A REAL-TIME SYSTEM FOR OBJECT MEASUREMENT WITH CCD CAMERAS

by
S.F. EL-HAKIM
Photogrammetric Research Section
National Research Council of Canada, Ottawa

ABSTRACT

A system for automatic measurement of 3-dimensional coordinates of object points is described. Data acquisition, image enhancement, feature extraction, target recognition, and target image-coordinates measurement to sub-pixel accuracy, are carried out in real time. Image matching and computation of object coordinates of the targets are then performed in near real-time. The accuracy achievable with this system is compared to that of a metric camera and a direct coordinate-measuring machine.

1. INTRODUCTION

Photogrammetry is ideally suited for precision non-contact measurements for industrial applications, particularly for objects which are difficult to be precisely measured by a direct or a contact method. It is unfortunately not widely used or even considered for most of these applications. The main reason for this is that traditional photogrammetric techniques, based on photographic images, usually require costly equipment and usually a trained operator. In addition, there is a significant time lag between the data acquisition and the final results which makes it unacceptable for many applications.

Recent advances in solid-state cameras and microcomputers provide the tools required for real-time photogrammetry to become a reality. The accuracy and stability of these cameras are rapidly improving, and the speed and capacity of microcomputers are constantly increasing. All this, combined with falling prices, makes the future of real-time photogrammetry very promising indeed. However, although the hardware is becoming available, the processing algorithms and techniques for fully automated measurements with high accuracy are still in an experimental stage. Along with the utilization of the existing developments in some dynamic and exciting fields such as digital image processing, machine vision, and artificial intelligence, photogrammetrists must develop algorithms for point measurement with sub-pixel accuracy and point matching of various images. These algorithms must be efficient, practical, very fast, and adaptable to different applications. Another area of development is the evaluation of solid-state cameras. There is currently little known about the geometric fidelity of these cameras, and suitable techniques for calibrating them for photogrammetric use are still to be developed.

In the past few years, some photogrammetrists have attempted to tackle the above mentioned areas of research. Real-time or near real-time systems, with features designed to overcome some of the problems, have been proposed by Lemmer, 1982, Haggren, 1984, Burner, et al., 1985, El-Hakim, 1985, and Wong and Ho, 1986, to name a few. Algorithms for measurement with sub-pixel accuracy have been developed, for example, by Mikhail, et al., 1984, and Havelock, 1984, while digital correlation techniques have been investigated by Foerstner, 1982, Ackermann, 1984, Pertl, 1984, and Gruen, 1985. Calibration of solid-state cameras for photogrammetric applications has been attempted by Guelch, 1984, and Curry, et al., 1985.

Since the CCD camera is an important component of the system presented here, an overview of these cameras is presented in section 2. Section 3 describes the real-time system while the approach for coordinate measurement is detailed in section 4. In the last section, test results obtained at a variety of conditions, and comparisons with results obtained by a large-format metric close-range camera and direct measurements with precision coordinate measuring machine are presented and analysed.

2. CCD CAMERAS

There are two major types of video cameras. The first is the vidicon camera based on vacuum-tube technology in which electronic beam scans a light-sensitive element to produce the image pattern. The scanning beam tends to drift with time and the vacuum tube may need frequent replacement and makes the camera bulky and fagile. The second type of cameras is based upon semiconductor technology (solid-state cameras). Most of these cameras are charge-coupled devices (CCD's) which were introduced in 1970. The CCD is composed of discrete light-sensitive elements (sensors) which can be either metal-oxide semiconductors (MOS), or photodiodes. The light energy falling on each element builds up a charge proportional to the integrated light intensity. These charges are collected in capacitors and then transferred from the array into an amplifier which outputs the image as a series of voltages. These voltages, which are analogue quantities, must then be converted into discrete numbers, for computer processing, via analogue to digital converters (A/D).

CCD cameras are compatible with digital hardware and processing techniques and they also have the advantages of being small, light-weight, virtually maintenance free, of longer life, and of higher shock resistance compared to vidicons. Most importantly, the sensors are stable over time, both in position and sensitivity, which means that the camera can be meaningfully calibrated. With the increasing resolution of new cameras, and for all the above mentioned advantages, CCD is the sensor of choice for photogrammetric applications. For more details about these cameras, see Kosonocky and Sauer, 1975 and Collet, 1985.

3. THE SYSTEM

Real-time phtogrammetry requires a system that is capable of digital image acquisition and processing at a very high speed. This means that image processing operations, such as noise reduction, image enhancement, segmentation, feature extraction and recognition, have to be performed by dedicated hardware and array processors. The system must also provide flexibility in program development, have sufficient storage for large amounts of data, and operate at a very high speed. There are already a few systems available commercially usually named "vision system" or "real-time image processing system", which provide the hardware necessary for real-time applications.

The system used at the Photogrammetric Research section of the National Research Council (NRC) of Canada is the IRI-D256 Vision System, made by International Robomation/Intelligence and consists of the following components:

1. The host computer has a Motorola MC68010 microprocessor with memory management for segmented virtual memory and floppy/Winchester controller interface. The main memory (DRAM) is 1-Mbyte and the operating system is a UNIX-look-alike (REGULUS by Alcyon Inc.) with a C-language compiler.
2. A 40-Mbyte Winchester disk and 1-Mbyte floppy disk drive.
3. Camera multiplexer, camera interface, A/D converter, and monitor output. The system accepts input from up to four different cameras simultaneously.
4. Image buffer memory for four images of 256×256 pixels each and 256 grey levels.
5. Dedicated processor for linear or non-linear point transformations and histogram calculations.
6. high speed multifunctional array processor for enhancement filtering and feature extraction operations in real time.

These components are typical of many vision systems, however, more powerful systems with higher resolution and speed are now becoming available.

Figure 1 shows the layout of the system in our laboratory. The system is used for two purposes:

1. Automatic measurement of three-dimensional coordintes of object points for industrial applications.
2. Evaluation of the performance of CCD cameras for photogrammetric use.

The automatic measurement application is described in the following section.

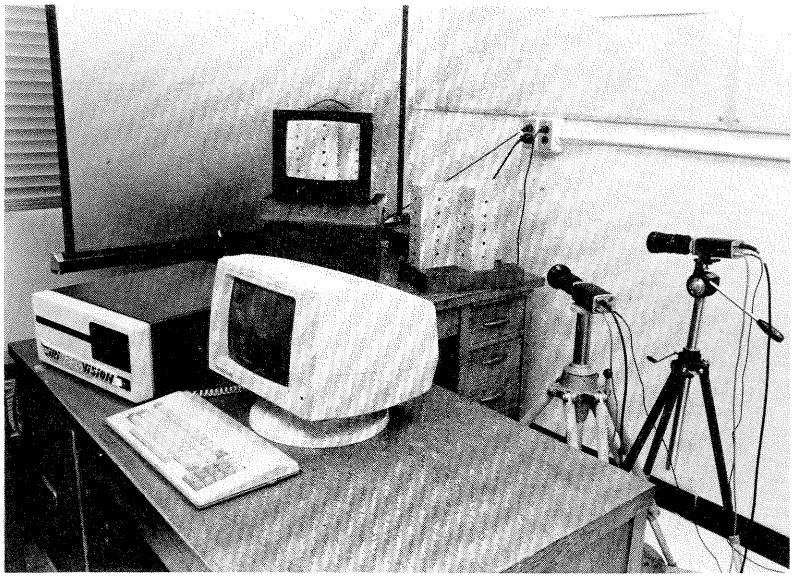


FIGURE 1. Components of the real-time system.

4. THE APPROACH

For the automatic measurement of three-dimensional coordinates of targetted object points, a program which combines image processing, pattern recognition, and photogrammetric algorithms has been developed for use with the above real-time system. The program, coded in C-language and named OBJECT, utilizes built-in functions with dedicated hardware, data files, and interactive input. A block diagram of the program is displayed in figure 2. There are eight processing steps, the first six of which are performed mostly by dedicated hardware and in real time. The system takes between 10 and 20 milliseconds to perform those six processing steps on each object point. These real-time steps result in two sets of image coordinates, one from each camera, ready for matching and photogrammetric processing. These last two steps require between 50 and 60 millisecond execution time for each object point. The photogrammetric processing solves for the orientation parameters for each camera and the object coordinates of all the identified points. Some of the program operations are described below.

4.1 Noise Reduction and Feature Enhancement:

For noise reduction, a sequence of images is taken, instead of only one, and averaged. The number of these images depends on the dynamics of the scene. Also for noise reduction, the bias frame (frame with no exposure) is subtracted from the scene image frame to remove built-in noise.

The enhancement process is needed to emphasize important features, such as targets or edges, and suppress others. Operations required for this purpose are a variety of linear and non-linear convolutions, and point-by-point mapping operations. These operations utilize about fifty built-in functions, however, a selection must be made by the user. The selection depends mainly on the environmental parameters, such as illumination and target and background texture and reflectance characteristics. Once the decision is made, usually after some trials, the linear or non-linear convolution, and/or the map-function identification number are either placed in a file or entered interactively via the keyboard. The file option is preferred to maintain real-time processing while the intractive mode is convenient when the environmental parameters are constantly changing. The enhanced image is then available for the next processing step, however, a copy of it is stored in one of the unused frame buffers for future use in target location (section 4.5).

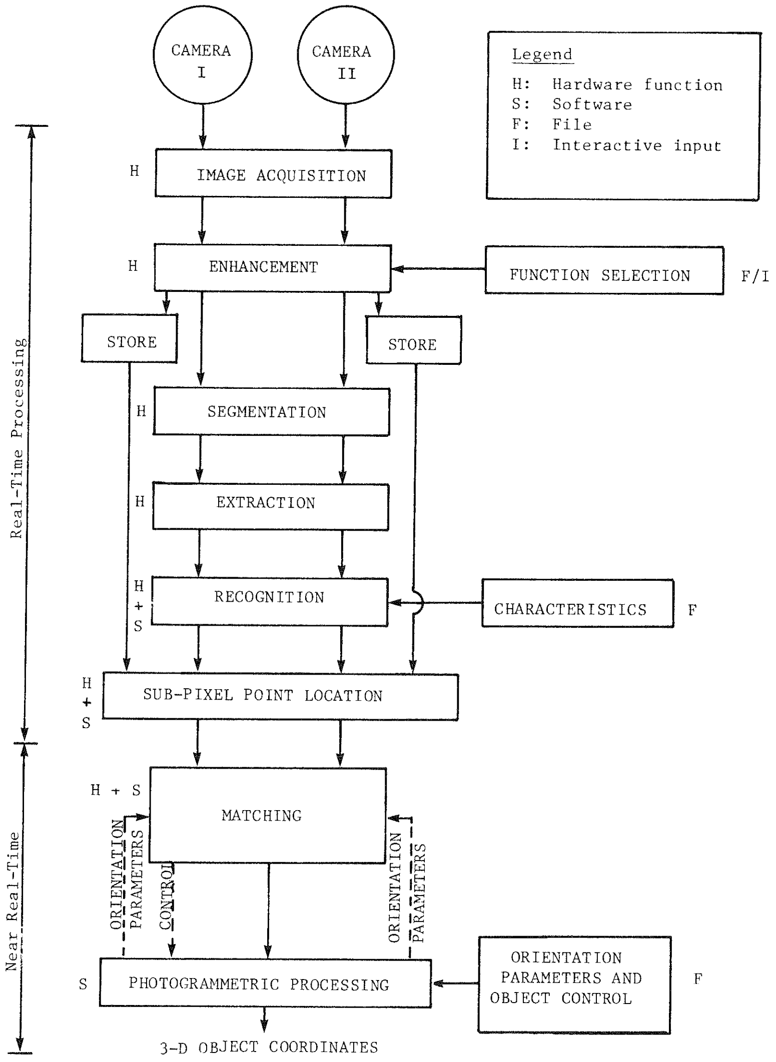


FIGURE 2. Block Diagram of Program OBJECT

4.2 Image Segmentation:

In this process, the enhanced image is converted into a binary image for easier feature extraction. The binarization of the image is a very critical step since all subsequent steps depend very much on this step. The threshold required for this operation may be different in different parts of the image, depending on the light conditions. Therefore, the image is divided into windows (up to 30 windows) and a threshold is selected for each window separately. This operation requires histogramming each window and is carried out entirely by dedicated hardware.

4.3 Feature Extraction:

The result of the above segmentation process is an image composed of white "blobs" on black background, or vice versa. The feature extraction process isolates each blob and gives it a unique label. This operation is carried out by built-in functions with dedicated hardware.

4.4 Target Recognition:

In order to distinguish the blobs representing the targets from other blobs, characteristic parameters are computed for each one. These parameters can be given by the ratio between width and height, the ratio between moments of inertia about the major and minor axes, and the ratio between each corner area enclosed between the axes and the total area of the blob. The differences between these parameters and a given set of parameters for the ideal target, which are stored on an input file, are computed and the blob is recognized as a target if the differences are within a pre-set tolerance. Once the target is recognized, it is relabelled and kept while the other blobs are eliminated from memory.

4.5 Target Location:

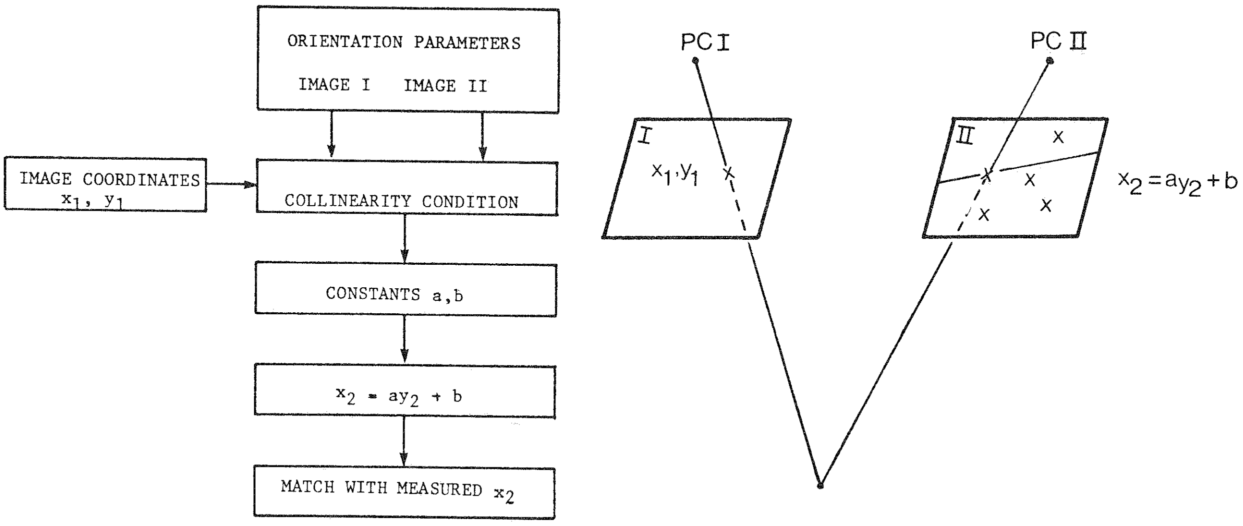
Each recognized target covers an area of several pixels and thus it is necessary to locate, with sub-pixel accuracy, the exact coordinates of the center of the target. For solid-colored targets of highly contrasting background, the coordinates of the centroid of the blob area are computed and considered to be the target coordinates. However, if the targets are designed with its center point clearly marked, for example a black target with a white point in its center, a second step of computation is carried out after the centroid has been computed. In this step, the enhanced image, from step 4.1, is recalled and the grey levels of the center pixel and the 3×3 or 5×5 pixels around this pixel are used to interpolate the coordinates of the center. Preliminary investigation showed that accuracy, when using the centroid alone, is very much affected by the degree of target-edge blur and this effect can be eliminated using the grey-level interpolation method. On the other hand, the latter adds significantly to the computation time.

4.5 Matching:

All the previous computations are applied to individual images, one from each camera. Now each point in one image must be matched with its corresponding point in the image taken by the other camera. The program handles this operation in two steps. The first is applied only to the control points, which are needed to determine camera orientation and calibration parameters, and requires a priori knowledge about the number of these points and the way they are arranged. The program makes use of this information and, by establishing the relative position of these points, it rearranges them, using their image coordinates, to match the given arrangement of the list of the object coordinates of these points. Another way of identifying the control points uniquely is that each may have a different shape. The control points are then used to determine the orientation parameters of each image by photogrammetric resection using the collinearity condition (in the case where the relation between the two camera locations is known and fixed, there is no need for control points and this step will be skipped). The second step is applied to all points other than the control points. To match a point in one image with its corresponding point in the other, the image coordinates of this point in the first image and the orientation parameters of the two images are used to determine the relationship between x and y coordinates of this point in the second image. This is a straight line relationship (the epipolar line) and is written as;

$$x_2 = ay_2 + b$$

where a and b are computed directly from the above mentioned parameters (see Figures 3 a and b). The image coordinates of all the targets in the second image are tested with the above equation and the target that satisfies the equation best is the best match. However, this best fitting must be within a preset tolerance, otherwise the point has no match and will be eliminated. Occasionally more than one target will satisfy all conditions, for example two targets in the second image matched one target in the first. In this case the epipolar line in the first image is computed to find a match for the second target in the first image. The two targets are then ordered in the same way in the two images. In the case where there is a redundancy of targets, when there are more than one target in the second image matching one in the first, these targets will be eliminated to save computation time, and the process will continue with another target.



(a) block diagram (b) target epipolar line

FIGURE 3. The Matching Process

4.7 Object Coordinate Determination:

Once all the targets are matched, their image coordinates and the orientation parameters are used to determine the object coordinates of these targets by photogrammetric intersection. This is the final output of program OBJECT.

5. TEST RESULTS AND ANALYSIS:

Several tests have been carried out with the real-time system and the program OBJECT. The objectives of the tests presented here are:

- to determine the achievable accuracy with the system and compare it with accuracy achieved with a metric camera and a high-precision direct coordinate-measuring machine.
- to study the effect of some environmental parameters and how the built-in enhancement functions could handle various situations, and
- preliminary modelling of systematic errors.

5.1 Test Objects and Cameras:

Three test objects were used in the experiments. Two of these objects are made of hard wood and are identical except for the finishing of their surfaces (fig. 4). One object is painted with a cream color opaque paint (object 1) while the second is coated with a clear polyester varnish (object 2). A number of round black targets are embedded into the wood and precisely measured by a direct coordinate-measuring machine (SIP model 560M). The dimensions of these two objects are 30 cm × 30 cm. The third object is a matt-grey steel with a dark grid printed on its surface (object 3 in figure 4).

Two solid state CCD cameras, Hitachi model KP-120 with a MOS two-dimensional sensor array with 320 horizontal and 244 vertical elements, are employed. The total image size is 6.6 mm × 8.8 mm with a square-pixel size of 27 micrometers. The lens provided with these cameras is a 12.5-75 mm F/1.8 6X zoom lens.

A metric close-range photographic camera, designed at NRC, has also been used in the tests. It takes 23 cm × 23 cm glass plates and has a focal length of 203 mm. The camera has a fixed focus at a scale of 1:3. Figure 5 shows this camera and its special mount during one of the experiments.

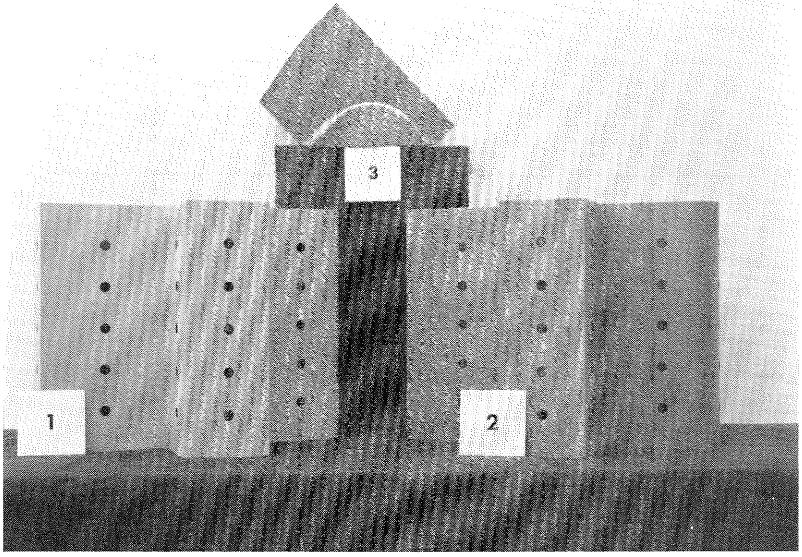


FIGURE 4. The Test Objects

5.2 The Achievable Accuracy:

The targets on the three test objects have been measured with three different methods:

A - with the coordinate measuring machine,

B - with the metric camera and conventional photogrammetry,

C - with the CCD cameras and the real-time system. In this case, illumination was controlled to produce the best results and to reduce the errors to mainly those attributed to the camera.

Table 1 displays the root mean square error (RMSE) of the differences in the coordinates computed by methods A, B, and C, with A as the reference method. The precision has been estimated in two different ways. The first, which is used for the coordinate-measuring machine, is by the repeatability, which is the variation from remeasuring the coordinates several times and computing the standard error, and the second, which is used for the photogrammetric adjustment, is by the standard error of unit weight for the residuals. The image scale for B and C is also displayed. These results were

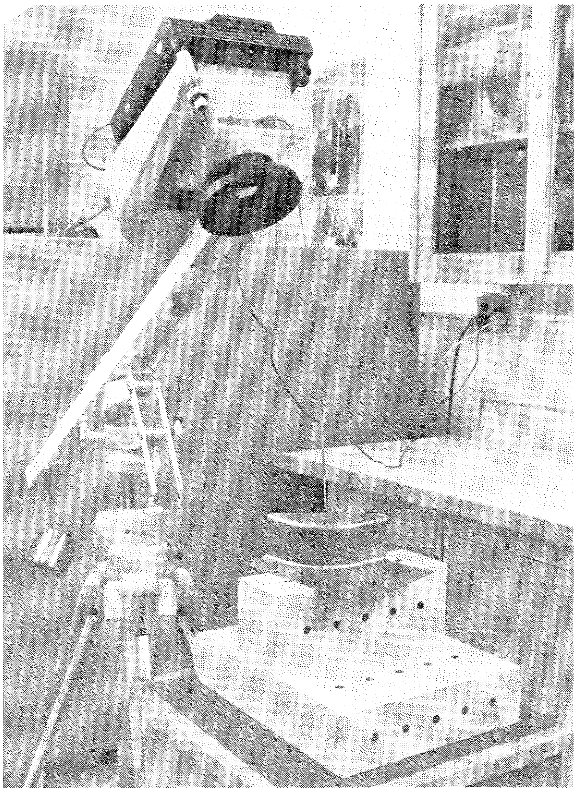


FIGURE 5. The Metric Camera

obtained after correction of some systematic errors (see 5.4).

For most cases, the results depend very much on the type of object and the shape of targets. This means, especially for the CCD cameras where the measurements are carried out automatically, that pointing seems to be a very significant contributing factor to the accuracy.

case	object	image scale	RMSE X	RMSE Y	RMSE Z	standard error X	standard error Y	standard error Z
A	1	-	-	-	-	0.003	0.003	0.002
B	1	1: 3	0.014	0.015	0.022	0.005	0.010	0.018
C	1	1:48	0.124	0.106	0.153	0.076	0.094	0.200
A	2	-	-	-	-	0.003	0.003	0.002
B	2	1: 3	0.015	0.015	0.024	0.005	0.010	0.018
C	2	1:42	0.158	0.146	0.220	0.077	0.130	0.220
A	3	-	-	-	-	0.054	0.052	0.012
B	3	1: 3	0.046	0.071	0.080	0.015	0.023	0.076
C	3	1:40	0.112	0.112	0.255	0.091	0.142	0.250

TABLE 1: Accuracy Estimates (in mm)

Since the direct measuring machine provided for objects 1 and 2 results with a high precision, the RMS values of the differences between coordinates measured by this machine and those measured by the cameras give a reliable estimate of the accuracy of the camera measurements. For object 3, the repeatability of the direct measuring machine has deteriorated significantly due to the difficulty in pointing at the large intersection points of the grid. Analysis of table 1, shows that the metric camera gave almost identical results for object 1 and 2, while the CCD cameras (case C) suffered from the lower contrast of object 2 and lost about one third of the accuracy compared to object 1. In absolute values, the CCD cameras yield an accuracy of about 0.12 mm in xy plane and 0.15 mm in z direction for the high contrast object at image scale of 1:48. This translates to about 3 micrometers at image scale, or 0.1 of the pixel size.

5.3 The Effect of Environmental Parameters:

Environmental parmeters are expected to have a major effect on the success of the automatic measuring procedure and on the final accuracy. These parameters are mainly;

-illumination (light intensity, stability, angle, and location),
-object and background texture and reflectance characteristics, and
-target and object contrast.

Another parameter, the temperature, which has an effect on the CCD radiometric performance, will not be studied here. The enhancement process is designed to compensate for the unfavourable effects these parameters may have on the recognition and positioning accuracy of targets. Many tests have been carried out under various conditions and some examples are presented below.

case	RMSE X	RMSE Y	RMSE Z
(a) Almost even illumination/no enhancement	0.19	0.19	0.28
(b) Uneven illumination/no enhancement	0.21	0.26	0.32
(c) Uneven illumination/trend removed	0.15	0.16	0.34
(d) Uneven illumination/Marr's convolution	0.17	0.16	0.25
(e) 1:70 scale/out of focus/no enhancement	recognition failed		
(f) same as (e)/min. non-linear filtering	0.24	0.23	0.42

TABLE 2: Effect of Image Enhancement (in mm)

Table 2 shows the results of various tests. All the results are uncorrected for systematic errors. Case (a) is a regular case with average room illumination. The illumination is then made uneven by creating some shades on the object, and case (b) shows the effect of this uneven illumination. The XY and Z accuracies deteriorated by about 25 to 15 per cent respectively. In case (c), the illumination trend, in both x and y directions, is removed in order to compensate for the unevenness in illumination. This is achieved by computing the accumulated profile of the grey level of the frame in each direction and determining its offset and slope. This improved the results by about one third in the XY plane, however, a slight deterioration has been noticed in Z direction. Various other filters have been applied and the most effective, for this type of targets, has been Marr's high-pass filter (Marr, 1982). The corresponding convolution, is given by the formula;

$$\nabla^2 G(r) = 1/\pi\sigma^4 (1-r^2/2\sigma^2) e^{-r^2/2\sigma^2}$$

where ∇^2 is the Laplacian operator and G is the two dimensional Gaussian distribution which has σ standard deviation and r is the radial distance from the origin. This is actually a combination of a Gaussian low-pass filter, which smooths the image, and a zero-crossing edge detection filter. Figure 6 shows the effect of 7×7 convolution with this filter. The improvement in the results is 30 and 22 per cent in XY plane and Z direction respectively.

Test case (e) (Fig. 7,a) is a situation that may occur in practice. Here, the targets appear small and out of focus and, after segmentation, it was not possible to recognize the target blobs from others in the binary image. However, by applying a non-linear filter, called a minimum filter, which replaces the grey level of the center pixel of a 3×3 matrix by the smallest of the grey levels of its neighborhood, the targets became very distinguishable (Fig. 7, b and c), and a reasonable accuracy has been achieved (case (f) in table 2).

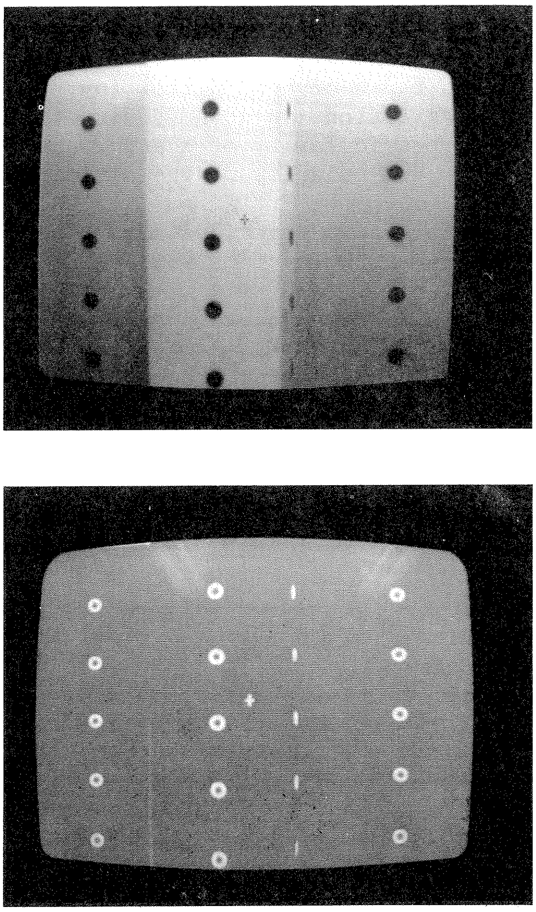
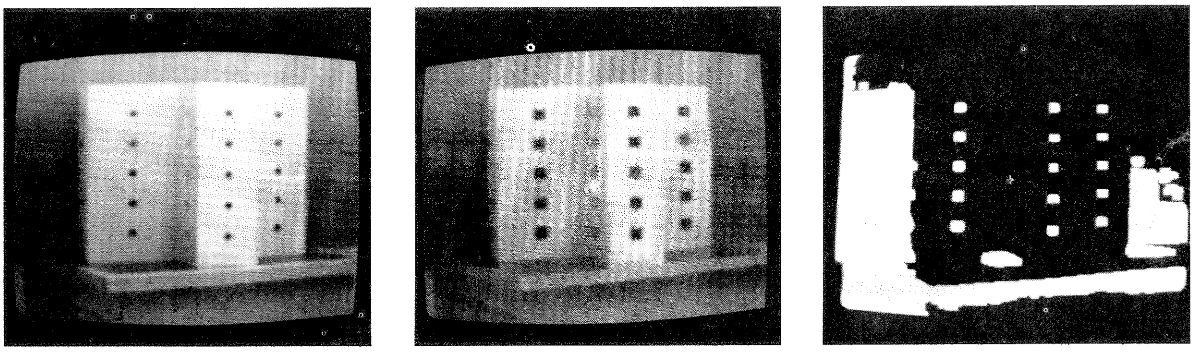


FIGURE 6. Marr's Convolution



(a) Original Image (b) Enhanced Image (c) Binary Image

FIGURE 7. Effect of the Minimum Filter

5.4 Preliminary Modelling of Systematic Errors:

The image coordinates of targets on test object 1 have been measured by the system then used in an off-line block bundle adjustment with self-calibration in order to study systematic errors in the camera system. The environmental parameters are chosen to have minimum effect on the results. This means that the illumination is uniform and the contrast between targets and object is very high. Also the orientations of the cameras are chosen to produce the best geometrical conditions for the solution. As a result of this set-up, the main source of error is now the camera. The added parameters are selected to compensate for:
-various types of lens distortion,
-scale variation, in x and y directions, in image plane,
-non-perpendicularity of x and y axes, and
-non-perpendicularity of the optical axis with image plane.

Several known functions have been tried and none performed satisfactorily, therefore a new function was formulated:

$$dx = x(a_4r^3 + a_5r\cos2\lambda + a_6r\sin2\lambda) + a_7 y$$

$$dy = y(a_1 + a_2\cos\lambda + a_3\sin\lambda + a_4r^3 + a_5r\cos2\lambda + a_6r\sin2\lambda)$$

where r is the radial distance to the principal point and λ is the angle whose tangent is y/x. There is only one parameter, a_4, for radial lens distortion, while a_1 and a_7 are for affine film deformation and non-perpendicularity of image axes. The remaining parameters are part of a harmonic function to correct for other systematic errors. Table 3 shows an improvement of about 30 per cent in accuracy when applying this function.

case	RMSE X	RMSE Y	RMSE Z
no added parameters	0.10	0.21	0.22
with added parameters	0.12	0.11	0.15

TABLE 3. Effect of Added Parameters (in mm)

6. CONCLUDING REMARKS

1. Automatic measurement of object coordinates with the real-time system and CCD cameras has produced some promising results. The achieved accuracy is about 0.1 of the pixel size at image scale. However, at the object scale, this accuracy is about 10 times worse than a good photographic metric camera can achieve due to the fact that the CCD camera has a very small format and thus needs much smaller scale to cover an object of the same size. CCD's with higher resolution than that used for the above tests are now available (such as 604×576 camera by VSP Labs Inc.) which means much better accuracy could be achieved. But even with the present level of accuracy, many applications that require real-time processing should be possible.
2. More extensive evaluation of the CCD cameras and improvement in the modelling of their systematic errors is needed and should result in higher accuracy.
3. The practical aspect of this approach faces a variety of problems. The environmental parameters, particularly illumination, are the dominant factors in the success or failure of the technique. Changes in these parameters require changes in many of the program inputs such as the enhancement function, target parameters, and the various tolerances needed for their recognition. Another problem is maintaining the fast speed for real-time processing in all situations. Some enhancement functions take more time than others and the total time is proportional to the number of blobs isolated and the number of targets. However, this problem will definitely be of less concern in the near future with faster microcomputers and parallel processing.

4. A very attractive aspect of this approach, other than the speed, is the cost. The total cost of all the hardware and the basic image processing software is about the same as of one good metric camera. There is no need for film processing, a comparator or an analytical plotter, or a trained operator. With the fast advanced field of electronics, which is resulting in better equipment at lower prices, this seems to be the system for the future, mainly for industrial and medical applications.

REFERENCES

Ackermann, F. (1984), "Digital Image Correlation: Performance and Potential Application in Photogrammetry", Photogrammetric Record, 11(64), pp 429-439.

Burner, A.W., Snow, W.L., and Goad, W.K. (1985), "Close-Range Photogrammetry with Video Cameras", Proceedings of ASP-ACSM Annual Convention, Washington, D.C.

Collet, M. (1985), "Make Way for Solid-State Imagers", Photonics Spectra, September, pp. 103-113.

Curry, S., Baumrid, S., and Anderson, J.M. (1985), "Calibration of an Array Camera", Proceedings of ASP-ACSM Annual Convention, Washington, D.C.

El-Hakim, S.F. (1985), "A Photogrammetric Vision System for Robots", PERS, 51(5), pp 545-552.

Foerstner, W. (1982), "On the Geometric Precision of Digital Correlation", Proceedings of ISPRS Commission III Symposium, Helsinki.

Gruen, A.W. (1985), "Adaptive Least Squares Correlation - A Powerful Image Matching Technique", Presented at ASP-ACSM Annual Convention, Washington, D.C.

Guelch, E. (1984), "Geometric Calibration of Two CCD-Cameras Used for Digital Image Correlation on the PLANICOMP C 100", International Archives of Photogrammetry and Remote Sensisng, Vol.XXV, part A3a.

Haggren, H. (1984), "New Vistas for Industrial Photogrammetry", International Archives of Photogrammetry and Remote Sensing, Vol. XXV, part A5.

Havelock, D. (1984), "Geometric Precision in Digital Images", International Archives of Photogrammetry and Remote Sensing, Vol. XXV, part A3a.

Kosonocky, W.F. and Sauer, D.J. (1975), "The ABCs of CCDs", Electronic Design, 8, April 12, pp 58-63.

Lemmer, J.F. (1982), "The Potential Synergy of Photogrammetry and Computer Vision", Proceedings of ASP-ACSM Annual Convention, Denver, Colorado.

Marr, D. (1982), "Vision", W.H. Freeman and Company, San Francisco.

Mikhail, E.M., Akey, M.L., and Mitchell, O.R. (1984), "Detection and Sub-Pixel Location of Photogrammetric Targets in Digital Images", Photogrammetria, 39(3), pp 63-83.

Pertl, A. (1984), "Digital Image Correlation with the Analytical Plotter PLANICOMP C 100", International Archives of Photogrammetry and Remote Sensing, Vol. XXV, part A3b.

Wong, K.W. and Ho, W.H. (1986), "Close-Range Mapping with a Solid State Camera", PERS, 52(1), pp 67-74.