OPTO-ELECTRONICS REVIEW 7(3), 215–221 (1999)

Simple video image segmentation system enabling real-time target tracking

J. SIUZDAK*

Institute of Telecommunication, Warsaw University of Technology
15/19 Nowowiejska Str., 00-665 Warsaw, Poland

The paper presents a simple system based on an IBM personal computer equipped with additional extension boards. It makes possible to process TV signal and track objects in real time. It consists of A/D converter, 2D filter, and FIFO memory interfaced to computer bus. Hardware and software description is given. The selection of 2D filter is also discussed. The system application is limited to the objects on a relatively uniform background.

Keywords: image processing, segmentation, 2D filters

1. Introduction

The segmentation is often a preliminary stage of further image processing. It enables to separate image areas belonging to different objects. Numerous methods of segmentation are known. Here, only a few are mentioned; thresholding, edge detection, fuzzy sets, and neural nets. A review is given in [1]. Only few methods can operate in real time and usually they require sophisticated equipment [2–5]. We describe a simple system based on an IBM personal computer equipped with additional extension boards. It makes possible to process TV signal and track objects in real time. The system operates on the edge detection principle. In the sequel, we will show both theoretically and practically that a single filter of special design can successfully replace two classical edge detectors. The system was designed and tested at the Telecommunication Institute of Warsaw University of Technology under the "Photonic Engineering" priority programme grant.

2. Principle of operation

A block scheme of the system is shown in Fig. 1. A CCD camera is a source of the standard black & white video signal (50 TV fields per second, 625 lines per frame, 1 $V_{p\text{-}p}$). A video signal is supplied to a fast A/D converter where it is converted into 8-bit binary data corresponding to image intensity. 512 samples are taken per each line. The binary data is fed to a two-dimensional filter (Plessey PDSP 16488) [6]. Here, the binary video data is multiplied by a group of coefficients arranged in 2D mask and then added.

Let us denote the pixel intensities by a_{ij} and the filter coefficients by $[p_{kl}]_{N\times N}$ where N can be as high as 8 [6]. Then the output s_{out} of 2D filter is given by

$$s_{out} = \sum_{k=1}^{N}\sum_{l=1}^{N} p_{kl}a_{i+k,j+l} \tag{1}$$

Due to the special architecture [6] the filtering process is performed in real time. Examples of the 2D filters are shown in Figs. 2 and 3. The output s_{out} is large when the filter is matched to the current part of the image. If the filter is matched to an edge, for example, it gives a high output when it encounters the edge in the image. This output is internally compared with a predefined threshold. In this way, the single bit output of the comparator is high when it encounters the edge and low in the opposite case. Such a binary edge image is depicted in Fig. 4. It is displayed on an auxilliary monitor for verification and control. This image is stored in a two-port FIFO memory where 256 lines each 512 pixels long may be saved. An interrupt request bit is set in ISA bus after the end of the TV field. An appropriate inter-

* e-mail: siuzdak@tele.pw.edu.pl

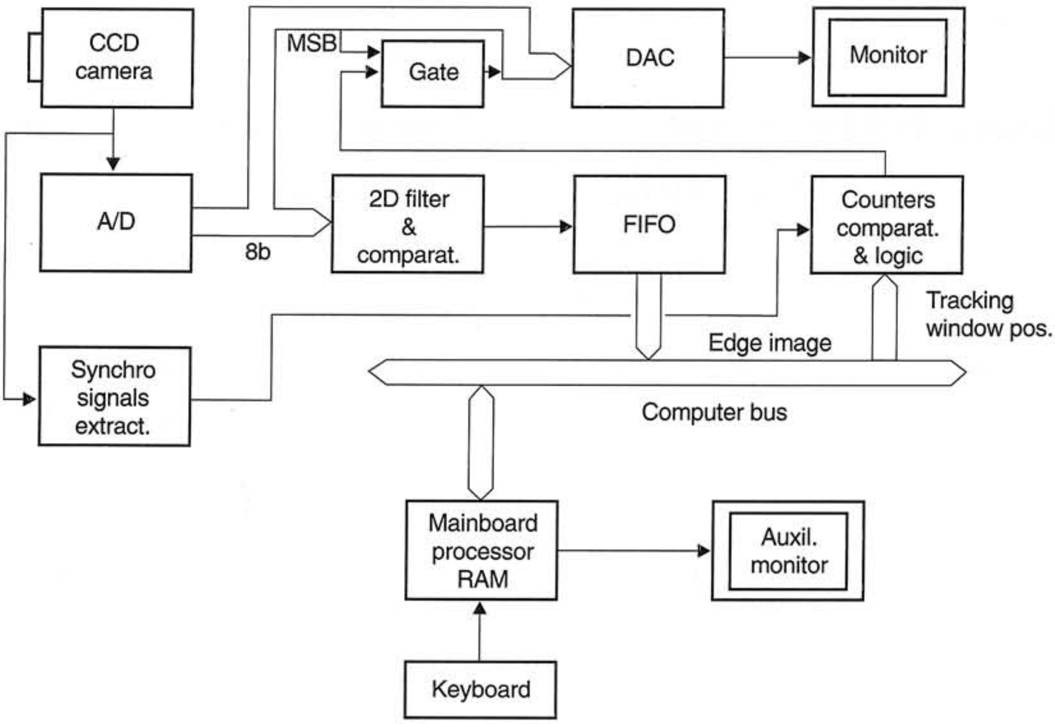

Fig. 1. Block scheme of the system.

rupt service routine initiates the data transfer from FIFO to the computer RAM. The binary image data is further processed by this routine and the target position is determined in the way described in the next paragraph. Values of the four coordinates of an area containing the target (x_{min}, x_{max}, y_{min}, y_{max}) are sent back to the extension card via the ISA bus. They determine the position of the tracking window on the screen as shown in Fig. 5. Since the 2D filter processes the entire image in real time, the tracking window size is not limited and may be almost as large as the image itself. To display the tracking window an MSB of the pixel intensity data is inverted in appropriate moments before the intensity is converted back to the analog form. During the object movement the coordinates x_{min}, x_{max}, y_{min}, y_{max} change what is shown on a screen as the tracking window movement or/and shape change. It enables the operator to monitor the tracking process.

1/3	0	– 1/3
1/3	0	– 1/3
1/3	0	– 1/3

1/10	1/10	0	– 1/10	– 1/10
1/10	1/10	0	– 1/10	– 1/10
1/10	1/10	0	– 1/10	– 1/10
1/10	1/10	0	– 1/10	– 1/10
1/10	1/10	0	– 1/10	– 1/10

Fig. 2. Examples of the 2D classic edge detectors.

0	0	– 1.5	– 1.5	– 1.5	– 1.5	0	0
0	– 1.5	– 0.25	– 0.25	– 0.25	– 0.25	– 1.5	0
– 1.5	– 0.25	1	2.5	2.5	1	– 0.25	– 1.5
– 1.5	– 0.25	2.5	2.5	2.5	2.5	– 0.25	– 1.5
– 1.5	– 0.25	2.5	2.5	2.5	2.5	– 0.25	– 1.5
– 1.5	– 0.25	1	2.5	2.5	1	– 0.25	– 1.5
0	– 1.5	– 0.25	– 0.25	– 0.25	– 0.25	– 1.5	0
0	0	– 1.5	– 1.5	– 1.5	– 1.5	0	0

Fig. 3. Example of a SED filter (to obtain unit edge response the filter coefficients should be multiplied by 0.1).

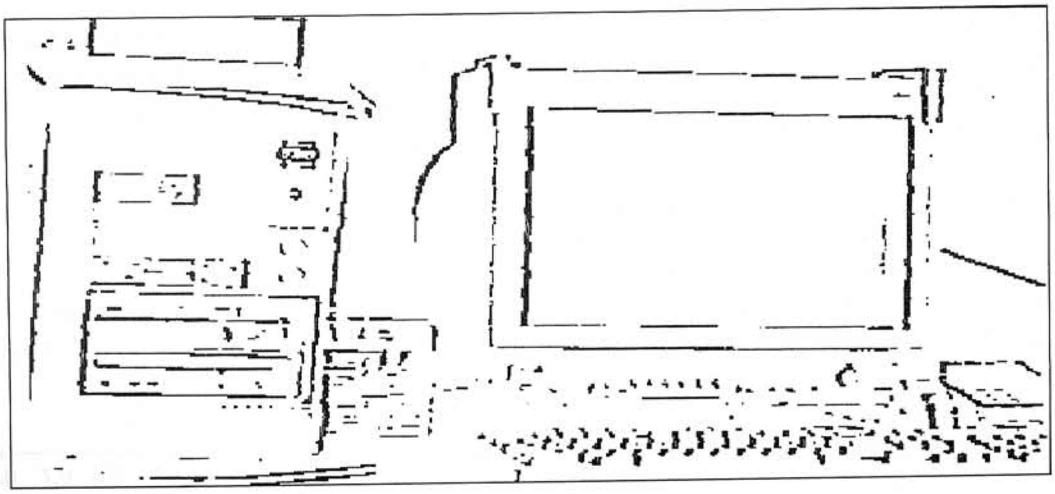

Fig. 4. Example of a binary edge image.

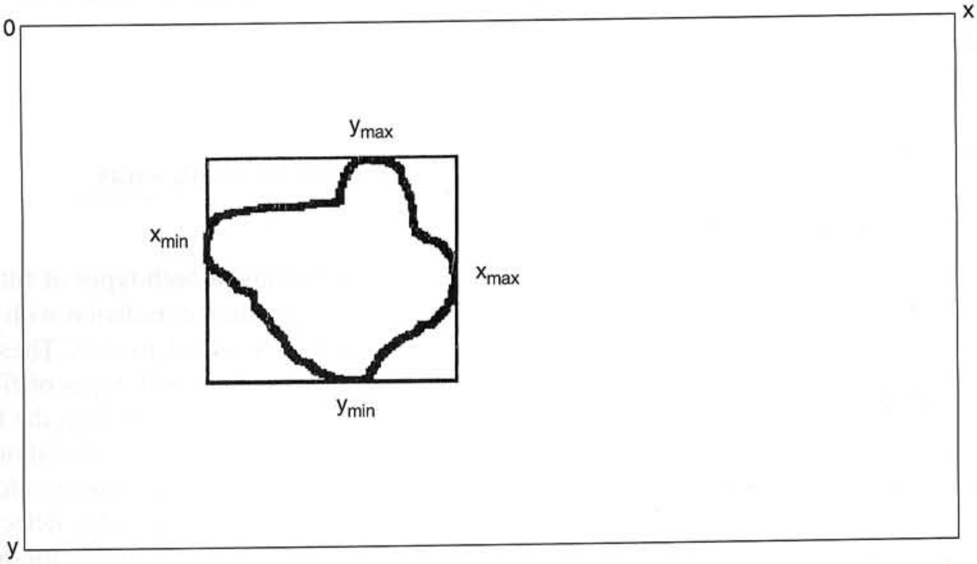

Fig. 5. Coordinates of the tracking window.

3. Software

After the main routine is started, the initial settings of the tracking window are sent to the card. The operator can change the window position from the keyboard. When the window contains at least a part of the selected object, the operator can initiate the automatic tracking procedure. This routine awaits the interrupt from the FIFO board. When it occurs the data transfer from FIFO to RAM is started. Once the transfer is finished the program begins searching for non-zero pixels in the area containing the tracking window. This area is determined by

$$x_{min} - d < x < x_{max} + d \qquad (2)$$
$$y_{min} - d < y < y_{max} + d$$

whereas the coordinates of the tracking window are

$$x_{min} < x < x_{max} \qquad (3)$$
$$y_{min} < y < y_{max}$$

The scanned area is slightly greater since the initial object size may exceed the tracking window size. An increase in the searched area enables the window to match the object size after a few iterations. Maximum and minimum x_{min}, x_{max}, y_{min}, y_{max} coordinates of non-zero pixels (shown in Fig. 4) are stored and they determine the tracking window coordinates during the next TV field. Since the data varies according to the target size and/or position, the operator can monitor the tracking process on a screen. If necessary, the tracking data may be recorded for further processing.

4. Filter selection

The system performance is mainly determined by the two-dimensional filter selection and the threshold level value. The latter one was chosen in a heuristic way to get the best results for a given class of images. A discussion of this problem may be found elsewhere [9]. Two types of filters were tested: classic edge detectors (Fig. 2) and single (filter) edge detectors (SED) of the type described below. The classic edge detectors are sensitive to the edge direction and they give zero output for edges orthogonal to the filter direction. Thus, at least two of such filters (perpendicularly directed) are necessary for correct operation.

The SED filter is optimised for edge detection [10]. Its correlation with an ideal edge is maximum or minimum. At the same time, it has good noise properties and its response to a constant background equals zero. Furthermore, its response is independent of the edge direction. It means that the filter coefficients should depend only on the distance to the filter centre (r). If we denote the filter (mask) coefficients by p_{ij},$(1 \le i \le N,\ 1 \le j \le N)$ the enumerated conditions lead to the following equations

$$J_1 = \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij} = 0 \tag{4}$$

zero correlation with a constant background,

$$J_2 = \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij}^2 = const \tag{5}$$

predefined value of the white noise response,

$$I = \sum_{i=1}^{N} \sum_{j=d+N/2}^{N} p_{ij} = \max/\min \tag{6}$$

maximum correlation with an ideal linear edge placed at the distance d from the filter centre.

The problem solution is given in the form of functional equation [10,11]

$$\nabla I + \lambda_1 \nabla J_1 + \lambda_2 \nabla J_2 = 0 \tag{7}$$

where ∇ denotes the functional gradient, and λ_1, λ_2 are the Lagrangian multipliers. Since the functionals I and J_1 are linear and the functional J_2 is quadratic, the gradients may be easily calculated. In [10] Eq. (7) is solved and solutions are given in an explicit form

$$f(r) = \begin{cases} -\dfrac{\lambda_1}{2\lambda_2} & for\ \ r \le d \\[2ex] \dfrac{\lambda_1 + 2\arccos\dfrac{d}{r}}{2\lambda_2} & for\ \ a \ge r \ge d \end{cases} \tag{8}$$

Here r is the distance from the filter centre, f(r) is the pixel value, a is the filter radius and d is the parameter. The examplary SED filter and the function f(r) are shown in Figs. 3 and 6, respectively.

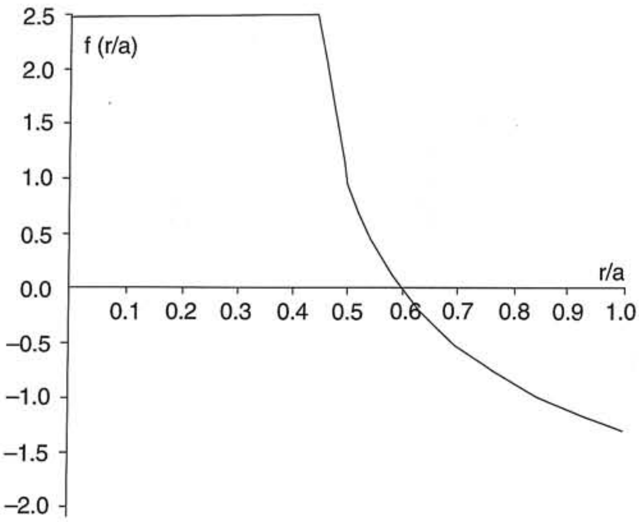

Fig. 6. Function f(r) for d/a = 0.45.

The coefficients of both types of filter are so chosen that the maximum correlation with an ideal edge (100% contrast) is equal to one. These correlations are shown in Fig. 7 for both types of filter as a function of the edge displacement from the filter centre. It is necessary to stress that the correlation maximum is slightly shifted from the edge position for the SED filters. Contrary to the classic edge detectors, the SED filters have the same sensitivity for horizontal and vertical edges and almost the same for slant edges. This is shown in Fig. 8. An important measure of performance of the filters is their noise resistance. The noise resistance increases with increase in the filter size. Furthermore, it follows from [10] that the classic edge detectors are less noise sensitive than the SED filters of the same size.

Another important feature of each filter is its ability to detect small objects. It is directly related to the noise resistance. In general, if the filter is more sensitive to noise it can detect smaller objects as well. Thus, the SED filters give better results for small targets. Another issue is tracking of mobile targets. Since the CCD camera integrates the incident light, the mobile objects have smeared edges as shown in Fig. 9. When the object shifts during TV field, it is difficult to find the precise position of its edge. Due to the smearing, the correlation of the object decreases for both types of filter. The maximum correlation is shown in Fig. 10 as a function of the object's shift. It is readily seen that it rapidly drops when the object

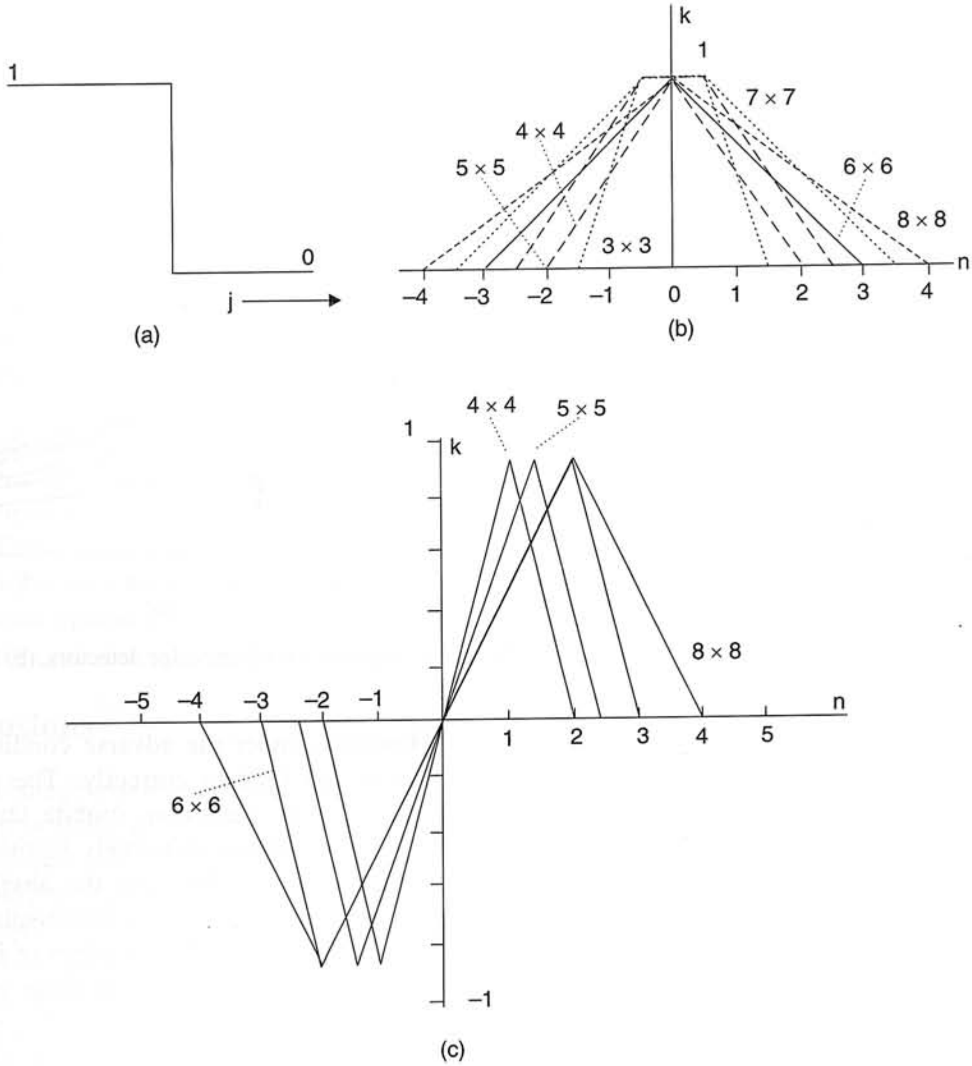

Fig. 7. Filter response to an ideal edge (a) as a function of the displacement n from the edge given in pixels: (b) classic edge detectors, (c) SED filters.

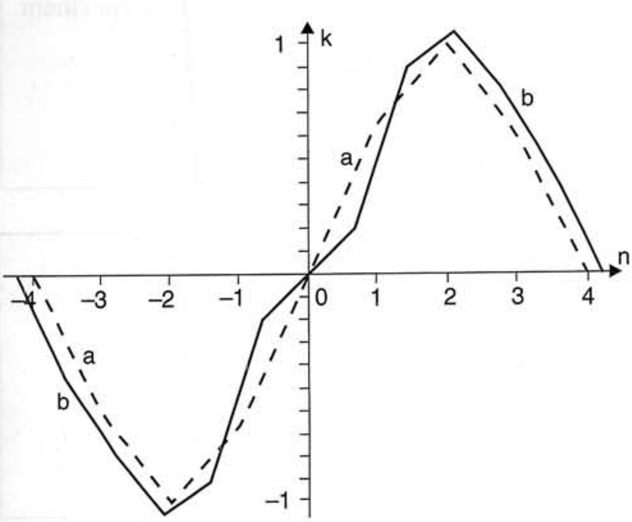

Fig. 8. SED filter (Fig. 3) response to edges of different orientations: (a) vertical or horizontal edges, (b) slant (45° or 135°) edges.

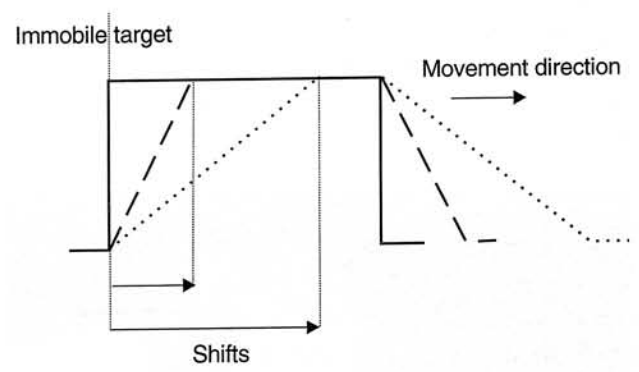

Fig. 9. Mobile object intensity.

shift increases. The correlation is greater for the filters of larger size.

In general, the filter selection is determined by the objects we want to detect and track. The edge filters are better suited for larger objects on (possibly) noisy images, whereas the SED filters give better results for

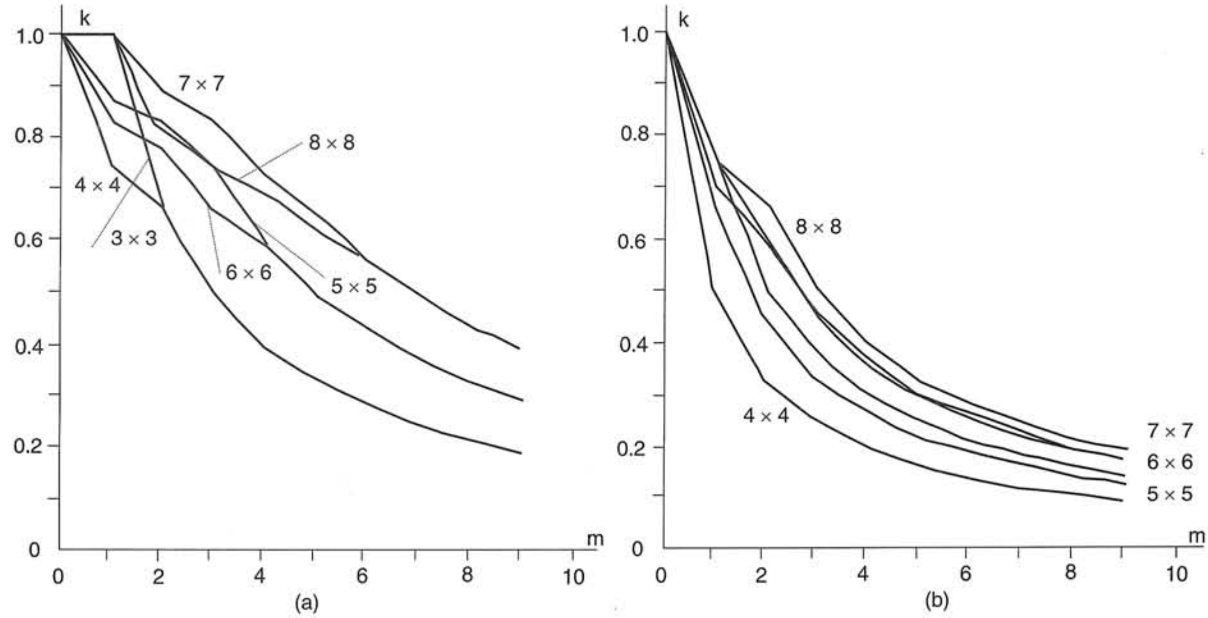

Fig. 10. Maximum correlation as a function of the object shift m given in pixels: (a) classic edge detectors, (b) SED filters.

smaller objects. We do not state that our filter always outperforms classical edge filters. We would like to stress that our (single) filter may successfuly replace a pair of classical edge filters.

5. Results

The photo of the system is shown in Fig. 11. A high noise level from the CCD camera prohibited employment of small size SED filters. Also the camera focusing and mean brightness adjustment turned out to be very important for correct operation. The equipment worked correctly (errorless detection and tracking of objects) under most conditions: when the background is relatively uniform and when the target does not move very fast in regard to the camera. Such conditions are satisfied for example during aircraft tracking.

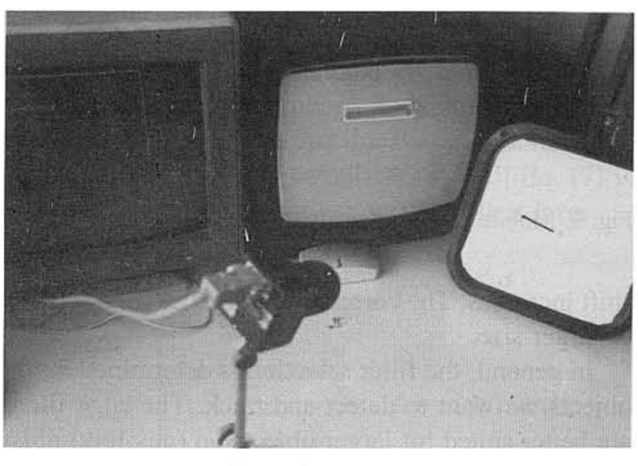

Fig. 11. System photograph.

However, under the adverse conditions the system does not operate correctly. The problems are listed below. In particular, mobile targets tracking caused the problems. Extremely fast moving targets practically disappeared from the auxillary monitor screen where the edge image was displayed. This effect is shown in Fig. 12. The edges of immobile diskette are clearly seen, whereas those of mobile one

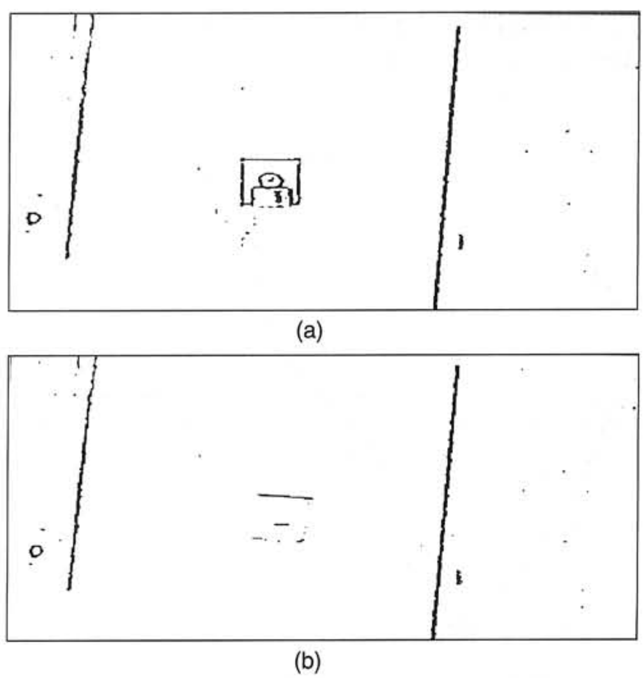

Fig. 12. Edge images: (a) immobile object, (b) fast moving object.

that are perpendicular to movement direction disappeared. This effect caused the problems with tracking of fast moving objects; the tracking window contained a part of the object only or even the target was lost. The target movement caused another problem. In a given TV field, the program scans a neighbourhood of the object position in the previous field. When the object position shifts from frame to frame the tracking window is delayed with regard to the actual target position. Some kind of movement prediction (like Kalman filtering [7]) would solve this problem.

On the other hand, when the object contrast was low or the background was not uniform some background fragments had the edges that exceeded the threshold. These points interfered with the object points leading to the miscalculation of a window size and to incorrect operation.

6. Conclusions

The presented system makes the target detection and tracking possible in real time. Its application is limited to objects on a relatively uniform background. Tracking objects with low contrast or partly obscured objects require application of matched filtering [8]. The 2D filter chips employed in the device may be readily cascaded to produce the filters of a size 16×16, 32×32 or even greater. The system was not designed for any specific application, but it may be employed in many areas including quality control, industry, military applications and other. Only minor parameters and program adjustments are necessary.

Acknowledgements

This work was performed under Warsaw University of Technology grant No 905/6 (Photonic Engineering Priority Programme).

References

1. P.K. Sahoo, "A survey of thresholding techniques", *Computer Vision, Graphics & Image Processing* **41**, 233 (1988).
2. X. Zhang and G. Sexton, "Automatic pedesterian counting using image processing techniques", *Electronics Letters* **31**, 863 (1995).
3. C.K. Lee and W.S. Ng, "Real-time image segmentation system for IC die patterns", *Int. J. Electronics* **78**, 959 (1995).
4. D. Kuan, "Autonomous robotic vehicle road following", *IEEE Trans. Pattern Analys. & Machine Intellig.* **10**, 648 (1988).
5. S.M. Smith and J.M. Brady, "ASSET-2: Real-time motion segmentation and shape tracking", *IEEE Trans. Pattern Analys. & Machine Intellig.* **17**, 814 (1995).
6. GEC Plessey, *Digital Video & Digital Signal Processing IC Handbook*, 1993.
7. B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, 1979.
8. D. Middleton, *An Introduction to Statistical Communication Theory*, McGraw-Hill Inc., 1960.
9. P.L. Rosin, "Edges: saliency measures and automatic thresholding", *Machine Vision and Applications* **9**, 139 (1997).
10. J. Siuzdak, "A single filter for edge detection", *Pattern Recognition* **31**, 1681 (1998).
11. K.W. Cathermole and J.J. O'Reilly, *Optimisation Methods in Electronics and Communication*, Pentech Press, 1984.