Segmentation and optical flow estimation in cardiac CT sequences based on a spatiotemporal PDM with a correction scheme and the Hermite transform

Leiner Barba-J[a,*], Ernesto Moya-Albor[b], Boris Escalante-Ramírez[a], Jorge Brieva[b], Enrique Vallejo Venegas[c]

[a] *Universidad Nacional Autónoma de México, Facultad de Ingeniería, C.U., México, D.F.*
[b] *Universidad Panamericana, Facultad de Ingeniería, México, D.F.*
[c] *Hospital Ángeles del Pedregal, México, D.F.*

Abstract

Purpose: The left ventricle and the myocardium are two of the most important parts of the heart used for cardiac evaluation. In this work a novel framework that combines two methods to isolate and display functional characteristics of the heart using sequences of cardiac computed tomography (CT) is proposed. A shape extraction method, which includes a new segmentation correction scheme, is performed jointly with a motion estimation approach.

Methods: For the segmentation task we built a Spatiotemporal Point Distribution Model (STPDM) that encodes spatial and temporal variability of the heart structures. Intensity and gradient information guide the STPDM. We present a novel method to correct segmentation errors obtained with the STPDM. It consists of a deformable scheme that combines three types of image features: local histograms, gradients and binary patterns. A bio-inspired image representation model based on the Hermite transform is used for motion estimation. The segmentation allows isolating the structure of interest while the motion estimation can be used to characterize the movement of the complete heart muscle.

Results: The work is evaluated with several sequences of cardiac CT. The left ventricle was used as object of evaluation. Several metrics were used to validate the proposed framework. The efficiency of our method is also demonstrated by comparing with other techniques.

Conclusion: The implemented tool can enable physicians to better identify mechanical problems. The new correction scheme substantially improves the segmentation performance. Reported results demonstrate that this work is a promising technique for heart mechanical assessment.

Keywords: Segmentation, Spatiotemporal point distribution model, Local image features, Optical flow, Hermite transform, Cardiac CT sequences

1. Introduction

Cardiac CT is currently one of the main types of radiological images used for heart analysis. Image slices showing the structural composition of the heart can be obtained with CT scanners [1]. The continued improvement of multidetector CT scanners has increased the potential of cardiac CT as clinical tool for heart imaging [2]. Since heart failure is one of the main health problems in developed and developing countries [3], tasks focused on cardiac analysis are of main concern for physicians. Several benefits of CT systems have been recognized to evaluate heart functions. Quantification of the ejection fraction, left and right ventricular functions, and wall motion evaluation are examples of typical uses of cardiac CT data [4].

The natural movement of the heart implies that its mechanical behavior must be evaluated as well. The spatiotemporal data obtained from cardiac CT studies can be used in computer-aided systems to evaluate the cardiac function, which has become essential over the past few years allowing faster assessments in the diagnosis process [5]. Since the left ventricle is vital for the proper functioning of the heart, it has become of major interest when analyzing cardiac images. In cardiac CT heart is commonly scanned at increments of 10 % of the cardiac cycle providing a 4D dataset.

Shape extraction for volume measurement and motion estimation are the most typical tasks for heart evaluation where computer-based algorithms are extensively used [6, 7, 8, 9]. In this sense, development of new and most efficient algorithms, methods and mathematical models to analyze cardiac structures in CT data are activities of great interest for researchers.

In a general sense, basic processes like segmentation and optical flow estimation are primary steps before applying higher processes like image interpretation. Cardiac CT se-

*Principal corresponding author:Universidad Nacional Autónoma de México, Edificio de Posgrado en Ingeniería, Departamento de Procesamiento de Señales, Laboratorio Avanzado de Procesamiento de Imágenes, C.U., México, D.F.

Email address: lebaji@gmail.com (Leiner Barba-J)

quences constitute great challenges for segmentation and motion estimation algorithms. A typical problem when working with medical images is that they may vary considerably from one patient to another, from an image processing point of view. These variations are perceived as changes of contrast, size and geometrical shape. Cardiac CT images do not scape from these kinds of problems. Even though development of segmentation and optical flow estimation algorithms imposes issues that researches have tried to solve for several decades, the problem remains open. Recent thorough reviews of segmentation techniques applied to cardiac images [10, 11] conclude that shape extraction in heart images remains a very challenging task.

Active Shape Models (ASM) [12] have gained enormous popularity during the last twenty years and have been extensively used for modeling 2D and 3D data in cardiac imaging [13, 14, 15]. We opted for this approach due to its ability to represent specific shapes of an image. Problems regarding the contrast and shape variability can be easily overcome with ASM-based algorithms. Related literature deals with active shape models as methods to analyze cardiac images [14, 16]. ASM have also been combined with other methods with the aim of segmenting heart images [6]. Modifications of the original approach become necessary for improving the segmentation performance in some cases when the training samples are scarce [14]. However, issues of ASM are evident when the number of training samples is small. It is therefore necessary to design new strategies to overcome these problems.

The dynamic nature of the heart has motivated researchers to design image tracking algorithms to process cardiac images [17, 9, 18]. Tracking heart structures like the left ventricle or the myocardial wall can be performed using optical flow estimation methods which also allow computing the displacements of the cardiac structures in a sequence of images. For this purpose we used a differential approach defined in the Hermite transform (HT) space. The HT is a bio-inspired human vision model that decomposes an image with a set of orthogonal functions defined by the Hermite polynomials. Image patterns and structures relevant to human vision perception such as oriented edges and textures can be efficiently represented with the HT. The proposed optical flow estimation approach using the HT allows defining local image constraints and a multiresolution strategy within differential scheme and are relevant in a perceptual sense as described in [8].

Our main goal is to build a tool that may help physicians evaluate heart mechanical functions. In order to achieve our objective, we implemented a framework with two main processes: 1) A segmentation stage based on a statistical shape model and a new correction scheme, and 2) An optical flow estimation approach based on the Hermite transform. For the first process we have designed a novel correction method that substantially improves the segmentation performance. The goal of the new correction method is to refine the segmentation previously achieved with the statistical shape model. It consists of a deformable scheme that combines three image parameters: histogram, gradient and a binary pattern. These parameters are locally computed for each point of the contour of the segmentation. This work is entirely focused on analyzing sequences of cardiac CT images (2D + time). The algorithms are specifically applied to the left ventricle because it is responsible for some of the most vital functions of the heart. Cardiac CT studies are analyzed using the original axial view. Nevertheless, the method can be extended without major problems to other views. Although short and long axis are the most accepted views used for cardiac analysis, the original axial view is also very important for this task [19]. Combined results of both algorithms are presented. Vectors indicating the motion of the left ventricle are jointly used with contours of the segmentation. Results are evaluated with several image sequences using quantitative and qualitative analysis.

The rest of the paper is organized as follows. Material used in this work is described in Section 2. Methods are depicted in Section 3. Here, segmentation and optical flow approaches are included. Results and discussions are finally presented in Sections 4 and 5 respectively.

2. Materials

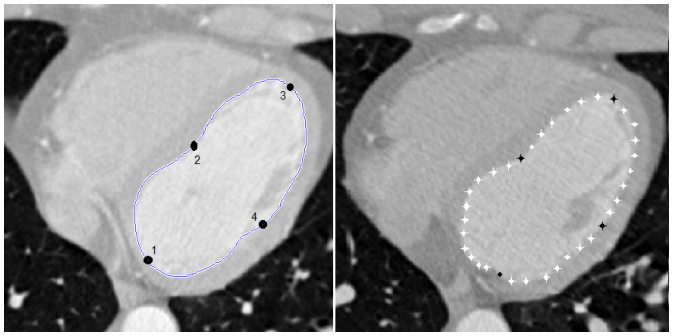
Our dataset consists of 40 sequences of cardiac CT images. Selected sequences used for evaluation show the left ventricle at half of the heart. The tomographic studies were acquired with a SIEMENS 16-slice CT system at 120 kVp of tube voltage and 900 mA. The scanner is composed of 128 detectors and is synchronized with the ECG signal. Each image has a size of 512 x 512 pixels, quantized to 12 bits per pixel. A contrast agent was also applied to each patient. Each sequence is composed by 10 frames showing the heart variation throughout the entire cardiac cycle from diastole to systole.

3. Methods

3.1. Statistical model of shape

Active shape models are one of the most powerful segmentation tools for medical image analysis. They consist of a statistical model that can be deformed within a specific range defined by a training set [12]. Here, shapes are represented using discrete points in the spatial domain. These points are commonly called landmarks when they are used to depict anatomical structures. Two main stages must be implemented in ASM algorithms: 1) Training of the statistical model, and 2) Segmentation of new images using the statistical model. An appearance model for each landmark is also required.

The trained model encodes the principal modes of variation of the landmarks. When these landmarks represent biological structures, they can be categorized as anatomical, mathematical and pseudo-landmarks [20]. In our model we used four mathematical landmarks and several pseudo-landmarks (see Fig. 1).



(a) (b)

Fig. 1 Landmarks of the shape. **(a)** Four mathematical landmarks, **(b)** Pseudo-landmarks (white points).

3.1.1. Spatiotemporal point distribution model

We adopted the method in [21] to build the statistical model. N samples, each one represented by r points, are used for training. Each sample corresponds to an image sequence. Landmarks of all the frames are concatenated in one shape vector. Let S_i be the vector describing shape i of the training set; it can then be obtained by concatenating the P_{ij} landmarks of the analyzed object: $S_i = (P_{i0}, P_{i1}, \ldots, P_{i,r-1})^\top$ with $i = 1, 2, \ldots, N$ and $j = 0, 1, \ldots, r-1$. $\top$ is the transpose of the vector.

A training set previously marked is needed in order to build the statistical model. For this task we created an interactive graphic interface that enables manual segmentation.

A sequence of cardiac CT images includes temporal and spatial information. In this method both types of information are represented using a unique shape vector. This shape representation is commonly called spatiotemporal [18]. All the landmarks are concatenated beginning with the first frame of the sequence (represented by the time t_0) and ending with the last one (denoted by the time variable t_9). Each shape vector is then computed as:

$$S_i = \{x_{i0t_0}, y_{i0t_0}, x_{i1t_0}, y_{i1t_0} \ldots, x_{i(n-1)t_0}, y_{i(n-1)t_0} \ldots, x_{i0t_9}, y_{i0t_9}, x_{i1_9}, y_{i1t_9} \ldots, x_{i(n-1)t_9}, y_{i(n-1)t_9}\}^\top \quad (1)$$

where $\{x_{ijt_f}, y_{ijt_f}\}$ is the set of landmark coordinates; t_f $(f = 0, 1, \ldots 9)$ indicates a frame of the sequence; $i = 1, 2, ..., N$ indicates a spatiotemporal shape, n is the number of points for each frame of the sequence, and N is the number of sequences of the training set. The statistical model used in this work is referred to as Spatiotemporal Point Distribution Model (STPDM) [21] and its construction can be achieved with the approach described in [12, 22, 21], which will be shortly explained in this section.

The training set is firstly aligned. The Generalized Procrustes Analysis GPA [20] is very efficient for this task. A final mean shape $\bar{S}$ is computed using the aligned shapes X_i:

$$\bar{S} = \frac{1}{N} \sum_i X_i \quad (2)$$

Principal Component Analysis (PCA) must be performed in order to assemble the STPDM. The set of eigenvectors e_k corresponding to the highest eigenvalues λ_k are obtained with the aim of coding the principal modes of variation for each landmark. Hence, the STPDM is computed as:

$$S = \bar{S} + Mb \quad (3)$$

where M is the eigenvectors matrix and b is the shape parameter. New shapes can be generated by varying the values of b. An appearance model must also be built for each landmark during training and is used to guide the STPDM in the segmentation process. In order to strengthen the edge-based adaptation of the algorithm, intensity and gradient profiles for each landmark are used. Once the statistical model and the appearance parameters have been calculated, new sequences can be segmented. With the STPDM we code the position of one landmark not only with respect to similar landmarks in the same frame, but also with respect to the rest of the landmarks of the other frames. The configuration of the spatiotemporal shape vector imposes an additional constraints in the deformation process. The algorithm followed in this work can be reviewed with more details in [22].

3.2. Segmentation correction algorithm

One of the main problems of ASM arises from the limited number of training samples used to build the statistical model. Ideally, the number of pre-segmented sequences needed to build a very efficient PDM must be at least the number of landmarks used to describe the shape sequence. The quantity of significant eigenvalues obtained with PCA depends on the number of shape samples (for N training samples, it is possible to obtain at least $N - 1$ significant eigenvalues). This causes that small details of the analyzed object can not be segmented because the deformation capabilities of the STPDM are poor. Davatzikos et al. [23] proposed a hierarchical model to overcome the problem of lack of data in the training set. They used a wavelet representation of the shape vectors to build the statistical model at the level of sub-bands. Substantively, the number of training samples is increased by using this framework. Following with the same idea, Nain et al. [24] extended the method to 3D data by using spherical wavelets. Despite these approaches are efficient in managing the set of training data, the performance depends on the number of decomposition levels used for the model construction.

Instead of improving the ASM formulation, in this work we focused our effort in correcting the segmentation errors by using a simple but an effective method to find better local positions of the shape landmarks and making them to be deformed independently of the general shape parameters. We assume that the segmentation with the ASM algorithm has reached a stable condition, it means that the active search has converged to a final solution in which global characteristics of the object were found.

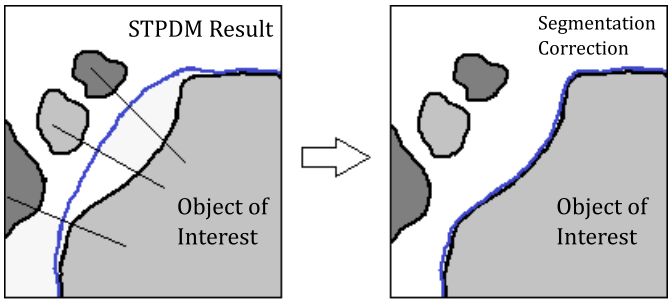


Fig. 2 Illustration of the deformation process carried out after the ASM segmentation. The blue line is the contour which is adapted to find the edges of the object (black lines). Since ASM algorithms lack of strong capacity of deformation when the STPDM is built with few training samples, small details are difficult to segment (see left image).

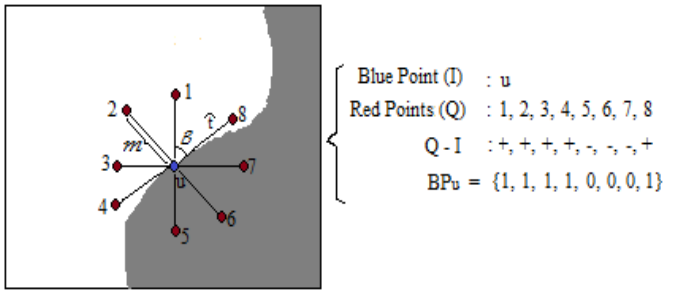


Fig. 3 Example of the binary pattern BP for a specific configuration. Red points are located at a distance m and angle β with respect to the tangent line from the blue point u which is the analyzed landmark. Q and I are the intensities of the red and blue points respectively. Depending on the intensity level, the difference $Q - I$ can be either positive or negative. For this example, we consider that intensity of the white region is higher than the dark one.

Small details of the object have to be segmented with our correction method. In figure 2 we graphically outline the deformation process to be followed after the ASM stage. Three image parameters are taken into account for the final adaptation.

Our correction algorithm aims to adjust the discrete contour to the boundaries of the object of interest. From figure 2 we can see that there are several object boundaries in which the blue contour can be deformed, however the correct one has to be identified. Intensity and gradient information, as well as a binary pattern are used as image features to adjust each landmark. In order to maintain a smooth contour, these parameters are embedded into a parametric active contour functional. The computation of these parameters is explained below. Moreover, better positions for the landmarks are found along the normal direction. Because we characterize the left ventricle boundaries using local edge and gray level features, we applied a Gaussian filter to each input image of the complete sequence. It allows improving the process of feature extraction to better describe the boundaries of the analyzed object.

3.2.1. Binary pattern

The first parameter used for edge characterization is a binary pattern that codes local information around each landmark. The goal is to find the relationship between a point and its neighborhood. Intensity points are sampled and compared with the analyzed landmark. For landmark j of frame t_f, we define the binary pattern as:

$$BP_{j,t_f} = \left\{ Hd\left(Q_{j,t_f}(m,\beta) - I_{j,t_f}\right)\right\} \tag{4}$$

with $0 \leq (Q, I) \leq G_{max}; m \in R; 0 \leq \beta \leq 2\pi$. Here, $Hd(y)$ is the Heaviside function, G_{max} is the maximum image intensity, $I_{j,tf}$ is the intensity of the analyzed landmark, and $Q_{j,t_f}(m,\beta)$ is the intensity at a distance m and angle β of the analyzed landmark. The pattern is therefore a set of binary elements obtained with a specific configuration given by the parameters m and β. Figure 3 illustrates an example of a particular configuration used to obtain the binary pattern.

A reference binary pattern BP_r for each landmark is also needed. This must be previously trained. Because the objective is to find better positions for the landmarks, the BP is computed for several points in the normal direction and afterwards compared with the corresponding trained reference. A similarity metric designed for binary data is used for comparison purposes. We selected the Jaccard distance JD since it constitutes an efficient method to compare binary data [25]. Let BP_r and BP_{X_n} be two sets of binary patterns where each element corresponds to an individual outcome. The Jaccard distance between them is computed as:

$$JD_{X_n} = \left(1 - \frac{DJ_{11}}{DJ_{01} + DJ_{10} + DJ_{11}}\right)_{j,t_j} \tag{5}$$

where,

DJ_{01} is the number of elements being 0 in BP_r and 1 in BP_{X_n}.

DJ_{10} is the number of elements being 1 in BP_r and 0 in BP_{X_n}.

DJ_{11} is the number of elements being 1 in BP_r and 1 in BP_{X_n}.

Here, BP_{X_n} is the binary pattern computed for a point X in the normal direction of the landmark (j, t_f). This metric ranges within the interval $[0, 1]$, being 0 the value obtained when both patterns are equal. The binary pattern helps identify the side of the object which a particular landmark belongs to.

3.2.2. Intensity parameter

The intensity parameter corresponds to the second image feature used to characterize the boundaries of the analyzed object. Here, a local histogram LH is associated with each landmark. Similarly, we need to train a local histogram which is subsequently used as reference in the final deformation process. Then, for each point in the normal direction of the landmark (j, t_f) we compute the local

histogram LH_{X_n} and is compared with the reference LH_r. As metric of comparison we used a vector space distance [26] defined as:

$$HD_{X_n} = \left(\frac{\sum_i LH_{X_n}(i) LH_r(i)}{\sqrt{\sum_i LH_{X_n}(i)} \sqrt{\sum_i LHr(i)}} \right)_{j,t_f} \quad (6)$$

where i indicates a bin of the histogram. When both histograms are very similar the distance HD approximates to 1. This parameter maintains the landmarks in the boundary of the left ventricle, preventing the deformation to other close boundaries.

3.2.3. Gradient parameter

Because the gradient is the standard operator to find edges inside an image, it is included as the third parameter in our segmentation correction algorithm. Therefore, the image gradient is locally computed and normalized for each landmark of the discrete contour. It is referred to as GI.

3.2.4. Embedding the parameters into an active contour

In order to provide a transparent way to deform the landmarks, we embedded the parameters as image energies into a parametric active contour functional. This also allows controlling the deformation through the internal energy of the active contour. We implemented a Greedy algorithm [27] which is more suitable for discrete approximations. The energy of the active contour is written as:

$$\begin{aligned} E(X) &= \mu E_{int}(X) + E_{ima}(X) \\ &= \mu E_{int}(X) + \lambda JD(X) + \alpha(1 - HD(X)) \\ &+ \beta(1 - GI(X)) \end{aligned} \quad (7)$$

where μ, λ, α and β are weight values that control the contribution of the parameters; E_{int} and E_{ima} correspond to the contour and image energies respectively; $X \in R^2$.

3.3. Differential optical flow estimation methods

The optical flow estimation is calculated in those situations where the displacement or corresponding of pixels into two images is required, e.g., image registration and reconstruction applications, video compression, motion–based segmentation, medical imaging, etc.

In [28] Gibson defined the optical flow (OF) as a 2D distribution of "apparent velocities" of the objects in the scene, that in most cases are associated with intensity variations of the objects into a image sequence. It is represented by a vector field that represents the displacement of pixels.

The classical differential optical flow methods assumes that the intensity value of the objects remain constant in two consecutive time instants of a image sequence $L(X,t)$. This assumption was proposed by Horn and Schunck in 1981 [29] and is known as the *Constant Intensity Constraint*:

$$L(X + W, t + 1) - L(X, t) = 0 \quad (8)$$

where $X = (x, y, 1)^\top$ is the pixel location, $W := (u, v, 1)^\top$ represents the horizontal and vertical pixel displacements respectively between two images at time t and $(t + 1)$.

If small displacements is assumed a Taylor expansion is considered and the *Optical Flow Constraint* equation is obtained:

$$W^\top (\nabla_3 L) = 0 \quad (9)$$

where $\nabla_3 L := (L_x, L_y, L_t)^\top$ and L_x, L_y, L_t are the derivatives of the intensity image $L(X)$ in x, y directions and time.

If only the *Constant Intensity Constraint* is used it is not possible to determine the two components of displacement u and v, and we are deal with an ill–posed problem known as *Aperture Problem* where only the normal component of the motion can be obtained [29]. Therefore, we need additional constraints to fully calculate the optical flow. Recent differential optical flow methods have proposed additional constraints to overcome the aperture problem and to improve accuracy of the displacement obtained, where add local image constraints, include spatial coherence, use statistical robust optimization functions and multiscale approaches have been suggested [30].

Horn and Schunck [29] proposed the *Smoothness Constraint* wich assumed that the flow is smooth, i.e., pixels within a neighborhood have similar magnitude and orientation. Thereby, the optical flow estimation method of Horn can be found minimizing, by an iterative approach, the following energy functional:

$$\int_\Omega \left(W^\top \left(\nabla_3 L \nabla_3 L^\top \right) W + \alpha |\nabla W|^2 \right) dX \quad (10)$$

where Ω is the image domain and α is a smoothness parameter.

The main disadvantage of the uniform smoothness in Eq. 10 is that it results in an excessive smoothing of edges in the flow. To avoid this problem a flow–driven smoothing approach can been used mainly in images strongly textured [31]:

$$\int_\Omega \left(W^\top \left(\nabla_3 L \nabla_3 L^\top \right) W + \alpha \Psi \left(|\nabla u|^2 + |\nabla v|^2 \right) \right) dX \quad (11)$$

where $\Psi\left(s^2\right)$ is a smooth function convex in s.

In most cases the constant intensity constraint and the small displacements assumption of Eqs. (8) and (9) are not fulfilled, as in the CT images case. Therefore, to overcome the problem that the intensity does not is constant in the image sequence, an additional term independent to intensity change is required, e.g., the image gradient. Secondly, to handle large displacements novel differential methods propose adopted multiresolution strategies and

use the original constancy assumptions in their non–linear form leaving the minimization for the numerical solution.

Papenberg et al. [32] proposes a method (OF–Warp) that takes into account the variability in intensity and the large displacements present in the images:

$$\int_{\Omega\times[0,t]} \Psi\Big(|L(X+W)-L(X)|^2+ \gamma|\nabla L(X+W)-\nabla L(X)|^2\Big)dX + \alpha\int_{\Omega\times[0,t]}\Psi\Big(|\nabla_3 u|^2+|\nabla_3 v|^2\Big)dX \quad (12)$$

where $\nabla L(X)$ is the image gradient, γ is a weight parameter and $\Psi\left(s^2\right)$ is the modified $\ell 1^*$–norm [33, 30]:

$$\Psi\left(s^2\right)=\sqrt{(s^2+\epsilon^2)} \quad (13)$$

the values of ϵ are set in the order of $1\text{x}10^{-3}$ such that $\Psi\left(s^2\right)$ is differentiable in $s=0$ [32].

Another approach to calculate the optical flow in an image sequence is by defining a property of conservation, e.g., the intensity, similarly to a physical processes. In [34] and [35] an analogy between the mass conservation of a fluid within a volume and the intensity of voxels in a 3D stack of images was presented.

The decrease in mass per time of a fluid with density ρ out of the volume V can be given as

$$-\int_V \frac{\partial\rho}{\partial t}dV \quad (14)$$

where the mass inside the volume is defined by $m=\int_V \rho dV$.

Furthermore, the flow leaving the volume per unit time to a velocity v in a surface S is give by $\int_S \rho v dS$. Using the divergence operator the last equation can be rewritten as a integral over the volume as follows:

$$\int_V div(\rho v)\,dV \quad (15)$$

Eq. (14) holds that the mass that leave the volume is equal to the mass in a surface S enclosing the volume (Eq. 15):

$$\frac{\partial\rho}{\partial t}+div(\rho v)=0 \quad (16)$$

In the image sequence the density of fluid ρ is related with the intensity value L and the velocity v of the fluid with the displacement of the voxels $W_*:=(u,v,w)^\top$. Therefore, the equation of mass conservation based optical flow is:

$$L_t+div(LW_*)=0 \quad (17)$$

In [35] they defined a energy functional using a non-quadratic penalization as Eq. 12 and the uniform smoothness of Eq. 10 is as follows:

$$\int_V \Psi_1\left(\left(L_t+div(LW_*^\top)\right)^2\right)dV+\alpha\int_V\Psi_2\left(|\nabla_3 W_*{}^2|\right)dV \quad (18)$$

where Ψ_i is the Charbonnier optimization function [36]:

$$\Psi_i(s^2)=2\beta_i^2\sqrt{\left(1+\frac{s^2}{\beta_i^2}\right)}\quad;\quad i=1,2 \quad (19)$$

and β is a scaling parameter.

For a 2D approach of Eq. (20) the energy functional is

$$\int_\Omega \Psi_1\left(\left(L(X)_t+\nabla(L(X)W^\top)\right)^2\right)dX+\alpha\int_\Omega\Psi_2\left(|\nabla W^\top|^2\right)dX \quad (20)$$

Optical flow estimation using the steered HT (OF–SHT)

The Eq. (12) include local characteristics of image, such as intensity and gradient as constant constraints. A way to represent important visual characteristics is using bio–inspired models. Some human visual system based model based are the Difference of Gaussians, Gabor filters and the Hermite transform.

The Hermite transform is an image model based in a polynomial decomposition from a perceptual standpoint into a multiresolution scheme [37, 38]. This image model emulate the local processing and the response of receptive fields, e.g., the Gaussian derivative model, of the human vision system [39, 40].

The Hermite transform is computed by a convolution of the image $L(x,y)$ with the analysis filters $D_{m,n-m}$ [40]:

$$\begin{aligned} &L_{m,n-m}(x_0,y_0)= \\ &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}L(x,y)D_{m,n-m}(x_0-x,y_0-y)\,dxdy \\ &n=0,1,\ldots,\infty \\ &m=0,1,\cdots,n \end{aligned} \quad (21)$$

where $L_{m,n-m}(x,y)$ are the cartesian Hermite coefficients, m and $(n-m)$ are the analysis order in the orthogonal directions x and y, T is a subsamplig factor and (x_0,y_0) represents the position of the sampling lattice S.

The filter functions $D_{m,n-m}(x,y)$ are defined by the analysis window $v^2(x,y)$ and the polynomials orthogonal to this window $G_{m,n-m}(x,y)$ by

$$D_{m,n-m}(x,y)=G_{m,n-m}(-x,-y)v^2(-x,-y) \quad (22)$$

For a Gaussian window $v(x,y)=\frac{1}{\sigma\sqrt{\pi}}\exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right)$ with unitary energy for $v^2(x,y)$ the polynomials of the expansion are:

$$G_{m,n-m}(x,y)=\frac{1}{\sqrt{2^n m!(n-m)!}}H_m\left(\frac{x}{\sigma}\right)H_{n-m}\left(\frac{y}{\sigma}\right) \quad (23)$$

where $H_n\left(\frac{x}{\sigma}\right)$ are the generalized Hermite polynomials by a Gaussian window with variance σ^2 and are given by Rodrigues' formula:

$$H_n\left(\frac{x}{\sigma}\right)=(-1)^n\exp\left(-\frac{x^2}{\sigma^2}\right)\frac{d^n}{dx^n}\exp\left(-\frac{x^2}{\sigma^2}\right) \quad (24)$$

As the Gaussian window is isotropic the Hermite filters $D_{m,n-m}(x,y) = D_m(x)D_{n-m}(y)$ are separable and can be obtained by

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) \exp\left(-\frac{x^2}{\sigma^2}\right) \tag{25}$$

A steerable version of the Hermite transform can be defined by rotating the cartesian Hermite coefficients $L_{m,n-m}(x,y)$ by the angular functions $g_{m,n-m}(\theta)$ by

$$l_{m,n-m,\theta}(x_0,y_0) = \sum_{k=0}^{n} \Big(L_{k,n-k}(x_0,y_0)\Big)\Big(g_{k,n-k}(\theta)\Big) \tag{26}$$

where $g_{m,n-m}(\theta)$ express the directional selectivity of the filter [38]:

$$g_{m,n-m}(\theta) = \sqrt{\binom{n}{m}}\Big(\cos^m(\theta)\Big)\Big(\sin^{n-m}(\theta)\Big) \tag{27}$$

The steered Hermite coefficients $l_{m,n-m,\theta}(x,y)$ allows adapt the analysis process to the local content of the image, e.g., using directional Gaussian derivatives filters [41].

The local orientation angle θ can be estimated by a criterion of maximum energy direction, where the cartesian coefficients L_{01} and L_{10} are a good approximation of the gradient of the image and they can give the orientation of the edges as follow:

$$\theta(x,y) = \arctan\frac{L_{0,1}(x,y)}{L_{1,0}(x,y)} \tag{28}$$

To illustrate the steering property of the Hermite transform in Fig. 4b we show the steered Hermite coefficients of the corresponding cartesian Hermite coefficients of Fig. 4a according to the gradient orientation.

The Hermite transform can be used to defined a energy functional using perceptive features as constant local constraints (see [8] to more detail). The proposed functional include in the data term a constant intensity constraint using the zero order coefficient L_0 and the steered Hermite coefficients $l_{n,\theta}$ up to order N as a constant high order constraint of local features [32], the last term allow dealing with intensity changes in the image sequence where the constant intensity constraint fails. For the smoothness term, that allows us to recover the flow in homogeneous areas of the image, a flow–driven regularizer [31] was used. Thereby, a bio–inspired energy functional that uses the Hermite transform to extract relevant perceptive features is defined as follow:

$$\begin{aligned}\int_\Omega \Psi\Bigg(&\Big|L_0(X+W) - L_0(X)\Big|^2 + \\ &\gamma\Big(\sum_{n=1}^{N}\Big|l_{n,\theta}(X+W) - l_{n,\theta}(X)\Big|^2\Big)\Bigg)dX \\ &+ \alpha\int_\Omega \Psi\Big(|\nabla u|^2 + |\nabla v|^2\Big)dX\end{aligned} \tag{29}$$

where γ determines participation of the constant intensity and high order features constraints, N is the maximum order of polynomial expansion, Ψ is the modified $\ell 1^*$–norm of Eq. (13) and α is a smoothness weight which affects the flow smoothing given a smoother flow for large values.

For find the minimum of Eq. (29) the Euler–Lagrange equations are obtained and an outer fixed point iteration process and successive over-relaxation (SOR) iteration approach is performed to compute the solution of the equations. The non–linear terms are overcome by using a 1st order Taylor expansion after the minimization of the functional.

Finally a multiresolution strategy is carried out in coarse levels by calculating small displacements, and thus not violate the Taylor expansion. Then the solution is propagate to the finer levels using a Gaussian pyramid.

4. Experiments and results

4.1. Segmentation

In this section we present results of the segmentation stage. To carry out the experiments we configured the algorithms as follows. A total of 50 points were used to represent the left ventricle in each frame of the sequence. It means that the spatiotemporal shape was built with 500 landmarks. A maximum of 25 iterations were enough to reach a stable solution in the ASM stage. Similarly, 5 iterations were used for the correction algorithm. The algorithm was initialized using the mean spatiotemporal shape. It was manually put very close to the object of interest. The first frame of the sequence was taken as reference to initialize the algorithm. With the aim of maintaining a standard experiment, we used the same initialization for each proof. The STPDM was trained using 35 samples. The complete dataset was validated using the leave-one out method.

The correction algorithm is executed when the ASM algorithm has converged. It was applied to each frame separately. The segmentation correction algorithm uses several weight parameters that need to be configured. Energies of the correction scheme were normalized to $[0,1]$. Although the weight parameters are difficult to select and depends on the particular image sequence, we experimentally found that good segmentation results for our dataset are achieved with $\mu = 0.3, \lambda = 0.3, \alpha = 0.2$ and $\beta = 0.2$.

One of the main advantages of statistical shape models is that they can perform efficient segmentations with noisy data. Since the PDM can only deform in the range specified by the training set, it can achieve acceptable results even in the presence of noise. Unlike, the correction algorithm can be affected by the image noise because it freely deforms the contour points without taking into account the shape of the object. We then reduce the effect of the image noise by previously applying a Gaussian filter to the sequence of images.

Qualitative and quantitative analysis are exposed in this section. The assessment is addressed by comparing with

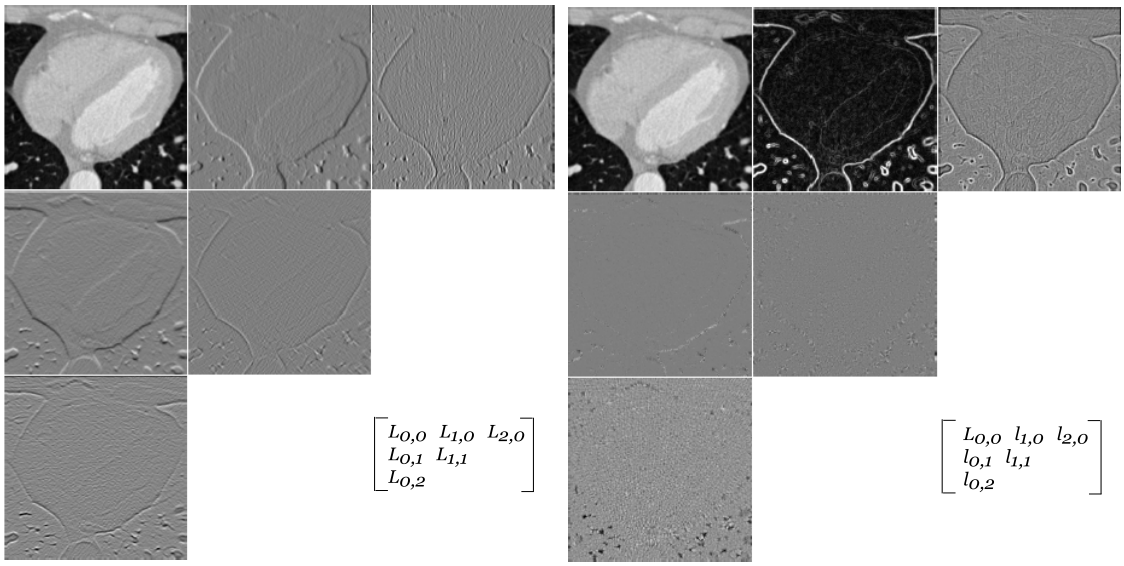


Fig. 4 a Cartesian Hermite coefficients and **b** the steered Hermite coefficients for $N = 2$ ($n = 0, 1, \ldots, N$ and $m = 0, 1, \cdots, n$) of a cardiac CT image at 20% of the cardiac cycle

the manual segmentation. With the aim of verifying the efficiency of the proposed approach, we compared the obtained results with the classical ASM [12] which was applied frame by frame.

Fig. 5 illustrates results of the segmentation obtained for 6 frames extracted from 2 different sequences of the dataset. For visualization purposes, we used linear interpolation to draw continuous contours. Images in Fig. 5a, 5b and 5c were taken at 0%, 30% and 60% of the cardiac cycle respectively.

Red, green and blue contours correspond to manual annotation, segmentation with the proposed method and segmentation with the classical ASM method respectively. As can be seen, the best results were reached with our method. This behavior is similar for the rest of the sequences of our dataset.

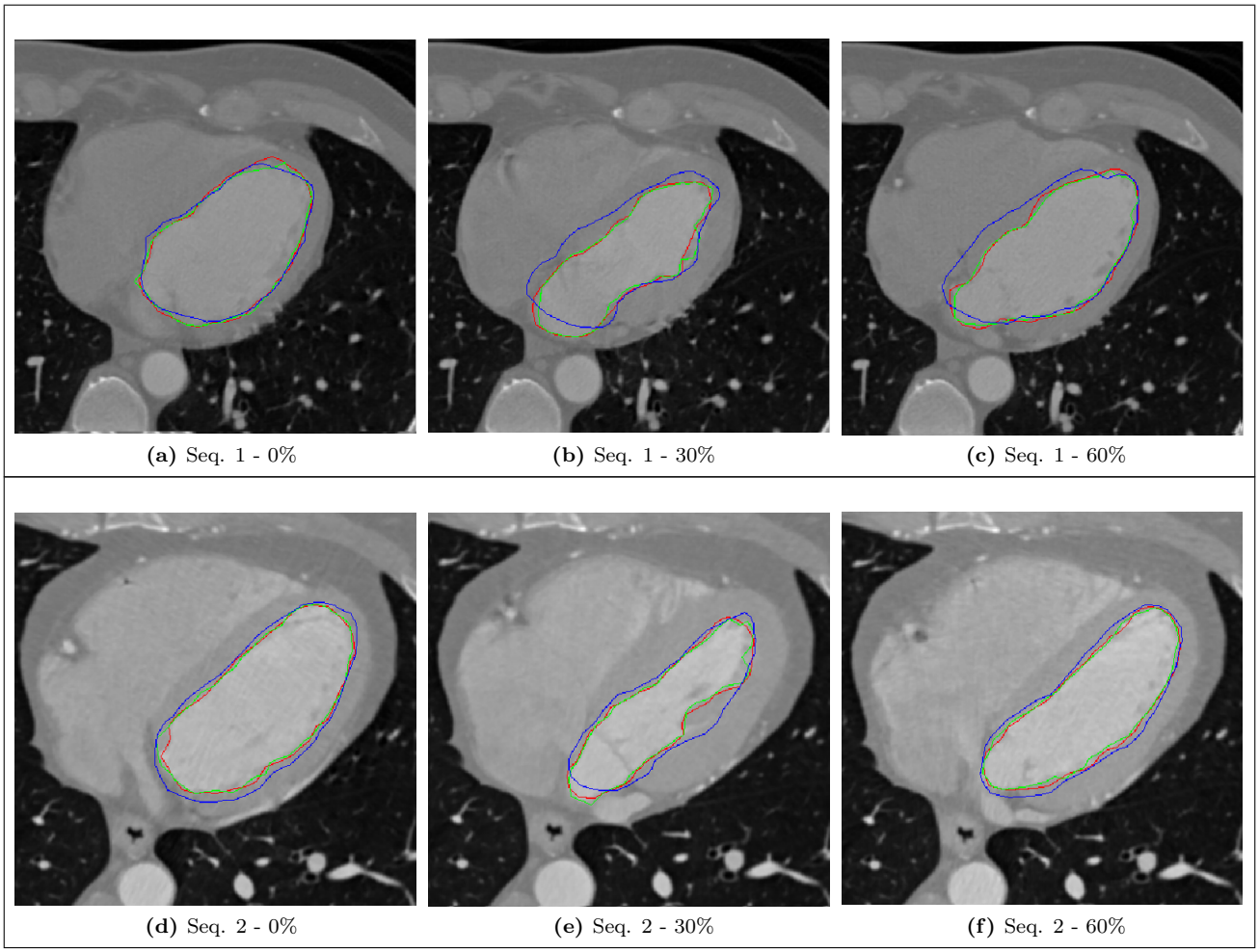
Two metrics were used to provide quantitative results of the segmentation: the Dice Similarity Coefficient (DSC) and the average point-to-curve distance. These metrics were calculated for each frame of the sequences. Fig. 6 presents the DSC obtained for 16 sequences using our segmentation approach. Results for three frames of each sequence at different cardiac phases are visualized. The performance reached for the proposed segmentation scheme is over 90% in most cases. This behavior is repeated for the rest of the sequences. A maximum DSC of 0.9855 and a minimum of 0.8141 were obtained, which represent the best and worst result respectively. Fig. 7 visualizes the areas of the best and worst result obtained. These areas correspond to regions enclosed by the contours. In the binary images, the white region is the object of analysis (left ventricle) and the black part is the background. The error region obtained by comparing the segmentation of the proposed method against the manual one is shown as well. This is the result of computing the difference between both binary images: $I_{err} = abs(I_m - I_a)$, where I_m and I_a are the corresponding binary images obtained through with the manual segmentation and the proposed method respectively, I_{err} represents difference where the black part is the overlapped region and the white region is the error.

We are also interested in evaluating how the correction algorithm improves the results of the STPDM. We then compared both stages of the segmentation method. In table 1 the point-to-curve distance values are reported for the proposed correction method, the STPDM and the classical ASM. Results are presented for all 10 frames of the spatiotemporal shape and averaged for all the sequences. It can be seen that the correction method improves the segmentation with respect to the STPDM in all cases.

Fig. 8 visualizes the final improvement carried out by the correction algorithm. Two images at different cardiac phases are evaluated. It can be seen that small details are subsequently segmented with our correction algorithm.

4.2. Optical flow estimation

Due to the very complex 3D motion pattern of the heart, the intensity of movement varies for different cardiac anatomies and within the cardiac cycle. In cardiac CT sequences, we can see that the strongest cardiac movement takes place during contraction of the atria and ventricles in systole, approximately between 0% and 30% of the cardiac cycle [42]. The short end–systolic rest phase is followed by a continuous filling phase of the ventricles during diastole that slows down towards mid- and end–diastole. The movement of relaxation during diastole can be seen



(a) Seq. 1 - 0% **(b)** Seq. 1 - 30% **(c)** Seq. 1 - 60%

(d) Seq. 2 - 0% **(e)** Seq. 2 - 30% **(f)** Seq. 2 - 60%

Fig. 5 Segmentation results for three frames at different cardiac phases: 0%, 30% and 60%. Results are illustrated for two sequences: **(a)-(c)** correspond to sequence 1. **(d)-(f)** correspond to sequence 2. The red contour is the manual segmentation, the green contour is the segmentation obtained with the proposed method, and the blue contour is the result using the classical ASM method.

between 40% and 60% of the cardiac cycle. According to Ohnesorge and Flohr [42] the least amount of movement is observed in end-systole and mid- to end-diastole of the cardiac cycle.

We show the resulting optical flow estimation for two sequences in Fig. 9. The first sequence was analyzed at 20% and 30% of the cardiac cycle (systole) to view the strongest cardiac movement. The second sequence was processed at 50% and 60% to view the movement of relaxation during diastole. In order to compare our algorithm, two optical flow results are shown. The first one was obtained applying the method of Papenberg et al. [32] and the second using the Steered Hermite transform (Eq. (29)).

In order to evaluate the optical flow estimation performance, we used an image reconstruction approach which consists of reconstructing sequential images from their adjacent images and their motion vectors, i.e., the computed optical flow for a particular image in a sequence is used to estimate the next image in that sequence.

For the reconstruction algorithm we used the *forward reconstruction* [43]. Let us consider two adjacent images in the CT sequence $L(x, y, t)$ and $L(x, y, t + 1)$, and the optical flow (u, v) between them. Given the pixel values of the first image, we can calculate the gray value at (x, y) in the second image using spline interpolation. With the reconstructed sequential images we then measured the Root Mean Squared (RMS) error of the prediction. We calculated the RMS error using the two optical flow methods. The RMS is defined as:

$$RMS\ error = \sqrt{\frac{\sum_x \sum_y \Big(L(x, y, t) - \hat{L}(x, y, t)\Big)^2}{\mathcal{M} \times \mathcal{N}}} \quad (30)$$

where $L(x, y, t)$ and $\hat{L}(x, y, t)$ are the true and reconstructed images of size $\mathcal{M} \times \mathcal{N}$ at time t.

In Fig. 10 we show the *forward reconstruction* results at 20% and 30% of the cardiac cycle for one sequence, and at 50% and 60% for a second sequence respectively. The first row corresponds to the slices at 20%, 30%, 50% and 60% of the cardiac cycle, and the vector field representing the motion of the left ventricle. The second row shows the *forward reconstruction* using the two optical flow estimation

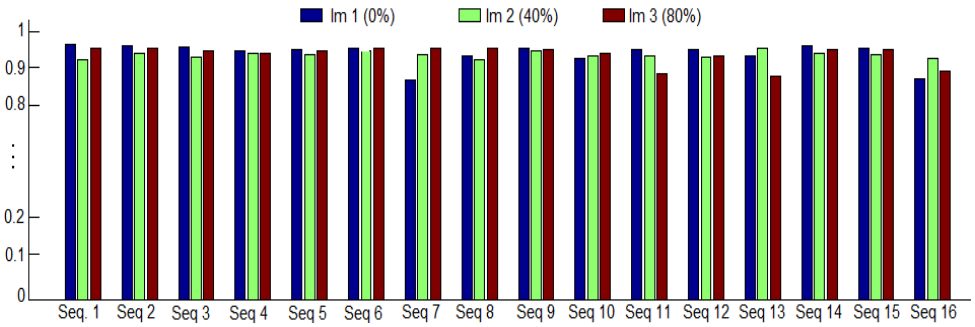


Fig. 6 Dice Similarity Coefficient (DSC) for 16 sequences. The coefficients are shown for 3 frames of each sequence. Image 1 is at 0% of the cardiac cycle, image 2 is at 40% and image 3 is at 80%.

Table 1 Point-to-curve distance averaged for all the sequences. Results are presented for the complete cardiac cycle.

Frame (Cardiac Phase)	**Correction Algorithm**	**STPDM** mean ± std (mm)	**Classical ASM**
Frame 1 (0%)	1.9328 ± 0.5300	2.436 ± 0.572	2.912 ± 0.669
Frame 2 (10%)	2.1397 ± 0.4655	2.684 ± 0.574	3.339 ± 0.835
Frame 3 (20%)	2.3003 ± 0.7580	2.972 ± 0.878	3.750 ± 1.153
Frame 4 (30%)	2.0529 ± 0.5877	2.965 ± 0.714	3.907 ± 0.928
Frame 5 (40%)	2.0814 ± 0.5864	2.927 ± 0.713	4.292 ± 1.199
Frame 6 (50%)	2.1613 ± 0.7988	2.864 ± 1.049	3.884 ± 1.070
Frame 7 (60%)	1.9208 ± 0.4868	2.471 ± 0.769	3.326 ± 0.983
Frame 8 (70%)	1.9221 ± 0.4621	2.382 ± 0.592	3.115 ± 0.719
Frame 9 (80%)	1.9528 ± 0.3579	2.486 ± 0.748	3.251 ± 0.702
Frame 10 (90%)	1.8493 ± 0.3201	2.452 ± 0.679	3.236 ± 0.973

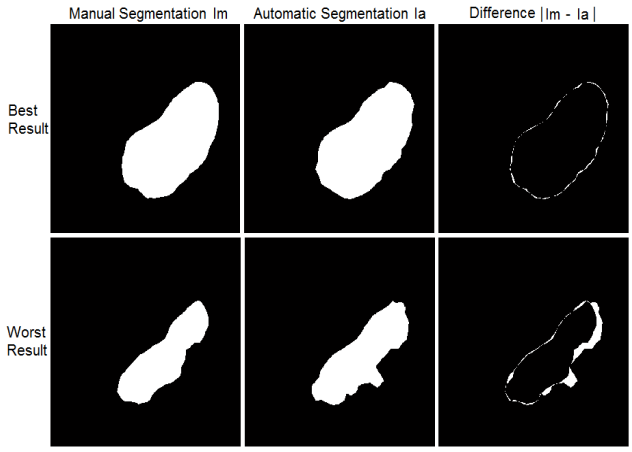


Fig. 7 Areas of the segmentation obtained for the best and worst result, considering individual frames of the complete dataset.

methods. The third row shows the absolute error between the true second image and the reconstructed image using the displacement vectors of the second row.

For a deeper analysis, the *RMS* error throughout the cardiac cycle for 10 sequences was performed. The *RMS* error was calculated in a region near to the manual annotated left ventricle. In Fig. 11 we show a comparison of the reconstruction performance using both methods. Here, we note that the RMS error is smaller using the SHT-based optical flow estimation approach.

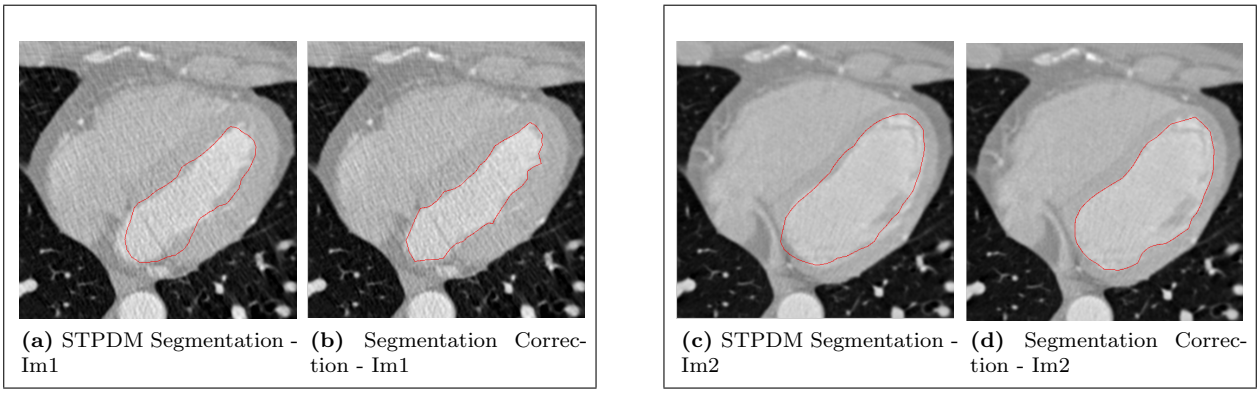
Fig. 12 shows the optical flow result for a frame of the *Dimetrodon* sequence to test the performance of the OF–SHT Method. This test sequence is available from the web site `http://vision.middlebury.edu/flow/data/` and it forms part of a database for evaluation of current optical flow methods.

Because of the presence of constant gradient and high order features constraints in Eqs. (12) and (29) we tested the performance of OF–SHT and OF-Warp methods in a first experiment. We added Gaussian noise of mean zero and different standard deviations. In table 2 we compare the average angular error of OF–Warp and OF–SHT methods for different noise levels and using optimised parameters γ, α and N.

The angular error (AE) was computed as was proposed by Barron et. al in [44]:

$$AE = \arccos\left(\overrightarrow{u} \cdot \overrightarrow{v}\right) \tag{31}$$

where $\cdot$ is the the dot product, and $\overrightarrow{u} = (u_0, u_1)$ and



(a) STPDM Segmentation - Im1 **(b)** Segmentation Correction - Im1 **(c)** STPDM Segmentation - Im2 **(d)** Segmentation Correction - Im2

Fig. 8 Correction scheme applied to two images at phases 40% and 90% of the cardiac cycle. **(a)-(b)** Results for image 1, and **(c)**-(d) Results for image 2.

Table 2 Average angular errors (AAE) computed for the *Dimetrodon* sequence with varying standard deviations σ_n of Gaussian noise.

Gaussian noise σ_n	**OF–SHT AAE±std (°)**	**OF–Warp AAE±std (°)**
0	2.7287 ± 5.8332	3.0387 ± 5.9236
10	7.4991 ± 10.3342	6.1124 ± 7.8844
20	9.8038 ± 14.169	7.5504 ± 9.1794
40	25.8927 ± 32.4171	8.5613 ± 10.283

Table 3 Average angular errors computed for the *Dimetrodon* sequence with parameter variation for OF–SHT method.

N	α	γ	**AAE±std (°)**
5	100	600	2.7287 ± 5.8332
4	100	600	3.0587 ± 5.9176
9	100	600	3.3247 ± 6.9388
5	50	600	3.1488 ± 7.0611
5	200	600	3.7013 ± 7.4547
5	100	300	2.909 ± 6.0738
5	100	1200	2.8359 ± 5.9213

$\overrightarrow{v} = (v_0, v_1)$ denotes the correct flow and the estimated flow, respectively.

In the second experiment we evaluated the robustness of the parameters in the OF–SHT method: the maximum order N of the Hermite expansion, the smoothness parameter α, and the weight parameter γ between the intensity value and the high order features assumption. The results shown in table 3 shows that the OF–SHT approach is robust under parameter variations.

4.3. Combining segmentation and motion estimation of the left ventricle

The objective of this work is to develop a method for the analysis of cardiac CT image sequences. This application can help physicians evaluate qualitatively several left ventricle parameters such as volume, myocardial deformation, ejection fraction and others [3]. Although the main effort is made on the left ventricle, the rest of cavities are also very important for the cardiac evaluation. In this application the segmentation is the first step and is used to isolate the left ventricle for further analysis. Optical flow vectors are only shown on the contour of the segmentation. This constitutes an efficient way to assist physicians to identify some failures of the cardiac function.

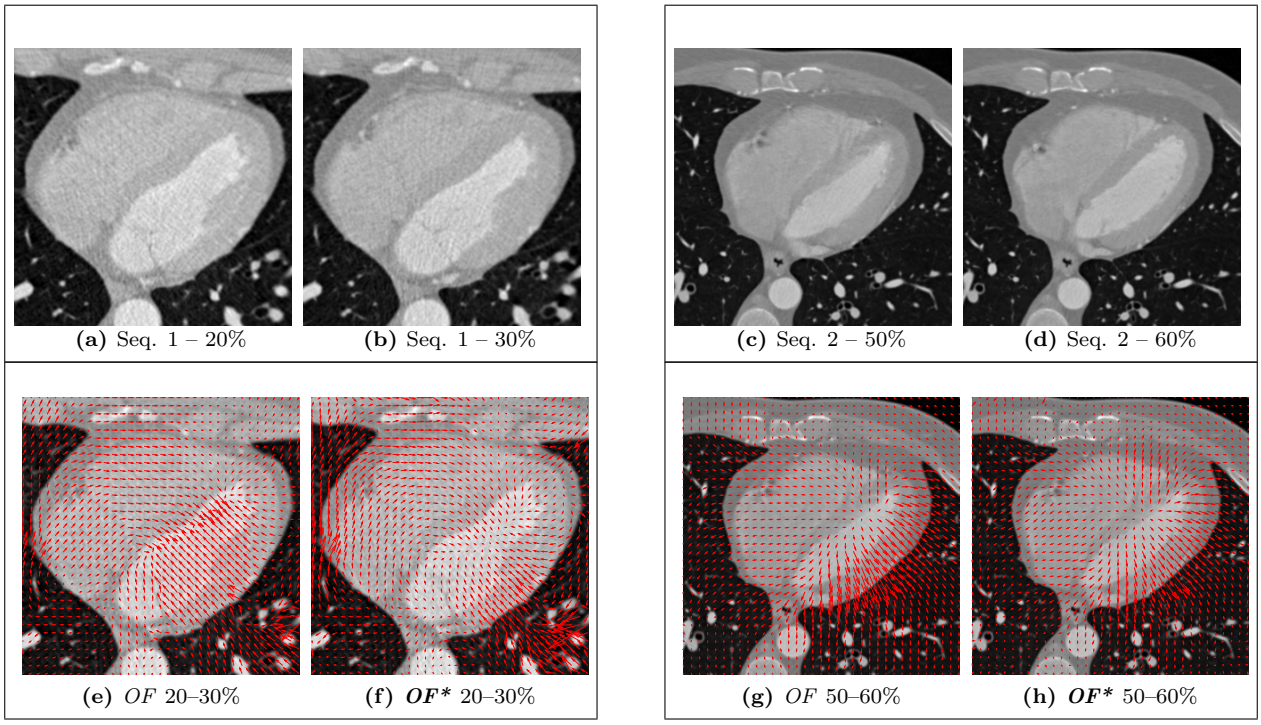
Fig. 13 illustrates the vector field for two sequences of the dataset. Four frames of each sequence are shown. The frames show the heart at 0%, 30%, 60% and 80% of the cardiac cycle. Vectors of the four frames were computed using images of the phases 0-10%, 30-40%, 60-70% and 80-90% respectively. The set of vectors in the first image pointing inwards the left ventricle indicates the beginning of the contraction or systolic cycle, meanwhile the last image outlines the end diastolic or relaxation period with the vector field pointing outwards. A complete visualization and a subsequent analysis of the whole cardiac cycle can be carried out by using this technique in the complete dataset. This method can then be used as a tool to assess the heart mechanical function.

4.4. Technical implementation and run-time

Both algorithms, segmentation and optical flow estimation, were implemented in Matlab R2013b using a 2.1 GHz machine composed by 12 processors. The machine also has a 16 GB memory. Each image sequence were selected from a 4D cardiac CT study with 10 volumes which describes the complete heart cycle. Sequences are then composed by 10 images with a size resolution of 512 x 512 pixels.

For the manual segmentation, we created an interactive graphic interface that enables the experts to mark the contour of the left ventricle in the images. The application allows selecting specific image sequences from the 4D cardiac CT to go through with the annotation process.

The segmentation algorithm was configured to stop until reaching 25 iterations for the ASM stage and 5 iterations



(a) Seq. 1 – 20% **(b)** Seq. 1 – 30% **(c)** Seq. 2 – 50% **(d)** Seq. 2 – 60%

(e) *OF* 20–30% **(f)** ***OF**** 20–30% **(g)** *OF* 50–60% **(h)** ***OF**** 50–60%

Fig. 9 Optical flow results for two sequences computed at phases 20–30% and 50–60% using [32] (*OF*) and the SHT-based method (*OF**). **(a)**-**(b)** and **(e)**-**(f)** correspond to sequence 1 at 20–30%. **(c)**-**(d)** and **(g)**-**(h)** correspond to sequence 2 at 50–60%.

for the correction one. The run-time for each iteration in the segmentation algorithm was about 2.8s for the ASM stage and 15.5s for the correction method.

The optical flow estimation using the Hermite transform algorithm uses a non optimized implementation using the convolution operator for compute the Hermite coefficients and the multiresolution resolution process of the Euler–Lagrange equations is not optimized. The process time was 40 seconds for two images of the sequence using 10 levels of decomposition, 5 outer fixed point iterations and 10 SOR iterations.

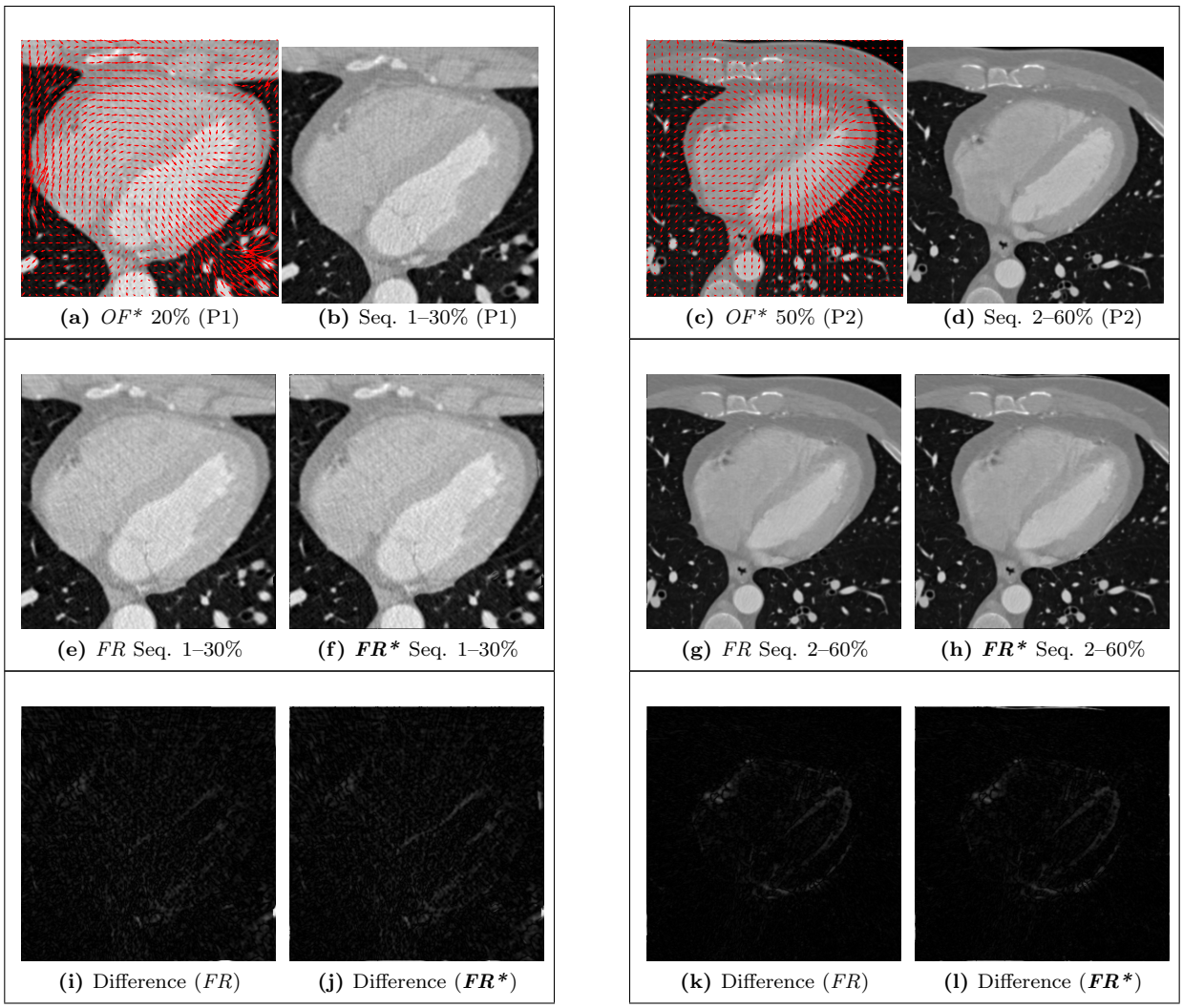
It's important to mention the the optical flow method does not focus on fast computation but if on high accuracy, thus, to improve time performance a fast Hermite transform (FHT) can be used [38], where only sums and butterfly operations are computed similarly to the used by the fast Fourier transform, additionally an efficient implementation of iterative processes is required to define a fast method.

5. Discussion and conclusions

We implemented a framework for the analysis of cardiac CT sequences using a shape extraction method and an optical flow estimation approach. The left ventricle was used as object of interest. We firstly performed the corresponding segmentation using a STPDM which consists of a trained statistical model that codes spatial and temporal information of the sequences. Errors of the segmentation were subsequently corrected using an algorithm that incorporates three image parameters for edge characterization. These parameters were embedded as image energies into an active contour model. Afterwards, an optical flow estimation method was calculated using a bio-inspired differential approach. Combining the boundaries of the segmented object and the displacements vector field obtained from the optical flow estimation technique enables physicians to carry out a better identification of mechanical problems. The proposed framework was validated with several sequences of cardiac CT images and compared with other techniques.

Results were individually evaluated for each frame of the sequence using several metrics. In general, the lowest performance achieved in our segmentation method was obtained for frames at half the cardiac cycle. These frames are acquired at the end of the contraction phase, i.e., when the heart is bombing flood to the body. The bar diagram (Fig. 6) with the DSC analysis and the metric distance reflect this interesting behavior. The poor definition of edges, the low contrast and the irregular shape of the left ventricle are causes for this result. Moreover, this can also be a consequence of the presence of the left atrium in the images at this cardiac phase when working with the original axial view of tomographic studies. The high irregularity of the structures found in this part of the cardiac cycle makes the segmentation task more complicated.

In table 1 we present comparative results of the two stages of the segmentation: STPDM and correction algorithm. As can be seen, the segmentation is substantially improved with our correction scheme. In all reported



(a) *OF** 20% (P1) **(b)** Seq. 1–30% (P1) **(c)** *OF** 50% (P2) **(d)** Seq. 2–60% (P2)

(e) *FR* Seq. 1–30% **(f)** ***FR**** Seq. 1–30% **(g)** *FR* Seq. 2–60% **(h)** ***FR**** Seq. 2–60%

(i) Difference (*FR*) **(j)** Difference (***FR****) **(k)** Difference (*FR*) **(l)** Difference (***FR****)

Fig. 10 *Forward reconstruction* (*FR*) results for sequences 1 and 2 at phases 20–30% and 50–60% of the cardiac cycle. First row: Images at 20 – 30% and 50 – 60%, and vector field obtained using the SHT approach (*OF**). Second row: *forward reconstruction* using the optical flow approach of [32] (*FR*) and the SHT-based method *FR**. Third row: Absolute error between the reconstructed image and the true second image using both methods. **(a)-(b)**, **(e)-(f)** and **(i)-(j)** correspond to sequence 1. **(c)-(d)**, **(g)-(h)** and **(k)-(l)** correspond to sequence 2.

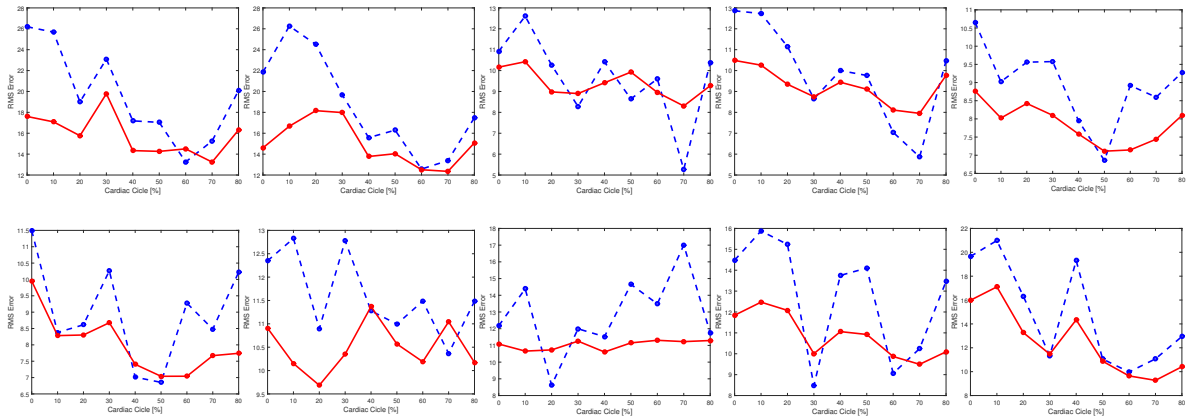
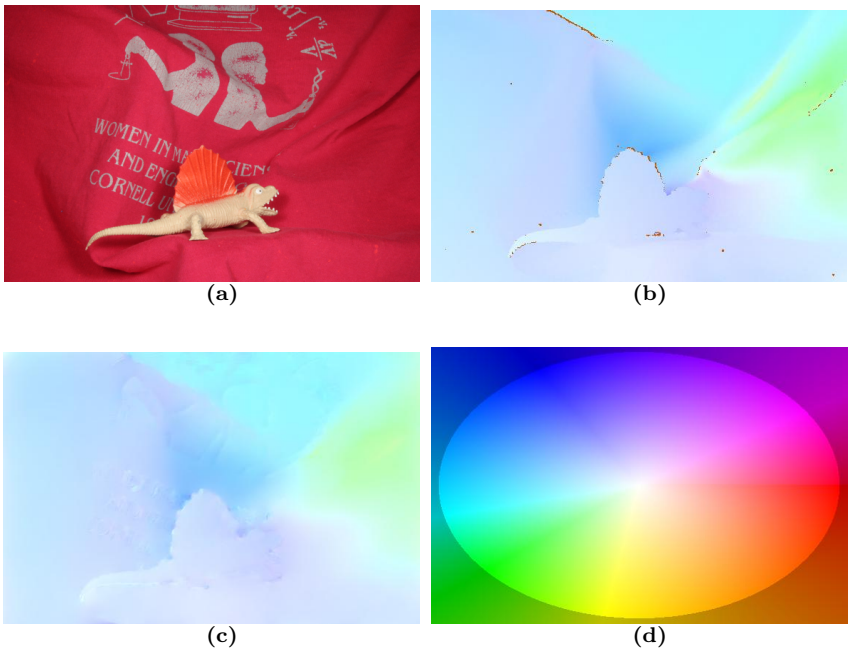
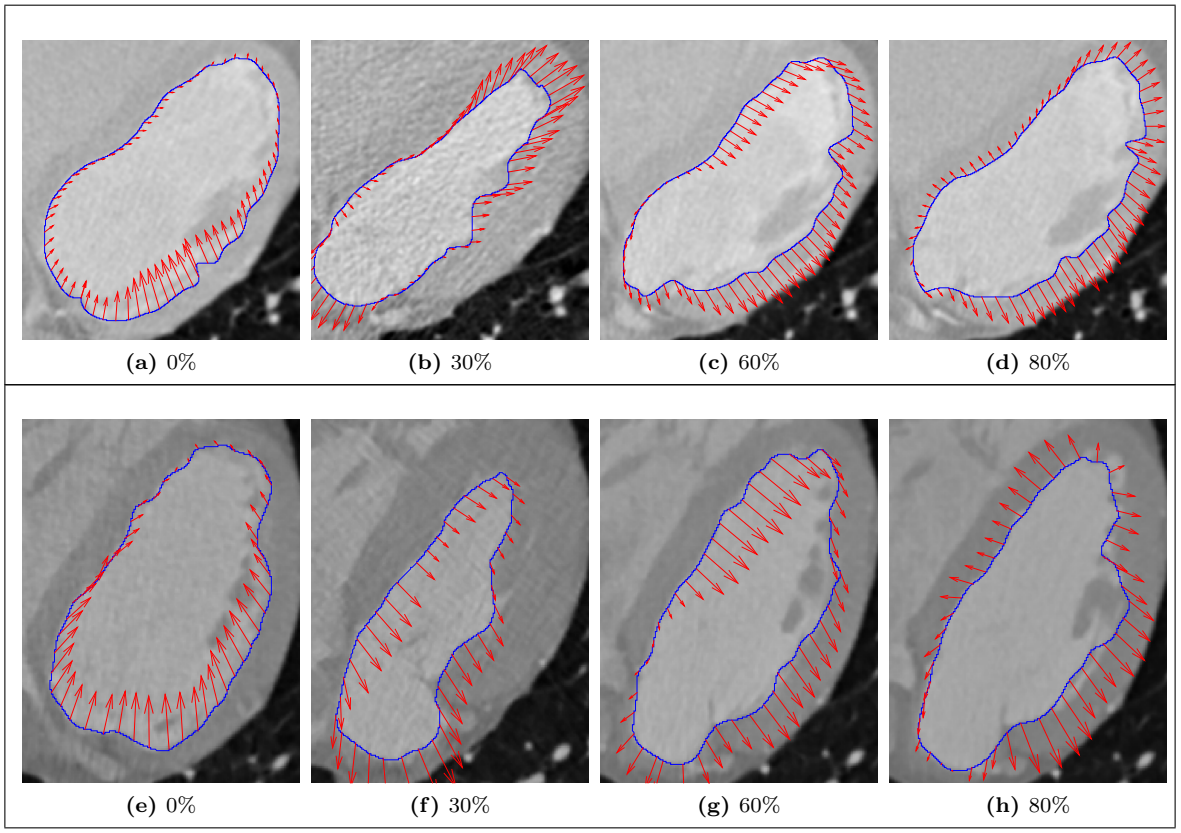


Fig. 11 *RMS* error (axis y) of the *forward reconstruction* using [32] (blue dashed line) and the SHT-based method (red solid line) for 10 sequences of the dataset. Axis x indicates a cardiac phase.



(a) **(b)** **(c)** **(d)**

Fig. 12 **(a)** Frame 10 of the *Dimetrodon* sequence. **(b)** Ground truth optical flow field. **(c)** Computed optic flow field using the OF–SHT method. **(d)** Reference color wheel.



(a) 0% **(b)** 30% **(c)** 60% **(d)** 80%

(e) 0% **(f)** 30% **(g)** 60% **(h)** 80%

Fig. 13 Combined results obtained with the segmentation and optical flow estimation methods for two sequences at 0%, 30%, 60% and 80% of the cardiac cycle. **(a)-(d)** correspond to sequence 2. **(e)-(h)** correspond to sequence 33.

cases, the segmentation error is reduced when applying the correction method. Because the correction method follows the segmentation of the statistical model, the level of correction naturally depends on the performance initially achieved with the STPDM. From Fig. 8 and table 1, it can be noted that bigger corrections are reached at the contraction phase of the cardiac cycle. It is a logical finding because the performance of the STPDM is lower at this phase. Therefore, fine details of the object boundaries are efficiently segmented with the proposed correction scheme, reducing the segmentation errors previously obtained with STPDM. Selection of optimal weight parameters is really a difficult task and there is not a standard way to do that. Since each image sequence presents different characteristics of contrast and noise, values of the parameters may require a specific configuration for each example. In this work we have found the weight parameters experimentally. However, it can be very interesting to design an automatic mechanism to select them as a future contribution. In cases when the contrast of the images is poor and the noise is very high, it is more convenient to set higher values to the parameter that attends to preserve the shape, as well as to the local intensity parameter. In the correction segmentation algorithm, these weight parameters correspond to μ and α respectively. When edges and contrast are well defined, it is preferable to give more relevance to the energies that act on the edge features directly (λ and β).

The vector field was obtained using a differential approach incorporated into the HT domain. Several local restrictions based on the SHT, non–linear constraints and a multiresolution approach for large displacements were adopted. Quantitative analysis was used to evaluate the performance. The forward reconstruction error obtained for the frames of the sequences shows better general performance with the proposed method. From Fig. 11 is clear that the best results (lowest error) were achieved during the relaxation phase of the cardiac cycle. This part of the cardiac phase presents smaller and more regular changes of the left ventricle motion. The worst result was consequently obtained in frames at the beginning of the systolic cycle.

The field vector in the optical flow estimation method is computed using the steered Hermite coefficients. Since the coefficients have been steered using the direction of maximal energy, the image noise is automatically filtered with this operation. It means, the steering property of the HT is indirectly filtering the image noise.

The proposed application, in which the vector field is visualized on the contour that encloses the left ventricle, is a promising technique for mechanical assessment of cardiac structures. Even though this technique was applied to the left ventricle, in future works it can be extended to others objects and others anatomical axes.

Since heart is a volumetric organ whose natural movement occurs in a 3D space, the logical task should be performing 3D evaluations using both segmentation and motion analysis. In the case of the segmentation, volumetric analysis can be provided by running the proposed algorithm slice by slice on each cardiac volume and then building the 3D model from the segmented slices. For the motion estimation, the task is a little more complex and implies to analyze more variables in order to achieve good 3D optical flow estimations. On the other hand, even though the heart is a structure that moves in a tridimensional space, most physicians are still evaluating the heart mechanical function using image sequences. Nevertheless, the future trend aim at providing 3D cardiac motion analysis [3] which also implies training physicians to diagnose based on moving volume and surface models. As future work, we will present an extension of the proposed framework to provide 3D evaluations.

Acknowledgements

This work has been sponsored by the following UNAM grant: PAPIIT IG100814. Leiner Barba-J thanks CONACYT for financial support, as well as Colciencias. Ernesto Moya-Albor and Jorge Brieva would like to thank the Faculty of Engineering of Universidad Panamericana (UP) for all support in this work.

Conflict of interest

None declared.

References

[1] J. M. Budoff, J. S. Shinbane, Cardiac CT Imaging, Diagnosis of Cardiovascular Disease, Springer-Verlag, London, 2006.

[2] J. P. OBrien, M. B. Srichai, E. M. Hecht, D. C. Kim, J. E. Jacobs, Anatomy of the heart at multidetector ct: What the radiologist needs to know, RadioGraphics 27 (6) (2007) 1569–1582.

[3] C. YF., The role of 3D wall motion tracking in heart failure, Nature Reviews Cardiology 9 (2012) 644–657.

[4] A. Mahnken, G. Mhlenbruch, R. Gnther, J. Wildberger, Cardiac ct: coronary arteries and beyond, European Radiology 17 (4) (2007) 994–1008.

[5] M. S. Krishnam, A. Tomasian, M. Iv, S. G. Ruehm, R. Saleh, C. Panknin, J. G. Goldin, Left ventricular ejection fraction using 64-slice ct coronary angiography and new evaluation software: initial experience, The British Journal of Radiology 81 (966) (2008) 450–455.

[6] O. Ecabert, J. Peters, H. Schramm, C. Lorenz, J. Von Berg, M. Walker, M. Vembar, M. Olszewski, K. Subramanyan, G. Lavi, J. Weese, Automatic Model-Based Segmentation of the Heart in CT images, IEEE Transactions on Medical Imaging 27 (9) (2008) 1189–1201.

[7] S. Faghih Roohi, R. Aghaeizadeh Zoroofi, 4d statistical shape modeling of the left ventricle in cardiac mr images, International Journal of Computer Assisted Radiology and Surgery 8 (3) (2012) 335–351.

[8] E. Moya-Albor, B. Escalante-Ramírez, E. Vallejo, Optical flow estimation in cardiac CT images using the steered Hermite transform, Signal Processing: Image Communication 28 (3) (2013) 267–291.

[9] J. Schaerer, C. Casta, J. Pousin, P. Clarysse, A dynamic elastic model for segmentation and tracking of the heart in MR image sequences,, Medical Image Analysis 14 (6) (2010) 738–749.

[10] D. Kang, J. Woo, P. J. Slomka, D. Dey, G. Germano, C.-C. Jay Kuo, Heart chambers and whole heart segmentation techniques: review, Journal of Electronic Imaging 21 (1).
[11] C. Petitjean, J.-N. Dacher, A review of segmentation methods in short axis cardiac MR images, Medical Image Analysis 15 (2) (2001) 169–184.
[12] T. Cootes, C. Taylor, D. Cooper, J. Graham, Active Shape Models-Their Training and Application, Computer Vision and Image Understanding 61 (1) (1995) 38–59.
[13] S. Mitchell, J. Bosch, B. P. F. Lelieveldt, R. van der Geest, J. Reiber, M. Sonka, 3-D active appearance models: segmentation of cardiac MR and ultrasound images, IEEE Transactions on Medical Imaging 21 (9) (2002) 1167–1178.
[14] H. C. Van Assen, M. G. Danilouchkine, A. F. Frangi, S. Ordás, J. J. Westenberg, J. H. Reiber, B. P. Lelieveldt, SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data, Medical Image Analysis 10 (2) (2006) 286–303.
[15] T. Heimann, H.-P. Meinzer, Statistical shape models for 3d medical image segmentation: A review, Medical Image Analysis 13 (4) (2009) 543–563.
[16] H. C. Van Assen, M. G. Danilouchkine, F. Behloul, H. J. Lamb, R. van der Geest, J. H. C. Reiber, B. P. Lelieveldt, Cardiac LV Segmentation Using a 3D Active Shape Model Driven by Fuzzy Inference, in: R. E. Ellis, T. M. Peters (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003, 6th International Conference, Montréal, Canada, November 15-18, 2003, Proceedings, Part I, Vol. 2878 of Lecture Notes in Computer Science, Springer, 2003, pp. 533–540.
[17] J. Montagnat, H. Delingette, 4d deformable models with temporal constraints: application to 4D cardiac image segmentation, Medical Image Analysis 9 (1) (2005) 87–100.
[18] G. Hamarneh, T. Gustavsson, Deformable spatio-temporal shape models: extending active shape models to 2d+time, Image and Vision Computing 22 (6) (2004) 461–470.
[19] L. Boxt, M. Lipton, Ct anatomy of the heart, in: M. Di Carli, M. Lipton (Eds.), Cardiac PET and PET/CT Imaging, Springer New York, 2007, pp. 94–109.
[20] I. Dryden, K. Mardia, Statistical Shape Analysis, Wiley series in probability and statistics, Jhon Wiley & Sons, Chichester [u.a.], 1998.
[21] B. Escalante-Ramrez, E. Moya-Albor, L. Barba-J, F. Arambula Cosio, E. Vallejo, Motion estimation and segmentation in ct cardiac images using the hermite transform and active shape models, in: Proc. SPIE 8856, Applications of Digital Image Processing XXXVI, 2013, pp. 88561E–88561E–15.
[22] B.-J. Leiner, J. Olveres, B. Escalante-Ramrez, F. Armbula, E. Vallejo, Segmentation of 4d cardiac computer tomography images using active shape models, in: Proc. SPIE 8436, Optics, Photonics, and Digital Technologies for Multimedia Applications II, 2012, pp. 84361E–84361E–11.
[23] C. Davatzikos, X. Tao, D. Shen, Hierarchical active shape models, using the wavelet transform, Medical Imaging, IEEE Transactions on 22 (3) (2003) 414–423.
[24] D. Nain, S. Haker, A. Bobick, A. Tannenbaum, Multiscale 3-d shape representation and segmentation using spherical wavelets, Medical Imaging, IEEE Transactions on 26 (4) (2007) 598–618.
[25] R. Real, J. M. Vargas, The Probabilistic Basis of Jaccard's Index of Similarity, Systematic Biology 45 (3) (1996) 380–385.
[26] C. Kurtz, P. Ganarski, N. Passat, A. Puissant, A hierarchical semantic-based distance for nominal histogram comparison, Data & Knowledge Engineering 87 (2013) 206 – 225.
[27] D. J. Williams, M. Shah, A fast algorithm for active contours and curvature estimation, CVGIP: Image Underst. 55 (1) (1992) 14–26.
[28] J. J. Gibson, The perception of the visual world, The American Journal of Psychology, 64 (1951) 622–625.
[29] B. K. P. Horn, B. G. Schunck, Determining Optical Flow, Artificial Intelligence 17 (1-3) (1981) 185–203.
[30] A. Bruhn, J. Weickert, C. Schnörr, Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods, International Journal of Computer Vision 61 (3) (2005) 211–231.
[31] J. Weickert, C. Schnörr, A Theoretical Framework for Convex Regularizers in PDE-Based Computation of Image Motion, International Journal of Compututer Vision 45 (3) (2001) 245–264.
[32] N. Papenberg, A. Bruhn, T. Brox, S. Didas, J. Weickert, Highly Accurate Optic Flow Computation with Theoretically Justified Warping, International Journal of Computer Vision 67 (2) (2006) 141–158.
[33] M. J. Black, P. Anandan, The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields, Computer Vision and Image Understanding 63 (1) (1996) 75–104.
[34] M. Qiu, Computing optical flow based on the mass-conserving assumption, Pattern Recognition, International Conference on 3 (2000) 7041.
[35] M. Dawood, F. Gigengack, X. Jiang, K. P. Schfers, A mass conservation-based optical flow method for cardiac motion correction in 3d-pet, Medical Physics 40 (1) (2013) –.
[36] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud, Two deterministic half-quadratic regularization algorithms for computed imaging, in: IEEE International Conference on Image Processing, Vol. 2, IEEE Computer Society Press, 1994, pp. 168–172.
[37] B. Escalante-Ramírez, J. L. Silván-Cárdenas, Advanced modeling of visual information processing: A multiresolution directional-oriented image transform based on Gaussian derivatives, Signal Processing: Image Communication 20 (9-10) (2005) 801–812.
[38] J. L. Silván-Cárdenas, B. Escalante-Ramírez, The Multiscale Hermite Transform for Local Orientation Analysis, IEEE Transactions on Image Processing 15 (5) (2006) 1236–1253.
[39] R. A. Young, The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line–weighting profiles, Tech. Rep. GMR-4920, General Motors Research Laboratories, Detroit, Mich, USA (1985).
[40] J.-B. Martens, The Hermite Transform–Theory, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (9) (1990) 1595–1606.
[41] A. M. van Dijk, J.-B. Martens, Image representation and compression with steered Hermite transforms, Signal Processing 56 (1) (1997) 1–16.
[42] B. Ohnesorge, T. Flohr, Principles of Multi-slice Cardiac CT Imaging, in: Multi-slice and Dual-source CT in Cardiac Imaging, 2nd Edition, Springer Berlin Heidelberg, 2007, pp. 71–126.
[43] T. Lin, J. L. Barron, Image Reconstruction Error for Optical Flow, in: Vision Interface, 1994, pp. 73–80.
[44] J. L. Barron, D. J. Fleet, S. S. Beauchemin, Performance of Optical Flow Techniques, International Journal of Computer Vision 12 (1) (1994) 43–77.