JKAU: Eng. Sci., **vol. 13 no. 1,** pp. 153-170 (1421 A.H. / 2001 A.D.)

Three Dimensional Digitizers with High Resolution

M. Hamed*, R. Massen** and H. Elhendy*
Faculty of Engineering, Suez Canal University, Port Said, Egypt
***Image Transfer Center, Constance, Germany*

Abstract. This paper presents a new approach for image digitizing in the Cartizian three dimension XYZ coordinates. It is defined as the displacement vector concept based on the hierarchical stereo vision technique depending on multi camera system, distributed around the tested object. Optical sensors without any mechanical parts under the control of a host computer with an integrated full automatic calibration process are introduced. Problems of this automated style are investigated. Different factors such as calibration, feature extraction, point to point correspondence, determination of cartizian coordinates, transformation into symbolic computer aided design form are included. Performance of the proposed system is analyzed and results of the application to a human foot model as a sample of testing are determined. The execution time and effort are dramatically decreased with increased measured accuracy.

Keywords: Photogrammetry, stereo vision, image, calibration, correspondence, transfer matrix, Epipolar line, Cartizian.

1. Introduction

Nowadays, both computer-aided design (CAD) and computer aided manufacturing (CAM) are the state of art in the modern technology where their applications are spread for high quality. So the accurate measurement of object surface as a need to serve the human on the globe will be the target. It is important to deal with objects having complex shapes specially that depending on their ergonomic and aesthetic properties. Therefore, the study of space determination of the real objects may be most required for help in our life in many fields of work.

A vision concept with an image processing hardware has been developed by the Transfer Center inside a dark room where two black and white cameras[1], expandable in even numbers, are attached to an analog/digital converter frame grabber. An analog monitor displays the image with a host computer Intel-486-DX-4/100 and 16 MB RAM, 540 MB hard disk as illustrated in Fig. 1. The deduced software (C language) is implemented with the operating system OS/2 based on the routine programs of the Transfer Center as a dynamic link library. The system is controlled using the host computer, where the analog monitor shows the original photos helping in image adjustment. The external CAD applications give the request and import the prepared data in tailored steps as shown in Fig. 2.

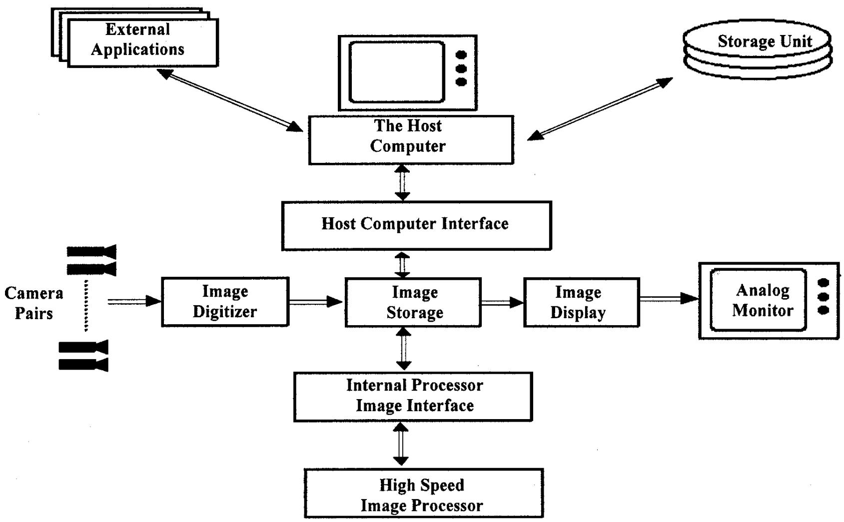

FIG. 1. The processing system.

2. Digitization

Different optical methods for digitizing the objects were appeared widely although they can be classified into Active and Passive principles. In Active (on line) methods, both light source and sensor are attached to the host computer directly while in Passive (off line) methods, one image or more would be taken and then later fed to the host computer as a second stage. The factors of flight time, depth of focus, triangulation, Moire', interferometery and stereo vision may be included[2-5].

However, stereo vision as a similar double eye biological sensor can be utilized and the distance between lens (eye) and their orientation image as epi-

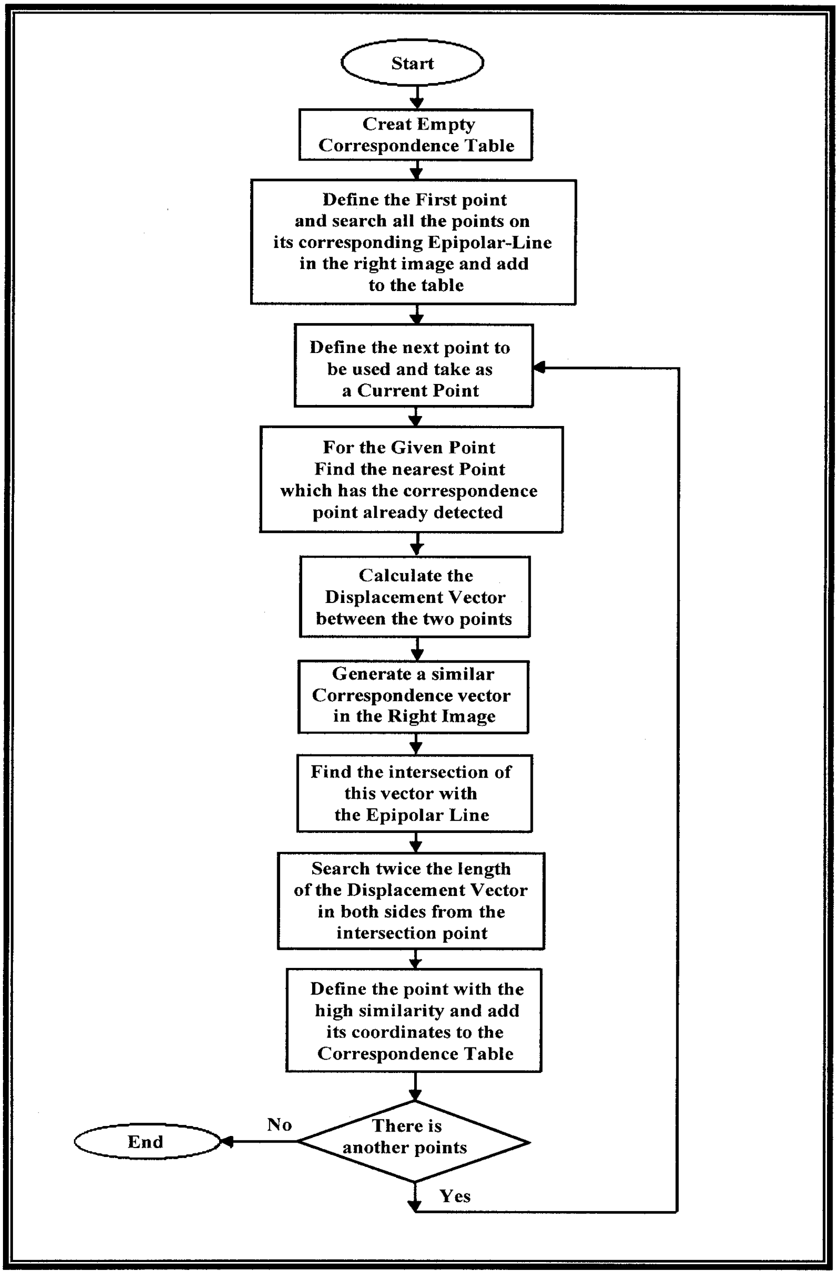

FIG. 2. The developed system flow chart.

polar lines may be introduced in the algorithm. Also the techniques of image processing such as template matching, single and double directional correlation or coarse-to-fine matching will be required for the same features in both epipolar lines. The use of multiple pairs of camera around the tested object expresses a general vision due to the possibility of determination of the internal characteristic[6 -7].

A procedure for merging to remove the redundant data in order to reduce the computational effort and minimize the consumed time should be introduced where the most important problems, as a calibration, point-to-point correspondence. The 3D coordinates calculations and transformation into a symbolic CAD system must be solved[8-9].

3. Mathematical Analysis

The stereo vision model as "Pin Hole Model" may be based on that: the lens is a point with a focal length "f" as the distance above the image plane. Thus, the line equation including the focus point and the vector from image point to the focus point may be expressed as:

$$v = v_c + a(v_i - v_c),\ a = \text{constant} \tag{1}$$

Since the object point (x_o, y_o, z_o) is an element of the line, the given object coordinates will verify it and so:

$$\{x_i, y_i, z_i\} = \{x_o\, f / (z_o + f),\ y_o\, f / (z_o + f),\ 0\} \tag{2}$$

These nonlinear equations will lead to a complicated process of calculations in 3D coordinates (x, y, z) for the object and thus a transformation concept into a linear system of equations would simplify the solution process in a mathematical form. Whatever, the use of homogeneous (h) coordinates (w_x, w_y, w_z, w) will be needed with the arbitrary constant "w" but the transformation matrix "t" must help in the formulation. Then, mapping the 3D objects on 2D plane would be the way in addition to the perspective transformation although the matrix should be reduced from (4 * 4) to (4 * 3) giving that

$$[wu, wv, w]_h = [x, y, z, 1]_h \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \tag{3}$$

Since the image coordinates (u, v) and that of object (x, y, z) are defined, the pre-specified coordinates for 6 non-coplanar points will be essential in order to deduce the unknown parameters t_{ij} on the basis of standard routines of Gaussian-Elimination or Gauss-Siedel iteration. Considering two cameras, left (L) and right (R), system with two transformation matrices, left and right for them, the equations of epipolar lines can be determined for each and the object projections on both images would be easily evaluated. Finally, after the elimination of "w" and the object coordinates, a general form for the equation may be achieved using the least square method as[10]:

$$[A]^t\ [A]\ x\ =\ [A]^t\ b \quad , \quad b\ = [A]\ \ t \tag{4}$$

4. Calibration

A calibration model is installed for the system where all its 3D coordinates are predefined exactly (Fig. 3) in order to find the transformation matrix elements for each eye. This is done automatically without any interference to make sure about the accuracy of measurement. The model has a white top points with black background to realize the required contrast and they lie on different levels (noncoplanar). The pins are located at 28 mm apart in the direction of x and 25 mm for y axis where z coordinate varies between 15 and 70 mm. Six points at least would be tested while their coordinates and the corresponding images were measured[6].

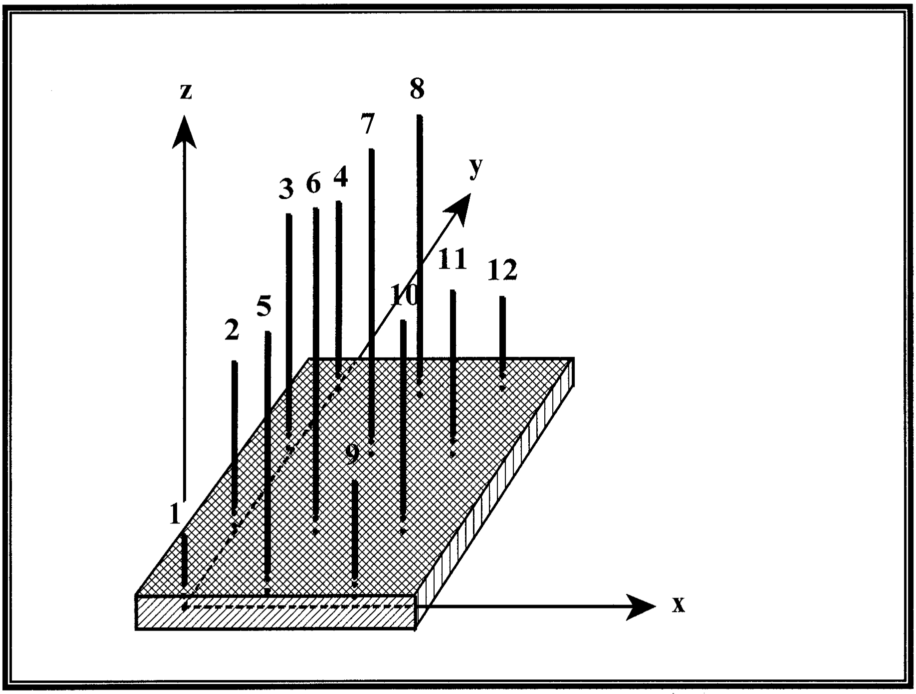

FIG. 3. Specimen with the global 3D reference coordinate system.

The model is positioned to image the points automatically without ambiguity in ordered manner according to the constraints:

$$\{x_i , y_i\} < \{ x_{i+1} , y_{i+1}\} < \{ x_{i+2} , y_{i+2}\} < \{x_{i+3} , y_{i+3}\} \tag{5}$$

where x_i and y_i are images of a point (i) in x , y coordinates, (i = 1, 5, 9) but for $x_i < x_j < x_{i+1}$, then y_j must be $< y_i$.

This constraint helps to uniquely and automatically defines for the image coordinates in both cameras. Whenever the high accuracy is reached (less than 0.4 mm) and consumes only 15 seconds including the two images, scanning and generating the two transformation matrices. This depends on converting the image into a binary shape[1] through a suitable dynamic threshold and the origin is taken at the bottom of maximum left point. This process is implemented with the proposed algorithm shown in Fig. 4 where the center (x_c, y_c) is computed from the binary image for the middle point as

$$(x_c, y_c) = \{(x_R + x_L) / 2 , (y_T + y_B) / 2 \} \tag{6}$$

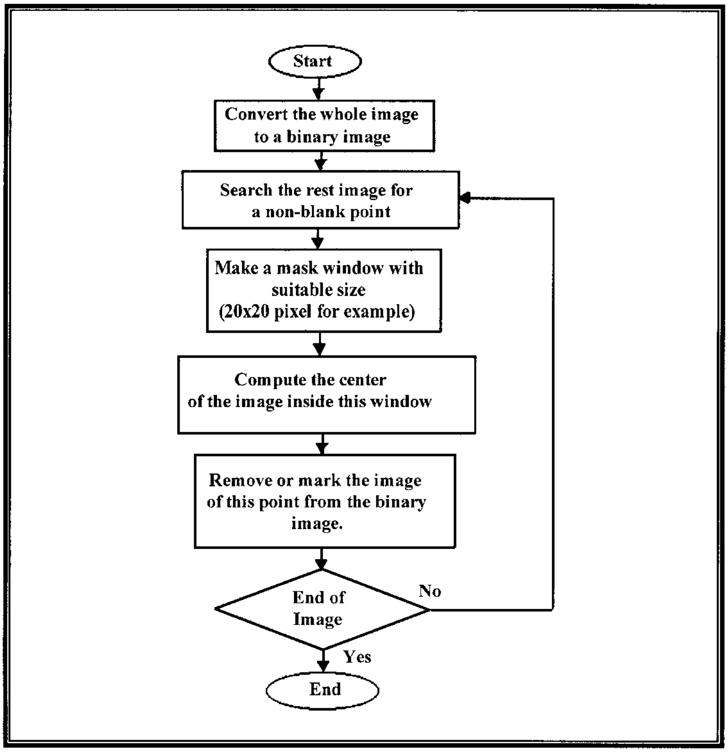

FIG. 4. Algorithm for detecting the calibration points coordinates by using the central point.

The deduced results should be modified (Fig. 5) and its results will be subjected to a third method as a center of gravity (Fig. 6) for accuracy where number of pixels with high intensity is poor compared to the other inside the tested window. A cut off intensity value as a minimum should be introduced. A low pass filter is applied to remove the high frequency components of calibration as well as noise consuming more time but for the aimed high accuracy.

Three different concepts (without, with fixed, and dynamic cut off) are included although the center of gravity is always shifted to the center of the computation window in the first concept. It is not always good in the second because of the difference in illumination while last one depends on four principles as the origin, and the three points with lowest and highest and average value of intensity. An algorithm for accurate coordinates is developed (Fig. 7) through a feedback process to overcome the problems of noise and bad illumination as well as to reduce the digitization effect. This may be reflected to the resolution of the measuring device, 400-800 pixel as a standard and the computed coordinates can be only integer which is not the real case causing a difference value of "d".

However, the effect of number of the calibrated points on the accuracy may be resulted as in Fig. 8 and noting that more calibrated points will raise the accuracy required for the digitization process. So, it is recommended to build a special calibration plate model with a large number of points, which must completely cover the measured area.

5. Correspondence

The problem here is created due to the appeared two images for the same object where they were specified as left and right images. This means that the single system of 3D Cartizian coordinates is transformed into double system of 2D images, which are simpler and consequently the relationship between both simple images becomes the target. Although the two images are deduced from the original object, the process of determination of the relative corresponding points in both will be managed.

However, the allocation of image coordinates in one plane of them and the corresponding in the second appears as a very complex problem. The area based stereo and feature based stereo methods beside the proposed concept here will reduce the computational time and effort during the process of noise and interference elimination. The window according to the area based stereo method as m columns and n rows can present the left and the right images on a similarity base looking for the highest value inside the factor of light intensity between pixels. A simple fast evaluation of the sum d of light intensity differences may be expressed as:

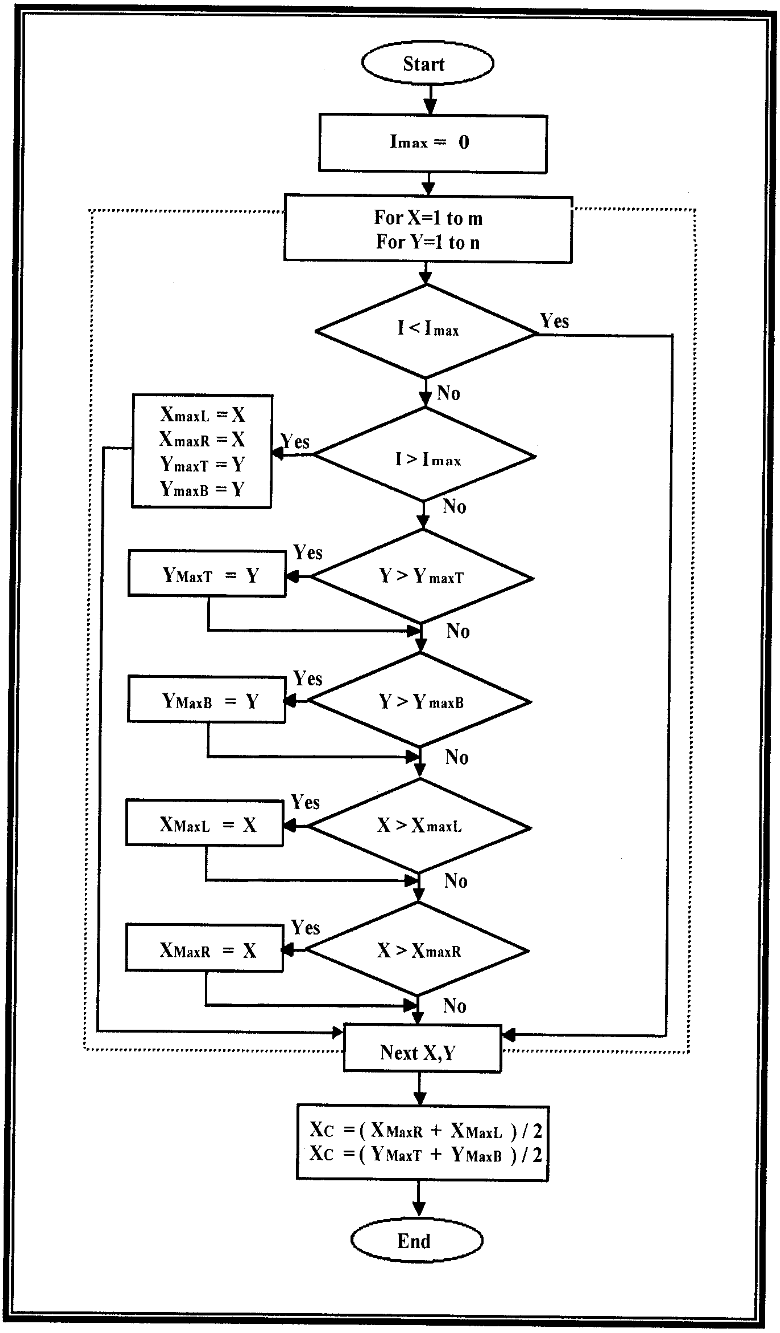

FIG. 5. Algorithm for detecting the calibration points coordinates by using the high intensity points.

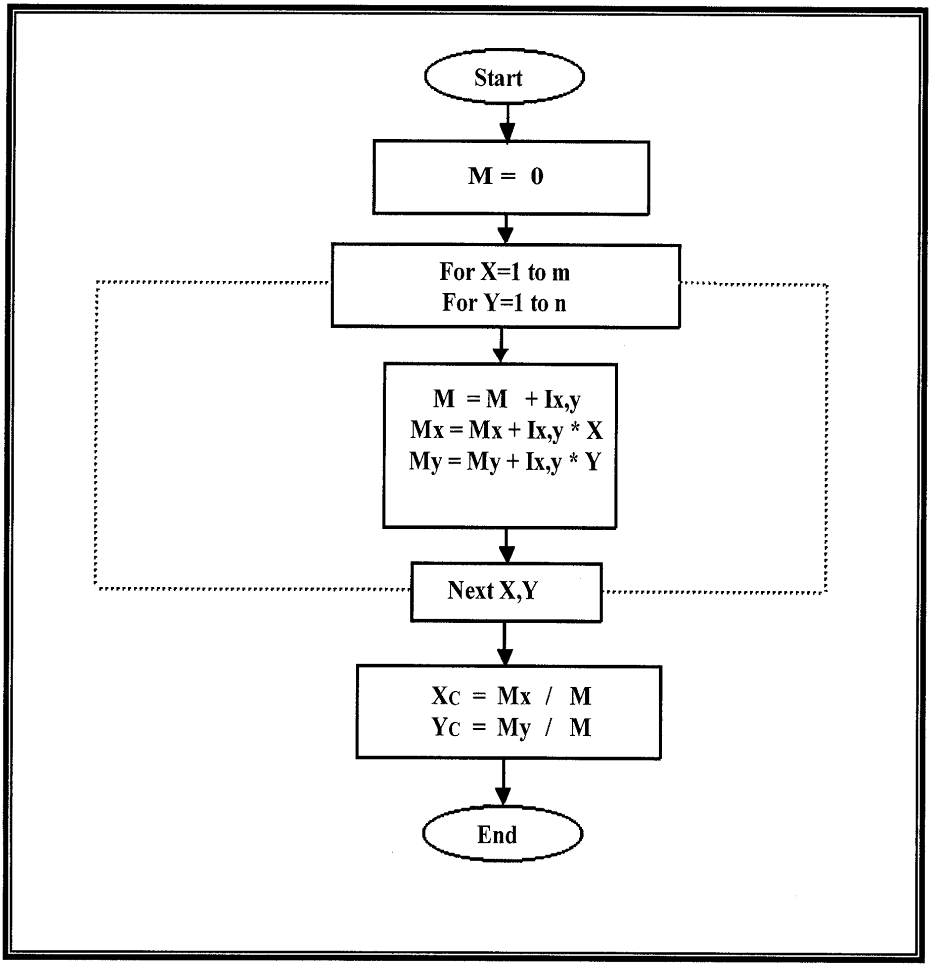

FIG. 6. Algorithm for detecting the calibration points coordinates by using the center of gravity approach.

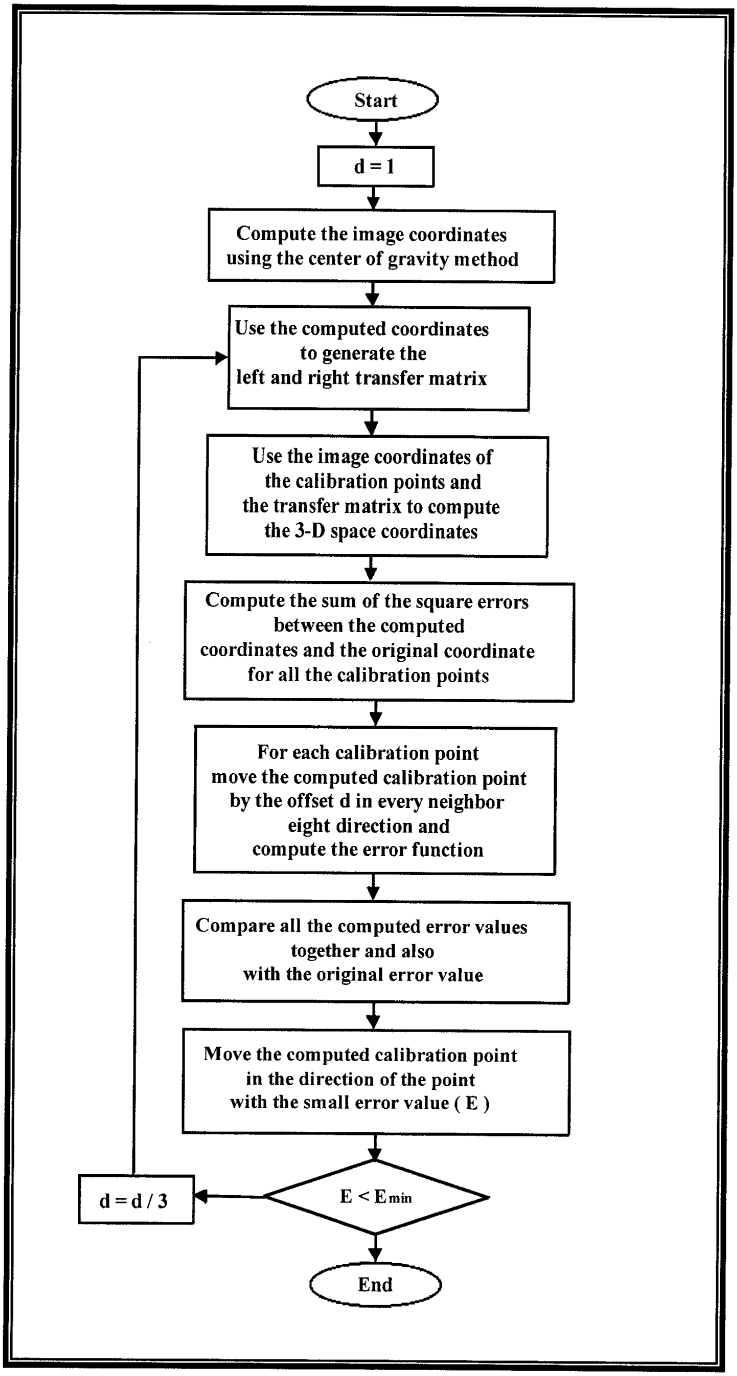

FIG. 7. Algorithm for improving the calculated center of gravity.

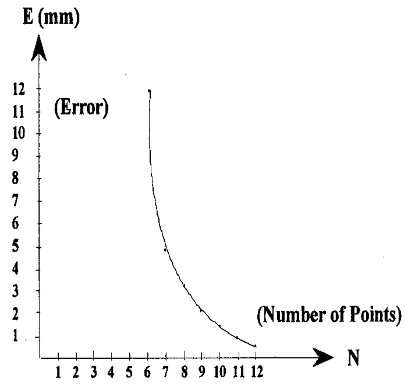

FIG. 8. Error.

$$d = \sum_{j=1}^{n} \sum_{i=1}^{m} [I(i,j)_L - I(i,j)_R] \tag{7}$$

The value of difference decreases with light intensity increase although it depends not only on the intensity but also on the level of contrasting. Then, a normalized image may be deduced to accelerate the operation with the variance σ_L because it must be calculated only once leading to an error "e" in the form:

$$e = \sum_{j=1}^{n} \sum_{i=1}^{m} \{[I(i,j)_L - I(i,j)_R] / (\sigma_L)^2\} \quad j=1 \;\; i=1 \tag{8}$$

On the other hand, the high pass filtering can remove the low frequency components in addition to the reduction of searching area to improve the performance of a tested object since the corresponding point can only lie on the epipolar line. So, the search may be only on the epipolar line instead of the window. Otherwise, the coarse to fine (for images with low resolution) or the multistage match technique will reduce the computational time. Images with low resolution can be obtained by dividing high resolution into small blocks of equal size and averaging the light intensity over each block, then, getting the matching blocks. Thus, fine matching will be rapidly realized.

Moreover, two problems appeared during searching when several points exist in the right image with a great similarity surrounding area. In this case the corresponding points are located within the uniform region. The second may be created when an object point is visible by only one camera of the pair because of their different viewpoints leading to a mistake in the maximum correlation. The thresholding would be a check for such condition rejecting any data below the specified level.

Moreover, the search for the corresponding points may be tailored into two successive steps where the feature points must be extracted from both images and then they should be matched. Only those points of similar class can be matched as edges with edges and corners for corners. As given above the historical information of previous steps will highly reduce the search area from the whole window to the interested epipolar line of 600 points as a standard value. Consequently, new reduction for the length of the epipolar line (600 points) to only some several points (about 15) can be achieved since the distance between a previous point and its next should be updated. Both displacement vector and epipolar line would be intersected at the required start point taking only a few points before and similar after (Fig. 9 a, b).

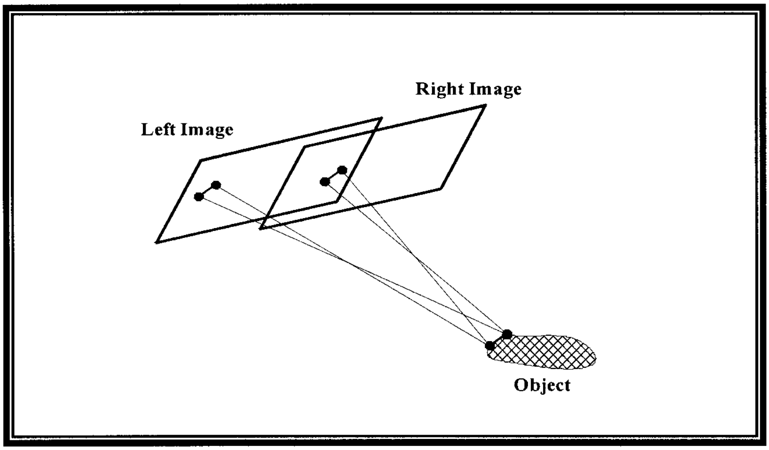

FIG. 9(a). The displacement vector method.

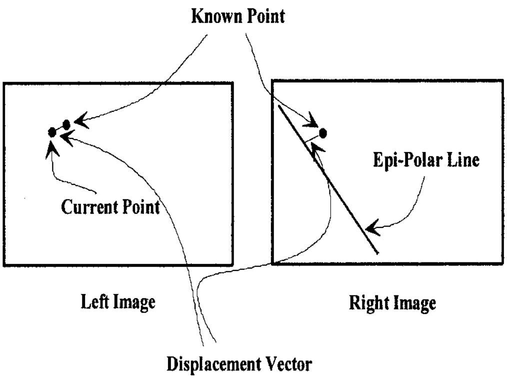

FIG. 9(b). The displacement vector.

A correspondence table must be constructed with zero elements, and then later, each zero element can be replaced by the new updated value obtained for each point. From the point of the nearest known point, the search may start (Fig. 10) on the basis of the correspondence table to ensure the detection of the nearest point. This will be terminated when the first point is obtained according to the algorithm of Fig. 11 (where shorter distance gives higher accuracy).

TABLE 1. Overview of measurement techniques.

System	Method	Range (mm)	Accuracy (micron)	Dim. of data	Speed (pt/s)	Price (1000$)
Time of light	Pulsed laser	10^3 - 10^{10}	2000 - 3000	1	10	6 - 18
Time of light	Modulated amplitude	10	50	1	10	15 - 30
Time of light	Frequency modulated	10^3 - 10^4	100 - 500	1	> 100k	15 - 30
Point of triangulation	Theodolite system	10^3 - 10^4	< 100	3	0.1	30 - 60
Point of triangulation	Laser projection	1 - 100	10 - 100	1	> 10k	0.3 - 1.5
Point of triangulation	Laser line	1- 100	10 - 100	2	> 10k	3 - 7.5
Depth of focus	Laser line	1 - 3	< 1	1	1000	15 - 75
Photogrammetry	Stereo on line	10 - 10^3	100	3	Low	30 - 300
Photogrammetry	Stereo off line	10^3 - 10^4	100	3	Low	30 - 300

The length of a displacement vector may not be the same in both images due to the existence of edges with different similarity. Normally, these edge points are adjusted to each other and the test path becomes only one or two points leading to a high reduction for searching time relative to the other case (searching along the length of the epipolar line). Thus, the execution time could be reduced to 10 % at least.

However, the two cameras are installed very close to get identical images for the best accuracy. A tolerance means that the point of intersection of a displacement vector and epipolar line in the right image is not taken exactly but some few near points should be checked selecting the highest similarity. The number of tested points depends on the length of displacement vector where twice its length is safe practically. A pair camera system can be adjusted to overcome the problem of accuracy but the calibration process must be repeated for each pair individually if moved.

TABLE 2. Comparison for different systems.

System	Opto shape	Eco- Scan	Opto- CAM	Hyscan	Replica	Proposed
Resolution (pixel)	700 × 512	768 × 512	N/A	N/A	N/A	768 × 512
Field of view (mm)	20 × 14	120 × 80	150 × 120	90	50	300 × 300
Accuracy (micron)	+/– 10	+/– 100	+/– 60	+/– 125	+/– 125	+/– 50
Range in Z (mm)	20	30	100	100	1000	200
Dim. (mm)	160/95/65	100/100/33	300/120/100	260/112/86	200/120/100	300/300/180
Weight (g)	1800	600	1500	2250	600	750
Speed (point/s)	3000	N/A	N/A	10 k	> 5k	6k
Price (1000$)	60	30	23	30	45	5
No. of axes (rotation-linear)	2 + 2	3 + 2	3	3	3	3
Merging of multiple range image	Yes	Only on CAD	Only on CAD	–	–	Yes
Automated and calib.	Yes	No	No	No	No	Yes

N/A (not available)

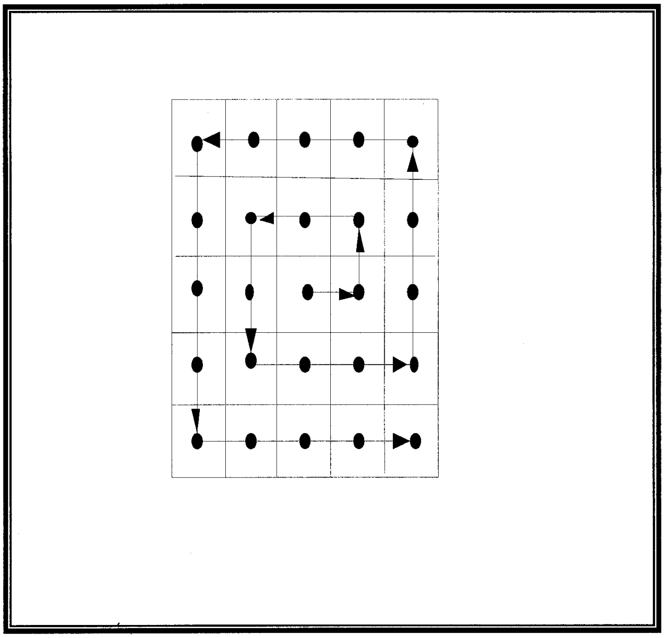

FIG. 10. The search path for the nearest point.

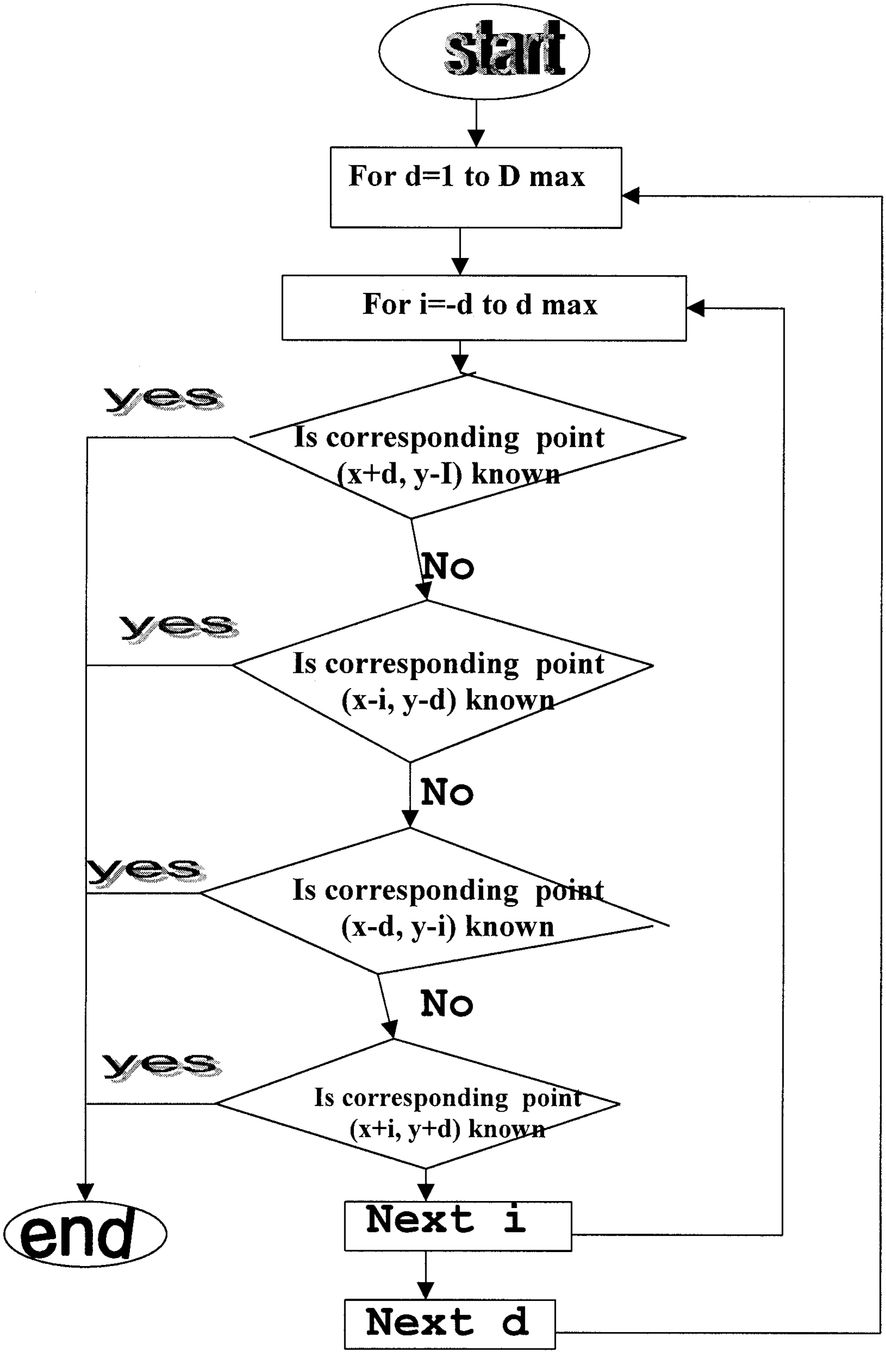

FIG. 11. Algorithm for the nearest known point.

6. System Performance

Both accuracy and speed are the most important factors in the process of image recognition because they reflect the characteristic of the overall system that should reply the real time needs. The accuracy depends on the sensor head (camera pair) while the mean error of the calibration model can be determined during the calibration operation. Results present a mean error of 0.2 pixel per image or 0.04 mm in the 3D coordinates. The grid may be scaled with 0.1 mm division and an area of 0.01 square millimeter. Using the sensor head with a pixel size of 28 by 28 microns in space, approximately 12-13 single measurement points correspond to one discrete surface point. Whatever with un-correlated measurements a reduction of noise leads to accuracy of 10-12 microns.

On the other hand, speed analysis for the experimental here at 100 MHz is applied to a calibration at about 15-20 seconds but this process of calibration will be affected if the ambient temperature is highly varied. Also, performance results of the proposed system can be tailored as:

1 – Four camera pairs need 20 ms for capturing / image with 160 ms overall.

2 – Normalization and smoothing for 4 images at 3*3 biased neighborhood requires 6 s.

3 – Detection for the edge point needs 15-20 µs per point.

4 – Corresponding process takes about 25-30 µs per point.

5 – Time for adapting different images for CAD applications consumes about 15-20 µs.

For example the digitization of a human foot model with an automated sequence of 8 images, 4 pairs of cameras, from different points with about 800000 measurement points takes about 60 seconds. For Pentium level of computers the improved characteristics give:

1 – With pentium processor clocked at 166 MHz, the execution time will be reduced to only 30-40 % of the above (486) system.

2 – Parallel processors decrease time as 16000 points consume 1 s.

3 – Special hardware for image smoothing and edge detection will save the computational time leading to the reduction of the required time. So 100000 points require only 8 seconds for the same host computer.

4 – Reference points can form a grid to raise the accuracy.

7. Conclusion

1 – A new range image based on stereo vision is developed to present a high accuracy measurements, reducing the computational time and effort.

2 – The calibration process can be fully automated and integrated with the proposed technology so that the suggested technique can be recommended for the application in different areas.

3 – The commercial production for wide utilization of the given new concept due to its low price is available.

4 – The suggested concept is reliable since it permits the possibility of extension for more than one pair of cameras.

5 – The variation in the image intensity can be normalized easily.

6 – The developed software modifies the performance of images through smoothing and edge detecting with low and high filters.

References

[1] **Hamed, M.,** A quick neural network for computer vision of gray images. *J. Circuits, Systems & Signal Processing,* **16**(1): 41-58 1997.

[2] **Gaessler, J.,** Full range images of 3 D components for CAD by optical digitization, *PhD thesis, Coventry University,* 1994.

[3] **Besl, P. J.,** Active optical range imaging sensor, *Machine Vision Application*, **1:** 127-152, 1988.

[4] **Nitzan, D.,** 3 D vision structure for Robot applications, *IEEE Trans. on Pattern Analysis and Machine Intelligence,* **10**(3): 1988.

[5] **Tiziani, H.J.,** Abstandsmessung beruhrungslos, *Industrie Elektrik + Electronik,* **4,** 1989.

[6] **Fishman, C.** and **Scheter, B.,** Computer display of height fields, *Computers and Graphics,* **5:** 53- 60, 1980.

[7] **Lovoy, M.,** Display of surfaces for volume data, *IEEE Computer Graphics and Applications*, **8**(3): 1988.

[8] **Lorenson, W.** and **Cline, M.,** A high resolution 3D surface construction algorithm, *Computer graphics (ACMSICGRAPH' 87 Conference Proc.,* 1987.

[9] **Vuylsteke, P.** and **Oosterlinck, A.,** Range image acquisition with a single binary-encoded light pattern, *IEEE PAMI,* vol. **12,** No.2, 1990.

[10] **Joanides, M.** and **Wehr, A.,** Reverse Engineering, rapid prototyping: Generation of surface information by using a 4D laser scanner for digitization, *Proc. of IMS, 31 Jan. - 2 Feb.* 1994.

ترقيـــم عـــالي الدقـــة ثــلاثي الأبعـــاد

محمد حامد* ، روبرت ماسن و هشام الهندي***

* *كلية الهندسة ،جامعة قناة السويس ، بور سعيـــد – مصر*

** *مركز معالجة الصور ، كـــونستانس – ألمـــانيا*

المستخلص . يقدم البحث طريقة جديدة لترقيم الصور في الأبعاد الثلاثية وتعرف باسم « متجه الإزاحة » وهو يعتمد على الرؤية المجسمة باستخدام العديد من آلات التصوير مزدوجة التوزيع ومنتشرة حول سطح الجسم المختبر . تدخل الأساليب الضوئية بدلا من الميكانيكية حتى يصبح التحكم في العمل من خلال أجهزة الحاسب الشخصي ممكنا وتتبعها المعايرة بأسلوب آلي تماما . تمت دراسة آليات الحسابات على الحاسبات الآلية ، وكذلك المعاملات الهامة والمؤثرة في النتائج ودقتها مثل المعايرة واستخراج الخواص وتحديد النقاط المتناظرة ، وأيضا إيجاد المحاور الثلاثية وتحويل النتائج إلى الاستخدام التصميمي على الحاسب الشخصي . قد تم تحليل مستوى الأداء للمنظومة المعروضة في البحث مع إعطاء النتائج الخاصة بقدم صناعي بشري مما يبشر بصلاحية النظام المقترح للتطبيقات الصناعية والطبية وغيرهما . يوضح البحث مدى الاختصار الزمني المتاح بهذا الأسلوب مع إعطاء الدقة العالية المطلوبة .