Three-dimensional gauging with stereo computer vision

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · September 1991

DOI: 10.1117/12.48233

Mohammed Taleb Obaidat
Jordan University of Science and Technology; Jadara University

PROCEEDINGS

Industrial Vision Metrology

Sabry F. El-Hakim
Chair/Editor

11–12 July 1991
Winnipeg, Manitoba (Canada)

Organized by
ISPRS—International Society for Photogrammetry and Remote Sensing
NRCC—National Research Council Canada

Cooperating Organizations
SPIE—The International Society for Optical Engineering
ASPRS—The American Society for Photogrammetry and Remote Sensing

SPIE
P
PROCEEDINGS SERIES

SPIE Volume 1526

Three-dimensional gaging with stereo computer vision

Kam W. Wong, Ying Ke, Michael Lew and Mohammed Taleb Obaidat

University of Illinois at Urbana-Champaign
2213 Newmark Civil Engineering Laboratory
205 N. Mathews Ave.
Urbana, Illinois 61801

ABSTRACT

Three-dimensional gaging involves the measurement and mapping of three-dimensional surfaces. Gaging accuracy depends on measurement accuracy in the images, image scale, and stereo geometry. Multiple cameras are often needed to provide adequate stereoscopic coverage of the object. This paper reports on an automatic 3-D gaging system that is being developed at the University of Illinois at Urbana-Champaign. A portable 3-D target field, consisting of 198 targets each identified with a bar code, is used to determine the interior and exterior orientation parameters of each camera. Image processing algorithms have been developed to identify conjugate image points in stereo pair of images, and the object-space coordinates of these points are computed by stereo intersection. Software have also been developed for analysis and data editing.

1. INTRODUCTION

Gaging is an important process in manufacturing, and computer vision has already gained widespread applications for gaging in two dimensions. The ability of vision systems to perform 2-D measurement for process and quality control is one of the major impetus for recent development of computer vision techniques. The challenge in 3-dimensional gaging is the development of algorithms for automated operation to achieve the accuracy and speed that is usually needed in a manufacturing environment.

This paper reports on an automated 3-dimensional gaging system that is being developed at the University of Illinois at Urbana-Champaign. A PC-based vision system, consisting of six 512 x 512 pixel CCD cameras forming three stereo pairs, was assembled from off-the-shelf components for program development and verification. Algorithms and computer software have been developed for the calibration of multiple cameras, automatic location of surface points on three-dimensional objects, integration of multiple stereo models, and computation of surface area, volumes, and cross-sections.

2. FACTORS AFFECTING ACCURACY

Figure 1 illustrates the basic geometry of stereo photogrammetry. An object to be measured is imaged from two cameras. The position and orientation of the two cameras with respect to the object-space coordinate system must first be determined. Then, by measuring the image coordinates of object point j on the two images, the object-space coordinates (X_j, Y_j, and Z_j) of point j can be computed. The accuracy of the computed coordinates depend on the following three factors:

1. accuracy of the image coordinates;
2. image scale; and
3. base/height ratio of the stereo camera setup.

2.1 Accuracy of Image Coordinates

In computer vision, there are basically two approaches for computing the image coordinates of an object point. If the object point is a well defined symmetrical target, such as a black circular target on a white background, the target must first be located in each image. The center coordinates of the target in each image are then computed by some centroiding algorithm. If the object has surface texture and/or markings that are represented by different levels of gray in the images, then an arbitrary point is first defined in one image. The corresponding point on the second image is then located by image matching or correlation algorithms. Both approaches are capable of computing image coordinates with an accuracy of a few hundredths

0-8194-0654-6/91/$4.00

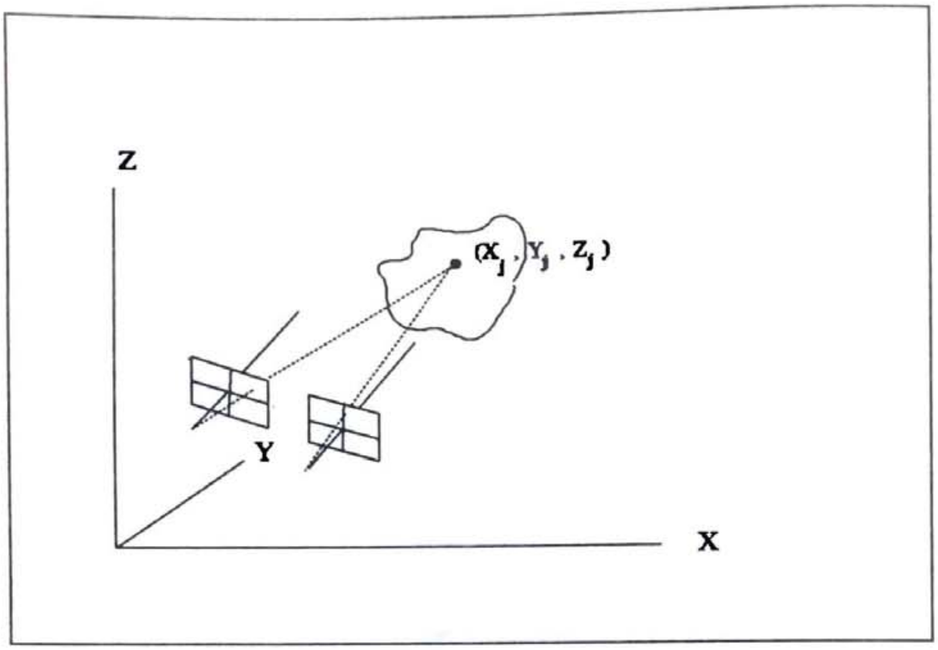

Figure 1. Geometry of stereo photogrammetry.

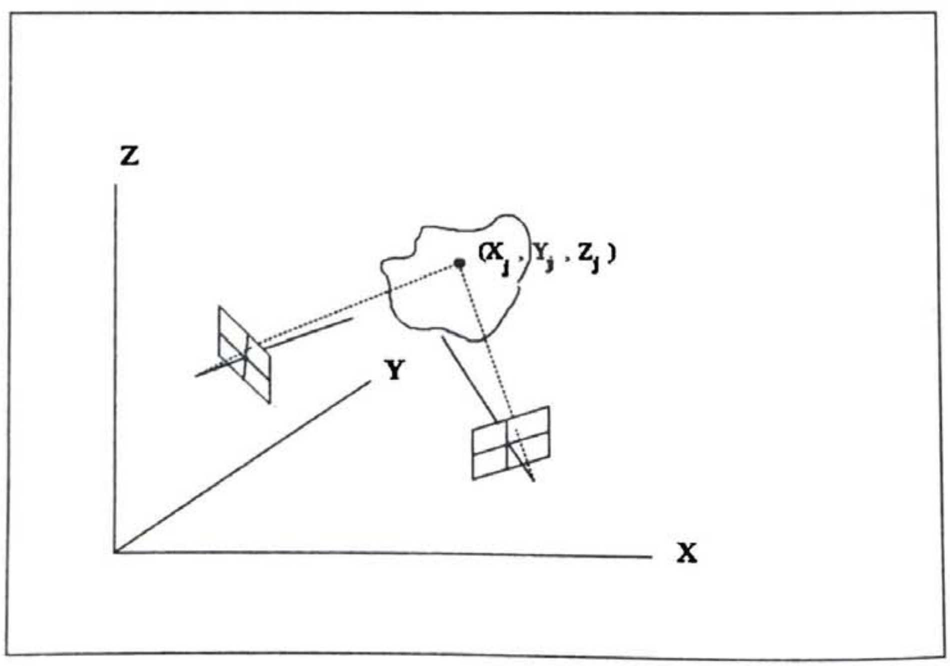

Figure 2. Convergent stereo imaging.

of a pixel.[1,2] However, accuracy of the computed image coordinates are also degraded by any uncorrected geometric distortions that can be caused by both the optical lens and the system electronics. Our research have shown that through proper calibration of the vision system, an accuracy of $\pm$ 0.1 pixel can be consistently achieved for the image coordinates.[3]

2.2 Image Scale

Measurement accuracy in the object space is directly related to image scale. For example, consider a CCD camera which has a focal plane consisting of 512 x 512 pixels, with each pixel measuring 0.02 x 0.02 mm. At an image scale of 1:1,000, a error of 0.1 pixel on an image coordinate would translate to an error of 2 mm in the object space. The same image coordinate error of 0.1 pixel would result in an error of only 0.02 mm in the object space if the image scale is enlarged to 1:10. However, as the image scale is increased, the area of the object that can be covered in an image is decreased. The camera in the above example will cover an object area of about 10 x 10 meters and 0.1 x 0.1 meter at an image scale of 1:1,000 and 1:10 respectively. Thus, increase in image scale can only be achieved at the expense of the area of coverage. For high measurement accuracy, multiple stereo models from different viewing angles are usually needed to provide adequate coverage. Algorithms for accurate integration of the different stereo models are then essential to preserve the overall accuracy of the measurements.

2.3 Base/height ratio

The base/height (B/H) ratio of a stereo camera system is defined as follows:

$$\frac{B}{H} = \frac{\textit{Base distance between the two cameras}}{\textit{Average perpendicular distance from base to object}}$$

The B/H ratio basically defines the geometry of intersection of the two rays of light from the two cameras to a common point on the object. The smaller the ratio, the smaller is the angle of intersection subtended by the base at the object point; and therefore, the larger is the uncertainty in the computed distance of the object point from the camera base.

Because of the small focal planes of CCD cameras, and the resulting small angular fields of view, the camera axes are often tilted with respect to each other in order to achieve the required B/H ratio. See Figure 2. Such a convergent arrangement, however, will result in larger perspective differences between the two images; which, in turn, will result in larger uncertainty in finding match points in the two images.

3. A STEREO GAGING SYSTEM

A PC-based stereo gaging system was assembled from off-the-shelf components for the primary purpose of program development and concept verification. The system consists of the following major components:

- 1 Zenith 16-MHz 386/AT personal computer, with 4 MB RAM, 80-MB hard drive, and a 5.25-inch high-density disk drive
- 2 EPIX silicon video MUX, 20 MHz, frame grabbers each equipped with 1 MB image memory
- 4 PULNiX TM-80 2/3-inch frame transfer CCD cameras, 800(H) x 490(V) pixels, Fujinon 12.5-75 mm F1.2 zoom lens
- 2 PULNiX TM-845 2/3-inch frame transfer CCD cameras, 800(H) x 490(V) pixels, with asynchronous variable shutter adjustable between 1/60 and 1/8000 second, Fujinon 12.5-75 mm F1.2 zoom lens

The six cameras are grouped into three stereo pairs. Each pair of camera is mounted on a common bar, which permits a base separation ranging from 300 mm to 760 mm between the two cameras.

The EPIX frame grabber has an internal video input signal switch. By using a single frame grabber and the EPIX SVIP software package, the system is presently capable of capturing images sequentially from the six cameras and downloading the images into the computer's hard disk in about 20 seconds of time.

At the present time, this system is being used primarily for image capture. Data processing and analysis are being performed on Apollo DN4000 workstations, which are interconnected to a network of over 100 Apollo workstations.

4. GEOMETRIC CALIBRATION

Geometric calibration of a vision system involves the determination of the interior geometry of the camera system, including: position of the principle point in the image plane, focal length, and geometric distortions caused by the optical lens and the system electronics. Extensive research have shown that significant geometric distortions exist in some CCD cameras.[3,4] Distortions amounting to as much as 10 to 15 pixels in the edge of the focal plane is not uncommon. In order to achieve sub-pixel measurement accuracy, it is essential that the geometric distortion characteristics be calibrated so that the distortions can be corrected during the computing process.

A variety of techniques have been developed for the geometric calibration of film cameras.[5,6] Recent literature have also reported on some techniques specifically developed for vision systems.[7,8,9,10] With the objective of developing a fully automated calibration procedure, we decided to employ a three-dimensional target field consisting of 198 targets located in 6 different planes. See Figure 3. Each target consists of a black circular target on a white background, and is identified by a bar code. The centers of the targets within each plane are positioned with an accuracy of $\pm$ 0.03 mm. To calibrate a vision system, an image of the target field is captured with the system. The interior geometric characteristics of the system can then be automatically determined by a computer program in the following sequential steps:

1. identify the targets and their corresponding bar code identification numbers in the image;
2. find the centroids of all the targets in the image;
3. use the computed image coordinates of the targets and the corresponding known 3-D spatial coordinates to compute the following parameters for the vision system:

   - image coordinates of the principal point (x_p, y_p);
   - focal length (f);
   - differential scale correction along the scan line;
   - parameters of radial lens distortions (l_1, l_2, l_3);
   - parameters of asymmetric distortions (p_1, p_2, p_3); and
   - position and orientation of the camera with respect to the target field.

The geometric corrections (dx, and dy) for an image point at coordinates (x, y) measured in the image are represented by the following expressions:

$$dx = \bar{x}(k + l_1 r^2 + l_2 r^4) + [p_1(r^2 + 2\bar{x}^2) + 2p^2\overline{xy}][1 + P_3 r^2]$$

$$dy = \bar{y}(\quad l_1 r^2 + l_2 r^4) + [2p_1\overline{xy} + p_2(r^2 + 2\bar{y}^2)][1 + P_3 r^2]$$

where $\bar{x} = x - x_p$, $\bar{y} = y - y_p$, and $r^2 = \bar{x}^2 + \bar{y}^2$. With the omission of the parameter k, the above equations

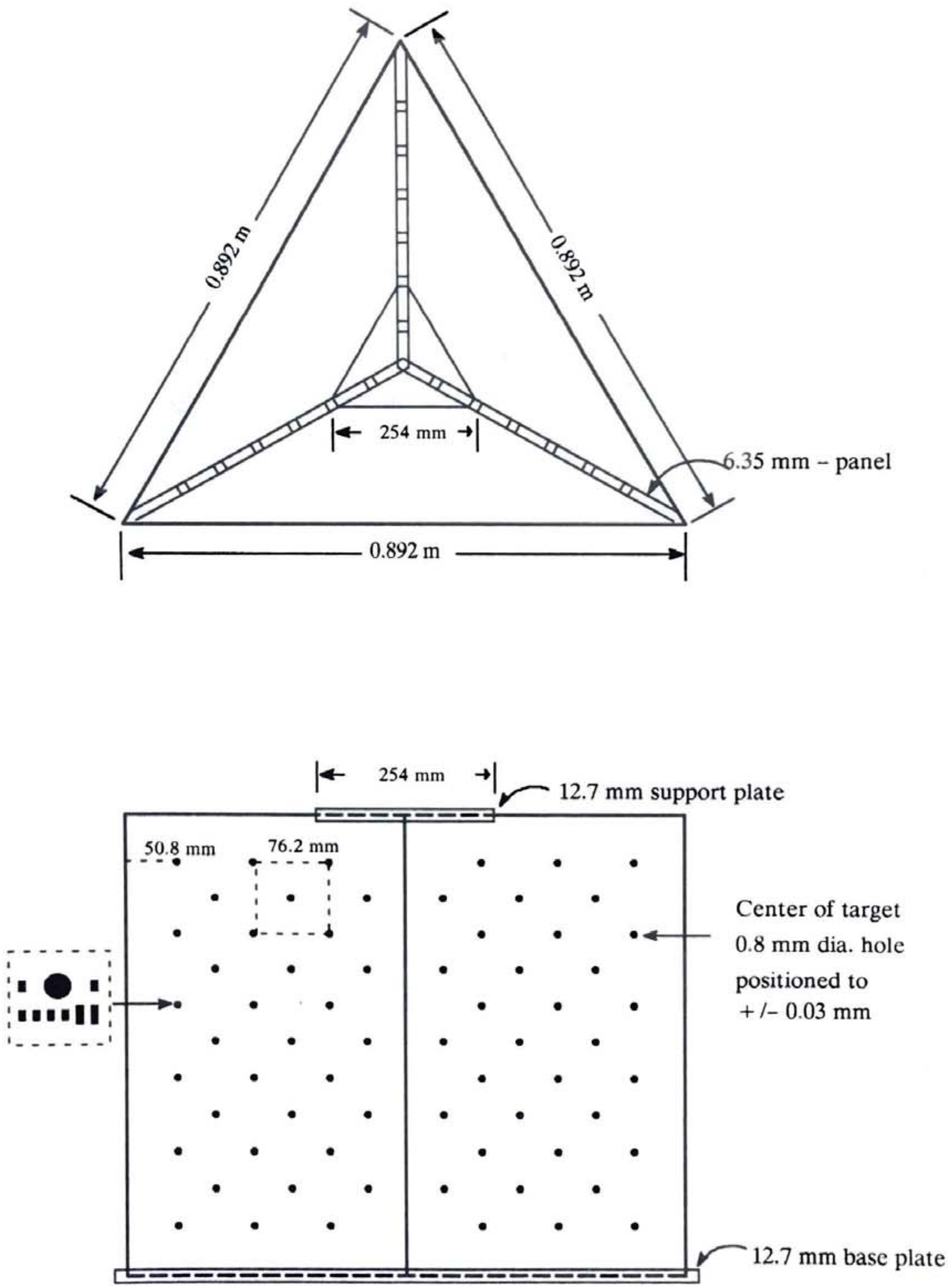

Figure 3. A three-dimensional target field for geometric calibration

would be identical to the mathematical model commonly used in photogrammetry for modelling lens distortions. The scale correction parameter, k, is needed in vision system to correct for the slight difference in scale between the horizontal (x) and vertical (y) coordinates in the image.

Figure 4 shows one arrangement of the target control field for system calibration just prior to performing 3-D gaging. With the six cameras arranged in three stereo pairs and the control field positioned near the center area, an image of the control field is acquired from each camera. Both the interior geometry and exterior orientation of each camera can be automatically computed. By removing the control field and putting in its place a three-dimensional object, the cameras are immediately ready for 3-D measurement.

Figure 5 shows an image of the control field. An accuracy of $\pm$ 0.1 pixel in residual root-mean-square (rms) error in the image coordinates has been achieved using this procedure.

5. 3-D POINT LOCATIONS

The computation of 3-D point locations in the object space involve three basic steps:

1. identification of an image point in the left image;
2. location of the conjugate image point in the right image; and
3. computation of object-space coordinates by spatial intersection using the image coordinates.

Steps 1 and 3 are relatively straight forward, while much of the problem lie in automating step 2.

Two different algorithms have been developed for finding conjugate image points in a stereo pair of images. One algorithm employs the approach of image correlation, and is intended for use in measuring objects surfaces that have well-defined textures and/or markings. A second algorithm depends on the projection of a dot pattern onto the object surface, and is intended for applications where surface texture or markings are lacking, such as clay models of automobiles.

The image correlation algorithm is based on the principle of image pyramid and cost minimization.[11] Given a pair of stereo images, the following basic steps are followed to locate the conjugate image points:

1. generate an image pyramid for each of the two images;
2. locate interesting points on the left image as points where intensity gradient exceed a pre-specified limit;
3. establish the genealogy of each interesting point on the left pyramid;
4. perform two dimension search along the right pyramid to find the conjugate points on the right image. A conjugate point is found by locating the path on the right pyramid that results in the smallest sum of intensity differences at all level of the pyramid.

This algorithm can compute a conjugate point in about 1 second of time on an IBM PC/AT-type personal computer. Thus, conjugate match points for 600 points can be computed in about 10 minutes. Preliminary tests have shown that the algorithm has an accuracy of about 0.5 to 1.0 pixel. If processing time is not constrained, further increase in matching accuracy may be possible by using the output from this algorithm as input to a least-squares correlation program.

The second algorithm for finding conjugate image points is based on the principle of targeting. A dot pattern is projected onto the surface of the object to provide a set of well-defined points for targeting. See Figure 6. The center of each dot in an image is determined by a centroiding algorithm. Several reference dots, which are 1.5 times larger in diameter, are used to help locate matching dots on the two images of a stereo pair. Depending on image quality, centroiding accuracy of $\pm$ 0.1 pixel or better is possible.

6. INTEGRATION OF STEREO MODELS

If the vision cameras are calibrated in an arrangement as shown in Figure 4, the positions and orientation of the cameras would already have been determined as a part of the calibration process. Object-space coordinates computed from the different stereo models would all be referenced to a single coordinate system, and no further computation step is needed to

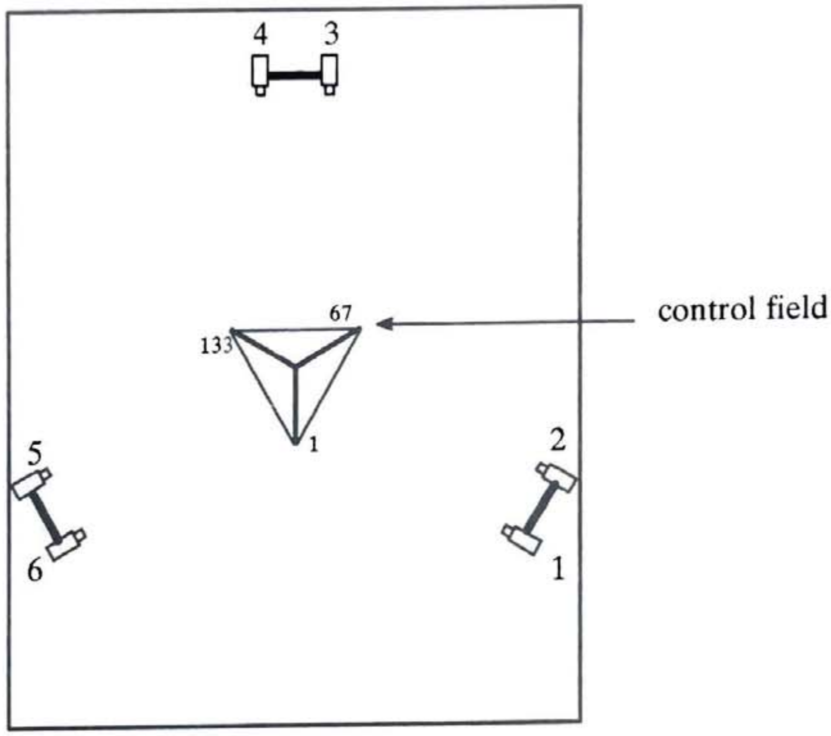

Figure 4. Arrangement for the calibration of a 3-D gaging system

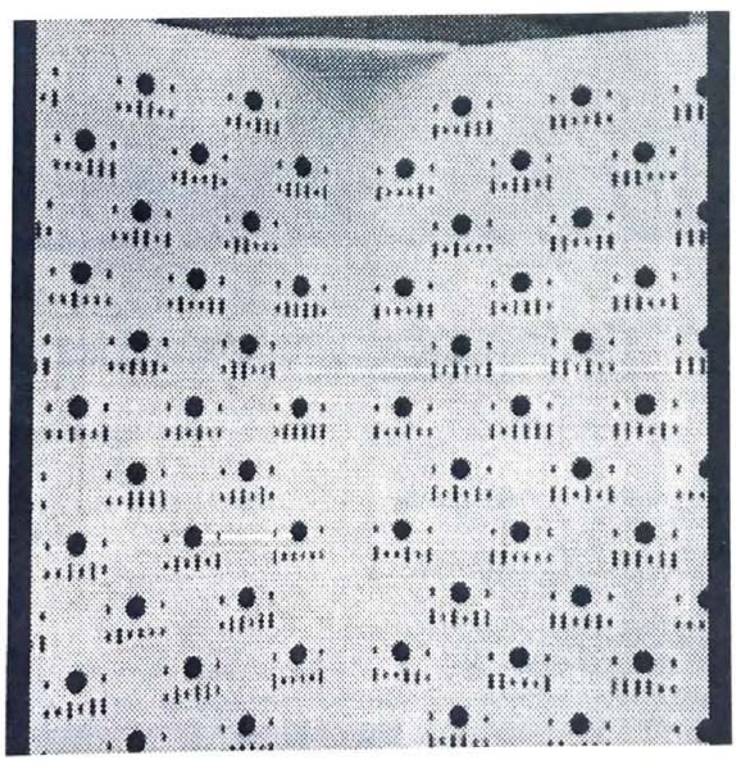

Figure 5. An Image of the 3-D target field

Figure 6. Projected dot pattern for targetting

integrate the stereo models. This approach can be expanded to use as many camera pairs as necessary to provide complete stereo coverage of the object to be measured.

If the number of available of cameras is limited and the object is too large to be measured in a single setup as shown in Figure 4, then target points must be provided for linking the stereo models from different setups. In Figure 6, the large dots near the subject's shoulder and waist areas are used for this purpose. In this manner, different sections of the body can be measured and then linked together to form a single digital model of the body by using common points between sections.

7. ANALYSIS AND DATA EDITING

Analysis and data editing are important functions in 3-D gaging using stereo visions. Computer software has been developed for the following functions:

1. computing volume, surface area, cross-sectional area, and circumferences;
2. display perspective views from different viewing positions;
3. display of images and interesting points; and
4. visual verification of match points.

Continuing efforts are being directed towards the development of software to expand the capability and simplify the processes of analysis and editing.

8. CONCLUSION

Complete automation of 3-D gaging using stereo computer vision is definitely possible. The problem is simplified when such a system is designed to perform repeated measurement on the same type of objects and in a controllable environment. However, the gaging system become significantly more complex if it is intended for application on a wide variety of objects of different sizes, shapes, and surface textures.

Measurement accuracy of $\pm$ 0.1 pixel on the focal plane is achievable. The accuracy in computing 3-D locations in the object space depends largely on image scale and the stereo geometry. Technically speaking, by using sufficiently large number of cameras, objects of any size can be measured to almost any level of accuracy.

9. ACKNOWLEDGEMENT

The research reported in this paper was supported in part by the University of Illinois Manufacturing Research center; and in part by the University of Illinois Advanced Construction Technology Center being sponsored by the U.S. Army Research Office under the DoD-University Research Initiative Program.

10. REFERENCES

1. F. Ackermann. "Digital Image Correlation: Performance and Potential Applications in Photogrammetry." *Photogrammetric Record*, II(64), pp. 429-439, 1984.

2. E. M. Mikhail, M. L. Akey, and O. R. Mitchell. "Detection and Sub-Pixel Location of Photogrammetric Targets in Digital Images." *Photogrammetria*, Vol. 39, pp. 63-83, 1984.

3. Anthony G. Wiley. "Metric Aspects of Zoom Vision." *Ph.D. dissertation*, University of Illinois at Urbana-Champaign, 1991.

4. Anthony G. Wiley and Kam W. Wong. "Metric Aspects of Zoom Vision." *Proceedings of ISPRS Symposium on Close-Range Photogrammetry Meets Machine Vision* (Zurich, Sept. 3-7), SPIE Proceedings Vol. 1395, pp. 3-7, 1990.

5. Duane C. Brown. "Close-range Camera Calibration." *Photogrammetric Engineering*, Vol. 37, No. 8, pp. 855-866, 1971.

6. H. Ziemann and Sabry El-Hakim. "System Calibration and Self Calibration-Part 1." *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 10, pp. 1617-1625, 1986.

7. Sean Curry, S. Baumrind and J.M. Anderson. "Calibration of An Array Camera." *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 5, pp. 627-636, 1986.

8. J. G. Fryer and S. O. Mason. "Rapid Lens Calibration of a Video Camera." *Photogrammetric Engineering and Remote Sensing*, Vol. 55, No. 4, pp. 437-442, 1989.
9. Sabry F. El-Hakim, "Real-Time Image Metrology with CCD Cameras," *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 11, pp. 1757-1765, 1986.
10. Tsai, Roger. "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision." *IEEE Proceedings of CVPR*, pp. 364-374, 1986.
11. K. W. Wong and Michael Lew. "Stereometric Vision." *Technical Papers 1990 ACSM-ASPRS Annual Convention*, Denver, Colorado, March 18-23, 1990, Vol. 27, Part B5, pp. 647-656, 1988.