

DYNAMIC CENTROID DETECTION IN OUTDOOR/INDOOR SCENES WITH DIFFERENT BACKGROUNDS

Alberto Soria-López
Departamento de Control Automático.

CINVESTAV-IPN. Av. IPN 2508, México D.F., México
email: soria@ctrl.cinvestav.mx

Pedro Mejía-Alvarez
Departamento de Ingeniería Eléctrica
Sección de Computación
CINVESTAV-IPN. Av. IPN 2508, México D.F., México
email: pmejia@cs.cinvestav.mx

ABSTRACT

Centroid detection is achieved using a dynamic thresholding contour detection. The algorithm is applied to a gray level image with different backgrounds in outdoor scenes obtained with a mobile camera using a RF video link. This technique allows the detection of more than one target.

KEY WORDS

Computer Vision, Centroid Detection, Robotics.

1 Introduction

In many works concerning visual feed-back for example see [1], [2], [3], [4], [5], centroid detection is performed using fixed binarization thresholding for centroid calculation. In this case, the environment is structured to allow a high contrast between the background and the robot end effector. Using binarization, two extreme gray levels for the foreground and the background are obtained allowing the end effector centroid calculation. Its main disadvantages are that it relies heavily on the illumination conditions and the binarization threshold; any change in these last two characteristics affects importantly the position calculation. The fixed binarisation/centroid calculation has the advantage that its calculation is simple and fast, hence its widespread use.

An alternative solution that allows tracking the target position is the use of robust vision. Vincze and Hager [6] present a compilation of various works concerning this solution. A large part of these works use a form of model based approach where the vision sensor data is used to update a model using cue integration, Hough transform methods or feature extraction. Other methods include hierarchy tracking or the use of the geometry of the visual sensors. These methods are in general complex, computational expensive and often require specialized hardware to attain frame rates.

Concerning the thresholding methods there has been quite a large interest in these methods giving a number of survey papers. Of the first papers, Weszka and Ronsenfeld [7] defined several evaluation criteria; Palumbo *et al.* [8]

compared three methods for document binarisation while Sahoo *et al.* [9] compared nine thresholding algorithms. More recently, Lee *et al.* [10] carried out a comparative analysis of five global thresholding methods. Glasbey [11] studied the performance of 11 histogram based algorithms based on a statistical data. Trier and Jain [12] evaluated 19 methods for character segmentation from complex backgrounds. In the most recent and complete survey Sankur and Sezgin [13] study 40 thresholding methods. The authors implemented each method and tested them with different inspection applications such as light, thermal, ultrasonic and eddy current images. Five complementary performance criteria are used to assess comparatively the various thresholding methods. In the present work we tested the methods finding similar results for our visual servoing application, leading us to use the method proposed by Kittler and Illingworth [14].

The contribution of this paper is to extend the thresholding/centroid detection method introducing a thresholding determination for each frame, allowing a dynamic thresholding followed by a contour following technique. This is first achieved by the selection an adequate thresholding method. Small changes in the scene, the presence of other objects or illumination changes will often introduce small areas that are not part of the target. Even a small amount of pixels will change drastically the calculated centroid position. For these reasons, a simple contour following technique is used that allows calculating the binarized objects length, area and form factor (area/length) allowing to distinguish the object of interest from other objects or noise in the binarized image. The proposed method could be situated between the simple fixed simple thresholding technique and the more evolved but computationally expensive robust vision techniques mentioned above. It should be noted that the results presented here do not use the “windowing technique” where only a small image windows is processed, and who’s location can be predicted from previous centroid calculations [15]. The proposed method is applied to the full image and does not need an explicit initialization.

The paper is organized as follows; section 2 describes the proposed dynamic binary thresholding and the contour extraction and selection. In section 3 we show experimen-

* This research work has been supported in part by CONACYT C03-42449

tal results concerning images from a camera using different backgrounds in outdoor/indoor scenes are presented to explore the limits of the proposed method. The paper ends with some concluding remarks.

2 Dynamic Binarization Thresholding

The most commonly used method in extracting objects from a picture is thresholding using a binarization threshold t . If the object is clearly distinguishable from the background, the gray-level histogram will be bimodal and the threshold for segmentation can be chosen at the bottom of the valley. However, gray-level histograms are not always bimodal. Methods other than valleyseeking are needed to determine an adequate threshold. In the evaluation of the different thresholding methods that we carried out in our application, each method will be refereed on the type of information it exploits followed by the author(s) that proposed the method [13]. For example, the shaped based method proposed by Rosenfeld [7] will be refereed as Shape_Rosenfeld[7].

Histogram shape-based methods. This category achieves the thresholding based on the shape properties of the histogram such as the distance from the convex hull of the histogram (Shape_Rosenfeld [7]), forced smoothed two-peak representation via autoregressive modelling (Shape_Guo [16]), rectangular approximation of the lobes (Shape_Ramesh [17]), search of peaks and valleys (Shape_Sezan [18],) or overlapping of peaks via curvature analysis (Shape_Olivo [19]).

Clustering based methods. The gray-level samples are clustered as background and foreground. The methods in this category include the search for the midpoint of the two peaks (Cluster_Riddler[20], Cluster_Yanni[21]), fitting of a mixture of Gaussians (Cluster_Lloyd[22], Cluster_Kittler[14], [23]), Mean-square clustering(Cluster_Otsu[24]) or fuzzy clustering(Cluster_Jawahar_a, Cluster_Jawahar_b[25]).

Entropy based methods. These methods use the entropy of the distribution of the gray levels. The maximisation of the entropy is interpreted as a maximum information transfer (Entropy_Pun_a[26], Entropy_Pun_b[27], Entropy_Kapur[28], Entropy_Yen[29], Entropy_Sahoo[30]), minimization of the cross-entropy between the gray-level image and the binary image as the preservation of information(Entropy_Li[31], Entropy_Brink[32]) or as a measure of fuzzy entropy (Entropy_Shanbag[33], Entropy_Cheng[34]).

Attribute Similarity. These methods are based on some attribute quality or similarity measure between the original image and the binarized image. These methods use edge matching(Attributue_Hertz[35]), shape compactness(Attribute_Pal[36]), graylevel moments(Attribute_Tsai[37]), stability of segmented objects(Attribute_Pikaz[38]) or fuzzy resemblance (Attribute_Huang[39]).

Spatial Thresholding. These methods use the dependency of pixels in a neighbourhood. These methods include cocurrence probabilities(Spatial_Pal_a, Spatial_Pal_b[40]), secondorder entropies(Spatial_Abutaleb[41]), fuzzy partitioning(Spatial_Cheng[42]) or local spatial dependency(Spatial_Beghdadi[43]).

Local Methods. In these methods, threshold is calculated at each pixel and is based on some local statistic like range or variance. These methods use local variance(Local_Niblack[44], Local_Sauvola[45]), local contrast(Local_White[46], Local_Bernsen[47], Local_Yasuda[48]), centre-surround scheme(Local_Palumbo[8], Local_Kamel[Kamel]) and graylevel landscape surface fitting(Local_Yanowitz[50]).

2.1 Thresholding method selection

Figure 1 show the four images used to evaluate the different methods. Due to the fact that the outdoor images allow to change the background, the employed outdoor images where chosen so that the thresholding task becomes difficult.

An image processing program [13] was used to test the four images from figure 1 with each of the thresholding methods allowing to rank the performance of each method using the arithmetic averaging of the normalized scores obtained from the following performance criteria [13]:

Misclassification error. Reflects the percentage of background pixels wrongly assigned to the foreground and foreground pixels wrongly assigned to the background.

Edge mismatch. This metric penalizes the differences between the edge map of the gray level image and the edge map obtained from the thresholded image.

Region nonuniformity. This measure evaluates how well an image is segmented. A well-segmented image with have a nonuniformity measure of zero while the worst case the value will be one. A value of one will correspond to an image which background and foreground are indistinguishable.

Relative foreground area error. This measure is obtained from the segmented image with respect to the reference image. For a perfect match of the segmented regions the value for this measure is zero while a no overlap condition of the images will have a value of one.

Shape distortion penalty uses an average of the modified Hausdorff distance. This metric is employed to asses the shape similarity of the thresholded regions to the groundtruth shapes.

Table 1 shows the results for our 4 test images(c.f. figure 1). The lowest score indicates the best segmentation quality, while the highest score indicates the worst segmentation quality.

The most important common characteristic of our results with the ones obtained by Sankur and Sezgin[13] is that the Cluster_Killter algorithm

[14] comes in first place. As for our obtained ranking of the methods, it only resembles marginally the results for the degraded document images in [13] for the Local_Saulova and Local_White methods that come in places five and six. Given this performance evaluation, we will in the following be using the Kittler method to calculate the threshold for every frame acquired by the camera in our application.

Clustering_Kittler	0.153
Entropy_Brink	0.280
Local_Yasuda	0.335
Local_Palumbo	0.346
Local_Saulova	0.389
Local_White	0.430
Entropy_pun_a	0.472
Local_Kamel	0.498
Spatial_Beghdadi	0.534
Attribute_Tsai	0.543
Attribute_Herz	0.543
Entropy_Li	0.554
Entropy_Shabag	0.554
Clustering_Ridler	0.555
Attribute_Pikaz	0.556
Spatial_Pal_a	0.561
Local_Niblack	0.562
Entropy_Yen	0.562
Attribute_Huang	0.574
Shape_Sezan	0.576
Clustering_Jawahar_b	0.586
Clustering_Otsu	0.588
Attribute_Guo	0.591
Local_Bernsen	0.605
Entropy_Shahoo	0.630
Clustering_Jawahar_a	0.632
Entropy_Kapur	0.642
Spatial_Abutaleb	0.644
Clustering_Loyd	0.646
Clustering_Yanni	0.660
Spatial_Pal_b	0.683
Local_Yanowitz	0.702
Shape_Ramesh	0.715
Shape_Rosenfeld	0.722
Shape_Olivio	0.731
Entropy_pun_b	0.750

Table 1. Thresholding evaluation ranking.

3 Threshold oriented contour detection

To successfully detect the object of interest, we apply a threshold oriented contour detection in the grey valued image using the threshold t determined previously. The object is supposed to be a 4-connected set of pixels having a grey

value greater or equal to t . Then the contour is the set of object pixels having at least one non-object 8-neighbor.

The contour following procedure uses a contour representation called “chain coding” that is used for contour following proposed by Freeman [51]. Each pixel of the contour is assigned a different code that indicates the direction of the next pixel that belongs to the contour in some given direction.

The algorithm begins searching for the first pixel of the contour, searching from left to right each line of the image. To this pixel the first direction code is given. The codes for the pixels that form the contour can be given analysing the neighbouring pixel grey level using 8-neighbor code clockwise. This process is repeated until the first pixel is reached. With the first pixel and the chain of codes the contour can be found. Using the invariant moment calculation [52] the centroid is found as well as the length, area and the form factor (length/area) of the contour. Each contour is stored in a list until no further contours are found. It should be noted as all the found contours are stored allowing more than one object to be detected. From the list of masked images, the target is detected using the contour attributes that match the target properties length, area and form factor (area/length). A contour is selected if the values of length, area and form factor are within manually determined intervals. These intervals will allow the selection of a contour that approaches the target characteristics. With the selected contour, the centroid is then calculated using moments m_{00} , m_{01} and m_{10} :

$$m_{pq} = \sum_R \sum X^p Y^q I(x, y) \quad (1)$$

Here, $I(X, Y)$ is 0 or 1, taking all the pixels within the contour, centroid is given by:

$$X_c = \frac{m_{10}}{m_{00}} \quad Y_c = \frac{m_{01}}{m_{00}} \quad (2)$$

4 Experimental Results

Image acquisition and processing is performed using a Pentium® based computer running at 3 Ghz under Windows XP®. For image acquisition we use a vt-100 type microcamera with a RF link for the outdoor/indoor scenes. The camera is connected to the vision computer through a National Instruments PCI 1408 framegrabber card. Image processing was done using C++ language and is based on the image processing library ICE and DIAS environment developed by the Image Processing Group of the Faculty of Mathematics and Informatics, Friedrich Schiller University, Jena Germany.

Figure 2 shows four successful centroid detection examples for outdoor/indoor images using the mobile camera. In these examples it can be noted that more than one contour is found and the contour of interest is marked in white. It can be noticed in image 2-b and 2-c that the contour of interest characteristics (length, area and form factor)

matching is rather large because in figure 2-b the contour length is shorter and in figure 2-c is larger than in figures 2-a and 2-d, where the contour approaches the circle.

Figure 3 shows four **unsuccessful** centroid detection examples for outdoor/indoor images using a mobile camera. In these examples it can be noted that there is not a match between the contour of interest that corresponds to the circle and the ones found from the threshold oriented contour detection. In figure 3-a the contour of interest can no be found because the sun illumination hits the target directly (the sun is behind the camera) making it brighter. In figures 3-b and 3-c the contour of interest is not found because the target “blends” itself with the background giving larger contours that do not match in characteristics that search for a circle. Finally figure 3-d shows an example of a contour that does match the specified target characteristics but does not correspond to the circle, this is caused by a rather large matching intervals for the length, area and form factor.

5 Conclusions

In this paper we presented the use of a threshold oriented contour centroid detection. Although the changing background, the dynamic thresholding compensates for the changes allowing a satisfactory detection when possible. Our experimental results show that for different backgrounds where the target of interest can be successfully detected and some examples where the proposed method reaches its limits. The proposed method requires the specification of detection intervals for the length, area and form factor for the contours found in the scene. In our implementation we have tuned these intervals manually since the target is unique of other objects present in the scene. Future work will concern the modification of the presented Binarization thresholding method and to propose a new one to solve the problems encountered in the failed experiments.

6 References

[1] Kelly, R.- “Robust Asymptotically Stable Visual Servoing of Planar Robots”.- *IEEE Transactions on Robotics and Automation*. Vol. 12, N° 5. October, 1996. pp. 759-766.

[2] Fässler, H.; Beyer, H. and WEN, J.- “A robot pong player”.- *Robotersystem*. Vol. 41, N° 12, 1990. pp. 161170.

[3] Andersson, R. *Real Time Expert System to Control a Robot PingPong Player*. PhD thesis, University of Pennsylvania, June 1987.

[4] Corke, P. and Paul, R. - “Video-rate visual servoing for robots”. *Experimental Robotics 1*. (Hayward, V. and Khatib, O. Eds.). Vol. 139. (Lecture notes in Control and Information Sciences), 1989. pp. 429-451.

[5] Hasimoto, K. & Kimulra, H. - “LQR Optimal and Non-Linear Approaches to Visual Servoing”. In Hashimoto, K. (Ed.).- *Visual Servoing*. Singapore: World Scientific, 1993.

[6] Vincze, M. and Hager, G. (Eds.).- *Robust Vision For Vision-Based Control of Motion*. New York, NY: IEEE Press, 2000.

[7] Weszka, J. and Ronsenfeld, A.- “Threshold Evaluation Techniques”. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-8, No.8, 1978. pp. 627-629.

[8] Palumbo, P.; Swaminathan, P. and Srihari S.- “Document image binarisation: Evaluation of Algorithms”. *Proceedings SPIE Application of Digital Image Processing*, SPIE Vol. 697, 1986. pp. 278-286.

[9] Sahoo, P.; Soltani, S.; Wong, A. and Chen Y.- “A Survey of thresholding Techniques”. *Computer Graphics and Image Processing*, Vol. 41, 1988. pp. 233-260.

[10] Lee, S; Chung, R. and Park.- “A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation”. *Graphical Models and Image Processing*. Vol. 52, 1995. pp.62-66.

[11] Glasbey, C.- “An analysis of histogram-based thresholding algorithms”. *Graphical Models and Image Processing* Vol. 55, 1993. pp. 532-537.

[12] Trier, O. and Jain, A.- “Goal-directed evaluation of binarisation methods”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-17, 1995. pp. 532-537.

[13] Sankur, B. & Sezgin, M.- “A Survey Over Image Thresholding Techniques And Quantitative Performance Evaluation”. *Journal of Electronic Imaging*, Vol. 13 No.1, 2003. pp.146-165.

[14] Kittler, J. and Illingworth, J.- “Minimum Error Thresholding. *Pattern Recognition*. Vol. 19, No.1, 1986. pp.41-47.

[15] Westmore, D. and Wilson, W.- “Direct dynamic control of a robot using an end-point mounted camera and Kalman filter position estimation”. *Proceedings of the IEEE International Conference of Robotics and Automation*, 1991. pp.2376-2384.

[16] Guo, R. and Pandit, S.- “Automatic threshold selection based on histogram modes and a discriminant criterion”. *Machine Vision Applications*. Vol. 10, 1991. pp. 331-338.

[17] Ramesh, N.; Yoo, H. and Sethi, I.- “Thresholding based on histogram approximation”. *IEE Proceeding on Vision Image Signal Processing*. Vol. 142, No. 5, 1995. pp. 271-279.

[18] Sezgin, I.- “A Peak detection algorithm and its application to histogram-based image data reduction”. *Graphical Models and Image Processing*. Vol. 29, 1985. pp. 47-59.

[19] Olivo, J.- “Automatic threshold selection using the wavelet transform”. *Graphical Models and Image Processing*. Vol. 56, 1994. pp. 205-218.

[20] Riddler, T. and Calvard, S.- “Picture thresholding using an iterative selection method”. *IEEE Transactions on*

- Systems, Man and Cybernetics*. Vol. SMC-8, 1978. pp. 630-632.
- [21] Yanni, M. and Horne, E.- "A new approach to dynamic thresholding". *Proceedings of EUSIPCO 94, 9th European Conference on Signal Processing*. 1994. pp. 34-44.
- [22] Lloyd, D.- "Automatic target classification using moment invariant of image shapes". *Technical Report RAE IDN AW126*. Farnborough, UK: 1985.
- [23] Kittler, J. and Illingworth, J.- "On threshold selection using clustering criteria". *IEEE Transaction on Systems, Man and Cybernetics*. Vol. SMC-15. 1985. pp.652-655.
- [24] Otsu, N.- "A threshold selection method from gray level histograms". *IEEE Transaction on Systems, Man and Cybernetics*. Vol. SMC-9, 1979. pp. 62-66.
- [25] Biswas, P. and Ray, A.- "Investigations on fuzzy thresholding based on fuzzy clustering". *Pattern Recognition*. Vol. 30, No.10, 1997. pp. 1605-1613.
- [26] Pun, T.- "A new method for gray-level picture threshold using the entropy of the histogram". *Signal Processing*. Vol. 2, No. 3, 1980. pp. 223-237.
- [27] Pun, T. - "Entropic thresholding: A new approach". *Computer Graphics and Images Processing*. Vol. 16, 1981. pp. 210-239.
- [28] Kapur, J; Sahoo, P. and Wong, A. - "A new method for gray-level picture thresholding using the entropy of the histogram." *Graphical Models and Image Processing*. Vol. 29, 1985. pp. 273-285.
- [29] Yen, J.; Chang F. and Chang S.- "A new criterion for automatic multilevel thresholding". *IEEE Transaction on Image Processing*. Vol. IP-4, 1995. pp.370-378.
- [30] Sahoo, P.; Wilkins, C. and Yeager, J. - "Threshold selection using Renyi's entropy". *Pattern Recognition*. Vol. 30, 1997. pp. 71-84.
- [31] Li, C. and Lee, C. - "Minimum cross-entropy thresholding". *Pattern Recognition*. Vol. 26, 1993. pp. 617-625.
- [32] Brink, A. and Pendock, N.- "Minimum cross entropy threshold selection". *Pattern Recognition*. Vol. 29, 1996. pp. 179-188.
- [33] Shanbag, A.- "Utilization of information measure as a mean of image thresholding". *Computer Vision and Graph Image Processing*. Vol. 56, 1994. pp. 414-419.
- [34] Cheng, H; Chen, Y. and Sun, Y.- "A novel fuzzy entropy approach to image enhancement and thresholding". *Signal Processing*. Vol. 75, 1999. pp. 277-301.
- [35] Hertz, L. and Schafer, R.- "Multilevel thresholding using edge matching". *Computer Vision, Graphs and Image Processing*. Vol. 44, 1988. pp. 279-295.
- [36] Pal, S. and Rosenfeld, A.- "Image enhancement and thresholding by optimization of fuzzy compactness". *Pattern Recognition Letters*. Vol. 7, 1988. pp. 77-86.
- [37] Tsai, W.- "Moment-preserving thresholding: A new approach". *Graphical Models and Image Processing*. Vol. 19, 1985. pp. 377-393.
- [38] Pikaz, A. and Averbuch, A.- "Digital image thresholding based on topological stable state". *Pattern Recognition*. Vol. 29, 1996. pp. 829-843.
- [39] Huang, L. and Wang, M.- "Image thresholding by minimizing the measures of fuzziness". *Pattern Recognition*. Vol. 28, 1995. pp. 41-51.
- [40] Pal, N. and Pal, S.- "Entropic thresholding". *Signal Processing*. Vol. 16, 1989. pp. 97-108.
- [41] Abutaleb, A.- "Automatic thresholding of gray-level pictures using two-dimensional entropy". *Computer Vision and Graphic Image Processing*. Vol. 47, 1989. pp. 22-32.
- [42] Cheng, H. And Chen, Y.- "Fuzzy partitioning of two-dimensional histogram and its application to thresholding". *Pattern Recognition*. Vol. 32, 1999. pp. 825-843.
- [43] Beghdadi, A.; Negrata, A. and DeLesegno, P.- "Entropic thresholding using a block source model". *Graphical Models and Image Processing*. Vol. 57, 1995. pp. 197-205.
- [44] Niblack, W.- *An Introduction to Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1986. pp. 115-116.
- [45] Sauvola, J. and Pietaksinen, M.- "Adaptive document image binarisation". *Pattern Recognition*. Vol. 33, 2000. pp. 225-236.
- [46] White, J. and Rohrer, G.- "Image thresholding for optical character recognition and other application requiring character image extraction". *IBM Journal of Research and Development*. Vol. 27, No. 4, 1983. pp. 400-411.
- [47] Bersen, J.- "Dynamic thresholding of gray level images". *Proceeding of ICPR'86 the International Conference on Pattern Recognition*. 1986. pp. 1251-1255.
- [48] Yasuda, Y.; Dubois, M. and Huang, T.- "Data compression for check processing machines". *Proceeding IEEE*. Vol. 68, 1980. pp. 874-885.
- [49] Kamel, M. and Zhao, A.- "Extraction of binary character/graphics images from grayscale document images". *Graphical Models and Image Processing*. Vol. 55, No. 3, 1993. pp. 203-217.
- [50] Yanowitz, S and Bruckstein, A.- "A new method for image segmentation". *Computer Graphics and Image Processing*. Vol. 46, 1989. pp. 82-95.
- [51] Freeman, H.- "On the Encoding of Arbitrary Geometric Configurations". *IEEE Transactions on Electronics and Computers*, Vol. Ec-10, 1961. pp. 260-268.
- [52] Haralick, R. and Shapiro L.- *Computer and Robot Vision*. Reading, MA : Addison-Wesley, (Vol. 2), 1992. pp. 247-430.



Figure 1. Test Images.

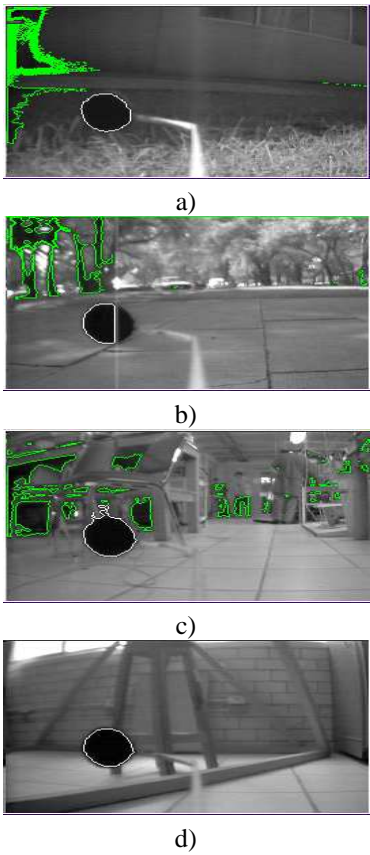


Figure 2. Outdoor/Indoor successful image centroid detection.

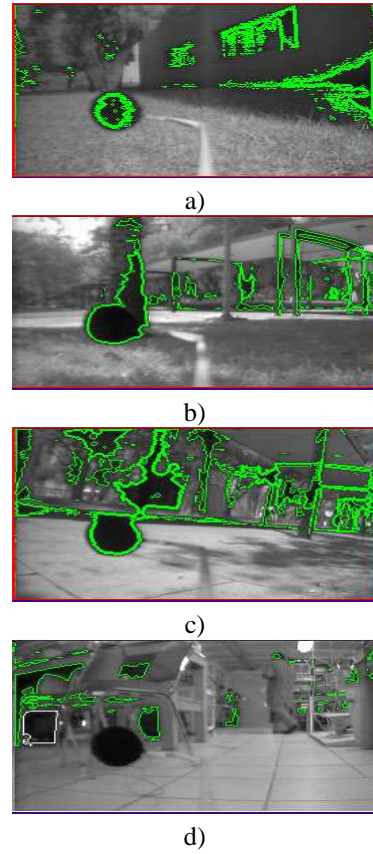


Figure 3. Outdoor/Indoor Unsuccessful image centroid detection.