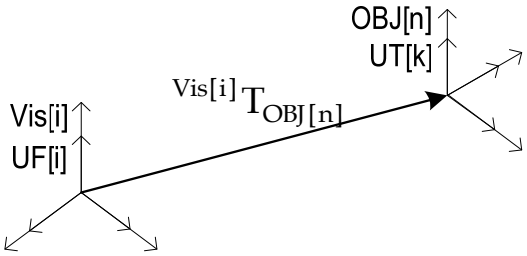
To integrate the vision-measured OBJ frame location of the object into the robot system, the robot programmer needs to setup a UF[i] that coincides with vision frame Vis as shown in Fig. 9a. The task requires the robot programmer to follow the UF Setup procedure provided by the robot system and teach three non-collinear points as shown in Fig. 8 by using the same camera calibration grid sheet that is at the same location in vision calibration as shown in Fig. 5. With the coincidence of frames UF[i] and Vis, object frame location OBJ[n] measured in vision frame Vis actually represents robot point P[n] as shown in Eq. (1):

$$P[n]^{UF[i]}_{UT[k]} = {}^{UF[i]}T_{UT[k]} = {}^{Vis}T_{OBJ[n]}\,. \tag{1}$$

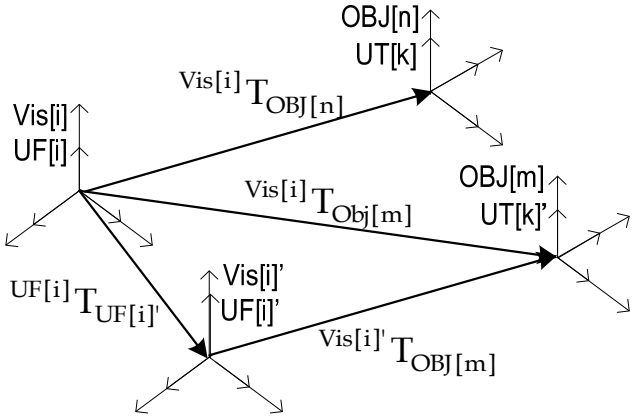
The robot programmer may directly use a vision measurement as a robot point as shown in Eq. (1). However, after reaching to the vision-defined point, the robot cannot use the robot points that are taught via the similar object located at a different position OBJ[m] within the camera view. To avoid this problem, the robot programmer usually determines the position offset of UF[i], ${}^{UF[i]}T_{UF[i]'}$, through two vision measurements ${}^{Vis}T_{OBJ[n]}$ and ${}^{Vis}T_{OBJ[m]}$ as shown in frame transformation Eq. (2) and Fig. 9b:

$${}^{UF[i]}T_{UF[i]'} = {}^{Vis}T_{OBJ[n]} \times ({}^{Vis'}T_{OBJ[m]})^{-1} \qquad \text{and} \qquad {}^{Vis'}T_{OBJ[m]} = {}^{Vis}T_{OBJ[n]}\,, \tag{2}$$

where frames Vis' and UF[i]' represent the locations of frames Vis and UF[i] after object frame OBJ changes its location from OBJ[n] to OBJ[m] and $({}^{Vis'}T_{OBJ[m]})^{-1}$ is the inverse of transformation ${}^{Vis'}T_{OBJ[m]}$. Usually, the vision system obtains ${}^{Vis}T_{OBJ[n]}$ when the object is found at a reference position within the camera view during the vision setup and acquires ${}^{Vis}T_{OBJ[m]}$ when the object is found from the actual view picture. With the vision-measured ${}^{UF[i]}T_{UF[i]'}$, the robot programmer is able to transform robot point $P[n]^{UF[i]}_{UT[k]}$ that is taught via the object at OBJ[n] into robot point $P[m]^{UF[i]}_{UT[k]'}$ for the similar object at a different location OBJ[m] in the robot program as shown in frame transformation Eq. (3):



(a) Coincided frames in robot and vision systems



(b) Determining UF[i] offset through two vision measurements

Figure 9: Frame transformations in integrated robot and vision systems

$$ {}^{UF[i]}T_{UT[k]'} = {}^{UF[i]}T_{UF[i]'} \times {}^{UF[i]'}T_{UT[k]'} \quad \text{and} \quad {}^{UF[i]'}T_{UT[k]'} = {}^{UF[i]}T_{UT[k]}, \tag{3} $$

where UT[k]' represents the location of UT[k] after UT[k] moves to object frame location OBJ[m].

In the FAUNU robot workcell in Fig. 1, the FANUC VisLOC vision system [7] measures ${}^{Vis}T_{OBJ[n]}$ via the object located at the reference position where the robot programmer also teaches the required robot points $P[1]^{UF[i]}_{UT[k]}$ and $P[2]^{UF[i]}_{UT[k]}$ via the same object. Executing the VisLOC micro function "Snap & Find" in the FANUC robot Teaching Pendant (TP) program [8] allows the vision system to obtain ${}^{Vis}T_{OBJ[m]}$ for the vision-found object within the camera view and calculate offset value ${}^{UF[i]}T_{UF[i]'}$ for the robot system as shown in Eq. (2). The vision system sends the determined ${}^{UF[i]}T_{UF[i]'}$ to the specified robot position register PR[x]. The robot programmer then utilizes the "offset" value in the following three FANUC robot TP program instructions to modify $P[1]^{UF[i]}_{UT[k]}$ and $P[2]^{UF[i]}_{UT[k]}$ into the corresponding ones for the vision-found object as shown in Eq. (3):

1: Offset Conditions PR[x], UFRAME[i];
2: J P[1] 100% Offset;
3: J P[2] 100% Offset;

where the motion instruction element J specifies the Joint (J) motion type, speed override 100% specifies the default speed of the robot joint, and motion option "Offset" applies the offset value to the robot points.

IV. Vision Calibration Accuracy and Pattern Distortion

The accuracy of vision calibration and pattern matching is important in vision-guided industrial robot applications. Due to the fact that during perspective calibration the camera lens in a 3D binocular stereo vision system cannot be perpendicular to the calibration grid sheet, certain degrees of distortion occur when the vision software records all the corresponding locations of the circles in the camera view. For example, Fig. 5b shows that the perspective calibration may result in an average error of 0.2 pixels, a maximum error of 0.45 pixels, and a scale of 0.76 mm/pixel. Based on these data, we can obtain the following error range of the vision measurement:

$$\textit{Average Error} = 0.2 \times 0.76 = 0.152 \text{ (mm)} \qquad \textit{Maximum Error} = 0.45 \times 0.76 = 0.342 \text{ (mm)}$$

Due to pattern distortion, the depth measurement (i.e. z value) of an object varies when the object is away from the center point of the 3D vision view plane as shown in the following test results. During the test, the block, as shown in Fig. 6, was moved 10 mm each time along the *x*- and *y*-axes of the camera view plane, respectively, and the corresponding depth value (mm) of the block was measured by using the FANUC VisLOC Snap & Find function. Then, the relationship between the average depth and the location of the block at each axis was statistically formulated.

The diagram in Fig. 10 shows that the Z – X relationship is a declining curve along the positive direction of the *x*-axis. Its model is:

$$z = -0.0096x + 36.108 \qquad \text{with } R^2 = 0.7948$$

The diagram in Fig. 11 shows that the Z – Y relationship is almost a non-fluctuating curve. Its model is:

$$z = -0.0002y + 35.912 \qquad \text{with } R^2 = 0.0018$$

These relationships illustrate that due to decreased identical features in both camera views the location variation of the object along the *x*-axis produces some degrees of pattern distortion and mismatch, resulting in an inaccurate depth measurement of the object. However, the depth measurement remains consistent in the *y*-axis due to more overlapped features in both camera views.

V. Conclusion

This paper provided the robot programmer with the method for using a 3D binocular vision system in an industrial robot application. The study has shown that the camera setup, calibration of robot and vision systems, vision measurement acquisition, and robot hand-eye system integration were the key design tasks in 3D vision-guided industrial robot applications and the developed method addressed the

solutions for successfully completing these design tasks. The Epipolar geometry allows the 3D binocular stereo vision system to conduct the perspective calibration by using two 2D camera view pictures of a common calibration grid sheet. The robot programmer integrates the vision measurement into the robot system by setting up a robot user frame UF[i] that coincides with vision frame Vis. The robot program uses the vision-determined "offset" value to transform all pre-defined robot operation points to the corresponding ones for the vision-found object. The results also show that the 3D binocular stereo vision system is able to provide the robot system with the accurate depth measurement for the 3D object located in the center of the camera view. The variation of depth measurement occurs as the vision-found object is away from the center of the camera view.

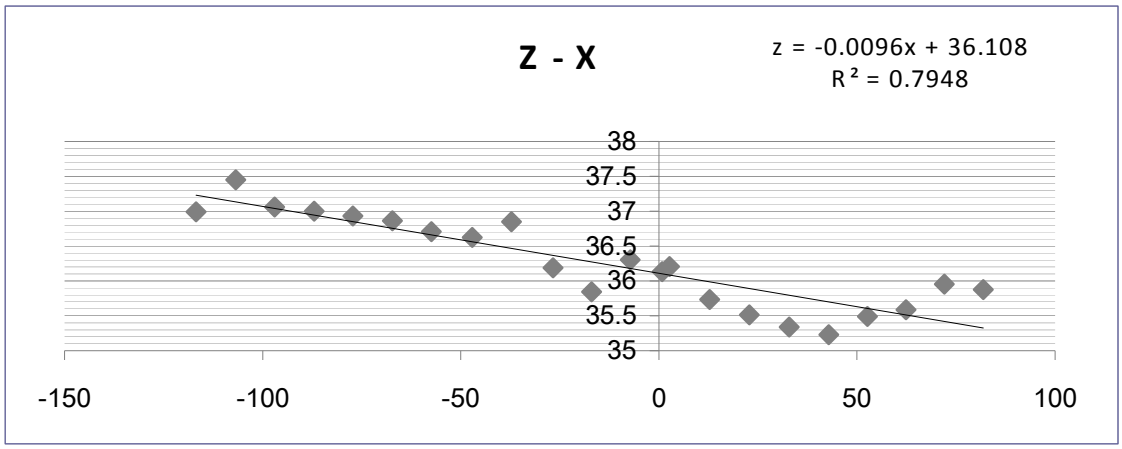


Figure 10: Depth variation (mm) along X-Axis

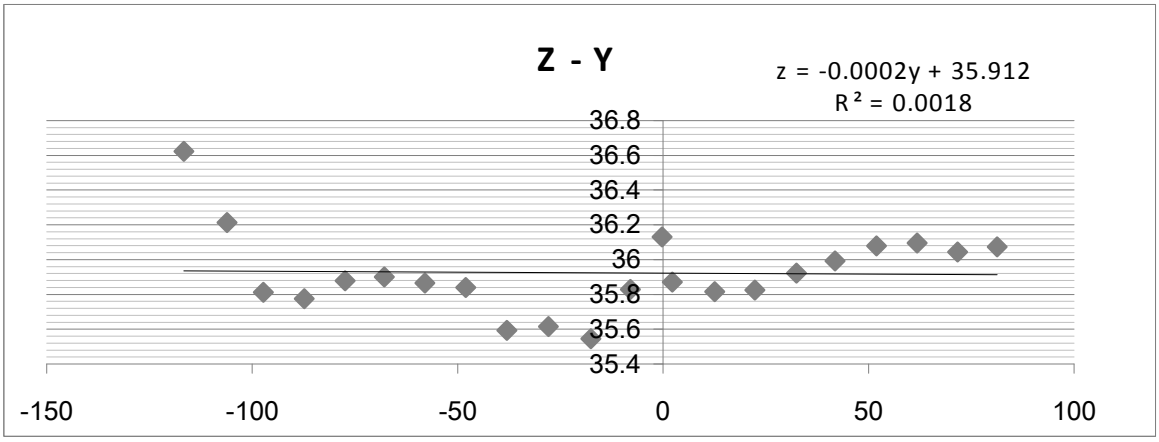


Figure 11: Depth variation (mm) along Y-Axis

References

[1] Woodill, J.I., Buck, R., Jurasek, D., Gordon, G. and Brown, T., "3D Vision: Developing an Embedded Stereo-Vision System," Computer, Vol. 40, No. 5, pp. 106-8, May 2007.

[2] Mure-Dubois, J. and Hugli, H., "Embedded 3D Vision System for Automated Micro-Assembly," Proceedings of SPIE Two- and Three-Dimensional Methods for Inspection and Metrology IV, Vol. 6382, pp. 63820J, 2006.

[3] Biegelbauer, G. and Vincze, M., "3D Vision-Guided Bore Inspection System," Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ISBN 0-7695-2506-7), 2006.

[4] Nguyen, M.C., "Vision-Based Intelligent Robots," In SPIE: Input/Output and Imaging Technologies II, Vol. 4080, pp. 41-47, 2000.

[5] Kress, S., "Machine Vision Makes Its Mark on the Automotive Industry," Automotive Design and Production, October, 2004.

[6] Craig, J. J., Introduction to Robotics: Mechanics and Control, Prentice Hall, 3rd Edition, 2003.

[7] FANUC VisLOC Help Manual, FANUC Robotics, 2001.

[8] FANUC HandlingTool Setup and Operation Manual, FANUC Robotics, 2001.

Biography

Frank S. Cheng received his Ph.D. degree from the University of Cincinnati in 1995. He is currently an associate professor in the department of Engineering and Technology at Central Michigan University. Dr. Cheng is the instructor for courses in the areas of robotics and automation for both engineering and engineering technology programs and is responsible for the Robotics and Automation Laboratory. His research interests include robotics, mechatronics, controls, and industrial automation. Dr. Cheng has published his research developments in refereed journals, proceedings, and book chapters. He is a member of ASEE, IEEE, and IAJC.

Xiaoting Chen is a graduate student in the department of Engineering and Technology at Central Michigan University. He holds a B.S. degree in Electrical Engineering from the Shanghai Institute of Technology, China. His research interests include robotics, control systems, and industrial automation.