The Edge Point Detection Problem in Image Sequences: Definition and Comparative Evaluation of Some 3D Edge Detecting Schemes[*]

L. Jetto[†]   G. Orlando[‡]   A. Sanfilippo

Dipartimento di Elettronica e Automatica

Università di Ancona

Ancona, ITALIA

Abstract

When dealing with image sequences it is important to take into account the temporal correlation among consecutive frames to improve the performance of image processing techniques. In this paper the edge detection problem is considered and some 3D edge detectors exploiting the information carried by the spatio-temporal correlation of the 3D signal are proposed. Their performance is quantitatively and qualitatively evaluated.

1 Introduction

In the last two decades, a growing interest has been observed in the field of three-dimensional (3D) signal processing. This because the analysis of 3D signals finds numerous applications in bio-medical imaging (Jaffe, 1982; Cappellini *et al.*, 1986; Digalakis *et al.*, 1993), computer vision (Ballard *et al.*, 1982), 3D TV video signals (Quellet and Dubois, 1981), image restoration and enhancement for geophysical data (Keskes *et al.*, 1982), motion estimation (Pardas and Salembier, 1994), processing of time-varying images (Huang, 1986). In this context, a particular important problem is the edge detection as a tool to improve the image understanding. In fact, edge preservation is essential in image processing due to the nature of human visual perception and the performance of many vision systems depends on the performance of the edge-detector. While in the past years a wide interest has been devoted to the two-dimensional (2D) edge detecting problem, (see e.g. (Rosenfield and Kak, 1982) and references therein), the area of the edge detection problem for 3D signals is still largely unexplored. The purpose of this paper is to define some 3D edge detecting schemes and to experimentally evaluate their performance. The starting point is a natural extension of some 2D edge detectors.

When analyzing a sequence of 2D images it is important to take into account the time correlation. In fact, the segmentation processing of each 2D frame is improved if the information carried by the time correlation with past and future frames is suitably exploited. To this purpose, a number of 3D edge detectors is here proposed by suitably modifying the structure of some among the most widely used 2D operators. The performance of the proposed 3D detectors has been quantitatively compared in terms of sensitivity to edge orientation, edge displacement and presence of noise. A qualitative evaluation has been also performed on a sequence of real 3D scenes.

[†]email: *L.Jetto@ee.unian.it*

[‡]email: *Orlando@ee.unian.it*

Besides the practical image segmentation problem, there is a theoretical reason calling for 3D edge detection schemes. The performance of a 3D minimum variance filter has been recently investigated in (Jetto, 1999). Following the lines of De Santis *et al.* (1994); Concetti and Jetto (1997), the 3D filter proposed in (Jetto, 1999) contains a structural information on the spatio-temporal signal discontinuities. The inclusion of this information into the filter equations allows one to prove the filter optimality and is useful to reduce the edge blurring phenomenon.

The paper is organized in the following way. Section 2 describes the basic structure of the 3D edge detectors, their performance is quantitatively evaluated in section 3, section 4 reports the results obtained on a real image sequence.

2   The 3D edge detectors

A 3D edge of an image sequence may correspond to the contour of an object in the space, to a shadow (illumination discontinuity) or to a surface mark (reflectance discontinuity). In any case a 3D edge is seen as an abrupt intensity change in the image sequence. Based on this definition, it seems natural to define a 3D edge detector as an operator which is able to identify sharp discontinuities of the 3D surface describing the distribution of the gray level. Two different basic approaches have been here defined and applied. These approaches are based on the use of gradient edge detectors and template matching operators. For the 2D problem, the above methods are extensively discussed in (Rosenfield and Kak, 1982).

2.1   Gradient Edge Detectors

Three (possibly orthogonal) linear operators M_x, M_y, M_t are defined. Convolving the 3D signal with the masks of the above linear operators results in the detection of the "edge weights", which are a measure of the rate of the change of the gray level distribution along the vertical, horizontal and time direction. For each pixel of coordinates (i, j, k), the relative edge weights are denoted by $G_x(i, j, k)$, $G_y(i, j, k)$, $G_t(i, j, k)$. The gradient magnitude $M_{i,j,k}$ and its directions $\alpha_{i,j,k}, \beta_{i,j,k}$ with respect to x axis and $x - y$ plane respectively, are computed as

$$
\begin{aligned}
M_{i,j,k} &= \sqrt{G_x(i, j, k)^2 + G_y(i, j, k)^2 + G_t(i, j, k)^2}, \\
\alpha_{i,j,k} &= \arctan \frac{G_y(i, j, k)}{G_x(i, j, k)}, \\
\beta_{i,j,k} &= \arctan \frac{G_t(i, j, k)}{G_{xy}(i, j, k)},
\end{aligned}
$$

where $G_{x,y}(i, j, k) = \sqrt{G_x(i, j, k)^2 + G_y(i, j, k)^2}$. If the value of $G_{x,y}(i, j, k)$ exceeds a suitably defined threshold, pixel (i, j, k) is identified as an edge point.

2.2   Template Matching

In the 2D case, eight alternative orientations of an edge around a pixel exist. The 2D template matching technique requires to define eight different linear operators (submasks). A convolution of the image with each one of the eight submasks is performed and the sharpest response from the set of the eight possible alternatives is selected as the best edge and compared with a suitably defined threshold.

In the 3D case, there are twenty-six possible orientations of an edge. The twenty-six directions are reported in figure 1.
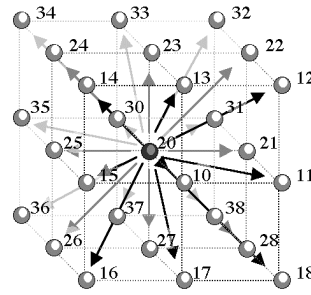
Figure 1: *The 26 possible orientations of a 3D edge*

Denoting by $P_{i,j,k}^{(l)}$, $l = 1, \cdots, 26$; the value of the convolution of the 3D signal with the twenty-six templates centered on pixel (i, j, k), then $\bar{P} = \max_{l=1,\cdots,26} |P_{i,j,k}^{(l)}|$ denotes the maximum rate of gray level variation, the value $\bar{l}$ for which $P_{i,j,k}^{(l)}$ is maximum identifies the direction of the edge.

2.3 The Linear Operator Masks

Both the gradient edge detectors and the template matching operators are based on the convolution of the 3D signal with linear operators represented by numerical matrices (masks). For analogy with the 2D case, the following operators have been considered: 3D Prewitt-like operator, 3D Sobel-like operator, 3D Cantoni-like operator, 3D extension of the noise-smoothed difference operator described in (De Santis *et al.*, 1994). Cantoni operator has been employed only for the template matching approach, the 3D smoothed difference operator only for the gradient technique, hence, only three masks are needed for this latter edge detector.

Extending the definition of the 2D (3×3) Prewitt detector, an operator defined by twenty-six masks (see figure 2) has been considered. Owing to the symmetric structure of the twenty-six matrices, only thirteen convolutions have to be performed, the others can be obtained with a change of sign.

For the sake of brevity, the masks of 3D (3×3) Sobel, Kirsch, ($3 \times 4 \times 3$) smoothed difference and Cantoni operators are reported in Figure 3 only for the direction $2, 1$ which corresponds to y axis. The other twenty-five masks of Sobel , Kirsch , Cantoni and the two others of smoothed difference can be easily obtained by rotation. The smoothed difference operator ($3 \times 4 \times 3$) shown in figure 3 is based on an extension of the 2D mask proposed in (De Santis *et al.*, 1994), which showed to be able to conciliate the two opposite requirements of noise reduction and edge preservation.

3 Quantitative evaluation of the linear operator masks

As in (Cantoni *et al.*, 1982), three figures of merit have been used for evaluating the performance of the 3D edge detectors described in the above section: sensitivity to edge orientation, sensitivity to edge displacement, sensitivity to noise. For a 2D surface representing an ideal step edge, the first figure measures the normalized amplitude of the edge detected as a function of the edge orientation, provided that the edge is passing through the center of the mask. The results obtained are shown in Figure 4. The second index measures the normalized amplitude of the edge detected as a function of the displacement of the actual edge from the center of the edge
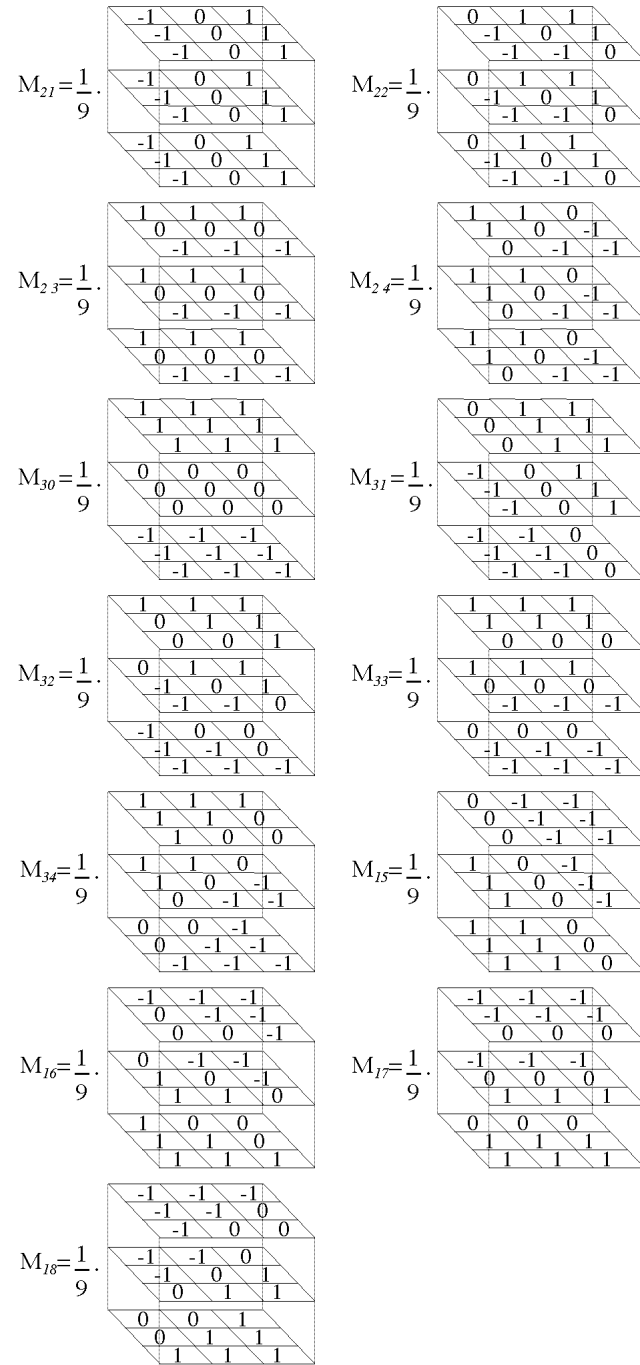
Figure 2: *This figure shows 13 basic masks of* (3×3) *3D Prewitt-like operator. The other 13 masks are obtained changing* -1 *with* 1 *and vice-versa.*
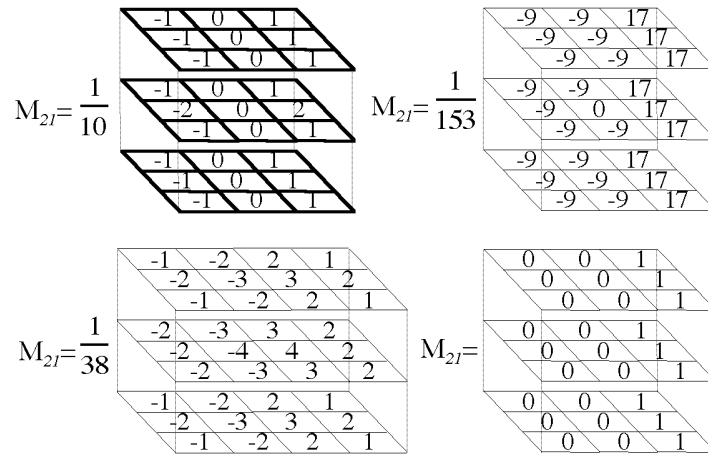
Figure 3: *Masks 2,1 of Sobel (a) , Kirsch (b) , smoothed difference (c) and Cantoni (d) operators for the direction* $2,1$ *(y axis).*
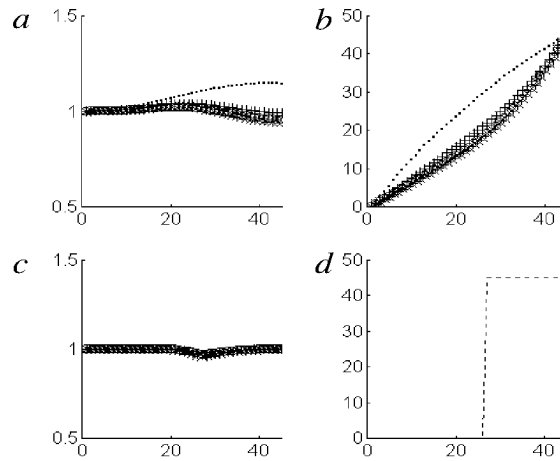

Figure 4: *Figure of merit relative to edge orientation. Continuous line:Cantoni, $+++$: Sobel, $\circ\circ\circ$: Kirsch, $\cdots$: smoothed, $\times\times\times$:Prewitt. The abscissa of the four plots is the edge orientation (expressed in degrees). The normalized amplitude and phase of the edge detected with the gradient approach are reported in diagrams a and b respectively. The normalized amplitude and phase (expressed in degrees) of the edge detected with the template matching approach are reported in diagrams c and d respectively.*

detector. The results obtained are reported in Figure 5. As in the noiseless case all the operators
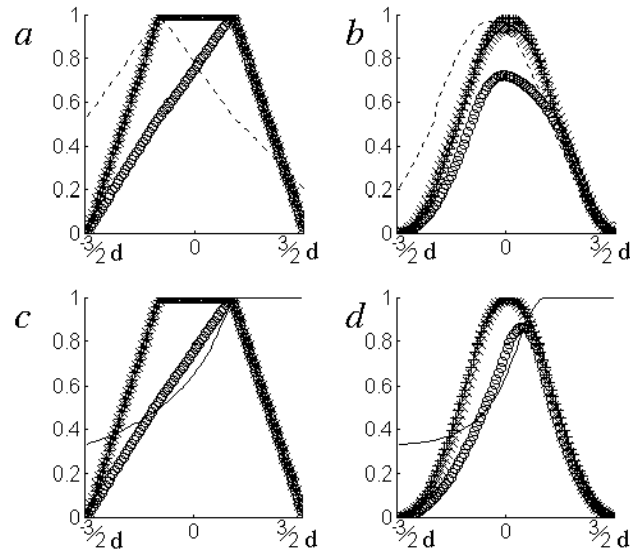

Figure 5: *Figure of merit relative to edge displacement. Continuous line: Cantoni, +++: Sobel, ∘ ∘ ∘: Kirsch, − − −: smoothed, × × ×:Prewitt. The normalized amplitude of the edge detected with the gradient approach for a vertical edge and for an edge inclined of 45^0 are reported in diagrams a and b respectively. The normalized amplitude of the edge detected with the template matching approach for a vertical edge and for an edge inclined of 45^0 are reported in diagrams c and d respectively. The abscissa of each plot is distance (positive or negative) between the edge and the central point of the mask.*

are roughly equivalent, the third figure of merit has been considered because it is useful to take into account the presence of noise. This figure is defined through the quantity

$$ I = \frac{1}{\max(N_l, N_d)} \sum_{i=1}^{N_d} \frac{1}{1 + \gamma d_i^2} $$

where d_i is the distance between a pixel considered as edge and the nearest ideal edge pixel, γ is a calibration constant, N_i and N_d are the number of ideal and detected edges respectively. The performance of an edge detector is proportional to the value of I. An idea of the possible situations of error in the 3D case is illustrated in Figure 6. The results obtained with this third figure of merit are reported in Figure 7. The indexes of merit plotted in Figures 4, 5 and 7 reveal the following. The response of the 3D detectors is relatively invariant to the edge orientation and the capability of producing thin boundary lines is well preserved. Save 3D Cantoni-like operator, all the responses of figure 5 are monotonically increasing before the edge and rapidly decreasing after it. Note, in particular, the sharp response of smoothed difference operator. Also the robustness against the presence of noise is well preserved. It may be concluded that under the different operating conditions simulated here, the proposed 3D edge detectors exhibit a response similar to that produced by the corresponding 2D operators applied to 2D sample test images (Cantoni *et al.*, 1982).
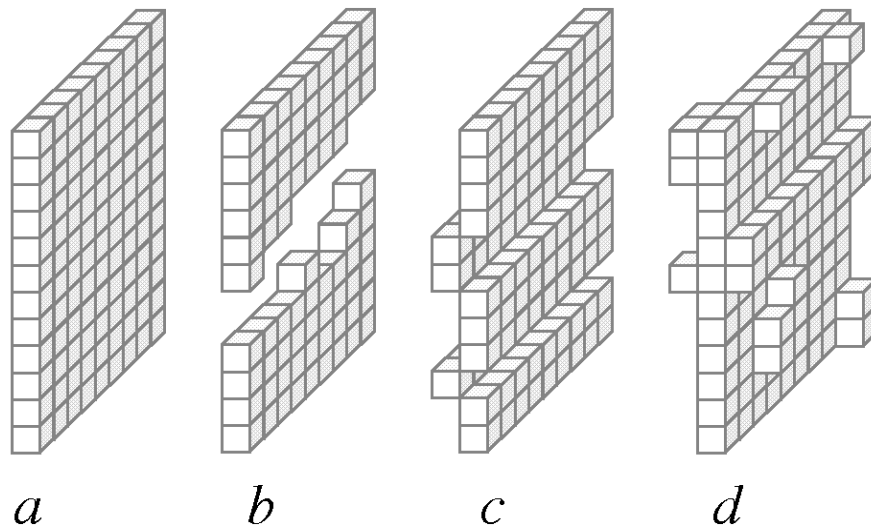
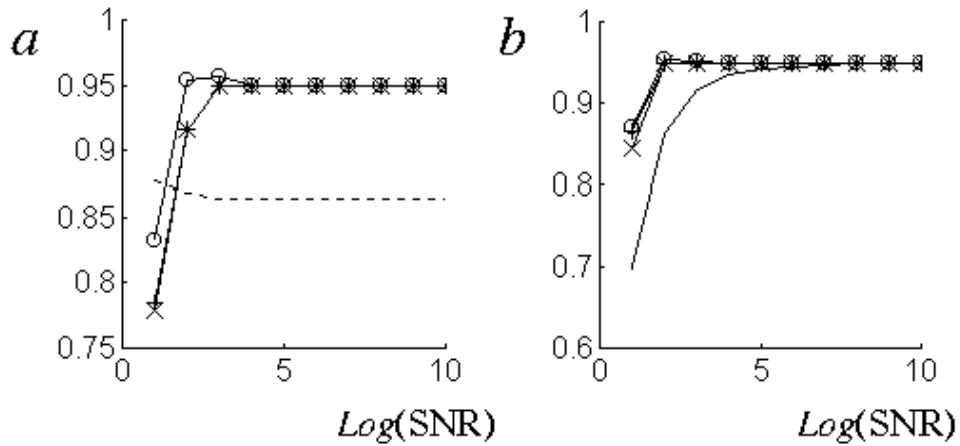Figure 6: *a) Ideal case, b) Fragmentation, c) Offset, d)Excess.*

Figure 7: *Figure of merit I. Continuous line: Cantoni, + + +: Sobel, ∘ ∘ ∘: Kirsch, − − −: smoothed, × × ×:Prewitt. Abscissa: values of the Signal-to Noise Ratio. Ordinate: values of I obtained with the gradient approach (diagram a) and with the template matching approach (diagram b).*

4   Numerical experiments on a real image sequence

The performance of the proposed edge detectors have been also visually evaluated since human eye performs some sort of edge detection. To this purpose a real 3D test image has been considered. The 3D signal was acquired with a handy camera Phonola 68VKR40 and a frame-grabber on PCI BUS of the National Instruments with a 8 bit analog-to-digital converter. The sampling frequency is 50 Hz, but a sampling period of 0.32 sec was simulated considering one frame every sixteen to accentuate the movement. In this way a sequence composed of nine frames has been obtained. The acquired sequence has been considered as a the true noise-free signal. The original signal has been corrupted by a zero-mean white gaussian noise with a variance such that the SNR (signal variance/noise variance) resulted 8. For the sake of brevity, only the results relative to fourth and fifth frame (see Figure 8) are here reported. The sequence contains
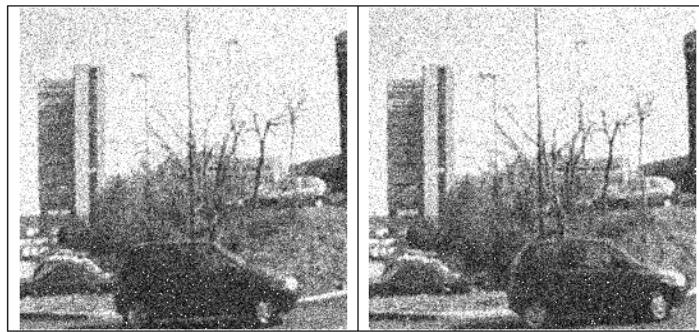

Figure 8: *Noisy signal. The sequence moves from the left to the right.*

a single object (a car) moving across a stationary textured background (the Engineering Faculty of Ancona).

Each sampled frame is composed of 256×256 pixels. The two following edge detectors have been used: $(5 \times 5 \times 5)$ Prewitt operator implemented according to the template matching approach, $(3 \times 6 \times 3)$ smoothed difference operator implemented according to the gradient technique. The masks of these operators corresponding to direction $2, 1$ (y-axis) are reported in Figure 9. The threshold has been computed as two times the standard deviation σ of the output of the edge detector. Parameter σ has been numerically estimated over a moving window centered around the current pixel. The results obtained with the gradient and template matching methods are reported in figures 10 and 11 respectively. Temporal edges images represent the discontinuity points in the time direction between two consecutive frames of the sampled sequence. The edge detection procedure has been implemented in C^{++} on a PC, with windows 95 O.S., 100 MHz processor, 16Mbyte Ram. To evidence the differences between 3D and 2D edge detection operators, a 2D edge detection has been performed considering each single frame separately, namely neglecting the temporal correlation. The 2D detectors have been obtained from the central masks of the 3D Prewitt and smoothed difference operators shown in Figure 9. The results obtained with the 2D Prewitt and smoothed difference operator are reported in Figure 12. A comparison with Figures 10 and 11 clearly reveals the improvement introduced by the use of 3D detectors.
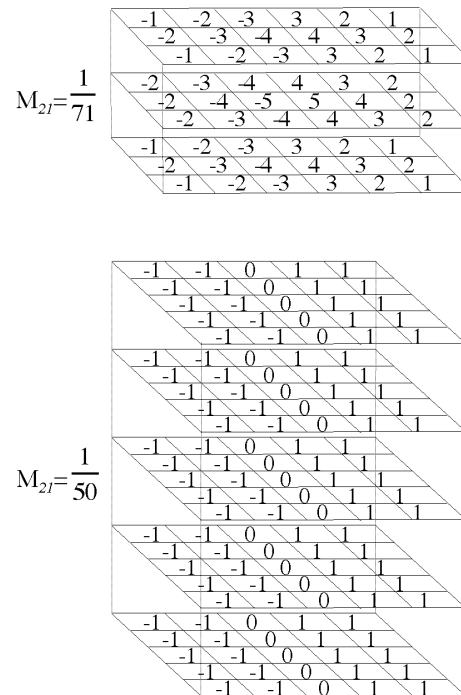
$$M_{21}=\frac{1}{71}$$

$$M_{21}=\frac{1}{50}$$

Figure 9: *Masks* $2,1$ *of smoothed difference (top) and Prewitt (bottom) operators used for the real sequence.*
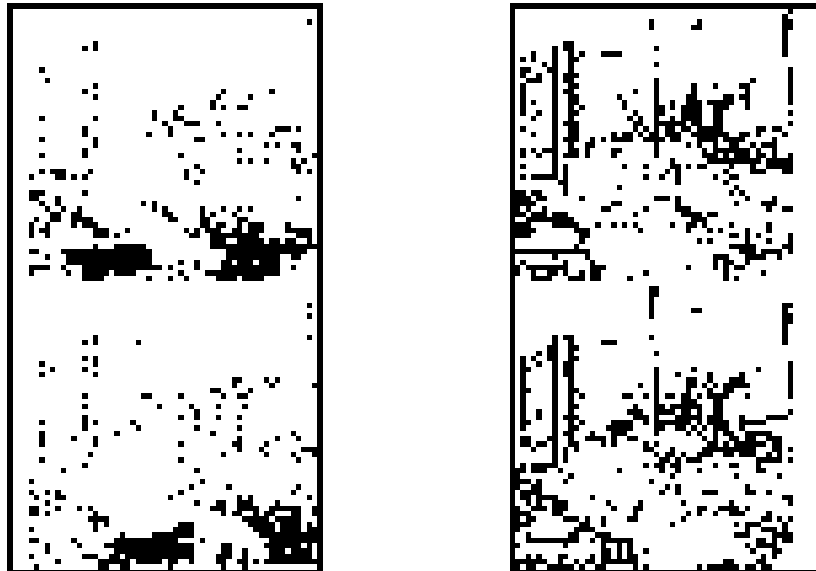
Figure 10: *Temporal (left) and spatial (right) edges detected with the edge gradient method implemented using the noise-smoothed difference operator.*
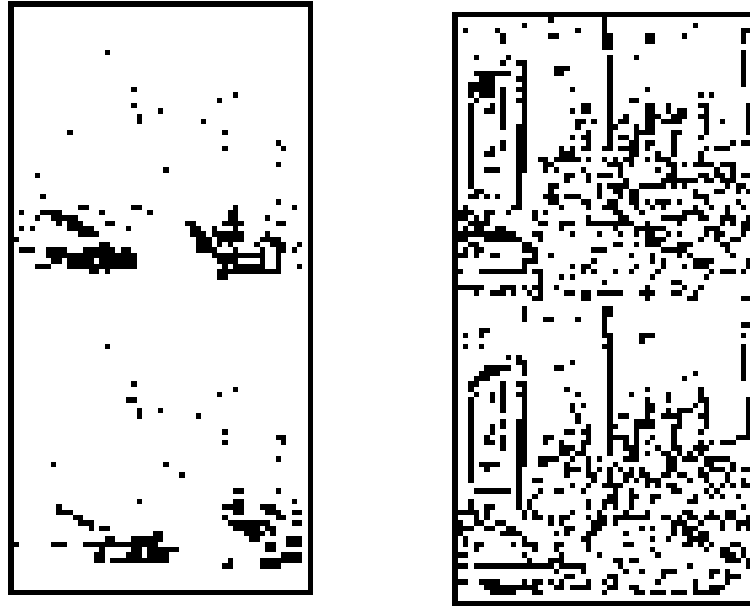
Figure 11: *Temporal (left) and spatial (right) edges detected with to the template matching approach implemented using the Prewitt operator.*
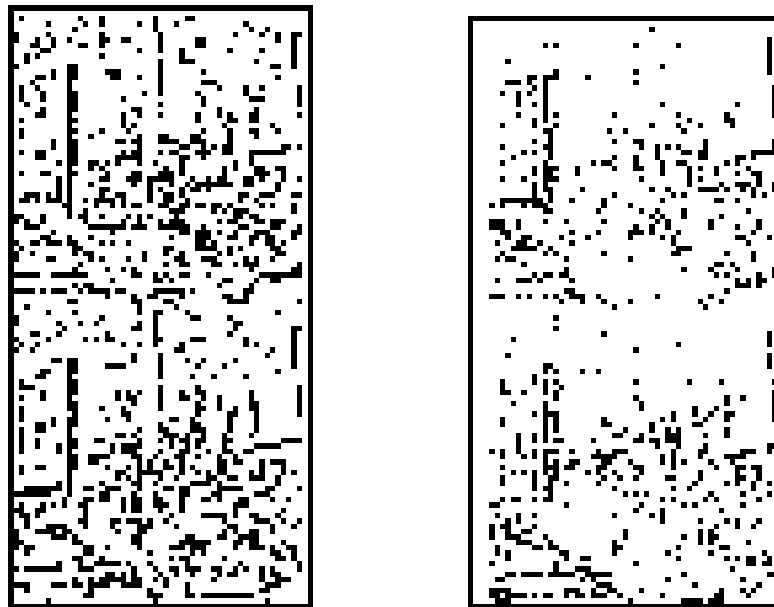

Figure 12: *Results of the 2D edge detectors. Left: Prewitt, right: smoothed difference. The sequence moves from the top to the bottom.*

5   Conclusions

Motivated both by practical and theoretical considerations, the 3D edge detection problem has been here considered. The proposed schemes are based on the definition of suitable 3D matrices and can be viewed as a generalization of the corresponding 2D detectors. A quantitative and visual evaluation of the performance of the proposed edge detectors has been carried out. The tests on the sensitivity to edge orientation and displacement and to the presence of noise have shown that the application of the 3D edge detectors to 3D test signals has produced a classification of performances similar to that obtained applying the corresponding 2D detectors to 2D test images. The application to a real sequence of noisy images has evidenced that significant improvements are introduced in the edge detection if the temporal correlation among consecutive frames is taken into account.

References

Ballard, D., C.M., and C. Brown (1982). *Computer Vision*, Prentice Hall, Englewood Cliffs.

Cantoni, V., I. De Lotto, and M. Ferretti (1982). "A template matching operator for edge points detection in digital pictures," *Signal Processing*, **4**, pp. 349–360.

Cappellini, V., R. Carla, and M. Melani (1986). "3-D digital filtering of biomedical images," *Proc. 1986 European Signal Processing Conference*, pp. 1383–1386.

Concetti, A. and L. Jetto (1997). "Two-dimensional recursive filtering algorithm with edge preserving properties and reduced numerical complexity," *IEEE Trans. on Circ. and Syst.*, **44**, pp. 587–591.

De Santis, A., A. Germani, and L. Jetto (1994). "Space-variant recursive restoration of noisy images," *IEEE Trans. Circ. and Syst.*, **41**, pp. 249–261.

Digalakis, V., V. Ingle, and D. Manolakis (1993). "Three-dimensional linear prediction and its application to digital angiography," *Multidimensional Systems and Signal Processing*, **4**, pp. 307–329.

Huang, T. (1986). *Image Sequence Processing and Dynamic Scene Analysis*, Academic Press, New York.

Jaffe, C. (1982). "Medical imaging," *American Scientist*, **70**, pp. 576–585.

Jetto, L. (1999). "Stochastic modelling and 3d minimum variance recursive estimation of image sequences," *to appear on MSSP Journal*.

Keskes, N., A. Boulanovar, and O. Faugeras (1982). "Application of image analysis techniques to seismic data," *Proc. IEEE Int. Conf. on Acoust. Speech , Signal Processing ICASSP*.

Pardas, M. and P. Salembier (1994). "3d morphological segmentation and motion estimation for image sequences," *Signal Processing*, **38**, pp. 31–43.

Quellet, J. and E. Dubois (1981). "Sampling and reconstruction of ntsc video signal at twice the color subcarrier frequency," *IEEE Trans. on Commun.*, **29**, pp. 1823–1832.

Rosenfield, A. and C. Kak (1982). *Digital Picture Processing*, vol. 2, Academic Press, New York.