International Journal of Computer Vision 26(1), 25–40 (1998)

Dense Non-Rigid Motion Estimation in Sequences of Medical Images Using Differential Constraints

SERGE BENAYOUN* AND NICHOLAS AYACHE
INRIA, B.P. 93, 06902 Sophia-Antipolis Cedex, France

Received April 3, 1995; Revised October 31, 1995; Accepted April 17, 1996

Abstract. We describe a new method for computing the displacement vector field in time sequences of 2D or 3D images (4D data). The method is energy-minimizing on the space of correspondence functions; the energy is split into two terms, with one term matching differential singularities in the images, and the other constraining the regularity of the field. In order to reduce the computational time of the motion estimation, we use an adaptive image mesh, the resolution of which depends on the value of the gradient intensity. We solve numerically the minimization problem with the finite element method which gives a continuous approximation of the solution. We present experimental results on synthetic data and on medical images and we show how to use these results for analyzing cardiac deformations.

Keywords: medical imaging, adaptive meshes, non-rigid motion, differential geometry, finite element method

1. Introduction

The problem of computing non-rigid motion from digital images is one of the most attractive challenges in computer vision. There are many applications including, for example, model-based image compression (Tang and Huang, 1994) or medical diagnosis (Ayache, 1995). Medical imaging produces sequences of 2D and 3D images which describe the time evolution of anatomic structures. This paper particularly deals with heart images. Cardiovascular diseases represent the "greatest scourge afflicting the population of the industrialized nations" (Braunwald, 1992). Medical imaging now produces a lot of modalities in order to visualize the cardiac cycle, like Magnetic Resonance Imagery (MRI), Nuclear Medicine Imagery (NMI) or Computed Tomography (CT). Because the heart is a nonrigid organ, it is difficult to assess its deformations. This objective, however, seems very important for a better understanding of this organ. For example, some pathologies characterized by ischemic or infarcted tissues are detectable because of an abnormal local motion. Image Processing enables an accurate computation of the motion of the cavities, which could help cardiologists to localize such tissues and/or to evaluate the effects of a medicinal treatment.

Many works have already been done on this subject. The most commonly used method consists of recovering the motion by tracking a deformable model (Terzopoulos et al., 1988; Cohen et al., 1992; Nastar, 1994; Park et al., 1995; Bardinet et al., 1995; McInerney and Terzopoulos, 1995). It needs a two-stage approach. First, one fits contour points with a deformable model. It can be either a free shape model like snakes (Kass et al., 1987; Cohen and Cohen, 1993) or a constrained model like superquadrics (Terzopoulos and Metaxas, 1991; Bardinet et al., 1995). In this latter case it is necessary to add local deformations in order to recover complex shapes. However, free shape models uses a smoothness constraint for controling the regularity of the fit. The second stage consists in tracking the model over the time sequence. It is based on the matching of some particular points based on metric

*Current address: Alcatel Alsthom Recherche, Route de Nozay, 91460 Marcoussis, France.

or geometric similarity measures. Our method mainly differs from these works by the fact we do not use any shape model. Several reasons explain this choice. First it can be very difficult to recover anatomic organ which may have complex shape and topology. Second, the motion information may be not only present in one structure but everywhere. The last reason concerns deformable models which present some specific problems, one of them being that they have to be initialized very close to the solution.

In this paper, we present a method which computes the displacement vector field between two successive frames without prior segmentation. This procedure is applied sequentially to the set of frames. We pursue the original idea introduced in a previous paper (Cohen et al., 1992) of using geometric singularities to anchor the motion field on a few but reliable characteristic points, and to propagate this sparse motion field, first to contours, and then to every image point. The idea of using salient features to guide the components of the nonrigid motion field was also suggested in different studies reported by Amini and Duncan (1992) and Benayoun et al. (1996). The implementation of our method is based on an energy minimization. We first consider a mesh whose resolution depends on the value of the gradient intensity computed on the data (Terzopoulos and Vasilescu, 1991). This adaptive mesh mainly allows to reduce the computational requirements without decreasing the accuracy of the displacement field.

The paper is organized as follows. First we describe the general formulation of our problem (Section 2). Next we present our model of adaptive image meshing (Section 3). The tracking stage is next modeled as an energy minimization problem (Section 4). It is solved by a mathematically rigorous finite element method. Finally we show experimental results on synthetic data, and on time sequences of 2D and 3D medical images (Section 6). We make use of analytical modal analysis (Nastar, 1994) as an instantaneous method for analysis and smoothing of the computed displacement field, and also show such results in this experimental section.

2. Problem Formulation

Our aim is to compute the displacement vector field between two successive frames of a sequence without prior segmentation. This means that we want to determine for every material point in a frame its new position in the following frame. The fundamental hypothesis we apply is that the mean displacement amplitude between successive frames is low. Under this constraint, it becomes possible to track a point just by searching for a *similar* point in the following frame. Two points are said similar if they share some features computed from the gray level data. The most straightforward choice consists of using the image intensity. However, in general, the hypothesis of a perfect conservation of the intensity is not robust. A most robust approach assumes the conservation of contours. Contour points are generally detected by extracting the local maxima of the intensity gradient amplitude in the gradient direction (Monga et al., 1991). They are also characterized by a high intensity gradient amplitude. But matching points is not sufficient neither to recover the tangential component of the motion along the contours, nor to compute a dense motion field.

In order to recover the tangential component of the motion, we look for high curvature points along contours. For example, the points belonging to ridge lines on surfaces correspond to the local maxima of the principal curvature in the associated principal curvature direction. They are anatomically significant in medical imaging and stable under small deformations (Monga and Benayoun, 1995). Thus they represent good landmarks for tracking.

A dense motion field is finally recovered by imposing a regularity constraint everywhere, which provides a propagation of the tangential component of the motion from high curvature points along the rest of the contours, and a propagation of the motion field from contours to the overall image.

To reduce the computational requirements, and because we do not need a very dense solution far from the contours, we introduce an adaptive image mesh. This mesh has a structure similar to the original frame, and keeps its topological and geometrical properties. But the number of nodes is reduced relatively to the original number of pixels or voxels. The density of nodes is increased near the contours where the underlying computed field is meaningful. The problem now reduces itself to searching for every node of the mesh its new position in the following image.

Having defined our constraints, we can now formulate the matching process as a minimization problem. Let f be the correspondence function which matches nodes in one frame with points of the following frame; then we define an energy $E_e(f)$ which measures the quality of the match. Because the problem of

minimizing this energy may have locally several equivalent solutions, we add a second energy $E_i(f)$ which constrains the regularity of the function f. It means that, when it is possible to match a node with several points, the most regular solution is favored. We call $E_e(f)$ the external energy because it depends on image information, and $E_i(f)$ the internal energy because it concerns only internal properties of f.

3. Adaptive Meshes

The use of adaptive meshes is a classical tool in computational fluid dynamics. They have already been used in computer vision, specifically for range data reconstruction (Terzopoulos and Vasilescu, 1991) or surface mesh decimation (Gourdon, 1995). One of the advantages of this technique is to decrease the numerical complexity of a structure while preserving its accuracy. In our case, it will be mainly used to reduce the original frame structure. In particular, for 3D images, it would be too long to compute the displacement of every voxel.

A *mesh* is a set of *nodes* connected by *springs*. Each spring models the connection between two neighboring nodes by the relationship:

$$S_{ij} = k_{ij} d_{ij} X_{ij} \tag{1}$$

The strength vector S_{ij} exerted by the node j on node i is in the direction of the unit vector X_{ij} which connects the two nodes. The intensity is proportional to the elongation of the spring d_{ij}, and also to its stiffness k_{ij}.

There are different techniques for building up an adaptive mesh. One generally makes the distinction between *unstructured* mesh and *structured* mesh. For unstructured meshes, the adjacency graph of the nodes is not invariant. For example, a 3D mesh is initialized at the higher resolution (voxel resolution) and irrelevant nodes are progressively eliminated. This technique is suited for example to data compression because no further computation follows the mesh adaptation. Since it is clearly not our case, we prefer to use structured meshes where connectivity between nodes is preserved. For 2D images, we use a rectangular mesh with a 8-neighborhood system. For 3D images, we use a 18-neighborhood system. At the beginning of the process, the mesh is regular and next the nodes move towards relevant points of the image (see Figs. 1 and 2).

Our aim is to increase the density of nodes near the high intensity gradient points of the image. This is because of our motion computation method which is well defined at these feature points. One way of doing this is to introduce an image dependent force. This force pulls the nodes towards high intensity gradient points. It is derived from a potential field computed over the whole image. In (Terzopoulos and Vasilescu, 1991), the adaptation is due to the spring stiffness which is made dependent on the gray level gradient. We also implemented this technique and results are very similar.

More precisely the adaptation of the mesh results from the solution of the system of dynamic equations. Each equation models the spatial behavior of a node when it is influenced by an external force:

$$m_i \frac{d^2 P_i}{dt^2} + \gamma_i \frac{d P_i}{dt} = F_i, \quad i = 1 \ldots n \tag{2}$$

where n is the number of nodes, $P_i \in \mathbb{R}^3$ is the location of the node i, m_i its mass, γ_i its damping coefficient, and $F_i \in \mathbb{R}^3$ the sum of the applied forces. Notice that we impose an infinite mass for nodes which are in the

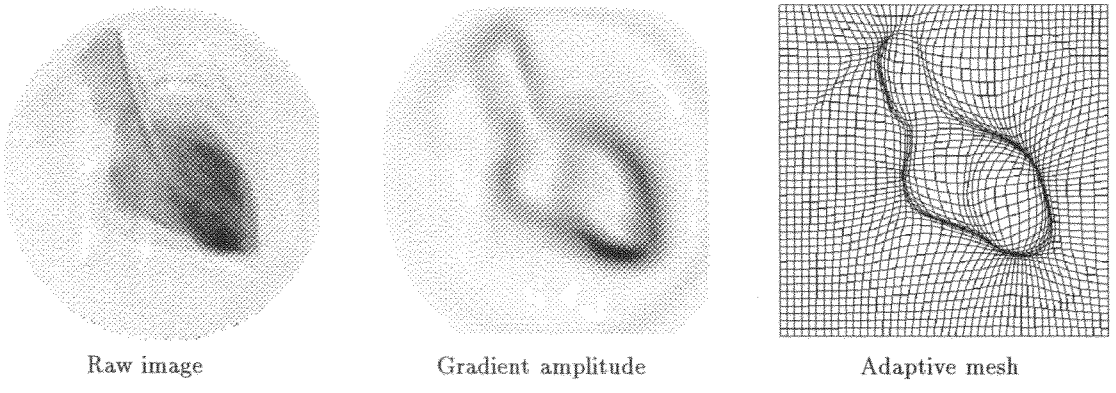

Figure 1. 2D adaptive mesh on the ventriculography data.

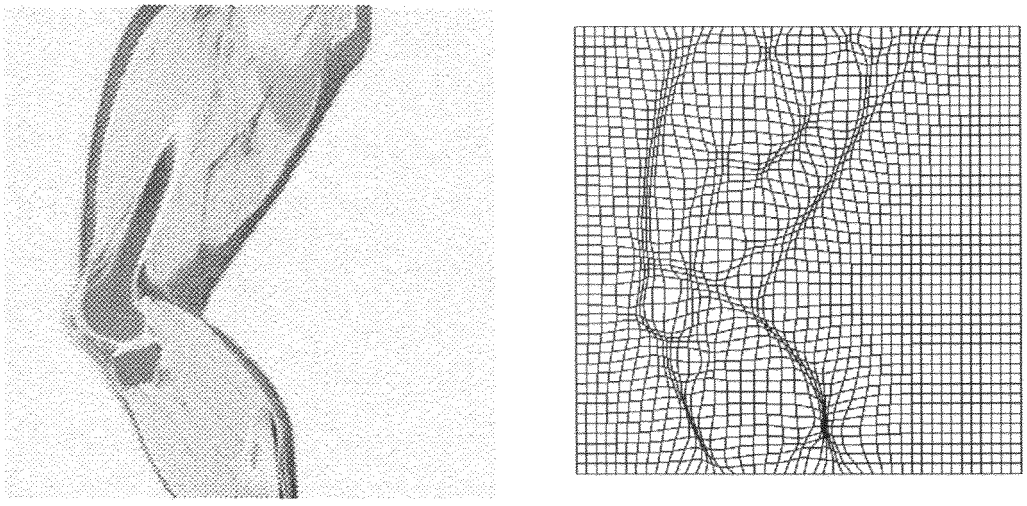

Figure 2. 2D adaptive mesh on the knee data.

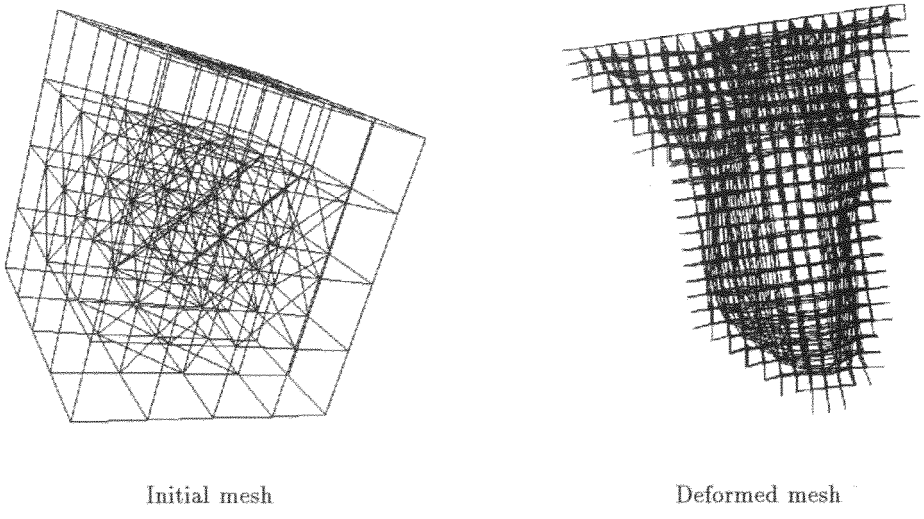

Figure 3. 3D adaptive mesh on the dog heart sequence (we cut non-relevant nodes for clarity).

border of the mesh. Thus they cannot move and we are sure that the mesh covers completely the image.

F_i is composed of two terms. First, a spring term $S_i = \sum_{j \in V_i} S_{ij}$ where V_i is the 18-neighborhood of node i (see Fig. 3). Second, an image term G_i which attracts the node towards high intensity gradient points.

Computation of G_i. Using an edge detector we obtain a binary image with 1 at contour points locations and 0 elsewhere (Monga et al., 1991). Next we apply a thinning operation to the edges so that they are all a single pixel wide. Finally, we smooth this binary image and we compute G_i by differentiation of the smoothed image. Because we use a convolution filter for differentiating (see Appendix A.1), the direction of the force is a combination of many directions but with a predominance of the direction(s) towards closest contour point(s). A similar result would have been obtained by directly differentiating the grey level function. However, it has the disadvantage that the amplitude of the force may vary along a contour because of its dependence on the intensity gradient.

We use a finite difference scheme to solve the system of Eqs. (2). The algorithm is the same explicit Euler method used in (Terzopoulos and Vasilescu, 1991). In our experiments, we choose to consider a node each four or five voxels. The mesh deforms through application of the image forces and converges to an equilibrium state.

4. Energy Minimization

Consider two successive images in a time sequence. We denote by Ω_1 (resp. Ω_2) the domain of the first (resp. second) frame. The correspondence function f is a vector function which associates each 3D point $(x_1, x_2, x_3) \in \Omega_1$ with another 3D point $(X_1, X_2, X_3) = f(x_1, x_2, x_3) \in \Omega_2$.

4.1. External Energy

The correspondences are achieved by minimizing an *external energy* E_e which preserves differential singularities computed on the images. This energy is split into two terms, a *curvature dependent energy* E_c and a *gradient dependent energy* E_g:

$$E_c(f) = \int_{\Omega_1} (C_1(x) - C_2(f(x)))^2 R_c(x)\, dx = \int_{\Omega_1} e_c^2(x) R_c(x)\, dx$$

where $x = (x_1, x_2, x_3)$ and C_1 (resp. C_2) is a curvature dependent function computed in the first (resp. second) image and R_c is a curvature dependent weighting function. Therefore, e_c is an energy which will measure the discrepancy between contour curvatures at points where R_c is different from zero.

$$E_g(f) = \int_{\Omega_1} (N_1(x) - N_2(f(x)))^2 R_g(x)\, dx = \int_{\Omega_1} e_g^2(x) R_g(x)\, dx$$

where N_1 (resp. N_2) is the gradient amplitude dependent function computed in the first (resp. second) image, and R_g is a gradient dependent weighting function. Therefore, e_g is an energy which will measure the discrepancy between the gradient amplitudes along contour points when R_g is different from zero.

The mathematical formula of the curvature dependent functions are given in the Appendixes A.3 and A.4. Concerning the gradient dependent functions, the formula can be found in (Monga et al., 1991).

4.2. Internal Energy

The problem of minimizing this energy is an ill-posed problem because there are many possible solutions. So, we add an *internal energy* E_i which constrains the regularity of the function:

$$E_i(f) = \int_{\Omega_1} \|\nabla f(x)\|^2\, dx \tag{3}$$

This also allows the displacement field to be propagated in image areas where the previous energies are not relevant. Finally we solve the minimization problem:

$$\mathrm{Min}_f\, E_c(f) + E_g(f) + E_i(f) \tag{4}$$

4.3. Weighting Functions

The internal and external energies are balanced through the two *weighting functions* R_g and R_c. R_g (resp. R_c) is a normalized increasing function of the gradient amplitude (resp. curvature):

$$R_g(x) = h(N_1(x))$$
$$R_c(x) = h(C_1(x))$$

where h acts like a high pass filter. Up to now we used the following form for h:

$$h(v) = -2\left(\frac{v - v_m}{v_M - v_m}\right)^3 + 3\left(\frac{v - v_m}{v_M - v_m}\right)^2, \tag{5}$$

where v_m and v_M are the minimal and maximal values. This choice is due to the following considerations. Weighting functions must locally enforce the external gradient and curvature energies at points where gradient and curvature values are significant. Conversely, they inhibit these external energies when a point has no significant gradient or curvature features (see Fig. 4). In that case, only the internal energy constrains the solution.

In fact, we have chosen to convert R_g into R_g':

$$R_g'(x) = (1 - R_c(x)) R_g(x)$$

in order to make the curvature energy dominate the gradient energy. This combination is justified for high

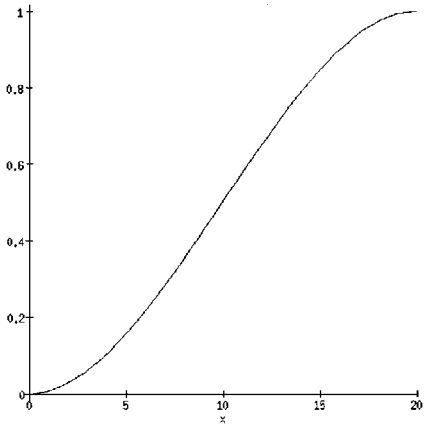

Figure 4. High pass filter ($v_{\min} = 0$, $v_{\max} = 20$).

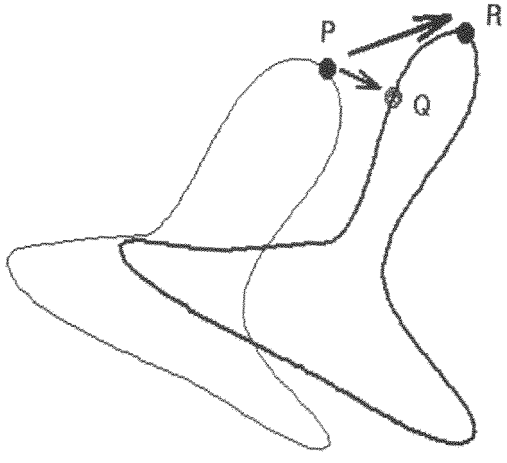

Figure 5. This figure shows the evolution of a 2D curve at two successive times. The high curvature point P is not attracted by the closest contour point Q but, thank to the weighted energies defined in the text, is attracted by the closest high curvature point R.

curvature points because it favors the attraction by another high curvature point. Its effect is illustrated in Fig. 5.

5. Partial Differential Equation

A necessary condition for the optimality of $f = (u, v, w)$ is the Euler-Lagrange equation $\nabla E(f) = 0$ which is equivalent to the partial differential equations:

$$\begin{cases} -\Delta u = R_g' \frac{\partial N_2}{\partial x}(u, v, w)e_g + R_c \frac{\partial C_2}{\partial x}(u, v, w)e_c \\ -\Delta v = R_g' \frac{\partial N_2}{\partial y}(u, v, w)e_g + R_c \frac{\partial C_2}{\partial y}(u, v, w)e_c \\ -\Delta w = R_g' \frac{\partial N_2}{\partial z}(u, v, w)e_g + R_c \frac{\partial C_2}{\partial z}(u, v, w)e_c \end{cases}$$

with boundary conditions $f|_{\partial\Omega_1} = f_b$. Consider the first equation; as its second member also depends on the solution, we study the associated evolution problem:

$$\frac{\partial u}{\partial t} - \Delta u = R_g' \frac{\partial N_2}{\partial x}(D)e_g + R_c \frac{\partial C_2}{\partial x}(D)e_c \qquad (6)$$

with

$$\begin{cases} D = (u, v, w) \\ (u^0, v^0, w^0) = (u, v, w)_{t=0} \\ f|_{\partial\Omega_1} = f_b \end{cases}$$

A solution of the original equation will correspond to a stationary solution of the evolution problem. Moreover, the introduction of a temporal variable allows us to reconsider the using of weighting functions. It becomes possible in particular to reduce little by little the influence of the curvature constraint. One way to do this is to make R_c dependent on time:

$$R_c'(x, t) = R_c(x) \frac{1}{1 + \alpha t}$$

where α controls the variation speed.

Indeed for high local deformations, the curvature is not invariant although it remains useful to guide the matching of singular points. At the end of the convergence process (6), the elimination of the curvature constraint ensures that all contour points are well matched (see Fig. 6). The numerical solution of the Eq. (6) can be found in (Ciarlet, 1987).

6. Experimental Results

We present experiments on sequences of 2D and 3D medical image.

6.1. A Synthetic Example

We first illustrate our algorithm by results on synthetic data. We build a 3D image representing a cube. The size of the image is $64 \times 64 \times 64$. By application of an affine transformation, we obtain a second image (the mean displacement issued from this transformation is almost 3.5 voxels). We next consider a $15 \times 15 \times 15$ mesh. After the adaptation stage, the nodes are concentrated near the high gradient points of the cube image. Next, we compute correspondences between the nodes and the points in the deformed cube image. In order to check the accuracy of the tracking, we then compute

Table 1. Synthetic example: Statistics on the distance between the displacement obtained with our algorithm and the exact displacement given by the affine transformation.

Distance (Voxel)	All nodes (1331 Nodes)	High gradient nodes (427 Nodes)	High curvature nodes (147 Nodes)
Minimum	0.0321	0.0321	0.0619
Maximum	1.4400	1.4400	0.8510
Mean	0.4720	0.4190	0.4380
Standard deviation	0.2380	0.1780	0.1540

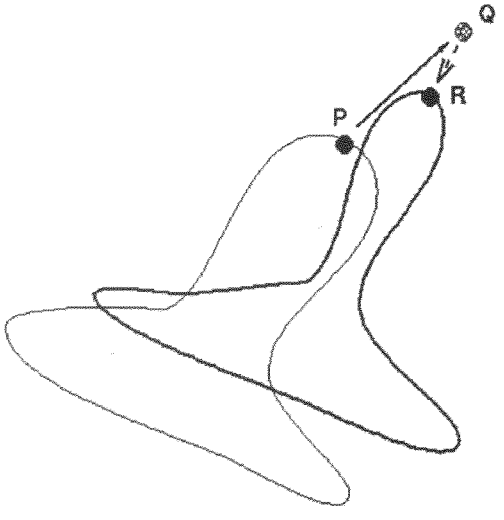

Figure 6. This figure shows the evolution of a 2D curve at two successive times. The curvature extremum point P is first "attracted" by a similar curvature point Q. Next, by relaxing the curvature constraint it matches the closest contour point R.

for each node the distance between the displacement obtained with our algorithm and the exact displacement given by the affine transformation. The results are presented on Table 1. The mean distance is about one half voxel. Notice that the accuracy is better for high curvature nodes, for which the standard deviation of the distance is the lowest.

6.2. *An Articulated Structure: The Knee*

We applied our algorithm to a knee sequence. Sagittal cross-sections have been obtained with a MRI device.[1] As the quality of these images is good, contour points may be well detected. However, the mean amplitude of the displacement between two successive images is quite high for a local approach (about ten pixels). The size of the images is 256×256 and we consider a mesh with 46×46 nodes. It means that the computational time of the motion field estimation stage is approximately reduced by a factor 25.

Figure 7 shows two successive images on this sequence, the mesh adapted to the first image and the motion field computed between them.

We used mapping to estimate the accuracy of our motion fields. We will notice $i \Leftarrow j$ the mapping after resampling of the jth image in a sequence onto the ith image. This mapping is based on solving for the mapping f using our proposed method. The mapping algorithm is as follows:

- Consider each pixel p on the ith image
- By application of the successive motion fields obtain a pixel q on the jth image (since the motion is only computed on nodes, the displacement on another location is obtained by linear interpolation)
- Set the gray level on q to p

By this way we obtain a new image which has to be very similar to the original one. The accuracy of the mapping may be checked by computing the difference between these two images. However, it would suppose that the gray level is invariant, which is not true. So we prefer to overlay into the resulting image the contours found from the original image.

Figure 8 thus presents the mapping of image 11 onto image 8. This mapping is issued from the several motion fields computed between these two moments. Contours of the original image 8 are overlayed on it.

6.3. *A Deformable Structure: The Heart*

During the beating of the heart we observe three phenomena: the blood fluid dynamic, the cardiac wall deformations and the interaction between the blood and the cavities. Our approach is dedicated to the numerical estimation of the wall deformations because we use at once a gradient information and a curvature information which are pertinent only near contour points.

2D Case. The 2D sequence of the heart motion was acquired by ventriculography with catheterization. This technique consists in introducing an X-ray contrast agent inside the left ventricle. The contrast agent allows to evaluate the motion of the left ventricle wall. Figure 1 shows one image produced by this device. The size of the images is 256×256. The cardiac cycle

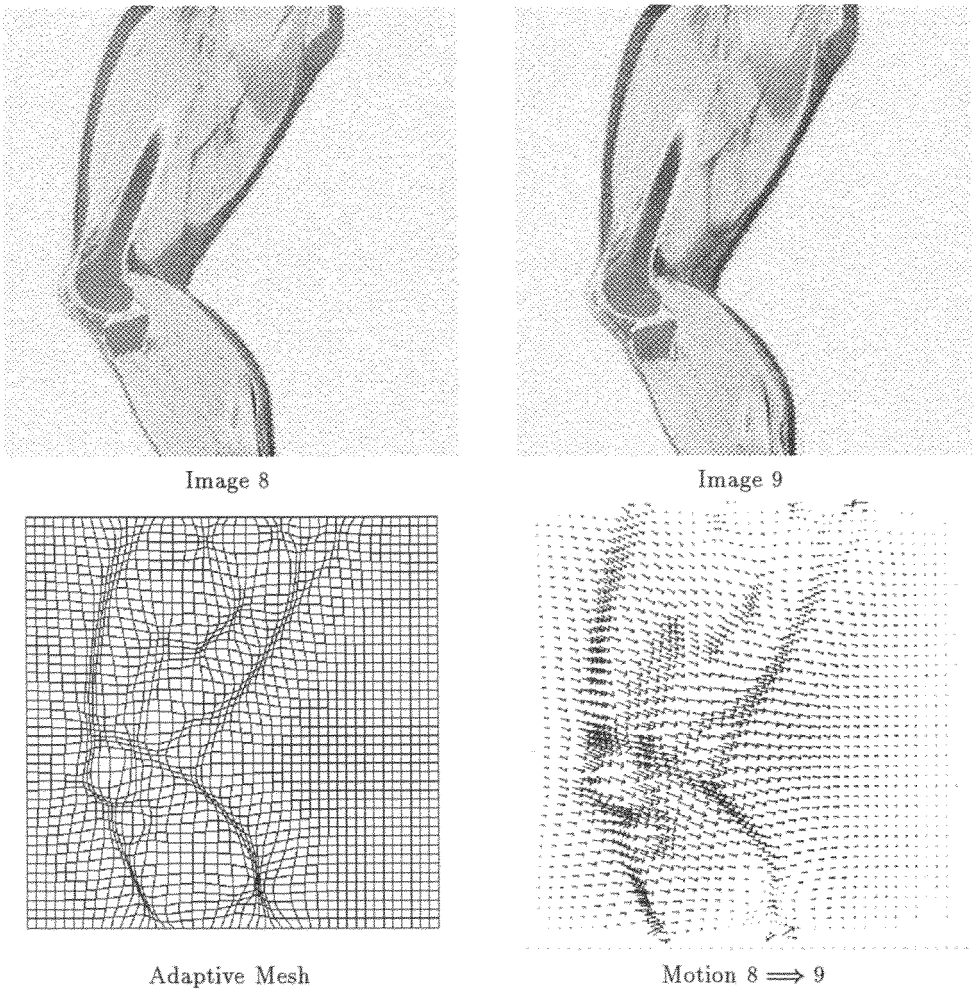

Figure 7. The top images are the gray level data of images 8 and 9 in the knee sequence. On bottom left is the mesh adapted to image 8. On bottom right is the motion field from image 8 to image 9 computed on each node of the adaptive mesh.

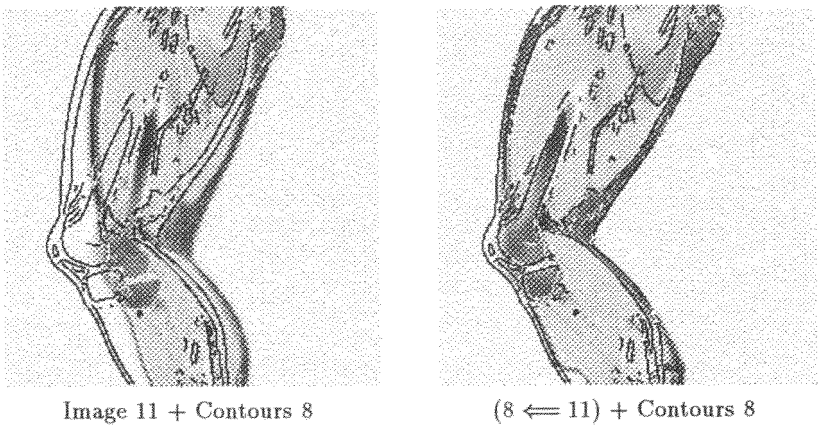

Figure 8. The image on the left is the gray level data of image 11 (bottom). The contours derived from image 8 are overlayed on it to show the motion amplitude. The image on the right represents the mapping after resampling of image 11 onto image 8. The contours found from the original image 8 are again overlayed as a check for accuracy in the mapping estimation.

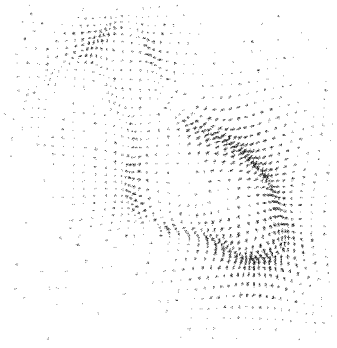

Figure 9. Motion field from image 5 to image 6 computed on each node of our adaptive mesh.

is sampled with 38 intermediate phases, which means that the time resolution is better than for the knee sequence. Thus, the maximal displacement between two successive frames is about six pixels. For our processing we use 46×46 meshes. Figure 1 also shows our mesh adapted to the contours of the image.

Figure 9 shows two motion fields computed on these data. One is the motion field computed between two successive frames. Mapping results are next presented in Fig. 10. We present the mapping of images 3, 6 and 9 onto image 2. They are presented twice without and then with contours of the original image 2.

We make use of analytical modal analysis as an instantaneous method for analysis and smoothing of the computed displacement field. This technique consists to decompose our output (displacement vectors field) in the analytical basis of the free vibration modes (for more details see (Nastar and Ayache, 1994)). The use of modal analysis is interesting first to characterize a motion field with a few number of parameters, next to quantitatively compare several motion fields.

The Fig. 11 shows an example of modal analysis. The original motion field corresponds to the overall contraction of the heart. We show the first two modes and the motion obtained by addition of the fifty more representative modes. With only 50 modes (out of 2116) the mean error between the original field and the approximated field is about one pixel. This means that the motion information may be compressed with a factor 44 without loss of accuracy.

3D case. We make use of a set of 4D (3D images plus time) data of a beating canine heart (see an alternative use of the same data in (Nastar and Ayache, 1994)). The images, obtained by fast cine-CT, come from the Dynamic Spatial Reconstructor (DSR) machine. It provides real-time data of the overall cardiac cycle without using synchronization, which is not the case until now for MRI. The size of the 3D images is $98 \times 100 \times 110$. We did experiments on the systole (heart contraction) which is divided into seven moments. We use 3D adaptive meshes of size $25 \times 25 \times 25$. Concerning the computational time, it takes about one minute to compute a motion field between two successive images with a workstation Dec Ultrix. This time only refers to the resolution of the Euler-Lagrange equation, and not to the preliminar computations, e.g., differential features and mesh adaptation. Indeed, since (i) our algorithm is $O(N)$ with $N =$ size of the stiffness matrix and (ii) the size of the stiffness matrix is reduced in that case by a factor 60, we can estimate that it would take more than one hour to compute the motion on each voxel with the same technique. Others techniques like octree-splines could certainly be used to reduce further the computing time (Szeliski and Lavallée, 1994).

First we show in Figs. 12 and 13 our motion field corresponding to the overall heart contraction. For display purposes, these views are presented with two surfaces of iso-intensity, one is the endocardium (internal wall) at the beginning-of-systole, the other is the endocardium at the end-of-systole. For a description of the algorithm used to extract the iso-intensity surfaces, one can refer to (Thirion and Gourdon, 1992).

For checking the accuracy of this motion field, we again present some image mapping results. We first show the endocardium contours at the beginning-of-systole overlayed on the end-of-systole image. It helps to evaluate the amplitude of the deformation during the systole. Next we show the mapping of the end-of-systole image into the beginning-of-systole image. We also overlayed on it the original endocardium contours at the beginning-of-systole. Since these are 3D images, we only show a few slices (Fig. 14).

We finally present some results concerning the modal analysis of our motion fields. In that case, it is too long to compute all the modes and next to classify them with respect to their relevance. However, it is possible to approximate the original field just by considering the lowest frequency modes. Table 2 presents the mean error (in voxel) between the original field and a few modal approximations. Thus, with only 343 modes (out of 15625), the mean error is already less than one half voxel.

The last result illustrates one potential capability of modal analysis. Because this technique allows to

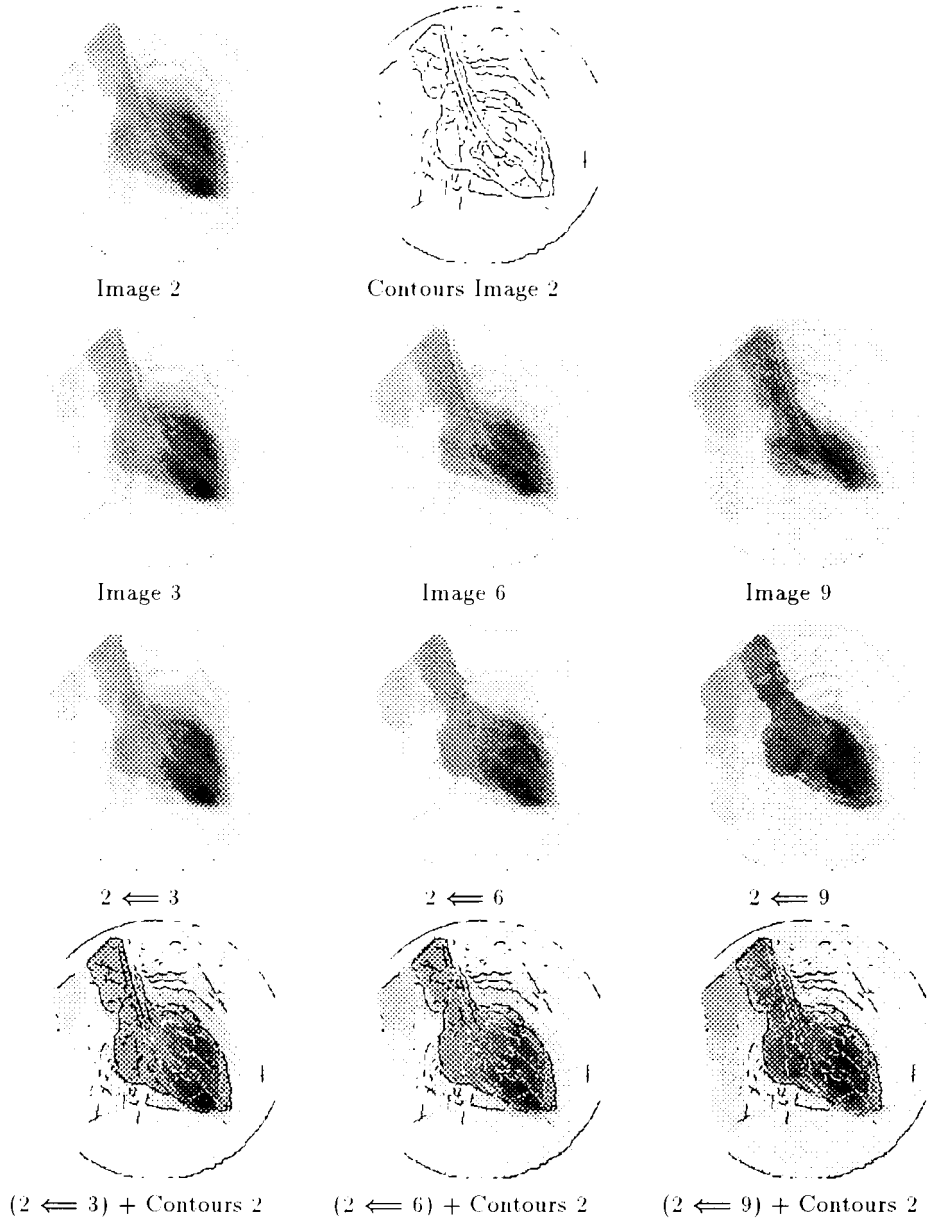

Figure 10. Third line: mapping after resampling of images 3, 6 and 9 onto image 2. Fourth line: same results but contours of the original image 2 are overlayed.

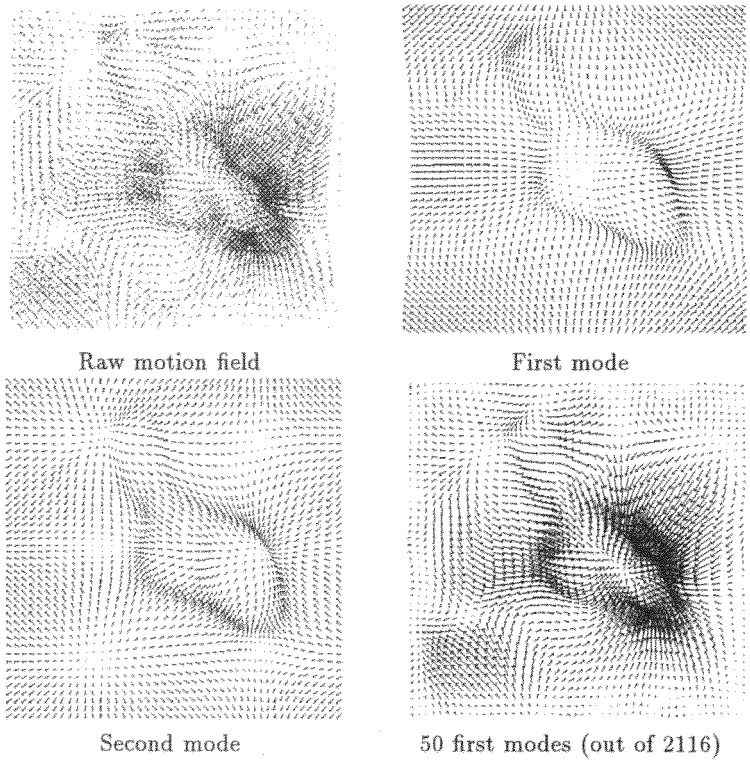

Figure 11. Modal approximation on 2D heart data. From up to bottom and left to right: the original motion field corresponding to the heart contraction, the two more representative modes and the field approximated by addition of the fifty more representative modes.

Table 2. Mean Distance Error between original motion field and some modal approximations by the lowest frequency modes (notice that all number of nodes are power of 3 because we consider all combinations of lowest frequency modes in each direction).

Lowest frequency modes	Mean approximation error (Voxel)
27	0.85
64	0.75
196	0.50
343	0.48
1000	0.38
15625	0.00

describe a field with a few number of parameters, it is possible to compare efficiently different kinds of deformation. In Fig. 15, we show three spectra which represent the modes representativity as a function of the modes frequency, for each direction. The mode representativity simply is obtained by computing the scalar product between the mode and the original field. The spectra of two different deformations are printed in the same graphic.

7. Conclusion

We have described a method for computing dense non-rigid motion fields in time sequences of 2D or 3D images. It consists of minimizing an energy which takes into account differential similarities for contour points like gray level gradient and curvature and imposes a regularity constraint for non-salient points. In order to reduce the numerical complexity, we have introduced adaptive meshes. Since the number of nodes is less than the original number of pixels (or voxels), the computational time is drastically reduced. However, because the nodes are concentrated in the relevant areas of the image, the result accuracy is preserved. Some preliminary results concerning a quantitative analysis of the deformations are performed via a straightforward analytical modal analysis (Nastar and Ayache, 1994; Benayoun et al., 1995).

Figure 12. Profile view of the heart systole. The endocardium iso-intensity surface is grey at the beginning-of-systole and red at the end-of-systole.

Directions for further work include improving our mesh model. We intend to use unstructured meshes (Vasilescu and Terzopoulos, 1992). This technique allows to reduce strongly the number of nodes and also to speed up the computation of motion. We also consider introducing other types of regularity constraints, for example in order to preserve some "natural" discontinuities of the motion, like those which occur at the boundary between healthy and ill tissues.

However, we think the most promising direction concerns the introduction of physical information about the motion. It can be very helpful in order to improve the accuracy of the results, in particular concerning the motion paths of individual points. Up to now we have considered only geometric features to recover the motion. By physical information, we mean for example the data provided by Tagged MRI (Gutman et al., 1994; Azhari et al., 1995; Park et al., 1995). This technique directly gives the true displacement of few 3D points during the heart systole. Because our method is really volumetric and tags are located everywhere, including such information would be quite straightforward. It would help to better recover the non-tangential part of the displacement, like the twisting component in the heart contraction.

Appendix

A.1. Partial Derivatives on Digital Images

Our tracking method is based on the comparison of differential features computed on the images. We already have shown in a previous study how to calculate these

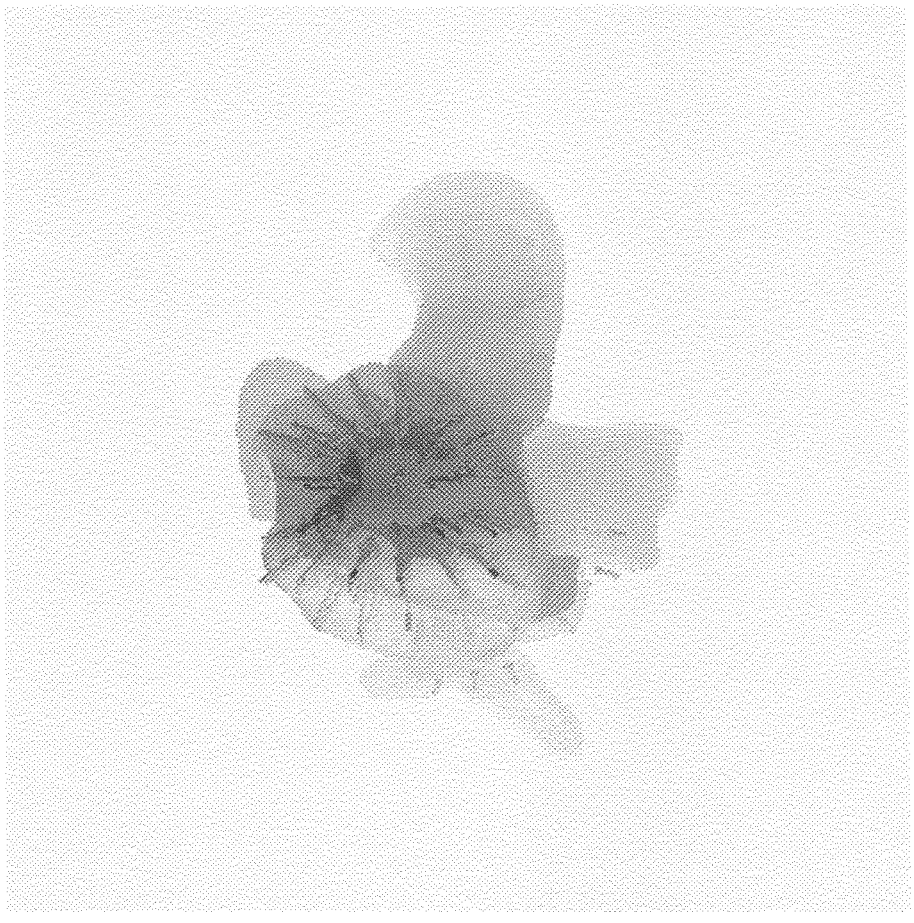

Figure 13. Back view (view from apex) of the heart systole.

features from the partial derivatives of the image function (Monga et al., 1992; Monga and Benayoun, 1995). We just recall here the main results.

Partial derivatives of the images are computed by application of a convolution filter. We used the Canny-Deriche filter which is optimal for detecting step edge contour model (Monga et al., 1991). For 1D input data, we recall here the smoothing operator corresponding to this filter:

$$F(x) = c_0(1 + \alpha|x|)e^{-\alpha|x|}$$

where α controls the width of the filter and c_0 is a normalization coefficient depending on α.

This filter has been extended to provide first, second and third order partial derivatives of the 3D images (Malandain, 1992).

A.2. Binary or Floating Features

For our experiments we used two categories of energies. Concerning the gradient dependent energy, we have considered two kinds of features:

- the gray level gradient norm
- the "contourness" criterion which is defined as the locus where the gradient norm is maximal in the gradient direction

The choice of feature may give appreciably different motion fields. For example, a contour point may have low gradient norm and consequently will be better tracked with the contourness criterion. For the curvature dependent energy we also have tested two features:

- the principal(s) curvature(s)

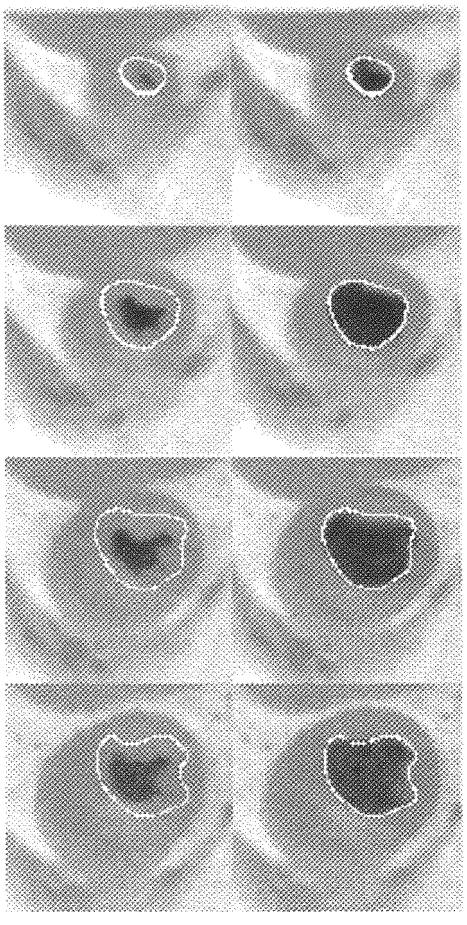

Figure 14. On left, a few slices of the original end-of-systole image. Endocardium contours at the beginning of systole are overlayed on them. On right, the corresponding slices of the mapping of the end-of-systole image into the beginning-of-systole image. Endocardium contours at the beginning of systole again are overlayed as a check of accuracy for the mapping. The mapping preserves the basic shape of the endocardium.

- the "cornerness" criterion, like the curvature extremality.

We now recall the mathematical formula which give these features. We focus on curvature(s), formula concerning contourness are given in (Monga et al., 1991).

A.3. 2D Features

We suppose that a curve C is locally defined as an iso-intensity line. Consider $C(s) = (x(s), y(s))$ a normal parameterization of this curve:

$$I(C(s)) = I_0, \quad \forall s$$

where I_0 is the iso-value. If we differentiate twice this expression with respect to s, we obtain an expression of the curvature K of the curve depending only on the partial derivatives of the image function I:

$$K(C(s)) = \frac{-\boldsymbol{t}^T \boldsymbol{H} \boldsymbol{t}}{\|\boldsymbol{g}\|} \tag{A1}$$

where $\boldsymbol{t}$ is the unit curve tangential vector, $\boldsymbol{g}$ is the curve normal vector and H is the Hessian matrix of I.

By differentiating again this expression with respect to s, we have an expression of the curvature extremality E (the curvature extremums corresponding to the loci where this expression cancels):

$$E(C(s)) = 3\,(\boldsymbol{t}^T \boldsymbol{H} \boldsymbol{t})(\boldsymbol{g}^T \boldsymbol{H} \boldsymbol{t}) - \|\boldsymbol{g}\|^2\, \boldsymbol{t}^T \partial H(\boldsymbol{t})$$

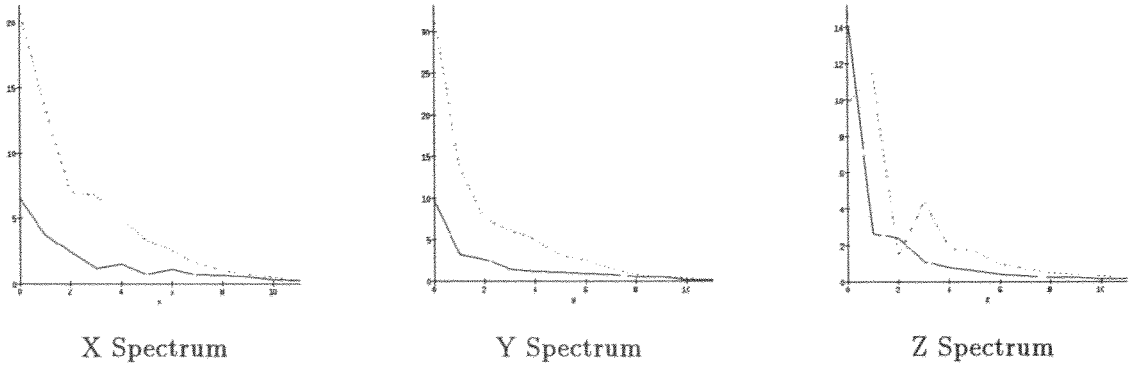

Figure 15. Modal Directional spectra: Modes representativity versus modes frequency in each direction. Two deformations are presented : a low deformation from beginning-of-systole to intermediate step during the systole (black), a high deformation corresponding to the overall systole (grey).

where ∂H is the vector operator:

$$\partial H(\boldsymbol{t}) = \begin{pmatrix} \boldsymbol{t}^T \dfrac{\partial \boldsymbol{H}}{\partial x} \boldsymbol{t} \\ \boldsymbol{t}^T \dfrac{\partial \boldsymbol{H}}{\partial y} \boldsymbol{t} \end{pmatrix}$$

A.4. 3D Features

We consider here an iso-surface $S(u, v) = (x(u, v), y(u, v), z(u, v))$:

$$I(S(u, v)) = I_0$$

where I_0 is the iso-value. On a point $P \in S$, we consider an orthonormal basis $(\boldsymbol{h}, \boldsymbol{f})$ of the tangent plane T_P. By a similar computation than in the 2D case, we obtain the expression of the normal curvature K_d where $\boldsymbol{d} \in T_P$:

$$K_d = \frac{-\boldsymbol{d}^T \boldsymbol{H} \boldsymbol{d}}{\|\boldsymbol{g}\|}$$

The principal curvatures $(K_i)_{i=1,2}$ correspond to the highest and to the lowest value of K_d:

$$K_i = \frac{\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{h} + \boldsymbol{f}^T \boldsymbol{H} \boldsymbol{f}}{2\,\|\boldsymbol{g}\|} \pm \frac{\sqrt{(\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{h} - \boldsymbol{f}^T \boldsymbol{H} \boldsymbol{f})^2 + 4(\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{f})^2}}{2\,\|\boldsymbol{g}\|}$$

We simultaneously obtain the corresponding principal curvature directions:

$$\boldsymbol{e}_i = \boldsymbol{h} + \lambda_i \boldsymbol{f}, \quad i = 1, 2 \qquad \text{(A2)}$$

with

$$\lambda_i = \frac{\|\boldsymbol{g}\| K_i - \boldsymbol{h}^T \boldsymbol{H} \boldsymbol{h}}{\boldsymbol{f}^T \boldsymbol{H} \boldsymbol{h}}$$

This also gives the mean curvature (K_m) and the Gaussian curvature (K_g):

$$K_m = \frac{\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{h} + \boldsymbol{f}^T \boldsymbol{H} \boldsymbol{f}}{2\,\|\boldsymbol{g}\|}$$

$$K_g = \frac{(\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{h})(\boldsymbol{f}^T \boldsymbol{H} \boldsymbol{f}) - (\boldsymbol{h}^T \boldsymbol{H} \boldsymbol{f})^2}{\|\boldsymbol{g}\|^2}.$$

By differentiating a principal curvature in the associated principal curvature direction, we finally obtain an expression of the curvature extremality:

$$E_i = \nabla K_i^T \boldsymbol{e}_i = 3\big(\boldsymbol{e}_i^T \boldsymbol{H} \boldsymbol{e}_i\big)(\boldsymbol{g}^T \boldsymbol{H} \boldsymbol{e}_i) - \|\boldsymbol{g}\|^2 \boldsymbol{e}_i^T\, \partial H(\boldsymbol{e}_i) + 2\alpha_i \|\boldsymbol{g}\| \boldsymbol{e}_j^T \boldsymbol{H} \boldsymbol{e}_i$$

with α_i depending on the cross-extremality $\nabla K_i^T \boldsymbol{e_j}$:

$$\alpha_i = \frac{1}{K_j - K_i} \nabla K_i^T \boldsymbol{e}_j$$

Acknowledgments

Special thanks are due to Chahab Nastar for his contribution concerning modal analysis and to Professor Mike Brady for a substantial help in the writing of this paper. We wish to thank also Isaac Cohen, Alexis Gourdon and Laurent Cohen for helpful discussions. Thanks are also due to Dr. Richard Robb at Biomedical Imaging Resource, Mayo Foundation/Clinic, for providing us the fast-CT heart data. This work was partially supported by Digital Equipment Corporation.

Note

1. Thanks to Dr. Boisgard from the hospital Montpied at Clermont-Ferrand (France).

References

Amini, A.A. and Duncan, J.S. 1992. Bending and stretching models for LV wall motion analysis from curves and surfaces. *Image and Vision Computing*, 10(6):418–430.

Ayache, N. 1995. Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, 13(4):295–313.

Azhari, H., Weiss, J.L., Rogers, W.J., Siu, C.O., and Shapiro, E.P. 1995. A noninvasive comparative study of myocardial strains in ischemic canine hearts using tagged mri in 3-D. *American Journal of Physiology*, 268:1918–1926.

Bajcsy, Ruzena and Kovacic, Stane 1989. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 46:1–21.

Bardinet, E., Cohen, L.D., and Ayache, N. 1995. Superquadrics and free-form deformations: a global model to fit and track 3D medical data. In *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, Nice, France.

Benayoun, S. 1994. Calcul Local du Mouvement, applications à l'imagerie médicale multidimensionnelle. Ph.D. Thesis, Université Paris Dauphine.

Benayoun, S., Ayache, N., and Cohen, I. 1994. Adaptive meshes and nonrigid motion computation. In *International Conference on Pattern Recognition*, Jerusalem, Israel.

Benayoun, S., Nastar, S., and Ayache, N. 1995. Dense non-rigid motion estimation in sequences of 3D images using differential constraints. In *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, Nice, France.

Benayoun, S., Kharitonsky, V., Zilberman, A., and Peleg, S. 1996. Local quantitative measurements for cardiac motion analysis. In *European Conference on Computer Vision*, Cambridge, U.K.

Braunwald, E. 1992. *Heart Disease: A Textbook of Cardiovascular Medicine*, Fourth edition, W.B. Saunders: London, vol. 1.

Ciarlet, P.G. 1987. *The Finite Element Methods for Elliptic Probems*. North-Holland: Amsterdam.

Cohen, Isaac, Ayache, Nicholas, and Sulger, Patrick. 1992. Tracking points on deformable objects using curvature information. In *European Conference on Computer Vision*, Santa Margherita Ligure, Italy, pp. 458–466.

Cohen, Laurent D. and Cohen, Isaac. 1993. Finite element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11).

Gourdon, A. 1995. Simplification of irregular surfaces meshes in 3D medical images. In *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, Nice, France.

Gutman, M.A., Prince, J.L., and McVeigh, E.R. 1994. Tag and contour detection in tagged MR images of the left ventricle. *IEEE Transactions on Medical Imaging*, 13(1).

Kass, Michael, Witkin, Andrew, and Terzopoulos, Demetri. 1987. Snakes: Active contour models. In *IEEE International Conference on Computer Vision*, London, pp. 259–268.

Malandain, G. 1992. Filtrage, topologie et mise en correspondance d'images médicales multidimensionnelles. Ph.D. Thesis, Ecole Centrale de Paris.

McInerney, T. and Terzopoulos, D. 1995. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *Computerized Medical Imaging and Graphics*, 19(1):69–83.

Monga, O., Deriche, R., Malandain, G., and Cocquerez, J.P. 1991. Recursive filtering and edge closing: two primary tools for 3D edge detection. *Image and Vision Computing*, 9(4).

Monga, O., Benayoun, S., and Faugeras, O.D. 1992. Using partial derivatives of 3D images to extract typical surface features. In *IEEE Conference on Computer Vision and Pattern Recognition*, Urbana Champaign, Illinois.

Monga, O. and Benayoun, S. 1995. Using partial derivatives of 3D images to extract typical surface features. *Computer Vision and Image Understanding*, 61(2).

Nastar, Chahab. 1994. Vibration modes for nonrigid motion analysis in 3D images. In *European Conference on Computer Vision*, Stockholm.

Nastar, Chahab and Ayache, Nicholas. 1994. Spatio-temporal analysis of nonrigid motion from 4D data. In *Proceedings of the IEEE Workshop on Nonrigid and Articulate Motion*, Austin, Texas.

Park, J., Metaxas, D., and Axel, L. 1995. Volumetric deformable models with parameter functions: a new approach to the 3D motion analysis of the LV from MRI-SPAMM. In *IEEE International Conference on Computer Vision*, Cambridge, Massachusetts, pp. 700–705.

Raviart, P.A. and Thomas, J.M. 1992. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Masson, Paris.

Szeliski, R. and Lavallée, S. 1994. Matching 3-D anatomical surfaces with non-rigid deformations using octree-splines. In *IEEE Workshop on Biomedical Image Analysis*, Seattle, Washington.

Tang, L.A. and Huang, T.S. 1994. Quantifying facial expressions: smiles. In *Proceedings of the Workshop Journee INRIA Analyse/Synthèse d'Images*, Paris, France, pp. 22–27.

Terzopoulos, D., Witkin, A., and Kass, M. 1988. Constraints on deformable models: recovering 3D shape and nonrigid motion. *Artificial Intelligence*, 36(1):91–123.

Terzopoulos, D. and Metaxas, D. 1991. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714.

Terzopoulos, D. and Vasilescu, M. 1991. Sampling and reconstruction with adaptive meshes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 70–75.

Thirion, J.P. and Gourdon, A. 1992. The 3D marching lines algorithm and its applications to crest lines extraction. Technical Report 1672, Institut National de Recherche en Informatique et Automatique.

Vasilescu, M. and Terzopoulos, D. 1992. Adaptive meshes and shells. In *IEEE Conference on Computer Vision and Pattern Recognition*, Urbana Champaign, Illinois, pp. 829–832.