

Features Extraction and Analysis Methods for Sequences of Ultrasound Images

Isabelle L. HERLIN and Nicholas AYACHE

INRIA - B.P. 105 - 78153 Le Chesnay - FRANCE

Abstract. Our principal motivation is to study time sequences of echocardiographic raw data to track specific anatomical structures. First, we show that the image processing can make direct use of the audio signal data, avoiding loss of information and yielding optimal results.

Secondly, we develop a strategy which takes a time sequence of raw data as input, computes edges, initiates a segmentation of a pre-selected anatomical structure and uses a deformable model for its temporal tracking. This approach is validated in a real time sequence of ultrasound images of the heart to track the left auricle and the mitral valve.

1 Introduction

1.1 Motivation and Objectives

There is a continuously increasing demand in the automated analysis of 2D and 3D medical images at the hospital[1]. Among these images, ultrasound images play a crucial role, because they can be produced at video-rate and therefore allow a dynamic analysis of moving structures. Moreover, the acquisition of these images is non-invasive and the cost of acquisition is relatively low compared to other medical imaging techniques.

On the other hand, the automated analysis of ultrasound images is a real challenge for active vision, because it combines most of the difficult problems encountered in computer vision in addition to some specific ones related to the acquisition mode:

- images are usually provided in polar geometry instead of cartesian geometry,
- images are degraded by a very high level of corrupting noise,
- observed objects usually correspond to non-static, non-polyhedric and non-rigid structures.

The geometric transformation (called scan correction) which transforms the data from a polar representation to the correct cartesian representation is usually applied through a bilinear interpolation.

We show in this paper the limitations of this scheme, which does not account for the varying resolution of the data, and we propose a new method, called sonar-space filtering, which consists of computing the scan conversion with a low-pass filtering of the cartesian image applied directly to the available polar data, and which can be used to optimally reconstruct the data with a chosen level of spatial linear filtering.

Furthermore, we develop a methodology to automatically track a physiological structure on an echocardiographic sequence. Interactivity is used to initiate the process on the first image of the sequence. Then edges are computed and an approximative segmentation of the structure is obtained by using deterministic algorithms. This information is finally combined with a deformable model to obtain the temporal tracking of the pre-selected structure.

Finally, to further demonstrate the efficiency of our approach, we apply it to a difficult time sequence of ultrasound images of the heart to track the left auricle and the mitral valve.

1.2 Previous work

Our approach is different from previous ones. This comes from the fact that we study directly the ultrasonic data. More commonly, feature extraction is applied to the cartesian video data. To our knowledge, there is only one study where all processing is performed on sector scans in polar coordinate form. This was published by Tact [13] and reports noise reduction and segmentation in time-varying ultrasound images. But a comparative study of scan correction methods to obtain cartesian images has apparently not been pursued yet. For cartesian images, the most commonly used approach to obtain the contour of left ventricle (in echocardiography) is radial search [5] [2] [7]: the procedure starts from a point inside the heart chamber and searches along different radial lines for edge points. The best-known dynamic approach is the one by Zhang and Geiser [15], who compute temporal cooccurrences to obtain both stationary points and moving points. The temporal information has also been used to filter images obtained at the same instant of the cardiac cycle [14].

2 Acquisition of an echographic image

The purpose of this section is to present some primary characteristics of image formation using echographic technologies.

A basic imaging system, called a pulsed system, is illustrated in Fig. 1. When the switch is in the transmit position, the pulse waveform $p(t)$ excites the transducer [9].¹ This results in a wavefront that is propagated in the body. The transducer produces a relatively narrow beam of propagation whose angular direction of propagation into the body is known. Immediately following the transmission, the system switches to the receive position, using the same transducer. The pulse is attenuated when it propagates through the body. When the wavefront hits a discontinuity, a scattered wave is produced. This scattered wave is received by the transducer and the resultant signal is processed and displayed along a line representing the direction of the beam.

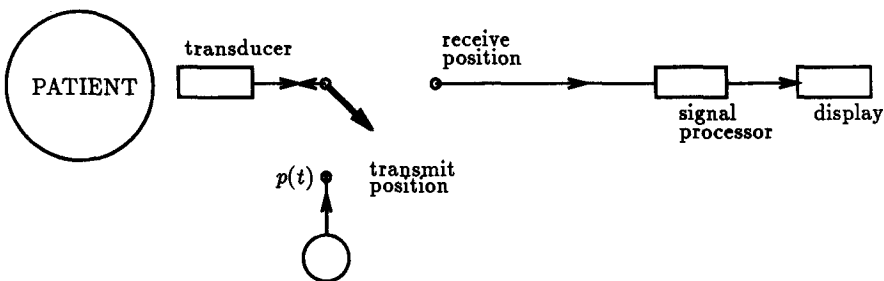


Fig. 1. Elementary pulsed ultrasonic system

¹ Other types of echographs using pseudo-random code correlation are studied in the literature [11]

3 Image representation in cartesian coordinates

The process of converting from the polar coordinates representation to the cartesian coordinates representation is necessary for the convenience of the users. Physicians are accustomed to viewing images in cartesian data and it would be difficult for them to interpret polar data. Moreover visualisation hardware and image processing algorithms are designed for data in cartesian coordinates.

Let us suppose that M different orientations are used to obtain an echocardiographic image, and that each return signal is digitized to L points. Fig. 2 shows an echographic image, with M rows and L columns, obtained with a commercial echographic machine, providing an image represented in polar coordinates. Fig. 3 shows the cartesian image corresponding to the same data.

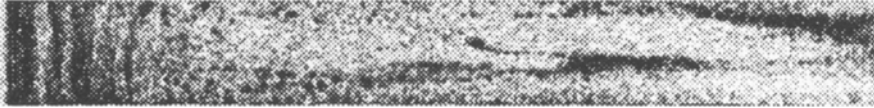


Fig. 2. Ultrasound image on raw data

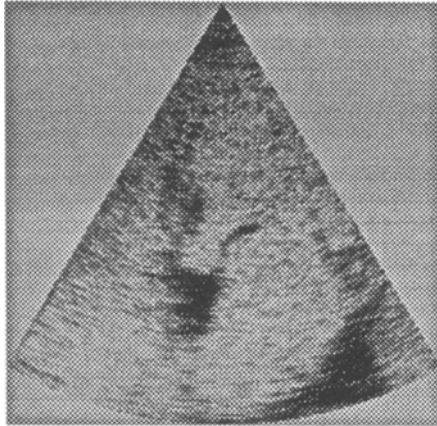


Fig. 3. Cartesian image after conversion

Scan conversion requires the knowledge of the following set of parameters (see Fig. 4):

- the angular extent of data acquisition wedge α ,
- minimal distance d for data acquisition,
- total distance D for data acquisition (these distances being calculated from the skin), and
- the number of rows, N , desired in the output cartesian image (The number of columns will be related to α , and will assume square pixels).

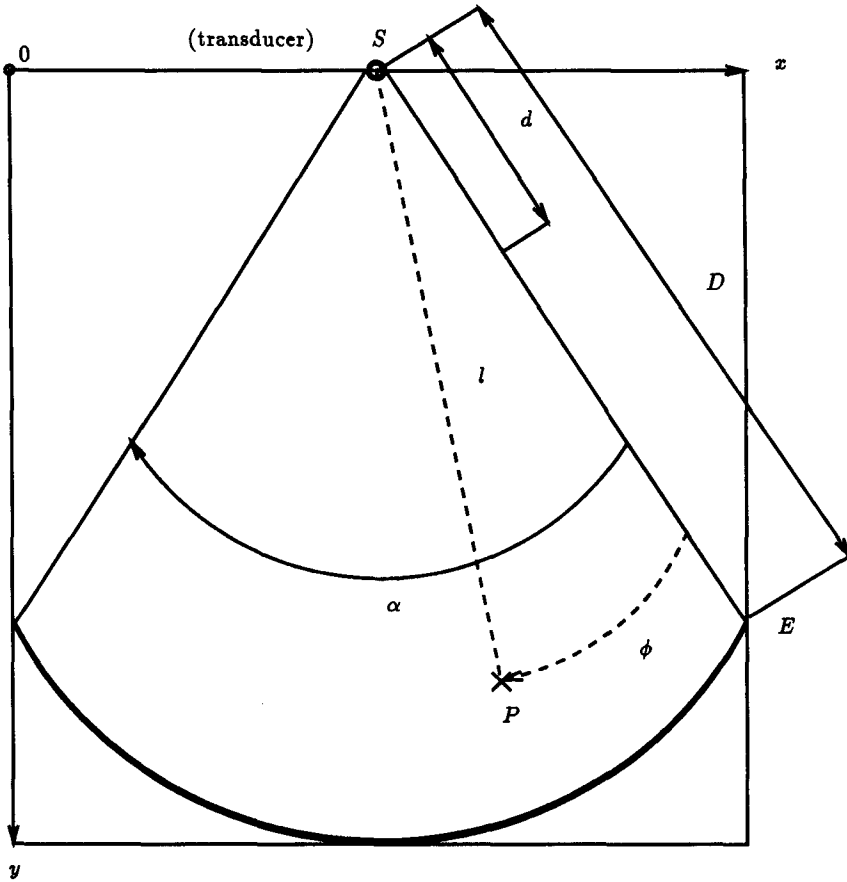


Fig. 4. Parameters of the conversion process

Several methods may be used for the conversion process. Usually, the video image on the echographic machine is obtained by assigning to a cartesian point the grey level of the nearest available point in polar coordinates, or the value of the bilinear interpolation of its four nearest points. In fact, we found that these methods do not make an optimal use of the available original data, and we introduced a new method, called sonar-space filtering, which can be used to optimally reconstruct the data with a chosen level of spatial linear filtering.

3.1 Conversion by sonar-space filtering

We assume that it is desired that the continuous input cartesian image $I(x, y)$ be filtered by the impulse response filter $f(x, y)$. The resulting image $R(x, y)$, in continuous space, is given by the convolution product:

$$R(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - u, y - v) \cdot I(u, v) \, du \, dv . \quad (1)$$

However, the input is only available in the polar coordinate space. We thus apply the following change of variables:

$$\begin{cases} \theta = (\arctan(\frac{y}{x-x_0}) - (\frac{\pi-\alpha}{2}))/\Delta_\alpha \\ \rho = (\sqrt{(x-x_0)^2 + y^2}) * e - \Delta_N . \end{cases}$$

where:

- $\Delta_N = d * \frac{(L-1)}{(D-d)}$ represents the distance from the surface of the skin where the acquisition process begins, measured in pixel units along a scan line of the raw data,
- $\Delta_\alpha = \frac{\alpha}{(M-1)}$ is the angular difference between two successive angular positions of the probe,
- $e = \frac{D}{(D-d)} \frac{(L-1)}{(N-1)}$ performs the change of pixel sampling rates along the axial direction of the beam, according to the desired height N of the cartesian image.

We obtain:

$$R(x, y) = \int_0^{2\pi} \int_0^\infty f(x - u(\rho, \theta), y - v(\rho, \theta)) \cdot I(\rho, \theta) \cdot |J(\rho, \theta)| \, d\rho d\theta . \quad (2)$$

Here $|J(\rho, \theta)|$ is the determinant of the Jacobian matrix corresponding to the inverse transformation of variables:

$$\begin{cases} x = (\frac{\rho + \Delta_N}{e}) \cos(\theta \Delta_\alpha + \frac{\pi - \alpha}{2}) + x_0 \\ y = (\frac{\rho + \Delta_N}{e}) \sin(\theta \Delta_\alpha + \frac{\pi - \alpha}{2}) . \end{cases}$$

It is easily seen that

$$|J(\rho, \theta)| = \frac{(\rho + \Delta_N) * \Delta_\alpha}{e^2} .$$

We have transformed the convolution in the cartesian coordinates to an infinite integral in polar coordinates, corresponding to the domain of the raw data.

Once the two-dimensional convolution filter f is chosen, we define its rectangle of essential support: say a rectangular window of width $2X$ and $2Y$. Outside this region of support the absolute value of the impulse response must be lower than a pre-selected threshold s , i.e.,:

$$|f(u, v)| < s \text{ if } ((|u| \geq X) \text{ OR } (|v| \geq Y)) .$$

Therefore the integral is approximated by a finite integral over the domain

$$(x - X \leq u \leq x + X) \text{ AND } (y - Y \leq v \leq y + Y) .$$

The filter is also sampled in this domain in order to approximate the integral by a discrete summation. Filtered numerical outputs are evaluated at original data point locations within the continuous domain.

We thus obtain the following equation:

$$R(x, y) = C \sum_k f(x - u(\rho_k, \theta_k), y - v(\rho_k, \theta_k)) \cdot I(\rho_k, \theta_k) \cdot |Jac(\rho_k, \theta_k)| , \quad (3)$$

where the summation is over the discrete collection of point (ρ_k, θ_k) in polar coordinates, where C is used to normalize the data and:

$$|J(\rho_k, \theta_k)| = \frac{(\rho_k + \Delta_N) * \Delta_\alpha}{e^2} .$$

We have transformed the computation of $R(x, y)$ into a discrete summation on a window of size $2X$ by $2Y$, making use of image data only at points where it is defined in the polar coordinate domain. Note that for the raw data, or more generally for any sonar-like data, the sampling of the filter is not regular along the x and y axes but it rather conforms to the sampling density of raw data. In this formula, the $|J(\rho_k, \theta_k)|$ value represents the surface area of the polar pixel patch in the cartesian domain. We will present in the following subsections the different values that we have chosen for the function f for different processing of the echocardiographic raw data.

3.2 Visualization application

In practice, the convolution filter $f(x, y)$ is typically separable and is denoted by $f(x)g(y)$. Then, classical 1-D smoothing and derivation filters can be used.

For visualization applications of sonar-space filtering, we used the Deriche's smoothing function [4] with $f(x) = g(x) = L(x)$:

$$L(x) = k_2(\alpha \sin(\omega |x|) + \omega \cos(\omega |x|))^{-\alpha|x|} . \quad (4)$$

The conversion algorithm performs simultaneously a conversion to cartesian coordinates and a smoothing of the data (whose amplitude can be adjusted with α) thus producing a cartesian image with a reduced speckle. Other smoothing functions could be used instead.

It will be noted that the visualization quality is not significantly different, or better, with sonar-space filtering than with classical bi-linear interpolation method. But our objective is not to improve visualization, but rather to improve automatic analysis of echocardiographic sequences.

3.3 Edge detection application

For further automatic boundary tracking, our goal is to use spatio-temporal approaches [10]. A time-varying edge may be represented as a surface in 3-D space, in which x and y are two spatial dimensions (in the cartesian coordinates space) and t is the temporal dimension. We modify Deriche's edge detector for this goal. Another approach could be to generalize Deriche's detector with spatio-temporal functions as in [6]. We denote G_x and G_y the two spatial components of the gradient vector and $I(x, y, t)$ the 3-dimensional grey level function. Let D be the Deriche differentiation filter and L the associated smoothing filter:

$$L(x) = k(\alpha \sin(\omega |x|) + \omega \cos(\omega |x|))^{-\alpha|x|}, \quad (5)$$

$$D(x) = k \sin(\omega |x|)^{-\alpha|x|} . \quad (6)$$

The two components of the gradient vector have the following expression:

$$G_x = (D_x L_y L_t) \otimes I(x, y, t)$$

$$G_y = (L_x D_y L_t) \otimes I(x, y, t)$$

where the subscripts are used to explain along which axis the corresponding filter is applied. Each component is obtained by differentiation in the associated direction and filtering in the other spatial direction and in the temporal direction.

The norm of the gradient is defined by:

$$N(x, y, t) = \sqrt{G_x^2 + G_y^2} .$$

The edges are obtained as local maxima of the gradient norm in the direction of the 2D gradient vector. The temporal dimension is only used to smooth the result. This produces a significant image enhancement in regions that are not moving too fast.

We denote α_x , α_y and α_t the filtering parameters of the Deriche filters (cf. Equ. 5 and 6) for the respective axes x , y and t . Since the 2D space is homogeneous, we can choose $\alpha_x = \alpha_y$. The value of α_t is independent and must be chosen according to the temporal resolution.

4 Temporal tracking

At this stage, we assume that we will work on ultrasound cartesian images and on edges represented in a cartesian space, whatever the methods used to obtain this information.

Our objective is to perform temporal tracking of a pre-selected anatomical structure by combining different kinds of information. First, we want to obtain an approximate segmentation of the structure by using simple deterministic processing. Secondly we want to use the edges computed directly on the raw data. These will be combined by a regularization process that takes an initial segmentation and deforms it from its initial position to make it better conform to the pre-detected edges. This approach is the idea behind the use of deformable models.

4.1 Estimation of the boundaries of the anatomical structure

To obtain a crude estimation of the boundaries of anatomical structures, we use techniques from mathematical morphology. The model of a cardiac cavity is very simple. This is an ovoid region with low intensity. These regions cannot be obtained by simple thresholding because of the speckle noise. But the fine structures of the speckle may be easily suppressed by the following morphological operations [12]:

- A first order opening eliminates the small bright structures on dark background.
- The dual operation (first order closing) suppresses the small dark structures.

After these operations, a simple thresholding gives an image C where all the cardiac cavities are represented in white. This detection can be refined by the use of higher level information. The specialist points out, using a computer mouse, the chosen cavity on the first image of the sequence. The whole cavity is then obtained by a conditional dilatation which begins at this point.

4.2 Use of a deformable model

The previous operations usually provide an approximately-correct but locally-inaccurate positioning of the structure boundaries. In order to improve this crude segmentation to an accurate determination of the boundaries, we use the deformable models of [3], in the spirit of [8].

The deformable model is initialized in the first image by the crude approximation of the structure boundary. It evolves under the action of image forces, which are counter-balanced by its own internal forces to preserve its regularity. Image forces are computed as the derivative of an attraction potential related to the previously computed spatio-temporal edges. Typically, the potential is inversely proportional to the distance of the nearest edge point.

Deformable models may be used independantly on each frame or iteratively on the sequence: once the model has converged in the first frame, its final position is used as the initial one in the next frame, and the process is repeated.

5 Experimental results for sonar-space filtering

This section gives the results obtained by bilinear interpolation and sonar-space filtering evaluated in terms of the visualization and edge detection capabilities of the methods.

A simple example concerns polar scanning of a thin dark structure (represented by horizontal lines in polar data) in a white background. Fig. 5 presents the polar data: 512 rays of 128 pixels. The dark structure has a width of 3 rays. Fig. 6 presents edges obtained on cartesian images reconstructed by bilinear interpolation and sonar-space filtering.

Because bilinear interpolation does not make use of all available polar data, information is lost which can never be retrieved by further processing such as edge detection or segmentation.

To conclude this section with ultrasound data, one can see in Fig. 7 the reconstructed image using sonar-space filtering with $\alpha=1$, and the same image with the detected edges superimposed. The value of the parameter α is determined experimentally but is the same for all ultrasound images provided by an echographic machine.

6 Experimental results for temporal tracking

We first note that temporal smoothing reduces some local distortions on the deeper edges of the left auricle (compare the bottom right cavity of Fig. 8). Simultaneously, temporal smoothing can cause a problem for the mitral valve (middle thin structure of Fig. 8), which is moving fast with respect to the temporal resolution. The strategy is thus to use temporal smoothing only to study cavities and to apply spatial gradient techniques to study fast moving structures like the valves.

Secondly, we present the use of deformable models to analyse echocardiographic data after the scan-correction process has been applied.

For structures moving slowly (heart cavities), deformable models may be applied iteratively using an initialization process and the results of edge detection. The software of L. Cohen and I. Cohen [3] requires three parameters. The elasticity and rigidity coefficients model the properties of the cavity boundary curve. The third coefficient is a weight representing the attraction of the edges. The results of this application of the deformable model may be seen in the first frame of Fig. 12 for the segmentation of left auricle. The

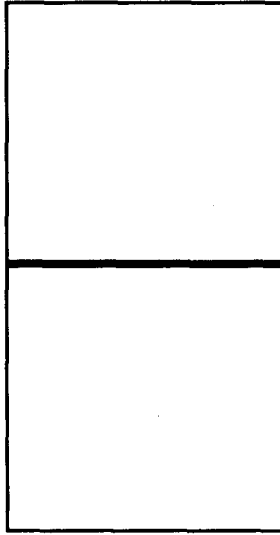


Fig. 5. Polar data

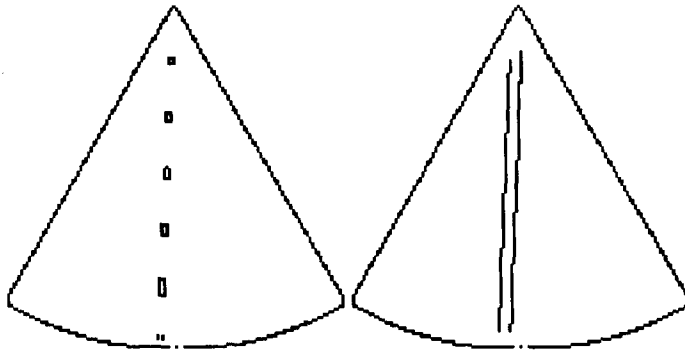


Fig. 6. Left: edges on bilinear interpolation image. Right: sonar-space filtering

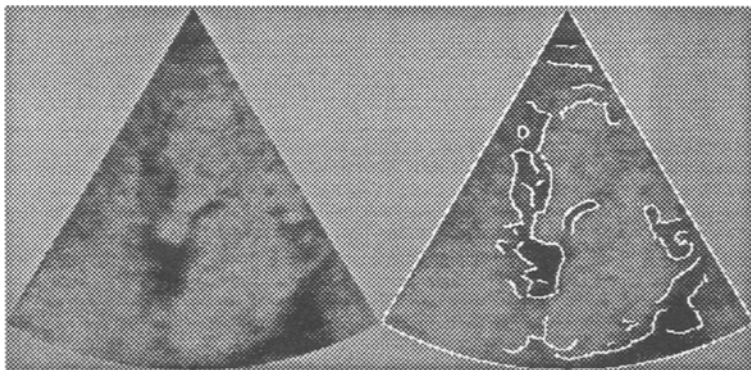


Fig. 7. Cartesian image and best edges obtained by sonar-space filtering

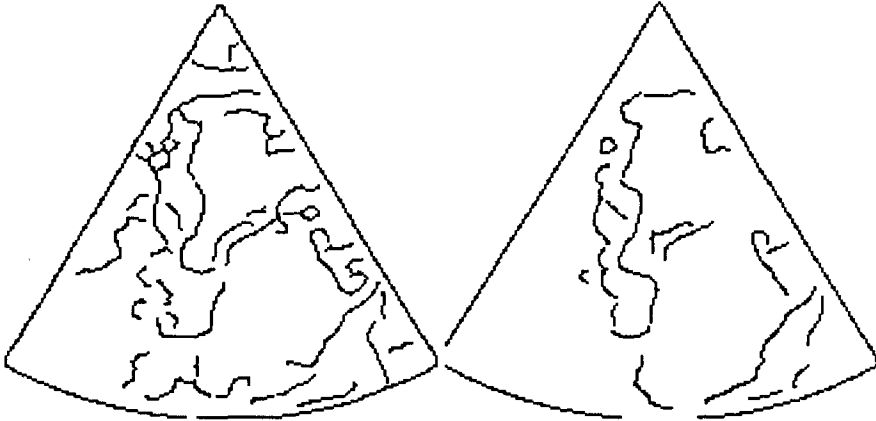


Fig. 8. Left: no temporal smoothing of edges, right: temporal smoothing of edges.

result is then used to initialize the second frame. The parameters are the same and the process is repeated sequentially through all frames, as it can be seen on Fig. 12.

For structures moving fast (mitral valve), deformable models are applied independently on each frame and results of these applications may be seen on Fig. 11.

We summarize the advantages of using deformable models to analyze echocardiographic data:

- Deformable models allow a compromise between an initial segmentation based on grey levels and texture properties and an edge detection process performed directly on raw data.
- The values of the parameters required by the deformable model are the same for both the regularization application on a single frame and for the tracking application on a sequence. They can be chosen interactively on the first image of the sequence.

The methods presented in this paper were applied to four different sequences obtained from two different echographs. The data presented here were obtained in a polar coordinate form on a VIGMED echograph at Henri Mondor hospital in Creteil, France. A sequence contains 38 images from a cardiac cycle. Fig. 9 shows a cartesian representation of the original data. (Only one image in four is displayed.) The left heart cavities (auricle and ventricle) and the mitral valve are visible in a typical image. Our aim is to track them. Tracking of the latter structure is successfully achieved in this example due to the fact that the edges were obtained from sonar-space filtering. Other methods (bilinear interpolation followed by edge detection) generally do not give accurate edges for the deep structures and cannot therefore be used for further temporal tracking. Edges are shown in Fig. 10, and temporal tracking is presented in Figs. 11 and 12.

7 Conclusions

We showed in this paper the importance of using an appropriate conversion method when dealing with images produced in polar coordinates. We introduced a new method which computes both the conversion and a convolution of the polar data with a smoothing filter in a single process. Using this approach, the quality of the edges and features that

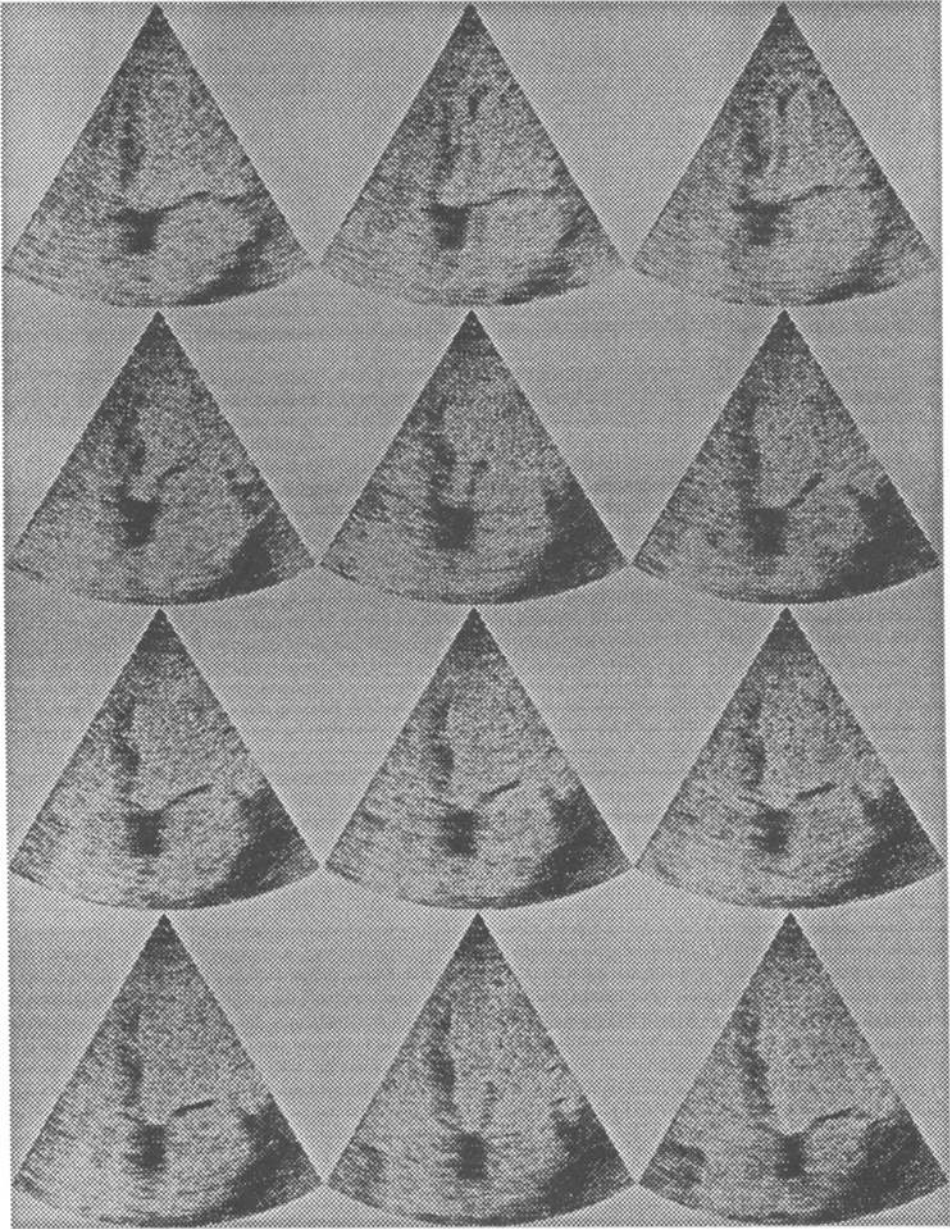


Fig. 9. data after scan-correction

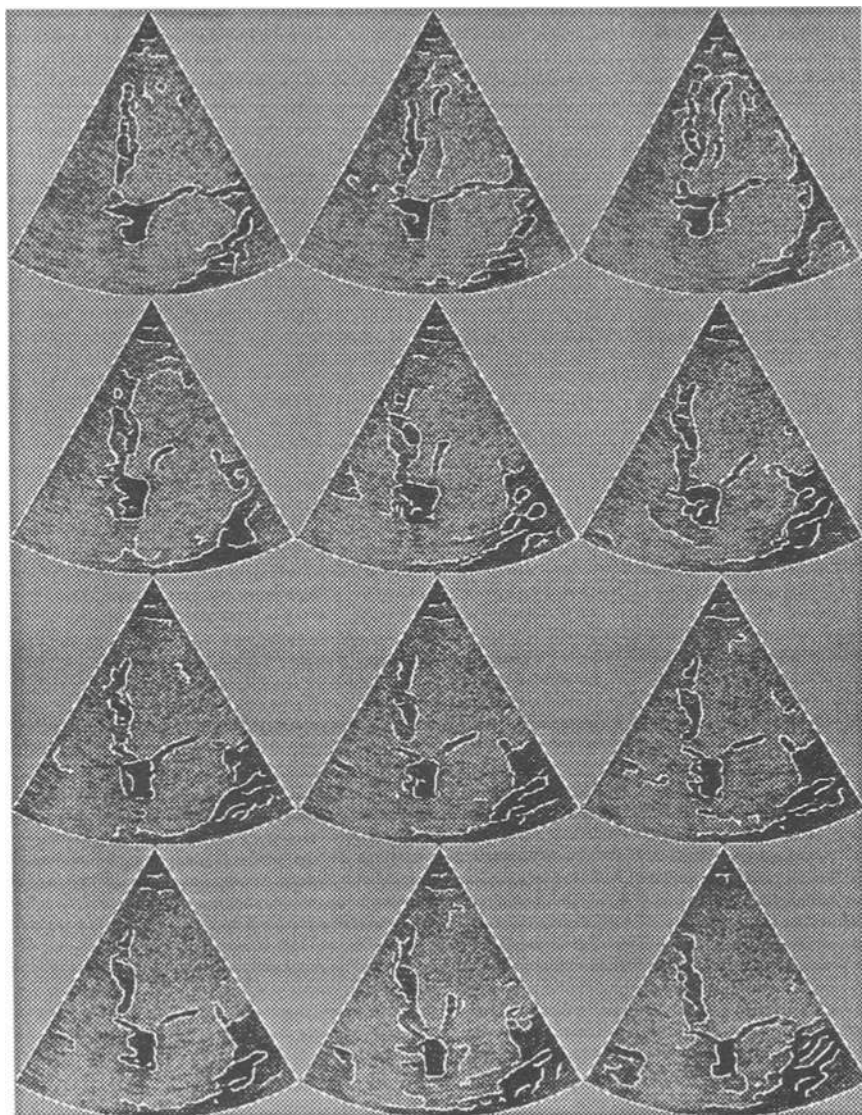


Fig. 10. Edges

are extracted can be enhanced. This approach is more flexible because it allows a variable level of smoothing to be chosen according to the actual resolution of the original data. This is not the case when an additional smoothing is required after a conversion by other algorithms. We showed the enhancement produced on edge detection by our approach.

Finally, we demonstrated the effectiveness of this approach by solving a complete application. We used morphological operators to initialize a deformable model in the first image of a time sequence. Then we applied our edge detector and we let the deformable

model converge toward the detected edges. Using the solution as an initialization in the following image, we tracked the left auricle boundary in a sequence of 38 images.

Our future research will concentrate on the generalization of these methods to be applied to 3-D ultrasound images produced in spherical coordinates.

8 Acknowledgements

We gratefully acknowledge Gabriel Pelle (INSERM, CHU Henri MONDOR, FRANCE) for providing the data and for helpful discussions, and Robert Hummel for a significant improvement of the final manuscript.

This work was partially supported by MATRA Espace and Digital Equipment Corporation.

References

1. N. Ayache, J.D. Boissonnat, L. Cohen, B. Geiger, J. Levy-Vehel, O. Monga, and P. Sander. Steps toward the automatic interpretation of 3-D images. In H. Fuchs K. Hohne and S. Pizer, editors, *3D Imaging in Medicine*, pages 107-120. NATO ASI Series, Springer-Verlag, 1990.
2. A.J. Buda, E.J. Delp, J.M. Meyer, J.M. Jenkins, D.N. Smith, F.L. Bookstein, and B. Pitt. Automatic computer processing of digital 2-dimensional echocardiograms. In *Amer. J. Cardiol.*, volume 51, pages 383-389, 1983.
3. L.D. Cohen and I. Cohen. A finite element method applied to new active contour models and 3D reconstruction from cross sections. In *Proceedings of the International Conference on Computer Vision*, Osaka, Japan, December 1990.
4. R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1 (2), May 1987.
5. F. Faure, J.P. Gambotto, G. Montserrat, and F. Patat. Space medical facility study. Technical report, ESA, 1988. final report, 6961/86/NL/PB.
6. T. Hwang and J.J. Clark. A spatio-temporal generalization of Canny's edge detector. In *10th International Conference on Pattern Recognition*, Atlantic City, New Jersey, USA, June 1990.
7. J.M. Jenkins, O. Qian, M. Besozzi, E.J. Delp, and A.J. Buda. Computer processing of echocardiographic images for automated edge detection of left ventricular boundaries. In *Computers in Cardiology*, volume 8, 1981.
8. Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. In *Proceedings of the First International Conference on Computer Vision*, pages 259-268, London, June 1987.
9. A. Macovski. *Medical Imaging Systems*. Prentice Hall, 1983.
10. O. Monga and R. Deriche. 3D edge detection using recursive filtering: application to scanner images. Technical Report 930, INRIA, November 1988.
11. V.L. Newhouse. *Progress in Medical Imaging*. Springer Verlag, 1988.
12. J. Serra. *Image analysis and mathematical morphology*. Academic Press, 1982. London.
13. T. Taxt, A. Lundervold, and B. Angelsen. Noise reduction and segmentation in time-varying ultrasound images. In *10th International Conference on Pattern Recognition*, Atlantic City, New Jersey, USA, June 1990.
14. M. Unser, L. Dong, G. Pelle, P. Brun, and M. Eden. Restoration on echocardiograms using time warping and periodic averaging on a normalized time scale. In *Medical Imaging*, number III, 1989. January 29 - February 3, Newport Beach.
15. L.F. Zhang and E.A. Geiser. An approach to optimal threshold selection on a sequence of two-dimensional echocardiographic images. In *IEEE Transactions on Biomedical Engineering*, volume BMB 29, August 1982.

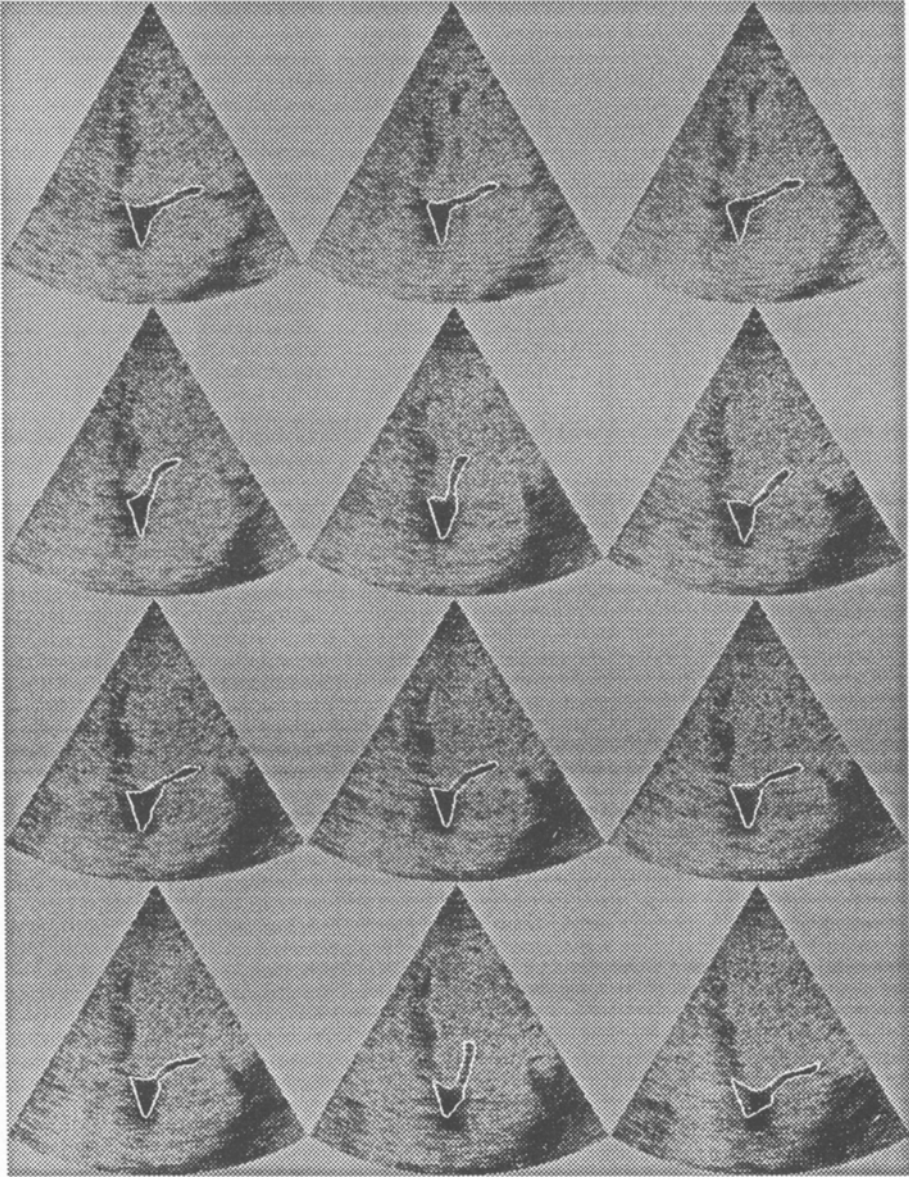


Fig. 11. Temporal tracking of the mitral valve

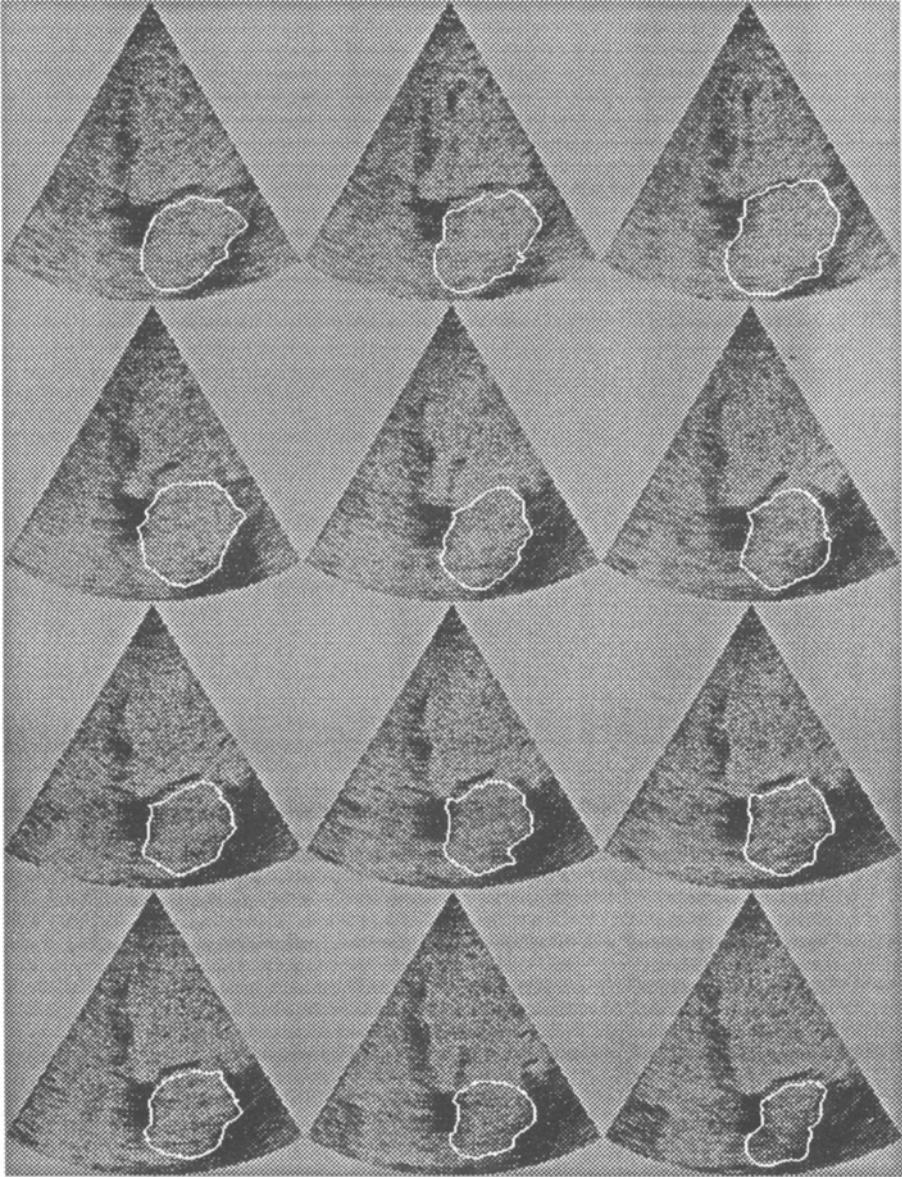


Fig. 12. Temporal tracking of the left auricle