

## **Integration of 3D Stereo Vision Measurements in Industrial Robot Applications**

Frank Cheng and Xiaoting Chen  
Central Michigan University  
[cheng1fs@cmich.edu](mailto:cheng1fs@cmich.edu)

### **Abstract**

Three dimensional (3D) vision systems are more and more used in recent industrial robot applications for enhancing robot flexibility and intelligence. This paper presents a method of integrating 3D binocular stereo vision systems into industrial robot systems. 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 stereo vision system is able to guide an industrial robot to manipulate a 3D object through accurate vision measurements and integrated robot programming.

### **Introduction**

Industrial robots are often equipped with vision systems for detecting and recognizing objects in a robot work environment. Among various vision approaches, a three-dimensional (3D) binocular stereo vision system utilizes two cameras to obtain the positioning information of a 3D object for an industrial robot. This technology is widely used in industrial robot applications due to the simple setup of lighting conditions and the efficient image process of the vision system. Fig. 1 shows an industrial robot workcell where the FANUC M6i robot is able to manipulate a given 3D part on the table by using the 3D measurements of the FANUC VisLOC vision system. However, in order to make the 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 <sup>[1], [2], [3], [4], [5]</sup> 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 a method of 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 discussion includes the Epipolar geometry and perspective calibration used by 3D binocular vision systems, the method for unifying the coordinates of

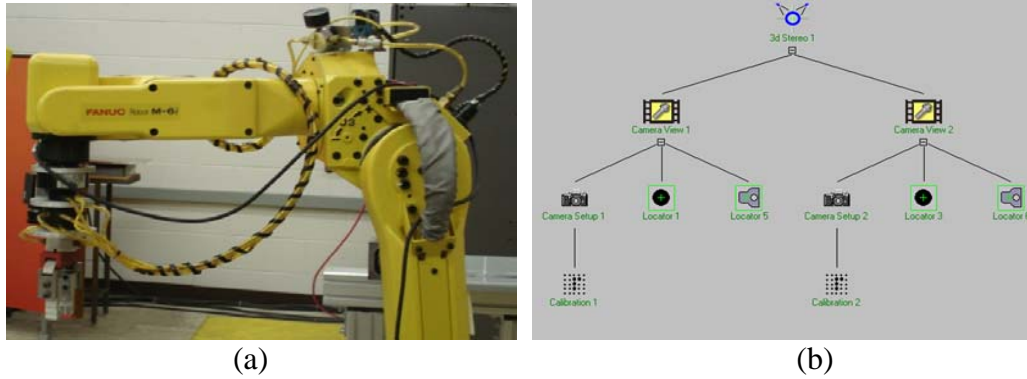


Figure 1: (a) Fanuc M6i Robot Workcell and (b) 3D Stereo Process in Fanuc VisLOC

vision and robot systems in robot programs, and the variation of vision 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.

### Binocular Stereo Vision and Calibration

The geometry of a 3D binocular stereo vision refers to the Epipolar geometry. When two cameras view a 3D scene from two distinct positions as shown in Fig. 2, there are a number of geometric relations between a 3D point and its projection on a 2D image as shown in Fig. 3.



Figure 2: Camera Setup for 3D Binocular Stereo Vision

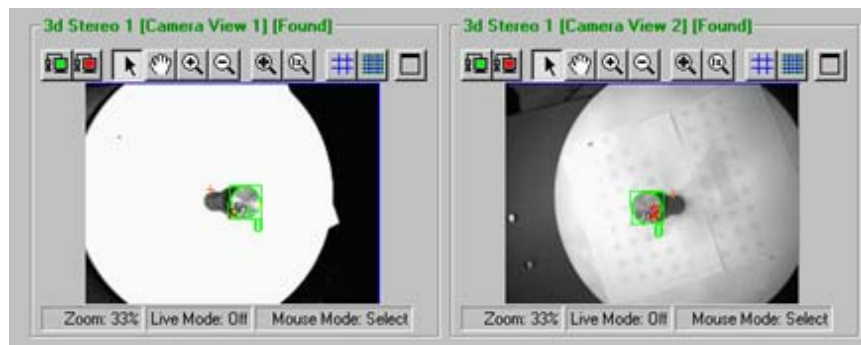


Figure 3: Target Views from Both Cameras in 3D Binocular Stereo Vision

The concept of a 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  $O_L$ -P as a point because it is directly aligned with its focal point, the right camera sees the same line as epipolar line  $E_R$ - $p_R$  on its image plane. Symmetrically, the left camera sees line  $O_R$ -P as epipolar line  $E_L$ - $p_L$  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,  $O_L$  &  $O_R$ . Thus, all epipolar lines intersect the corresponding epipoles regardless of P point locations.

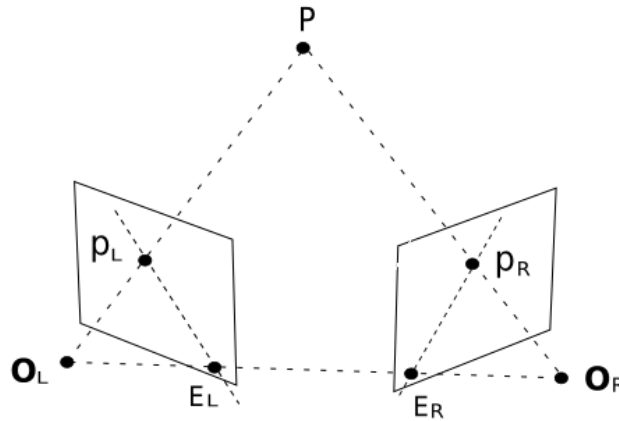
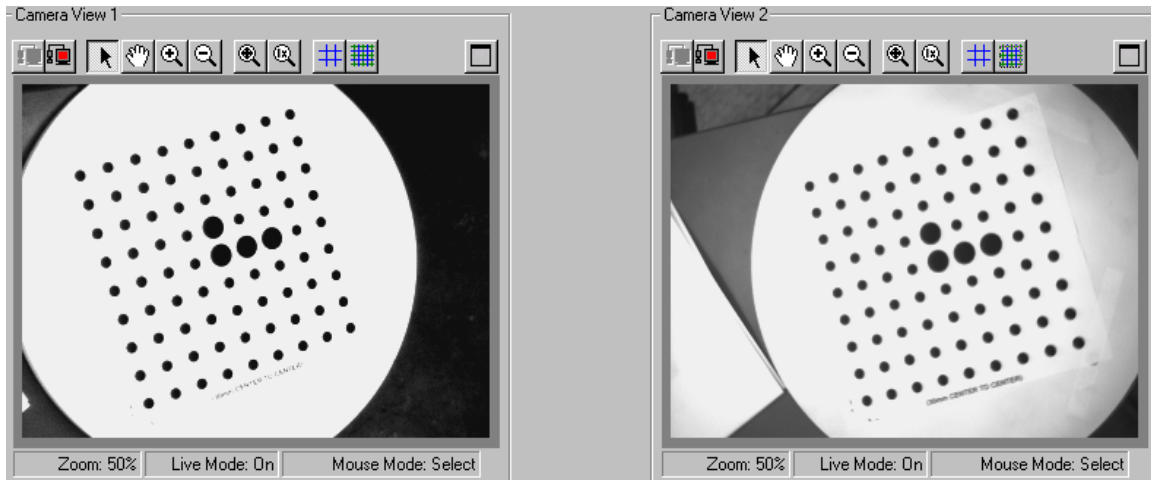


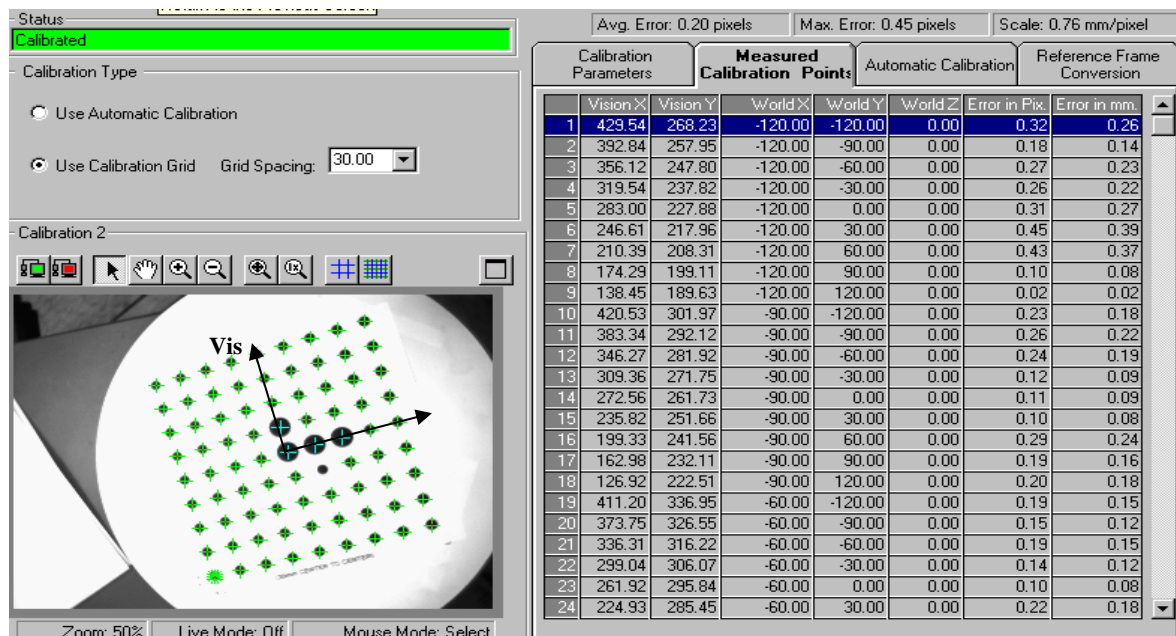
Figure 4: Epipolar Line of Binocular Vision

If projection point  $p_L$  is known, then corresponding epipolar line  $E_R$  -  $p_R$  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.

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 perspective calibration by using the two 2D camera view pictures of a common calibration grid sheet as shown in Fig. 5, where the large circles define vision frame  $Vis$  and the small circles calibrate the camera pixel values measured in vision frame  $Vis$ .



(a) Calibration View Planes from Both Cameras



(b) Calibration Accuracy

Figure 5: 3D Stereo Vision Calibration

A 3D stereo vision system also utilizes the user-trained pattern of an actual object to identify similar objects that appeared on the camera view picture. Fig. 6 shows a real block and the corresponding trained patterns from 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 automatically searches the whole camera view picture and finds the OBJ frame position of any similar object as shown in Fig. 7.

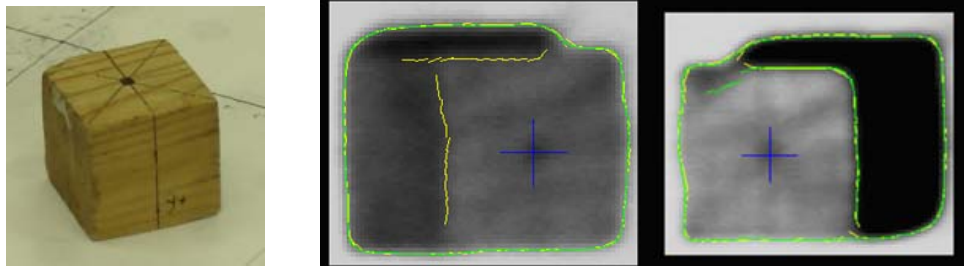


Figure 6: An Object and its Trained Patterns for Two Camera Views

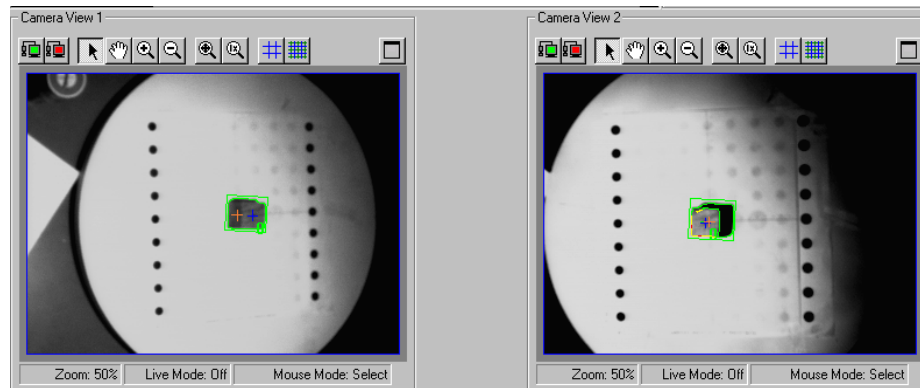


Figure 7: Vision-Identified Object and its OBJ Frame Position

### Robot Hand-Eye Integration

An industrial robot system employs Cartesian reference frames, as well as frame transformations, to represent the static positions of the robot hand as shown in Fig. 8 [6]. The robot world reference frame,  $R(x, y, z)$ , is a fixed one in the robot world space and the robot default tool-center-point frame,  $Def\_TCP(n, o, a)$ , is a moving one that represents the end-point of the robot hand. As shown in Fig. 8, the robot programmer may also define different user frames,  $UF(x, y, z)$ , relative to robot base frame  $R$  within the robot work volume and different user tool frames,  $UT\_TCP$ , relative to robot default TCP frame  $Def\_TCP$  for an actual robot end-effector. After the robot programmer sets both an available robot user frame and a user tool frame to be active in the robot system, frame transformation  ${}^{UF}T_{UT\_TCP}$  represents the actual robot TCP frame (or hand) position.

To unify the coordinates of robot TCP frame position and the vision measurement of an object, the robot programmer must establish a robot user frame,  $UF$ , in the robot system by teaching three non-collinear points as shown in Fig. 8, and make frame  $UF$  coincident with vision frame  $Vis$  as shown in Fig. 9a. For that, the programmer must establish user frame  $UF$  by using the three non-collinear circles on the same camera calibration grid sheet that define vision frame  $Vis$  at the vision calibration position as shown in Fig. 5. With the coinciding of frames  $UF$  and  $Vis$ , OBJ frame position measured in vision frame  $Vis$  actually represents robot TCP frame position measured relative to robot user frame  $UF$  as shown in Eq. (1):

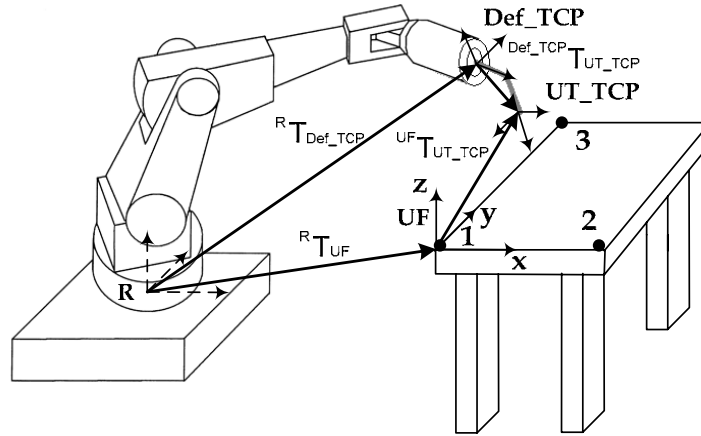
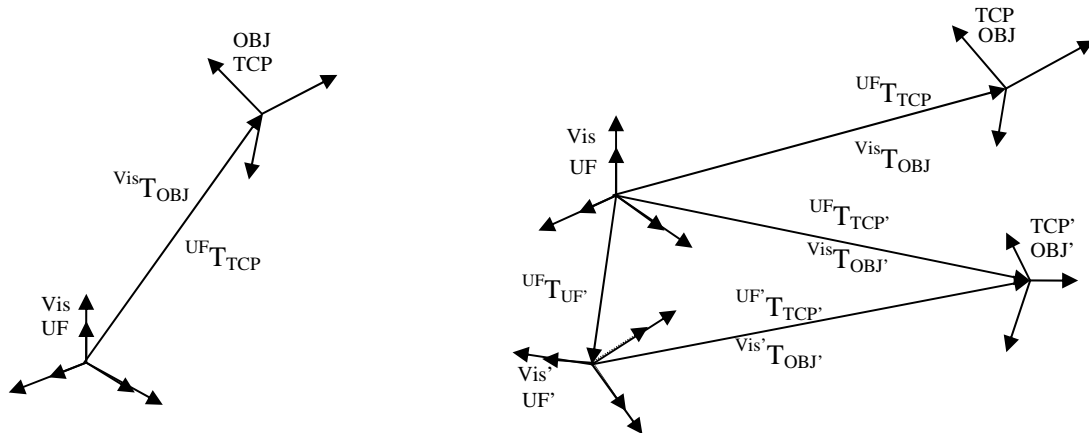


Figure 8: Robot Reference Frames and Frame Transformations



(a) Coincidence of Frames UF and Vis

(b) Transformations Under Vision Guidance

Figure 9: Frame Transformations in Robot Vision and Systems

$${}^{Vis}T_{OBJ} = {}^{UF}T_{TCP} . \quad (1)$$

There are two ways of applying the vision measurements (e.g.,  ${}^{Vis}T_{OBJ}$  or  ${}^{Vis}T_{OBJ'}$  in Fig. 9b) in an industrial robot program. The robot programmer may directly use a vision measurement as a robot TCP frame (or hand) position. However, due to the position differences of the vision-identified parts in a camera view, the robot program cannot perform the same robot operations that are defined via a standard part to a new vision-identified part after the robot hand reaches the new OBJ frame position. To avoid this problem, the robot programmer usually obtains two different vision measurements (e.g.,  ${}^{Vis}T_{OBJ}$  and  ${}^{Vis}T_{OBJ'}$ ) and determines the position change of robot user frame UF,  ${}^{UF}T_{UF'}$ , as shown in Eq. (2) and Fig. 9b:

$${}^{UF}T_{UF} = {}^{Vis}T_{OBJ} \times ({}^{Vis'}T_{OBJ})^{-1} \quad \text{and} \quad {}^{Vis'}T_{OBJ} = {}^{Vis}T_{OBJ} . \quad (2)$$

With vision-measured offset value  ${}^{UF}T_{UF}$ , the robot program is then able to apply Eq. (3) to transform all pre-defined robot operation points into the corresponding ones for any vision-identified part:

$${}^{UF}T_{TCP} = {}^{UF}T_{UF} \times {}^{UF'}T_{TCP} \quad \text{and} \quad {}^{UF'}T_{TCP} = {}^{UF}T_{TCP} . \quad (3)$$

In the case of the FAUNU robot workcell in Fig. 1, the FANUC VisLOC vision system [7] automatically measures the OBJ frame position (i.e.  ${}^{Vis}T_{OBJ}$ ) of the object when the robot programmer trains the object pattern. Executing the VisLOC micro function “Snap & Find” in the FANUC Teaching Pendant (TP) program [8] allows the vision system to identify the OBJ frame position (i.e.  ${}^{Vis'}T_{OBJ}$ ) of another object in the camera-view picture and calculate offset value  ${}^{UF}T_{UF}$  for the robot system as shown in Eq. (2). The vision-determined “offset” value is saved in the specified robot position register,  $PR(1)$ . The robot program then utilizes the “offset” value in the following three TP program instructions to modify all pre-taught robot TCP positions (e.g.,  $P(1)$  and  $P(2)$ ) in user frame UF (e.g.,  $UFRAME(4)$ ) into the corresponding ones (i.e.  $P(1)'$  and  $P(2)'$ ) for the vision-identified object as shown in Equation (3):

- 1: *Offset Conditions*  $PR(1)$ ,  $UFRAME(4)$
- 2: *J*  $P(1)$  100% *Offset*
- 3: *J*  $P(2)$  100% *Offset*

where the motion instruction elements  $J$  and 100%, respectively, specify the motion type and speed override of the TCP frame.

### 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 the camera lens in a 3D binocular stereo vision system cannot be perpendicular to the calibration grid sheet for conducting perspective calibration, 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:

$$\begin{aligned} \text{Average Error} &= 0.2 \times 0.76 & \text{Maximum Error} &= 0.45 \times 0.76 \\ &= 0.152 \text{ (mm)} & &= 0.342 \text{ (mm)} \end{aligned}$$

Due to pattern distortion, the variation of height measurement occurs 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, a block was moved (as shown in Fig. 6) 10 mm each time along the  $x$ - and  $y$ -axis of the camera view plane, respectively, and measured the corresponding height value of the block by using the VisLOC Snap & Find function as shown in Fig. 8. Then, the relationship between the average height 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 x-axis. Its model is:

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

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

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

These relationships illustrate that the position change of the object along x-axis produces some degrees of pattern distortion and mismatch, which results in an inaccurate measurement of the block height. The height distortion is caused by not only the location variations of the object but also the decreasing of the identical features in both camera views. However, a higher consistency of the measurements remains in y-axis due to more overlapped features in both camera views.

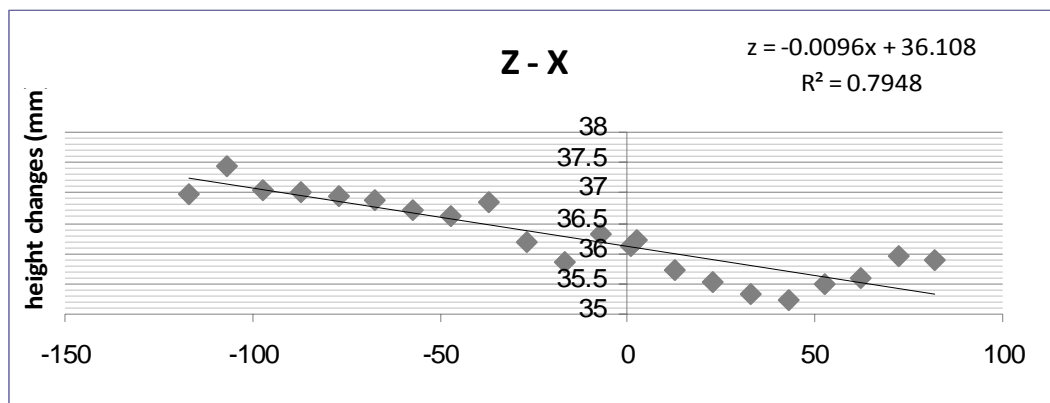


Figure 10: Height Changes along X-Axis

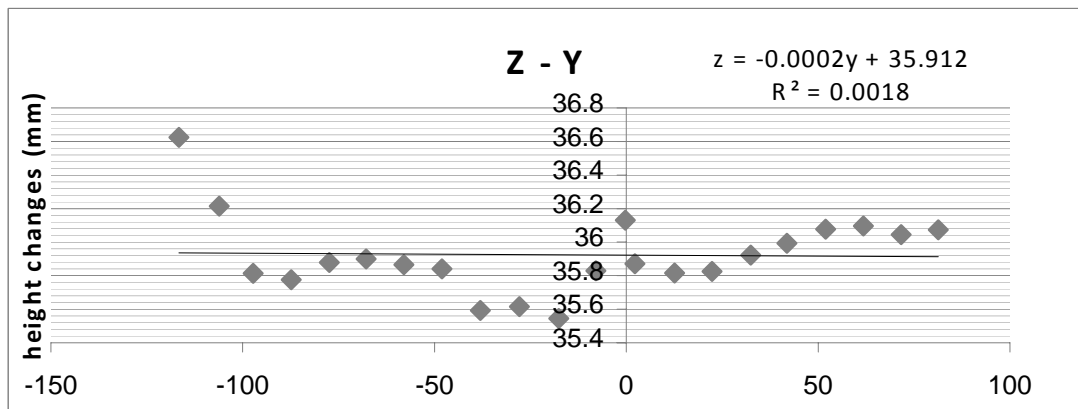


Figure 11: Height Changes along Y-Axis



## Conclusion

This paper provided the robot programmer with the knowledge and method of using 3D binocular vision systems for enhanced flexibility and intelligence of industrial robot operations. 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 instituting a robot user frame and making it coincide with the vision frame. The robot program uses the vision “offset” value to transform all pre-defined robot operation points to the corresponding ones for the vision-identified object. The results also show that the 3D binocular stereo vision system is able to provide the robot system with the accurate position and height measurements of a 3D object appeared in the center of the camera view. The variation of height measurements occurs as the vision-identified object is away from the center of the camera view.

## Reference

- [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 is an associate professor in the department of Engineering and Technology at Central Michigan University. He holds a M.S. degree in mechanical engineering and a Ph.D. in industrial engineering from the University of Cincinnati. Dr. Cheng has taught various courses in robotics and industrial automation. His research interests include robotics, control systems, and factory automation. He is a member of Institute of Electrical Electronics Engineers (IEEE) and American Society for Engineering Education (ASEE).

Xiaoting Chen is a graduate student studying 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. His research interests include robotics, control systems, and industrial automation.