PERIODICA POLYTECHNICA SER. EL. ENG. VOL. 41, NO. 1, PP. 63-74 (1997)

PROCESSING OF IMAGES IN PASSIVE MARKER BASED MOTION ANALYSIS

Ákos JOBBÁGY, László GYÖNGY, Ferenc MARTIN, György ÁBRAHÁM

Department of Measurement and Instrument Engineering
Technical University of Budapest
H-1521 Budapest, Hungary
Tel/fax: +36 1 463 4938
e-mail: jobbagy@mmt.bme.hu

Received: Nov. 27, 1995

Abstract

The majority of non-contacting motion analysers process images of CCD cameras. Hardware based feature extraction can be utilised to alleviate the necessary computation: markers are attached to the landmark points of the moving objects and the marker images are separated from the environment based on their high luminosity. Data reduction can be achieved by characterising the extracted marker images with the co-ordinates of a single point. When a spherical marker is used then the centre point of its image must be determined. Real-time motion analysers usually can afford simple data processing only. This explains that a simple geometric centroid calculation of binarised images is so popular. The height and width of the marker images are used to characterise image distortion: these measures of ideal circular images are equal. This paper summarises the results of our theoretical research work. For the centre estimation of binarised images a new method with high accuracy, the ring fitting is recommended. Two methods are shown for the processing of distorted images. Both methods can recognise if a distorted image derives from the overlapping of two (or more) marker images. Furthermore, the methods give good estimates of the centres of the individual marker images. The features of multiple level thresholding for the video/digital conversion has also been studied. The results show that the achievable accuracy does not increase proportionally with the increase in resolution.

Keywords: motion analysis, marker image processing, ring fitting, multiple level thresholding.

Introduction

The location of an object with non-contacting methods is usually done with the help of video cameras. If some points of an object are marked, then the positions of the markers should be determined in order to characterise the actual location of the object. If three dimensional location is required, then the markers are usually spherical. In this case their projection to a camera sensor does not depend on their actual position, it is always a disk. If two dimensional location is required, then flat, disk shaped markers can also be used. Motion analysis (FURNÉE, 1989) is a good example for the 3D problem, while positioning of masks on printed circuit boards is a 2D problem (BOSE and AMIR, 1990). The images of the markers on the camera sensor will be disks and their exact location requires that the centre of the image be determined. The most frequently used method is geometric centroid calculation (DINN et al., 1970, TAYLOR et al., 1982). Although this method is simple, allowing easy real-time application, its accuracy is moderate. Accuracy becomes especially important in motion analysis when we need not only the position versus time function of marked points of an object but the derived parameters (velocity, acceleration, jerk) are also important. The *noise amplification factor, NAF* expresses the increase in the uncertainty of the value of a derived parameter as compared to the uncertainty of the input data. Even with apparently accurate position data ($\Delta s/s = 0.1$ %) the uncertainty of acceleration, based on numerical differentiation might be unacceptable (in excess of 100 %) without special low-pass type filtering. This justifies the efforts invested in increasing the accuracy of determining the position of a marker. The *ring fitting method* has substantially better accuracy in centre estimation of circular binary images than the conventional methods, geometric centroid calculation, circle fitting or template matching.

The marker images to be processed may be distorted by occlusion. A simple occlusion deteriorates the accuracy of the centre estimation methods. If the occluding object is a marker itself, then the phenomenon is called overlapping. Having clustered the marker images none of the above mentioned methods is able to separate overlapping images, though while tracking several markers their trajectories may seem to intersect in the view of a camera. Two methods, based on the Hough transform have been studied to resolve the ambiguity caused by distorted images. Both methods evaluate distorted and overlapping images reliably.

The accuracy of marker centre estimation can be increased substantially by increasing the resolution of digitisation while converting the analogue video output signal of a camera. Multiple level thresholding results in several bit data representation instead of the one bit representation of a binary image. A more accurate centre estimation can be reached based on grey-scaled data with less calculation as the application of the simple geometric centroid calculation is sufficient. We have found, that increasing the resolution of video/digital conversion does not mean a proportional increase in the achievable accuracy.

The research work has been based on simulation (JOBBÁGY, 1993, JOBBÁGY et al., 1994) but some results have also been tested using the *Precision Motion Analysis System, PRIMAS.* This passive marker based, non-contacting motion analyser is based on the video/digital conversion method first used at Delft University of Technology and first reported by (FURNÉE, 1967).

Ring Fitting Method for Centre Estimation

The analogue image of the CCD sensor is digitised before any further processing. The digitisation is usually done by comparing the analogue output signal of a CCD camera to a threshold level. Pixels with a corresponding above threshold level video signal are qualified as covered (binary representation: 1) otherwise as uncovered (binary representation: 0). This is one bit A/D conversion, the digitised data can be further processed with relatively small computational effort, thus making real-time operation possible. (Increased resolution during A/D conversion will be analysed in the last section.) The centre of a marker image based on its binary image can only be estimated, as marker images with different radii and centres can cause the same binary image. Consequently the quantisation error of one-bit digitisation causes a deviation of the estimated centre of a marker image from the actual one. While using conventional centre estimation methods this deviation might be substantial as these methods do not take into account that *quantisation errors along the circumference of the marker image are not independent of one another.* The ring fitting method (JOBBÁGY, 1993) gives an estimation of the centre of a binary marker image by considering all the marker images that might have caused it.

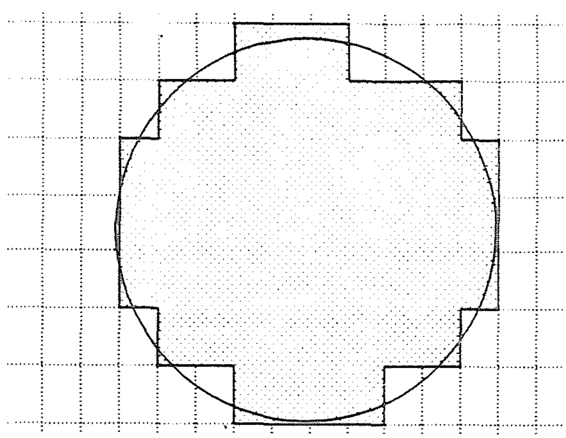

Fig. 1. A marker image on a CCD sensor surface (circle) and the resulting binary image (hatched pixels)

Fig. 1 shows a circular marker image on the surface of a CCD sensor. The hatched pixels are qualified as covered during digitisation. *Fig. 2* summarises all the circular marker images that would result in the binary image of *Fig. 1.* These marker images are called *covering images.* A part of the CCD sensor surface (the fraction of a pixel) is enlarged, the centres of possible covering images are within this part of the sensor. The radius of a covering image varies between the smallest possible value, $r_{\min}$ and the longest possible value, $r_{\max}$. If the radius is fixed at a value r_i in between $r_{\min}$ and $r_{\max}$ then the centres of the covering images with radius r_i can be anywhere along a region. *Fig. 2* relates the geometric centroids of these regions to the different radii. It is clear that the estimated centre point varies as a function of the supposed radius. It is also evident that the geometric centroid estimate of the binarised image in *Fig. 1* is poor.

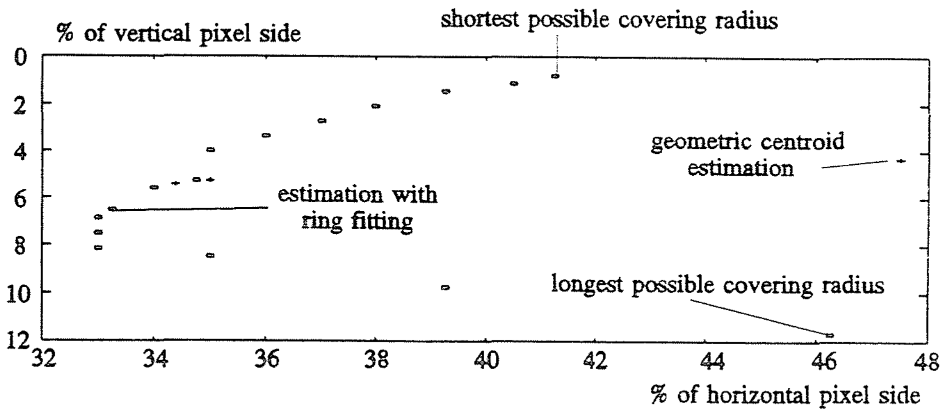

Fig. 2. Estimations for the centre of a circular image with different radii that might have caused the binary image in *Fig. 1*

The ring fitting method considers all the possible covering images and selects the centre point that might have caused the marker image with the highest probability assuming that all the possible covering images are equally likely. This is the correct assumption if there is no a priori information about the parameters of a marker image, which is the general case in processing marker images. Further details of the ring fitting method are given in (JOBBÁGY, 1993).

Processing of Distorted Images

The data from a CCD sensor is read out serially, one line after the other. The adjacent pixels in the same line, which are qualified as covered constitute a *segment*. Centre estimation requires the clustering of the segments, that belong to the same marker image. There are two different conditions for segment clustering. One makes decision based on the segment midpoint

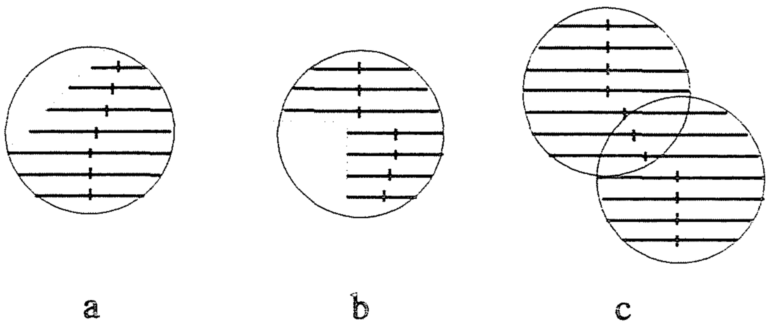

Fig. 3. Segments of distorted (a, b) and overlapping (c) images

co-ordinates. The other checks whether the segments in two consecutive lines have covered pixels with the same horizontal co-ordinate.

Fig. 3 shows three distorted marker images. Using the clustering method based on the segment midpoint co-ordinates splits the occluded images (a, b) into two and the two overlapping images (c) into three. Clustering according to the second condition does not split the (a, b) markers but takes the overlapping markers of (c) as if they were one marker.

The Hough transform has been applied to process distorted images. Two (or more) overlapping marker images can be resolved and separate estimates given for their centres. Consequently, the application of the second clusterisation algorithm is advantageous. When this algorithm clusters two (or more) overlapping marker images as if they were one marker, the processing of the image will indicate that the image is composed of two (or more) marker images.

A survey on the Hough transform is given by (ILLINGWORTH and KITTLER, 1988) and (V. F. *Leavers*, 1993). If the radius of a marker image is known, then circles with this radius must be drawn around the points (pixel midpoints) along the circumference. If such a circle intersects a pixel, then the value in the accumulator, related to this pixel is increased by one. (The accumulators need to be zeroed prior to the application of the algorithm.) The method is demonstrated in *Fig. 4*.

Circles with the known radius are drawn around all the pixels along the circumference and then the contents of the accumulators are evaluated. The pixel with the *largest number in its corresponding accumulator is the estimate for the centre of the original marker image.* The method can be explained so that each point along the circumference of a marker image votes for the possible midpoint and finally that point is selected as the estimated centre that has the largest number of votes. This method is able

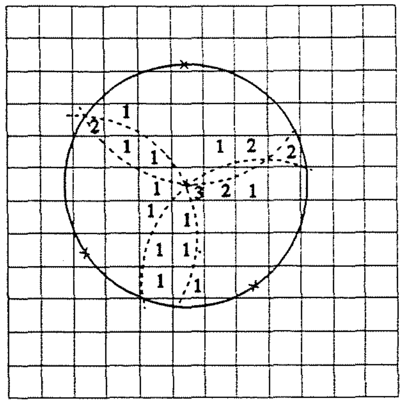

Fig. 4. The application of the Hough transform for centre estimation of a circle

to estimate the midpoint of a distorted image with high accuracy. Subpixel resolution can be achieved by dividing the pixels around the midpoint into subpixel areas and counting the votes for these areas. However, a drawback of the method is that if the radius of the image is unknown (as is the case with marker images) then the votes must be collected assuming different radii. This means that instead of circles rings are drawn around the pixel midpoints along the circumference. The minimal and maximal assumed radius values, $r_{\min}$ and $r_{\max}$ are application dependent. The method can effectively be applied if the possible range of r, the ring width is small. Too wide ring width deteriorates the accuracy. If one bit A/D conversion is used then the minimal ring width should not be smaller than one pixelside so that the real centre point can be covered from a point on the circumference.

The radius information can be eliminated by making use of the fact that the perpendicular bisector of a chord in a circle crosses its centre. The intersecting point of two such bisectors can be considered as an estimate of the centre of the circle. The area around the centre point should be divided into sub-areas and that sub-area is considered to be the centre point that contains the maximum number of intersecting points. In case of a distorted marker image the intersecting points *vote for the centre point.*

Not all the possible chords are used, only those of medium length. Too short chords (e.g. connecting adjacent pixel midpoints) introduce substantial distortion as a result of the quantisation errors. In case of distorted images pixels far from each other not necessarily designate a chord: one pixel may be on the circumference of the ideal circular image while the other may be on the borderline of the distorted part.

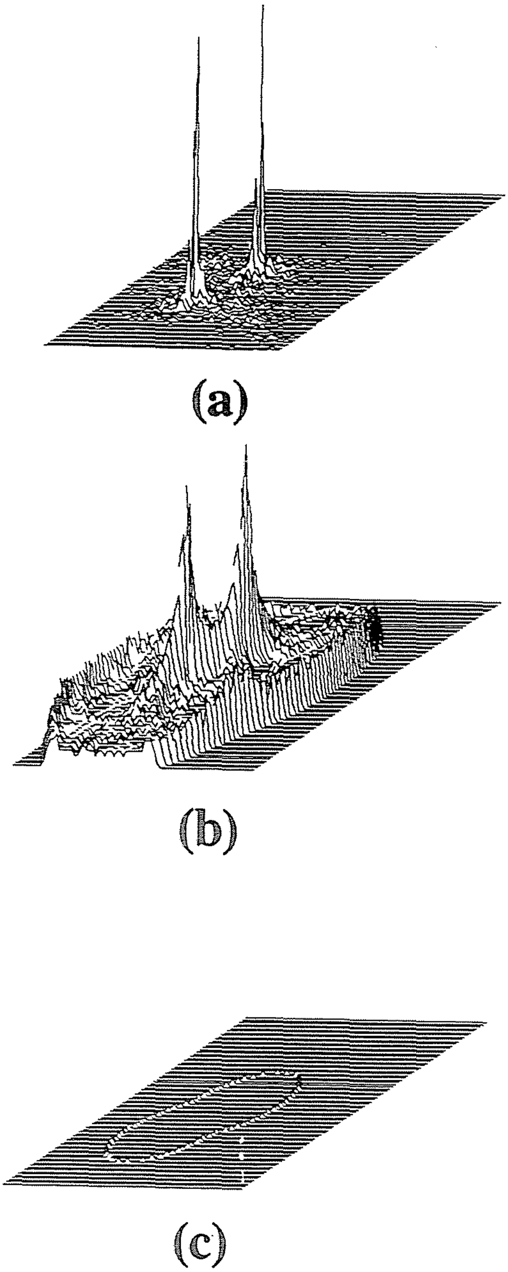

Fig. 5. Estimates of the centres of two overlapping images: the original image (a), estimates using the classical version of the Hough transform (b) and estimates using the perpendicular bisectors (c)

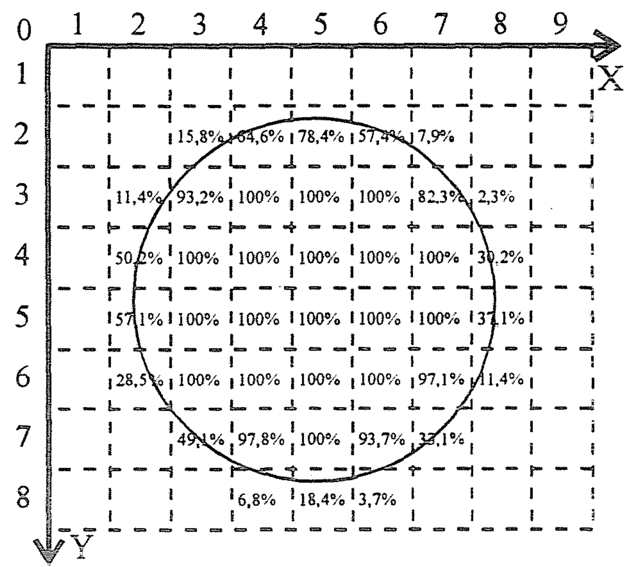

Fig. 6. The coverage of pixels on the CCD sensor surface

The results are demonstrated in *Fig. 5.* The horizontal plane is the plane the camera perceives. The vertical axis shows the votes for the centre(s) of the marker images along the horizontal plane. The figure demonstrates that both methods can process distorted images very effectively, with approximately the same accuracy.

Increasing the Resolution of Digitisation

The accuracy achievable with real time on-line systems can be improved by introducing the several bit A/D conversion of the analogue video signal. To check whether several bit A/D conversion of the video signal is suitable for image processing needed in marker based motion analysis, we have developed a set of simulation programs. The core of the simulation programs is a routine which provides the exact coverage of pixels of the CCD sensor (an example is shown in *Fig. 6*, where the coverage of each pixel is expressed in percentage) when a marker image projected onto it. We call coverage the following fraction:

$$\text{Coverage} = \frac{\text{area\_covered\_by\_image}}{\text{total\_area\_of\_the\_pixel}}$$

The simulation is based on the physical model of the CCD sensor, that is why the coverage of an actual pixel is proportional to the charge gathered on the surface of that pixel, and that is why it is also proportional to the analogue video signal corresponding to that pixel. This program can simulate the analogue video signal if the position of a marker image is given. The next routine of the simulation is the one which simulates the A/D conversion. The quantisation made with resolution 'n', which means that the digitisation has '$n-1$' threshold levels, so $n = 2$ means 1 bit, $n = 4$ means 2 bit, $n = 8$ means 3 bit etc. A/D conversion respectively. The value we get after quantisation can also be interpreted in different way. Let us divide the surface of a pixel to 'n' equal subparts. The number of covered subparts of the pixel is equal to the digitised value of the corresponding coverage.

The mechanism of simulation is shown in *Fig. 7.*

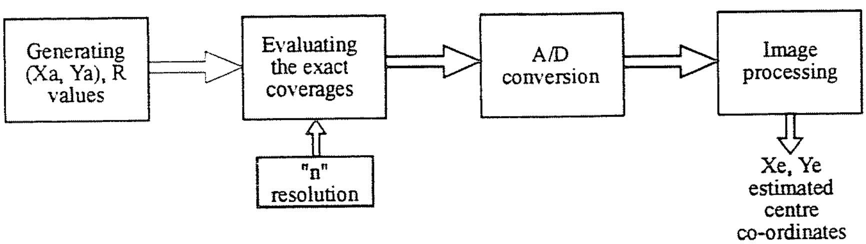

Fig. 7. The elements of the simulation programs

X_a, Y_a and R are the parameters of the *actual* marker image, X_e, Y_e are the estimated centre co-ordinates. Maximum value of coverage corresponds to a pixel totally covered by a marker image, the minimum value corresponds to a pixel not covered at all. Values between maximum and minimum correspond to pixels partly covered by a marker image. The greater is the value we get, the higher is the covered area of that pixel. We used weighted averaging for the centre estimation, where the weights are the coverages of the pixels. The X, Y co-ordinates of the estimated centre P_e are:

$$X_{\text{estim}} = \frac{\sum_{i=1}^{N} \text{Weight}_i^* X_i}{\sum_{i=1}^{N} \text{Weight}_i} \tag{1}$$

$$Y_{\text{estim}} = \frac{\sum_{i=1}^{N} \text{Weight}_i^* Y_i}{\sum_{i=1}^{N} \text{Weight}_i} \tag{2}$$

where the weighting factor Weighti is the corresponding coverage of a pixel, N is the total number of partly or fully covered pixels, X_i, Y_i are the midpoint co-ordinates of the corresponding pixel. If we use 1 bit A/D conversion, i.e. the weights can only be 1 and 0, we get the well known geometric centroid estimation. Several bit A/D conversion of the analogue video signal can provide input data for other estimation methods as well.

To characterise the accuracy estimation with weighted geometric centroid the expected value of the distance between the actual and the estimated centre points was used. It means that the smaller is the above mentioned expected value, the more accurate is the centre estimation.

$$\mathrm{dev}(r, \mathrm{resolution}) = E\{d(P_a, P_e)\} \tag{3}$$

where P_a means the actual, and P_e means the estimated centre of marker image. The accuracy of the estimation increases if the radius of the marker image increases because the greater image means that more pixels are involved in the averaging. The increase in resolution during digitisation can be interpreted as the use of smaller subpixels instead of the real ones. That is similar to the case when we use marker images with bigger radius i.e. with the increase of 'n' the decrease of 'dev(r, resolution)' is expected.

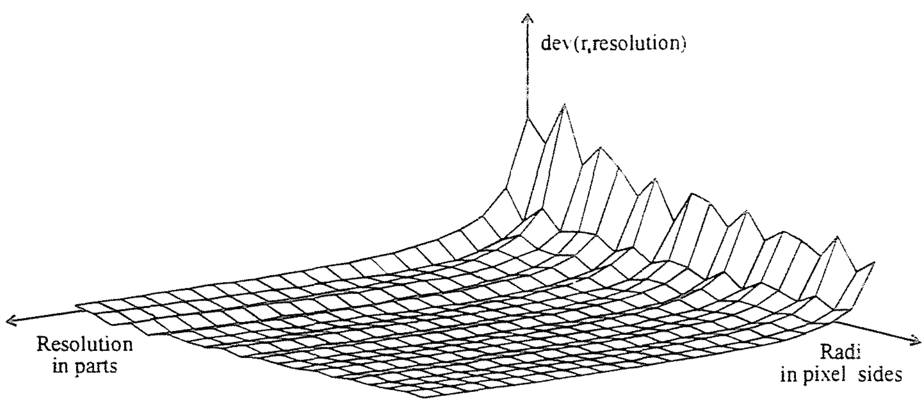

Fig. 8. Both the marker image radius and the resolution of digitisation influence the accuracy of the estimation

The result of the simulation (*Figs. 8* and *9*) shows that the accuracy improves with the number of bits used in video signal quantisation. There is a big improvement when changing from 1 bit to several (2,3) bit A/D conversion and it slows down using more (5,6..) bit A/D conversion (*Figs. 9* and *10*). *Fig. 10* shows the accuracy in the range of 1 to 4 bit digitisation. This range seems to be suitable for use in image processing. The simulation also provided the theoretical limit (*Fig. 11*) of accuracy achievable with

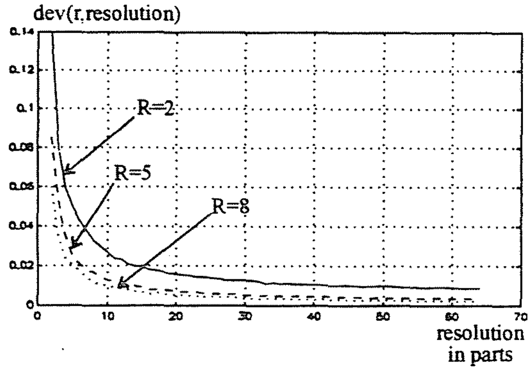

Fig. 9. The dependence of 'dev' on the resolution of digitising $n = 2 \dots 64$

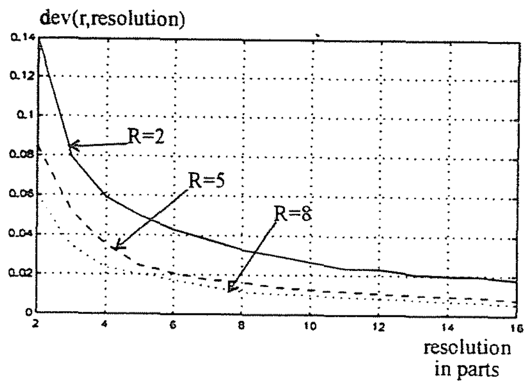

Fig. 10. The dependence of 'dev' on the resolution of digitising $n = 2 \dots 16$

weighted centroid estimation. The theoretical limit is the case when the coverages are exactly known, they are not approximated by digitisation.

Acknowledgement

The research work reported here was partly supported by the Copernicus program of the Commission of the European Communities, project No. 7845: 'Precision Image-based Real-time Motion Analysis System at Economical Cost' and by the contract with The National Committee for Technological Development (OMFB) No. 84-97-47-0731: 'The Application of New Algorithms and Methods to Improve the Accuracy of Motion Analysis Systems Based on the Tracking of Passive Markers'.

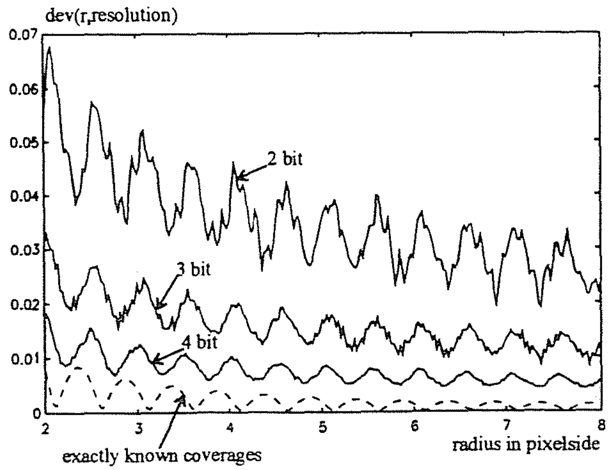

Fig. 11. The accuracy versus radius of marker image. The resolution of digitising also influences the accuracy

References

1. BOSE, C. B. – AMIR, I. (1990): Design of Fiducials for Accurate Registration Using Machine Vision. *IEEE Trans. on PAMI*, Vol. 12, No. 12, pp. 1196–1200.
2. DINN, D. F. – WINTER, D. A. – TRENHOLM, B. G. (1970): CINTEL – Computer Interface for Television. *IEEE Trans. on Computers*, Vol. C-19, pp. 1091–1095.
3. FURNÉE, E. H. (1967): Hybrid Instrumentation in Prosthesis Research. *Digest 7th Int. Conf. Med. & Biol. Eng.*, Stockholm, p. 446.
4. FURNÉE, E. H. (1989): TV/Computer Motion Analysis Systems: The First Two Decades. *PhD Thesis*, Oct. 1989. Delft Univ. of Technology, ISBN 90-9003095-6.
5. ILLINGWORTH, J. – KITTLER, J. (1988): A Survey of the Hough Transform. *Computer Vision, Graphics, and Image Processing*, 44, pp 87–116.
6. LEAVERS, V. F. (1993): Which Hough Transform? *CVGIP: Image Understanding*, Vol. 58. No. 2. September. pp. 250–264. 1993.
7. JOBBÁGY, Á. (1993): Centre Estimation in Marker Based Motion Analysis. *Technical Report of the Technical University of Budapest*, ISBN 963-421-509-2.
8. JOBBÁGY, Á. – FURNÉE, E. H. (1994): Marker Centre Estimation Algorithms in CCD Camera-based Motion Analysis. Medical and Biological Engineering and Computing, Vol. 32. pp. 85–91.
9. TAYLOR, K. D. – MOLTIER, F. M. – SIMMONS, D. V. – COHEN, W. – PAVLAK, R. JR. – CORNELL, D.P. – HANKINS, B. (1982): An Automated Motion Measurement System for Clinical Gait Analysis. *J. Biomechanics*, Vol. 15, No. 7. pp. 505-516.