Active Shape Model based on a spatio-temporal a priori knowledge: applied to left ventricle tracking in scintigraphic sequences

Said Ettaieb *settaieb@gmail.com*
Research laboratory of image, signal and information Technology - University of Tunis El Manar ROMMANA 1068, Tunis-B.P. n° 94, Tunisia

Kamel Hamrouni *kamel.hamrouni@enit.rnu.tn*
Research laboratory of image, signal and information Technology - University of Tunis El Manar ROMMANA 1068, Tunis-B.P. n° 94, Tunisia

Su Ruan *su.ruan@univ-rouen.fr*
LITIS-Quantif University of Rouen 22 boulevard Gambetta 76183 Rouen, France

Abstract

The Active Shape Model – ASM is a class of deformable models that relies on a statistical a priori knowledge of shape for the segmentation of structures of interest [5]. The main contribution of this work is to integrate a new a priori knowledge about the spatio-temporal shape variation in this model. The aim is to define a new more stable method, allowing the reliable detection of structures whose shape changes considerably in time. The proposed method is based on two types of a priori knowledge: spatial and temporal variation of the shape of the studied structure. It was applied first on synthetic sequences then on scintigraphic sequences for tracking the left ventricle of the heart. The results were encouraging.

Keywords: Active shape model, a priori knowledge, spatio-temporal shape variation, scintigraphic sequences.

1. INTRODUCTION

The deformable models [1-6] are certainly the most popular approach in the field of medical images segmentation, due to their flexibility and ability to integrate a priori knowledge about the anatomical structures. The basic idea is to start with an initial coarse segmentation that will evolve gradually, according to several constraints, towards the target contours. These models have the advantage of segmenting an image by integrating a global vision of the shape of the structure to be extracted. They are widely studied and applied to the static segmentation of rigid structures, whether in the 2D case or the 3D case [7]. However, in some medical applications, it is sometimes necessary to follow up the spatio-temporal variation of non-rigid structures, whose shape varies over time. In this aim, several extensions of deformable models were proposed. For example, in [8], the authors propose to track anatomical structures in sequences of images by active contour [1] whose initialization in the image i is deduced automatically from the previous result in the image $i-1$. In several other works [9-13], the sequence of images is treated in a global way and the studied shape variation is described by a single model that evolves over time. The majority of these works is focused mainly on the spatio-temporal tracking of cellular structures [7, 9] and the left ventricle of the heart [10, 11, 14, 1 5].

However, despite the success obtained in some cases, the quality of the results depends on the initialization step and the choice of propagation parameters. In addition, the used a priori knowledge has generally a global criterion.

The active shape model (ASM) is a particular class of deformable models, introduced by Cootes et al. [5] in order to extract complex and non-rigid objects. This model has two major advantages compared to the other classes of deformable models. On one hand, the initialization is a mean shape of the structure to be segmented. Thus, it will be very close to the target structure during the localization step, which affects advantageously the accuracy of the result. On the other hand, the progressive evolution of the initialization is guided by a statistical shape model that describes the geometry and the authorized deformation modes of the aimed structure. This reduces the solutions space and leads always to acceptable shapes. However, if the structure to be segmented changes considerably over time, these two advantages lose much of their interest. Because, if the shape variation is very important, the mean shape becomes more general and the statistical shape model becomes less precise. Thus, there might be a generation of shapes that is far from the target structures. In order to improve the precision of the active shape model in the case of segmentation of structures whose shape changes significantly over time, we suggest incorporating a new a priori knowledge about the spatio-temporal variation of shape into this model. Indeed, we propose to model the spatial variation of the studied structure over time in order to define a statistical spatio-temporal shape model. This model, which has to describe precisely the shape and the deformation modes of the studied structure at every moment, will be then used to guide a spatio-temporal localization stage to segment a sequence of images.

In this paper, we will explain first the steps of the proposed method. Then we will show its application on synthetic sequences and on real sequences of scintigraphic images. A comparative study between a ground truth drawn by an expert, the ASM and the ASMT, will also be established in order to deduce the interest from the integration of an a priori knowledge on the spatio-temporal shapes variation.

2. PROPOSED METHOD

Given a structure whose shape changes over time. At each instant t, it may take a different shape from that taken at an instant t_1. We suppose that this variation according to time is represented by a sequence of images. The aim is the automatic localization, in the most reliable way, of this structure in all images of the sequence at the same time.
The proposed method can be described by figure 1.

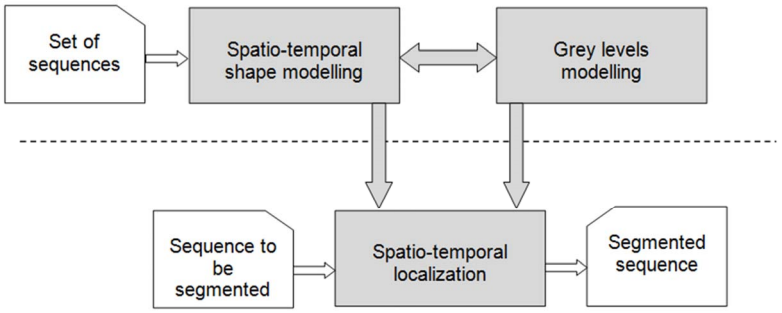


FIGURE 1: Proposed method.

This method requires three main stages: a stage of spatio-temporal shape modelling, a stage of grey levels modelling and a stage of spatio-temporal localization which is based on the results of both first ones to locate the target structure in a new sequence.

2.1 Stage of spatio-temporal shape modelling

The objective of this stage is to build a statistical spatio-temporal shapes model which describes exactly the variation over time of the non-rigid structure to be segmented.

It requires, first of all, the preparation of a spatio-temporal training set, which includes all the possible configurations of this structure.

Step 1: Preparation of a spatio-temporal training set
First, we have to collect a set of sequences of different images, reflecting the possible variations of the studied structure. Every sequence must contain the same structure in the same period of time and with the same number of images. Then, we have to extract the spatio-temporal shapes by putting, carefully, on the contour of every sequence, a sufficient number of landmark points on the wished contour. Each sequence is well modelled by a spatio-temporal shape which describes both the spatial and temporal variation of the studied structure. Given F the number of images by sequence and L the number of landmark points put on each image, the spatio-temporal shape that models a sequence i can then be represented by a vector $\mathrm{S_i}$, constructed by concatenating the coordinates of the points defined on the contours of the studied structure through all the images in the sequence:

$$\mathrm{S_i} = [\mathrm{u_{i1}}, \mathrm{u_{i2}}, \mathrm{u_{i3}}, \ldots\ldots \mathrm{u_{iF}}] \qquad (1)$$

with $\mathrm{u_{ij}} = [\mathrm{x_{ij1}}, \mathrm{y_{ij1}}, \mathrm{x_{ij2}}, \mathrm{y_{ij2}}, \ldots . \mathrm{x_{ijL}}, \mathrm{y_{ijL}}]$ is a vector that models the J^{th} shape in the i^{th} sequence. Thus, the spatio-temporal training set will be represented by a set of spatio-temporal shapes: $\{S_i \text{ with } i = 1 \ldots . N \ \ (\mathrm{N\ number\ of\ sequences})\}$

Step 2: Aligning spatio-temporal shapes
After extracting spatio-temporal shapes from samples of sequences, an alignment step of shapes is required in order to put the corresponding vectors $\{S_i\}$ at a centered position. This allows to eliminate the problem of variation in position and in size and to study only the most important variation in shape between the various configurations of the studied structure. The alignment procedure of the spatio-temporal shapes has the same idea of shapes alignment in the ASM [16]. First, it consists in taking, randomly, a spatio-temporal shape on which are aligned all the others. Then, in every iteration, a mean spatio-temporal shape is calculated, normalized and on which the others will be realigned. This process is stopped when the mean spatio-temporal shape reach some stability.

Step 3: Generation of the statistical spatio-temporal shapes model

The aligned vectors$\{S_i\}$, resulting from the two previous steps, can be arranged in an observation matrix whose size is $(2\mathrm{lF}, \mathrm{N})$. This matrix describes both the spatial and temporal variation of the shape of the studied structure. The columns represent the temporal variation while the lines represent the spatial variation at each instant t. The aim is to deduce from this matrix, the modes and the amplitudes of the spatio-temporal variation of the studied structure. Using the same principle of the ASM, this can be done by applying principal component analysis (PCA) on the raw data. Indeed, the main modes of spatio-temporal variation of the studied structure will be represented by the principal components deduced from the covariance matrix $\mathrm{C_s}$ associated with the observation matrix (equation 2).

$$\mathrm{C_s} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{dS_i dS_i^t} \qquad (2)$$

where $\mathrm{dS_i} = \mathrm{S_i} - \overline{\mathrm{S}}$ is the deviation of the i^{th} spatio-temporal shape S_i compared to a mean spatio-temporal shape, that is calculated:

$$\overline{\mathrm{S}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{S_i} \qquad (3)$$

These principal components are given by the eigenvectors of the matrix $\mathrm{C_s}$, such as:

$$\mathrm{C_s P_k} = \lambda_{\mathrm{k}} \mathrm{P_k} \qquad (4)$$

$\mathrm{P_k}$ is the K^{th} eigenvector of $\mathrm{C_s}$ and λ_{k} is the corresponding eigenvalue. Each vector represents a variability percentage of the variables used to build the covariance matrix.

The variability percentage represented by each vector is equal to its corresponding eigenvalue. In general, we can notice a very fast decreasing of the eigenvalues, which is used to classify the corresponding vectors in decreasing order.

Therefore, we can choose the first t eigenvectors, which represent the important variability percentage, as principal components. Every spatio-temporal shape S can be simply represented by the mean spatio-temporal shape and a linear combination of principal components (main deformation modes):

$$S = \bar{S} + Pb \qquad (5)$$

where $\bar{S}$ is the mean spatio-temporal shape, $P = (p_1, p_2, p_3 \ldots\ldots, p_t)$ is the base of t principal components and $b = (b_1, b_2, b_3 \ldots\ldots, b_t)^t$ is a weight vector representing the projection of the spatio-temporal shape S in the base P. Generally, the amplitude of the allowable deformation following a principal component P_k is limited as follows:

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \qquad (6)$$

As a result, from the basic equation (Equation 5) we can deduce infinity of shapes describing the spatio-temporal studied structure by choosing correctly the b_k values (equation 6). Equation 5 defines, then, the statistical spatio-temporal shapes model, which defines an allowable deformation space for spatio-temporal studied structure.

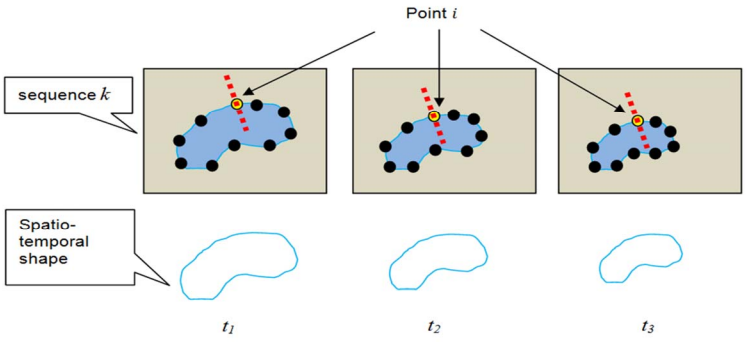
This model will be used in the spatio-temporal localization stage to guide the evolution in such a way that it is only in the allowable space. Finally, this first step can be defined by the following functional algorithm:

Algorithm 1 : Spatio-temporal shape modelling

1. Enter the initial parameters:
   Number of sequences of the training set : N
   Number of images by sequence: F
   Number of landmark points : L
   Variability percentage to be represented: P
2. Extract manually the spatio-temporal shapes: $\{S_i\}$ / *i=1... N*
4. Align of the spatio-temporal shapes
5. Generate of the statistical spatio-temporal shapes model by PCA : $S = \bar{S} + Pb$

2.2 Stage of grey levels modelling

As for the ASM, in the stage of spatio-temporal localization, the proposed method is based on intensities information of the treated sequence. It is about finding an optimal correspondence between the properties of luminance of the treated sequence with information of luminance collected from sequences samples. For that purpose, in addition to the spatio-temporal shape modelling, it is necessary to model the grey levels information from the training sequences. For example, k is a sequence in the training set that is composed of three images. This sequence represents the variation over time of a structure. We suppose that ten landmark points are sufficient to extract the shape of the structure presented in every image (Figure 2).



GURE 2: Grey levels modelling

The grey levels modelling consists in extracting, for each landmark point i (yellow point) on each image j of the sequence k and then through all the training sequences, the grey levels profile g_{ijk} from a segment of length n (red segment), centered in this point i and carried by its normal:

$$g_{ijk} = [g_{ijk0}, g_{ijk1}, g_{ijk2}, \dots, g_{ijkn-1}] \qquad (7)$$

g_{ijkt} $(t = 0 \dots n-1)$ is the grey level of the t^{th} pixel of the examined segment.
The derivative of this grey levels profile is defined by the expression 8:

$$dg_{ijk} = [g_{ijk1} - g_{ijk0}, g_{ijk2} - g_{ijk1}, \dots, g_{ijkn-1} - g_{ijkn-2}] \qquad (8)$$

dg_{ijk} is a vector of size $(n-1)$, including the differences in grey levels between two successive points of the examined segment. The normal derivative of this profile is defined by the expression 9:

$$y_{ijk} = \frac{dg_{ijk}}{\sum_{k=0}^{n-2} |dg_{ijkt}|} \qquad (9)$$

Through all the images in a sequence and then through all the training sequences, we can define for each landmark point i, a mean normal derivative of the grey levels given by the expression 10:

$$\bar{y}_i = \frac{1}{FN} \sum_{j=1}^{F} \sum_{k=1}^{N} y_{ijk} \qquad (10)$$

This mean normal derivative related to the point i , will be used in the stage of spatio-temporal localization to move the same point towards a better position. The stage of grey levels modelling can be summarized in by the following functional algorithm:

Algorithme 2 : Grey levels modelling

Enter the initial parameters :
  Length (in points) of the grey levels profile: n
  Number of sequences of the training set : N
  Number of images by sequence: F
  Number of landmark points: l
Calculate the mean normal derivative for each landmark point :
For i from 1 to l do
  For j from 1 to F do
    For k from 1 to N do
      Extract the profile : $\mathbf{g_{ijk}}$
      Calculate the derivative **:** $\mathbf{dg_{ijk}}$
      Calculate the mean derivative : $\mathbf{y_{ijk}}$
      Add $\mathbf{X_i = X_i + y_{ijk}}$
    End for
  End for
  Calculate the mean normal derivative : $\mathbf{\bar{y}_i = \frac{X_i}{FN}}$
End for

2.3 Stage of spatio-temporal localization

The objective now is to bound the studied structure in a new sequence. A way to achieve this is to start with an initial spatio-temporal shape, which will gradually evolve towards the contours of the studied structure in all images simultaneously. This idea can provide a procedure of spatio-temporal localization, which consists in repeating iteratively the following four steps:

Step 1: Initialization
This step consists in putting an initial spatio-temporal shape S_i on the treated sequence. This shape can be built from a spatio-temporal shape S_a belonging to the training set:

$$S_i = M(k_i, \theta_i)[S_a] + t_i \qquad (11)$$

with $M(k_i, \theta_i) = \begin{bmatrix} k_i \cos\theta_i & -k_i \sin\theta_i \\ k_i \sin\theta_i & k_i \cos\theta_i \end{bmatrix}$ a matrix (2*2)
$t_i = (t_{xi}, t_{yi}, t_{xi}, t_{yi}, t_{xi}, t_{yi} \ldots\ldots t_{xi}, t_{yi})$ a translation vector of size 2*F*N
k_i, θ_i and t_i are respectively the homothety, rotation and translation to be applied to every point of S_a in order to build the initialization S_i.

Step 2: Search for the elementary movement
Having fixed an initial spatio-temporal shape S_i, the objective is to determine the elementary movement dS in order to slide the landmark points towards a better position; by using the grey levels characteristics. We will first address the problem of moving a single landmark point. Then, we will show how to calculate the elementary movement of the initial estimate dS_i.
Indeed, A is a particular landmark point of S_i (Figure 3).

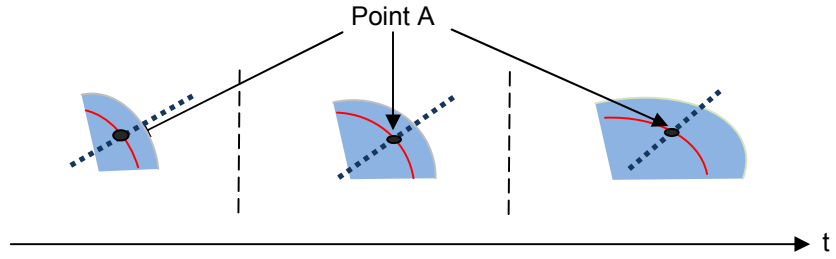


FIGURE 3: Movement of a landmark point A. The structure to be located is in blue. The initial spatio-temporal shape is the red curves.

To move the point A to the borders of the studied structure, the idea is to extract from each image j of the processed sequence, a search grey levels profile of length m pixels (with $m >> (n-1)$) which is centered in A and supported by the normal to the edge passing through this point (black segment).
Then, the point A will be represented by a matrix H_A, defined as follows:

$H_A =$

	1	**2**	**…**	**F**
1	g_{A11}	g_{A21}	$\cdots$	g_{AF1}
2	g_{A12}	g_{A22}	$\cdots$	g_{AF2}
3	g_{A13}	g_{A23}	$\cdots$	g_{AF3}
$\vdots$	$\vdots$	$\vdots$	(g_{Ajk})	$\vdots$
m	g_{A1m}	g_{A2m}	$\cdots$	g_{AFm}

H_A is a matrix of $m * F$ combinations where each column represents the search profile on the image j. (g_{Ajk} is the level grey of the k^{th} pixel on the segment passing by the point A in the j^{th} image). Knowing that each landmark point A is defined by a mean normal derivative of the grey levels $\bar{y}_A$ (information calculated from sequences samples during the stage of grey levels modelling), we can calculate, from the matrix H_A, a new matrix $H_A'(j, l)$ which represents the difference between the grey level information surrounding the current point A and that related to the same point during the grey levels modelling.

This matrix can be defined by the expression (12):

$$H_A'(j,l) = (g_{Aj}(l) - \bar{y}_A)^t I_{n-1}^t I_{n-1} (g_{Aj}(l) - \bar{y}_A) \quad (12)$$

with j from 1 to F, l from 1 to m and I_{n-1} is the identity matrix $(n-1)$.
$g_{Aj}(l)$ is the sub-profile of length $(n-1)$ centered at the l^{th} position of the search profile g_{Aj} that contains the normal derivative of the intensities. (It is necessary to remind that $g_{Aj}(l)$ and $\bar{y}_A$ have the same size $(n-1)$). The best positions to which has to slide the point A on the treated sequence are given by the expression (13):

$$\{ p_{Aj} = [\min(H_A'(j,l))] \text{ with } j \text{ from 1 to F and } l \text{ from 1 to m} \} \quad (13)$$

p_{Aj} : is the position to which has to slide the point A on the j^{th} image.
Therefore, we can calculate the elementary movements of the particular point A in all the images of the sequence, such as:

$$\{ dp_{Aj} = \mathbf{distance}(A, p_{Aj}) \text{ with j from 1 to F } \} \quad (14)$$

dp_{Aj} : is the elementary movement of point A in the j^{th} image.
distance : is the Euclidean distance between the point A and the position p_{Aj}.

By applying the same principle for the other landmark points of the initial spatio-temporal shape S_i, we can deduce finally the elementary movement dS_i :

$$dS_i = [dp_{Aj}] \text{ avec A from 1 to L and j from 1 to F} \quad (15)$$

where:
L : Number of landmark points.
F : Number of images by sequence.

Step 3: Determining the parameters of position and shape
After determining the elementary movement dS_i, we must now determine the parameters of position and shape to make this movement, while respecting the constraints of spatio-temporal deformation imposed by the modelling stage.

- Determining the position parameters:

We suppose that the initial estimate S_i is centered in a position (x_c, y_c) with an orientation θ and an homothety k. Determining the position parameters means determining the parameters of geometric operations $1 + dk$, $d\theta$ and $dt = (dx_c, dy_c)$ to be applied to each point of S_i in order to reach the new position $(S_i + dS_i)$. A simple way to determine these parameters is to align the two vectors S_i and $(S_i + dS_i)$ [2].

- Determining the shape parameters:

Once the position parameters ($1 + dk$, $d\theta$ and dt) are known, it remains to determine the shape parameters. That is to say, if we suppose that the initial estimate S_i is defined in the base of the principal components by a weight vector b, we seek to determine the variation db in order to trace $(S_i + dS_i)$ in the same base. Given that the initial estimate is built from a spatio-temporal shape S_a belonging to the training set ($S_i = M(k_i, \theta_i)[S_a] + t_i$), determining the shape parameters db is to solve first in dx the following equation:

$$M(k_i(1 + dk), \theta_i + d\theta)[S_a + dx] + t_i + dt = S_i + dS_i \quad (16)$$

which means

$$M(k_i(1 + dk), \theta_i + d\theta)[S_a + dx] = S_i + dS_i - (t_i + dt) \quad (17)$$

but we have $S_i = M(k_i, \theta_i)[S_a] + t_i$

If we replace S_i by its value in the equation (17), we find

$$M(k_i(1+dk), \theta_i + d\theta)[S_a + dx] = M(k_i, \theta_i)[S_a] + \cancel{t_i} + dS_i - (\cancel{t_i} + dt) \quad (18)$$

But we know that

$$M^{-1}(k, \theta)[\ldots.] = M(k^{-1}, -\theta)[\ldots.] \quad (19)$$

By applying this rule to the equation (18), we obtain

$$S_a + dx = M((k_i(1+dk))^{-1}, -(\theta_i + d\theta))\ [M(k_i, \theta_i)[S_a] + dS_i - dt] \quad (20)$$

what means that

$$dx = M((k_i(1+dk))^{-1}, -(\theta_i + d\theta))\ [M(k_i, \theta_i)[S_a] + dS_i - dt] - S_a \quad (21)$$

dx is determined in $2 * L * F$ size. However, we have t modes of variation. Then, we have to calculate dx', the projection of dx in the base of principal components P. This can be done by adopting the approach of least squares [17]. Indeed, $dx' = wdx$ with $w = P(P^tP)^{-1}P^t$ is a projection matrix. However, the principal components of P are pairwise orthogonal, meaning that $P^tP = I$. This, then, gives $dx' = PP^tdx$. We know that $dx' = Pdb$, if we multiply both sides of this equation by P^t, we can deduce finally the shape parameters $db = P^tdx'$.
$db = (db_1, db_2, db_3, \ldots. db_t)$ is a weight vector allowing to build and to limit the new vector $(S_i + dS_i)$ in the base of principal components (main modes of deformation).

- Movement of the spatio-temporal shape and the limitation of the shape parameters:

This last step consists in moving S_i to the new position ($S_i + dS_i = S_i^1$), by using the already calculated parameters.
We obtain

$$S_i^1 = M(k_i(1+dk), \theta_i + d\theta)[S_a + Pdb] + t_i + dt \quad (22)$$

We should note that the shape parameters $db = (db_1, db_2, db_3, \ldots. db_t)$ must be limited in the allowable intervals of variation defined by the equation (6), to produce acceptable spatio-temporal shapes. Indeed, if for example a value db_k $(1 \le k \le t)$ exceeds the maximum value in a component k, it will be limited as follows:

$$\begin{cases} \text{if } db_k > v_{max_k} \\ \text{then } db_k = v_{max_k} \\ \text{if } db_k < -v_{max_k} \\ \text{then } db_k = -v_{max_k} \end{cases} \quad (23)$$

with $v_{max_k} = 3\sqrt{|\lambda_k|}$ is the maximum value of allowable variation following the component k. λ_k is the eigenvalue related to the component k. Now, from S_i^1, we will repeat the same steps to build S_i^2 then S_i^3… and so on, until no significant change is detected or the maximum number of iterations is reached. The stage of spatio-temporal localization can be described by the functional algorithm 3.

Algorithme 3 : Stage of spatio-temporal localization

Initialize of a spatio-temporal shape : S_i
While (convergence==false and $i < nbr\_max\_iterations$)
  Search of elementary movement: dS_i
  Determine the parameters of position and shape :
    $1 + dk, d\theta, dt$ and db
  Movement of the spatio-temporal shape and limitation of the shape parameters:
    $-v_{max_k} \leq db_k \leq v_{max_k}$
    $S_{i+1} = M(k_i(1 + dk), \theta_i + d\theta)[S_i + Pdb] + t_i + dt$
  Convergence=compare (S_i, S_{i+1})
  i=i+1
End While

3. EXPERIMENTAL RESULTS

The proposed method is designed mainly for tracking the spatio-temporal variation of the left ventricle in scintigraphic sequences of images of the heart. But before moving on to this real application, we chose to test the performance of our method on synthetic sequences, in order to deduce its effectiveness in an ideal case.

3.1 Validation on synthetic data

First, we built a database of synthetic sequences, which will serve as a spatio-temporal training set. We synthesized ten sequences disturbed by Gaussian noise which parameters are: m = 0.2 and v = 0.1. Each sequence contains six images of 256 * 256 pixels, simulating the variation over time of a simple shape (figure 4.a). During the stage of spatio-temporal shape modelling, thirty landmark points were put on each image to extract the studied shape. Each sequence is modelled by a spatio-temporal shape which size is 2*30*6 (figure 4.b). The variability percentage to be represented is fixed to 95%. The length of the grey levels profile in the modelling stage is 7 pixels, and in the localization stage, the length of search profile is 19 pixels. The maximum number of iterations is fixed to 40. Figure 5 shows an example of a result of spatio-temporal localization obtained on a synthetic test sequence.

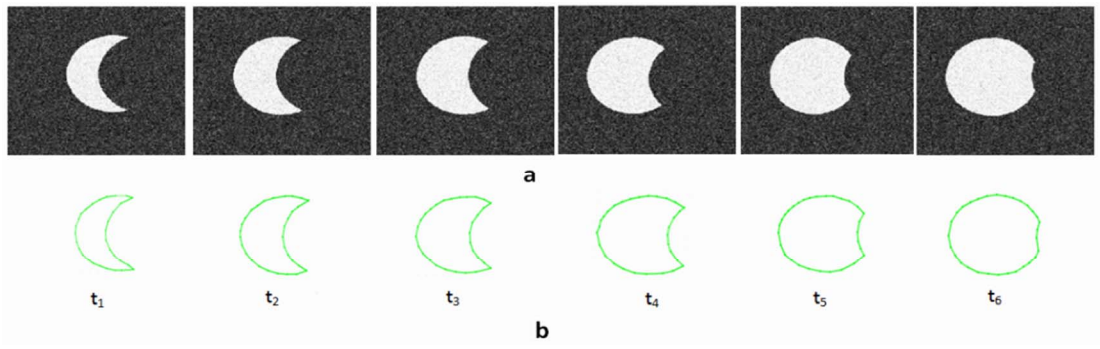


FIGURE 4: (a) Example of a sequence of the synthetic set simulating the variation over time of a simple shape (moon). (b) Corresponding spatio-temporal shape.

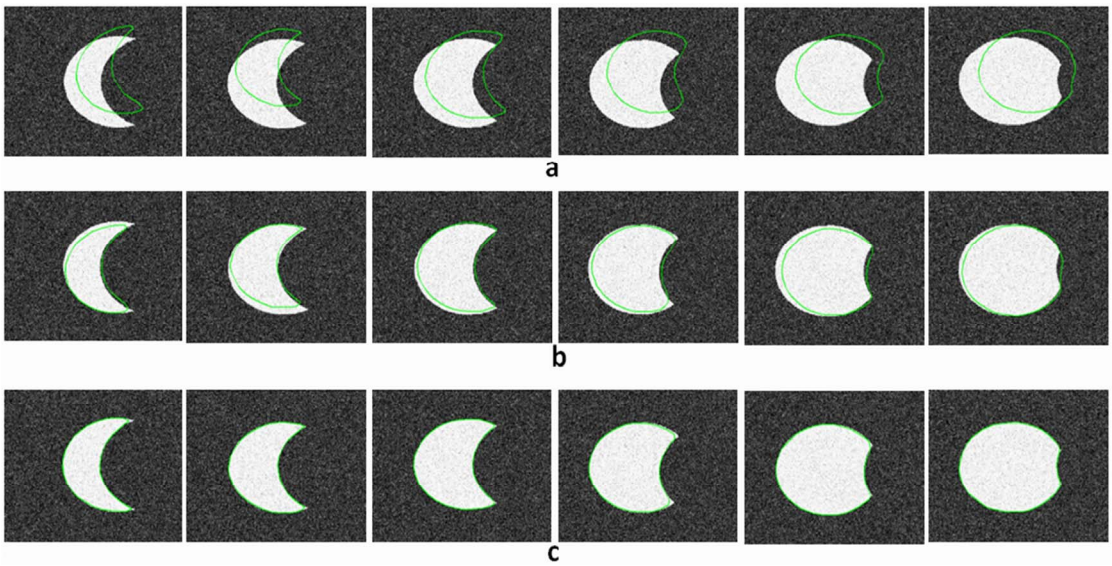


FIGURE 5: Result of spatio-temporal localization obtained on a synthetic sequence. (a) Initialization of the mean spatio-temporal shape. (b) Result obtained after 10 iterations. (c) Final localization result after 15 iterations.

Not surprisingly, we note that the spatio-temporal shape arrived to correctly locate the target shape in all images of the sequence. The accuracy of this result can be explained by two points. On the one hand, the target contours are quite clear and thus easily detectable.
On the other hand, we can say that the spatio-temporal shape modelling provided very accurate information about the studied shape at each moment. The mean spatio-temporal shape used as initialization is very close (in terms of shape) to the target structure on each image, which improves the accuracy of the result. This result shows clearly that our method works in a simple synthetic case. Further, we will apply it in a real case where even the manual tracing of the target contours is difficult.

3.2 Tracking of the left ventricle in scintigraphic sequences of images of the heart

- Background:

The heart is a hollow muscle in the middle of the chest, whose role is to circulate cyclically the blood in the body. In particular, the left ventricle is considered as the main pumping chamber of the heart, because of its great pushing force of the blood through the body against the body pressure. In clinical practice, the study of the function of the heart pump then requires necessarily to follow up the ventricle contraction, during the cardiac cycle in order to estimate the quantity of blood pumped during the corresponding time interval. However, this task is not easy to accomplish, especially for the scintigraphic images. This medical imaging modality is characterized by a low contrast and a low resolution, where even the manual tracing of the target contours is difficult. In this context, several methods are proposed to mark out the left ventricle on scintigraphic images. We can distinguish two types of methods: region-based segmentation [18, 19, 20, 21, 22] and contour-based segmentation [23, 24, 25, 26, 27, 28, 29]. All these works show that the marking out of the left ventricle in scintigraphic images is a difficult task. The results often depend on the parameters of the used method. We can conclude the importance of using a strong a priori knowledge about the studied organ's physiology and anatomy, in order to develop an effective segmentation method. Another important finding is that the developed methods treat the scintigraphic images sequences that represent the variation of left ventricle over time image by image. This fact is severely affecting the quality of the overall result on an entire sequence. For that reason, we thought to exploit a priori knowledge about the spatio-temporal shape variation for the marking out of the left ventricle.

- Experimentation:
The application of our contribution for tracking the left ventricle requires first the preparation of a spatio-temporal training set, which includes all possible configurations. The image database used in this work was provided by the department of nuclear medicine at the Institute Salah Azeiz in Tunis. This database contains 25 scintigraphic sequences of images of the heart from 25 different patients. Each sequence shows a heart beat cycle, represented by 16 images of 128 * 128 pixels. After showing the collected sequences to a specialist doctor, we concluded that the details of the left ventricle can be represented by 20 landmark points. We noticed that if we work on the 25 sequences, the doctor has to put manually 20 * 16 * 25 = 8000 landmark points! This is a tedious task. So, we have to reduce then the number of sequences in the database without losing the concept of variability.

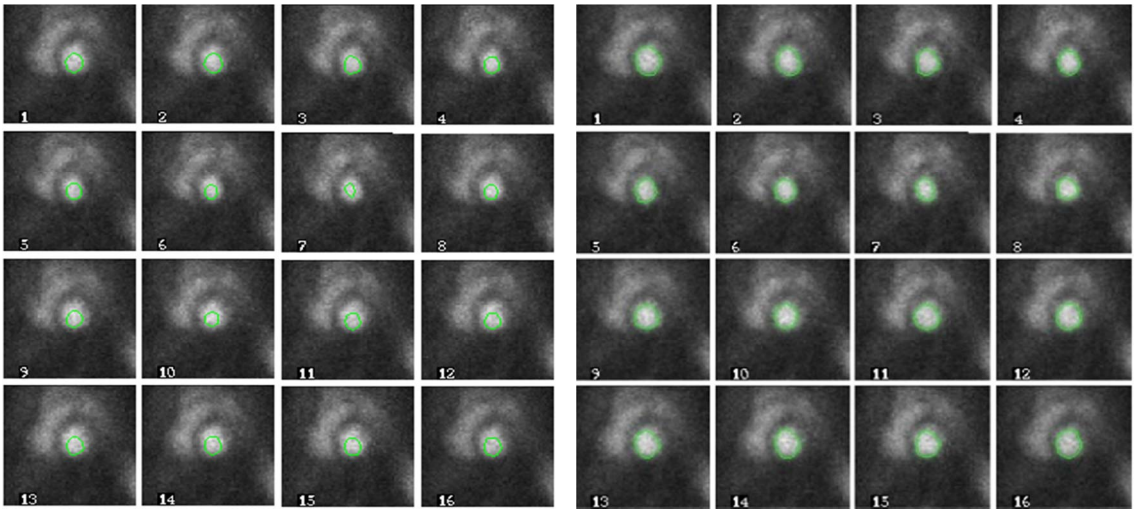
To solve this problem, we propose to apply a selecting strategy of sequences to build an optimal spatio-temporal training set. We thought then to apply a classification of all the collected sequences. This classification aims to group sequences according to the shape of the left ventricle. To do this, we selected the same image of each sequence. Each of these images represents then a sequence. The entire population can be then classified into a small number of classes. Each class represents a mode of shape variability. To achieve this goal, we propose to move on from the field of parametric curves (contours of objects) to a field of invariants, where the invariance by translation, rotation and scaling factor is maintained. Thus, every shape will be represented by a set of invariants. The difference between the invariants of two different instances represents the difference between the natures of the shapes themselves. Therefore, to classify the population, it is sufficient to classify the vectors of invariants corresponding to each shape. For the invariants, we chose the Fourier descriptors known for their performance in the field. And for the classifiers, we chose the K-Means classifier, both for its simplicity and its performance. The results of the classification of the 25 collected sequences are given in the table 1:

Sequence	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
Class	2	2	2	2	2	2	2	3	1	4	3	2	3
Sequence	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	
Class	4	2	3	4	3	2	1	2	2	3	4	1	

TABLE 1: Results of the classification of 25 database sequences

The population was thus reduced to four classes. Then, we have just to choose arbitrarily two sequences from each class. Therefore, the used spatio-temporal training set is composed of 8 sequences.

The next step consists in extracting from each sequence, the spatio-temporal shape which represents both the spatial and temporal variation of the left ventricle of a heart cycle. This is done by placing 20 landmark points on the left ventricle contour through all images in all sequences. The result of this step is to obtain 8 vectors, each of size 2x20x16 = 640. After aligning these vectors and using 95% as a variability percentage parameter, the application of PCA on these data provided five main modes of variability. The length of the grey levels profile in the modelling stage is 7 pixels, and in the localization stage, the length of profile search is 19 pixels. The maximum number of iterations is fixed at 60 iterations. The used test sequences are selected from the 17 remaining sequences in the original database. Figure 6 shows an example of the result of the spatio-temporal localization that is obtained on a test sequence.

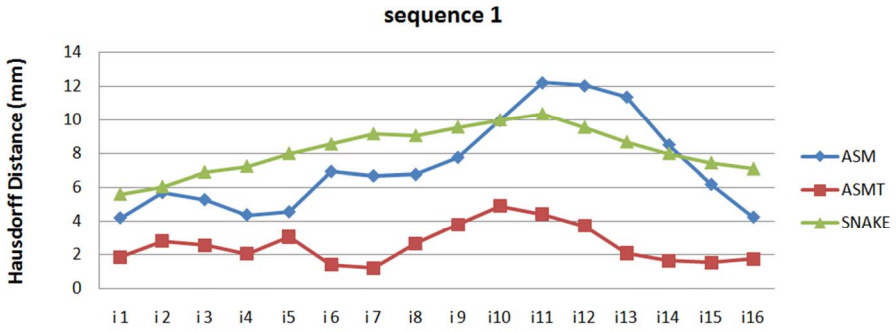


(a) (b)

FIGURE 6: Result of the spatio-temporal localization of the LV. (a) Initialization of the mean spatio-temporal shape on a treated sequence. (b) Final result of the localization.

We note that the final spatio-temporal shape succeeded generally in locating the shape of the left ventricle. This can show the performance of the method even in the presence of contours that are difficult to identify. This result is qualitatively considered satisfactory by the medical specialists. However, we should establish a quantitative precise evaluation of the results. In our case, we have four sequences that are manually segmented by a radiologist. In order to deduce the interest from the integration of an a priori knowledge about the spatio-temporal variation of shape, we chose to compare our method ASMT with the ground truth, the basic model ASM and with another method that is proposed by Fekir and al. [8]. This method allows the tracking of non-rigid objects in sequences of images using active contour SNAKE [1] whose initialization in the image i is automatically deduced from the result in the image $i-1$. Since the compared methods are contour-based methods, we chose the Hausdorff distance as a measure of segmentation quality [30]. This metric is widely used in multiple applications of the medical field. In our case, we use this distance to measure the similarity between two shapes.

Figure 7 shows the Hausdorff distance between each method (ASM, ASMT and SNAKE) and the reference segmentation of the four sequences.



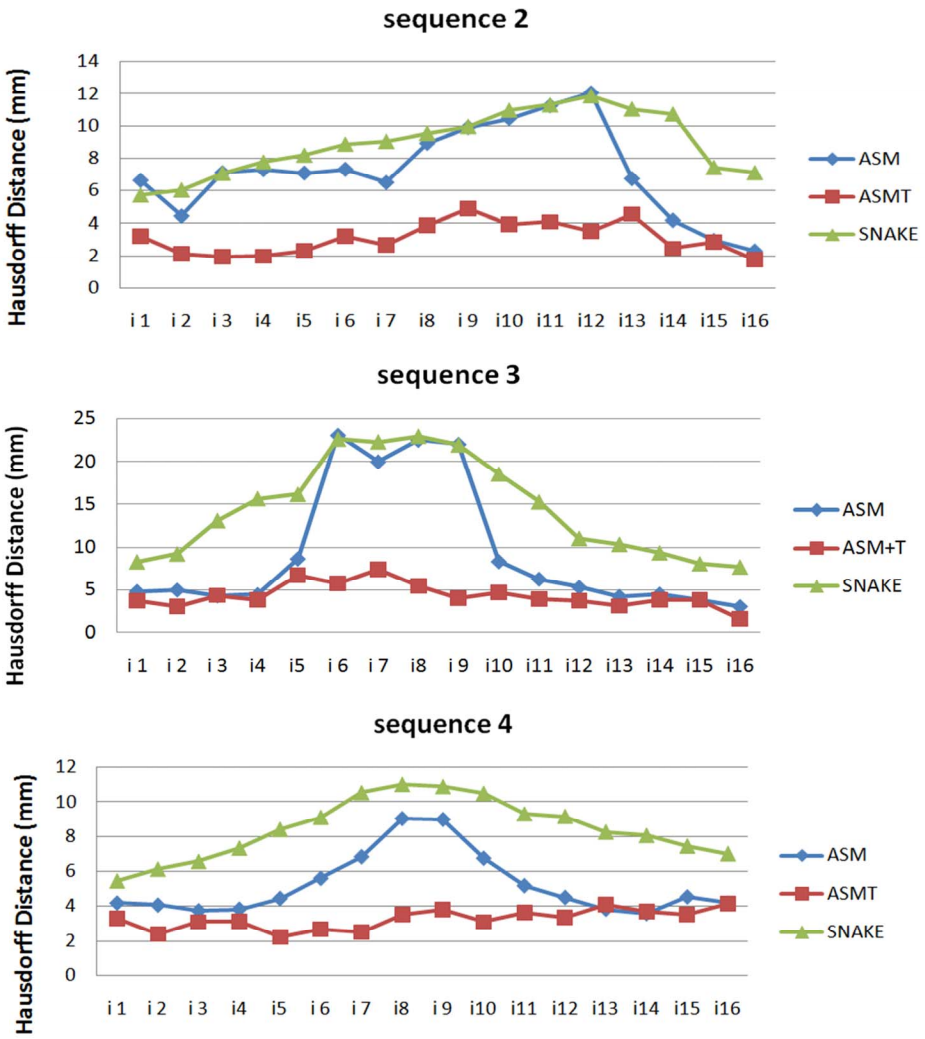


FIGURE 7: Hausdorff distance between each method (ASM, ASMT and SNAKE) and the reference segmentation of the four sequences.

On each diagram of figure 7, the horizontal axis represents the images of the sequence and the vertical axis represents the values of the Hausdorff distance. The red curve represents the values of the Hausdorff distance between the manual segmentation and the automatic segmentation obtained by our method ASMT. The blue curve represents the values of the Hausdorff distance between the manual segmentation and the automatic segmentation obtained by ASM. The green curve represents the values of the Hausdorff distance between the manual segmentation and the automatic segmentation obtained by the method based on SNAKE [8].

Looking at the four diagrams, we can see clearly that the red curve has some stability compared to the other curves (blue and green). Indeed, for the red curve and through the four diagrams, the values of the Hausdorff distance are between 1.18 and 7.38 (mm). By cons, for both blue and green curves, the values of the Hausdorff distance often represent great variations, which rose from 2.3 (mm) and reach 23.06 (mm). Through these measures, and although that in some cases the ASM and SNAKE provide acceptable results (especially in diastole images), we can deduce that our method provides for all images in each sequence an overall result that is more stable and closer to manual segmentation.

This proves the effectiveness of the integration of a priori knowledge about the spatio-temporal shape variation of the left ventricle. Indeed, the stage of spatio-temporal shape modelling provides more precise information on the spatial shape variation of the left ventricle at every moment of the cardiac cycle. This is what influences, consistently and in each image of the sequence, the accuracy of the results of the localization stage. The poor results obtained by ASM and SNAKE may be explained by the imprecision of the initialization in some images of the sequences and the generality of the a priori information about the shape. Besides, these results are mainly obtained in systole images (contraction stage) where the size of the left ventricle becomes very small and difficult to detect.

Another interesting finding is that the value of the Hausdorff distance, whether for ASM, SNAKE or ASMT compared to manual segmentation varies considerably from one sequence to another. For example, in sequence 3, this distance exceeds 23 (mm) for the ASM and SNAKE, while in sequence 4, this distance doesn't exceed 11 (mm). This may be related to the quality of the processed sequence, which affects then the result. That makes us ask: should we start with a pre-treatment stage to improve the quality of sequences before moving on to the segmentation stage?

In conclusion, it is clear that the integration of a priori knowledge about the spatio-temporal shape variation of the left ventricle improved significantly the results of segmentation. This increases the reliability of diagnostic parameters such as the activity-time curve and the ventricular ejection fraction, whose calculation is based on these results. However, we should know that these findings and results may be enriched to include more sequences in the process of quantitative validation. We should also note that the most delicate stage in our approach is the spatio-temporal shape modelling. This stage which is based on a manual process, affects widely the quality of results. It must be therefore made carefully with the help of an expert.

Figures 8, 9, 10 and 11 illustrate a qualitative comparison of the obtained results. Table 2 presents a report on the execution time of our approach. This time is divided into three stages: the spatio-temporal shape modelling, the grey levels modelling and the spatio-temporal localization.

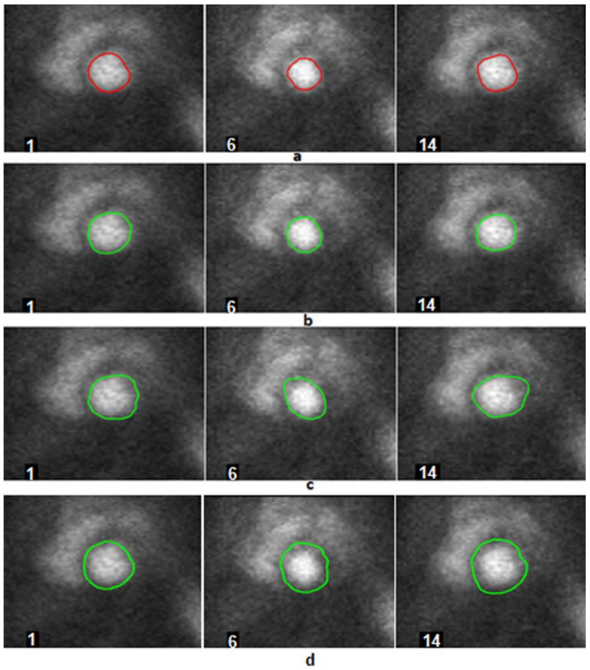


FIGURE 8: Qualitative comparison of results on sequence 1 (Images 1, 6, 14). (a) Manual segmentation, (b) Segmentation by ASMT, (c) Segmentation by ASM and (d) Segmentation by SNAKE.

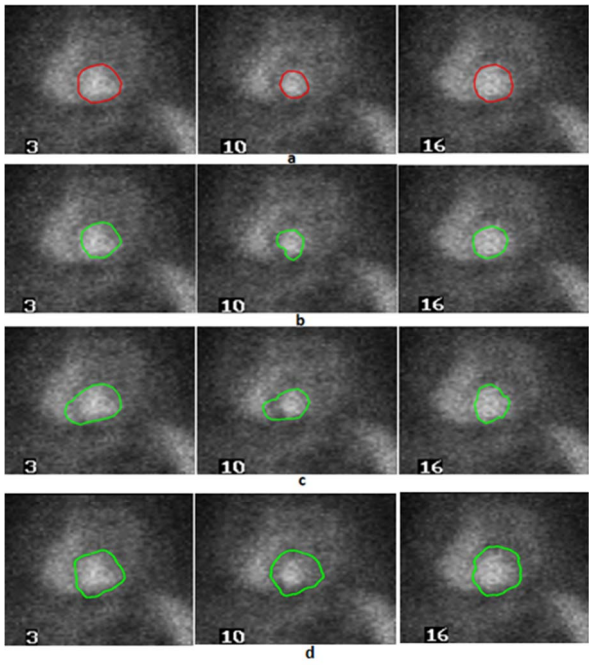


FIGURE 9: Qualitative comparison of results on sequence 2 (Images 3, 10, 16). (a) Manual segmentation, (b) Segmentation by ASMT, (c) Segmentation by ASM and (d) Segmentation by SNAKE.

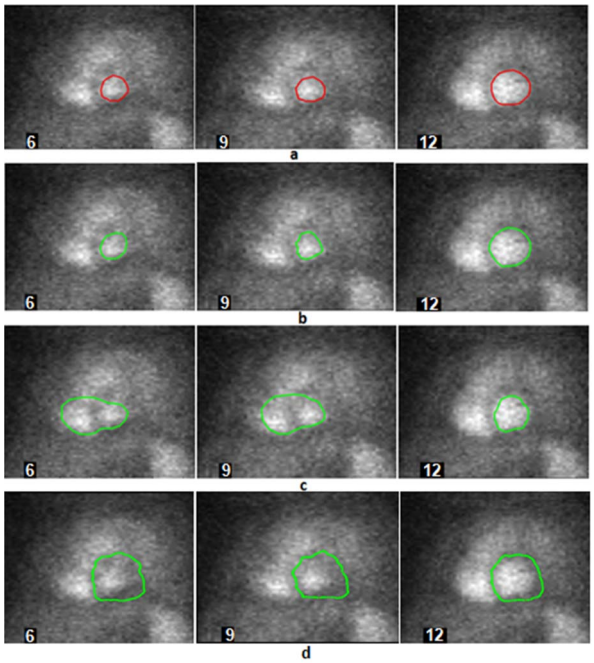


FIGURE 10: Qualitative comparison of results on sequence 3 (Images 6, 9, 12). (a) Manual segmentation, (b) Segmentation by ASMT, (c) Segmentation by ASM and (d) Segmentation by SNAKE.

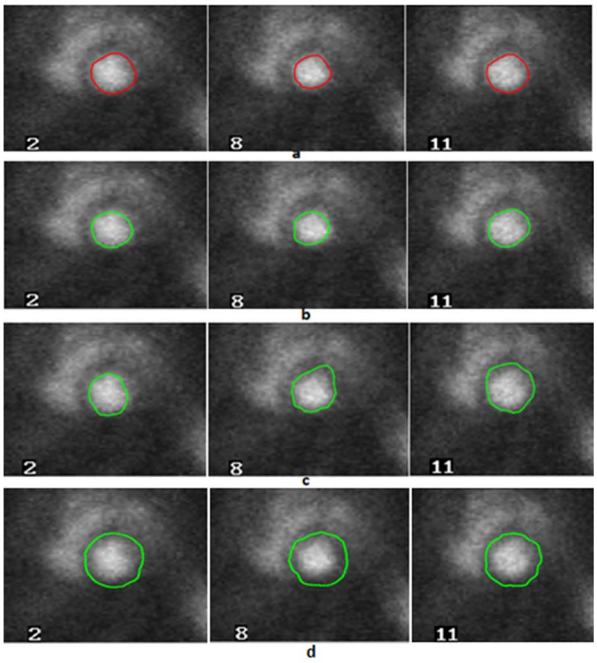


FIGURE 11: Qualitative comparison of results on sequence 4 (Images 2, 8, 11). (a) Manual segmentation, (b) Segmentation by ASMT, (c) Segmentation by ASM and (d) Segmentation by SNAKE.

Spatio-temporal shape modelling	Grey levels modelling	Spatio-temporal shape localization			
		Seq 1	Seq 2	Seq 3	Seq 4
26.27(s)	5.24(s)	29.12(s)	29.21(s)	29.56(s)	29.09(s)

Table 2: Report on the execution time of our method (Matlab 7.0.1, Processor: Intel ®, Core™, i3, 2.53 GHz × 2.53 GHz and RAM: 4 GO).

4. CONCLUSION

In this paper, we proposed to incorporate a new a priori knowledge about the spatio-temporal shape variation in the active shape model in order to define a new simple and more stable method for detecting structures whose shape change over time. The proposed method is based on two types of a priori knowledge: the spatial and temporal variation of the studied structure. It has also the advantage of being applicable on sequences of images. The experimental validation of this method, whether it is on simple synthetic sequences or on scintigraphic sequences for the left ventricle tracking, shows the interest of integrating a priori knowledge of the spatio-temporal shape variation. Indeed, having accurate information (geometry and deformation modes) about the shape of the studied structure at every moment provides more stable results, uniformly on all images of the processed sequence. In the training stage, the proposed optimization step, which is based on Fourier descriptors and K-Means classifier, helped to reduce the labelling step without losing the concept of variability.

We are convinced of the relevance of the used method, however, some improvements can be added and the validation should be pursued. Indeed, the most difficult step in our approach is the labelling step. It consists in manually extracting the spatio-temporal shapes from training sequences. That is why, it is usually performed by an expert.

The complexity of this step is in function of the number of training sequences, the number of images by sequence and the number of landmark points needed to represent the target structure details. Once, these parameters become important, this task becomes tedious and time consuming. Then, we should think to make this task semi-automatic or fully automatic. A way to make it semi-automatic is to consider that the shape of the studied structure at instants $t-1$, t and $t+1$ has a low variation. The manual training can be thus done on a reduced number of images which correspond to well chosen moments of the sequence. Then, the result of this training will be used for the automatic segmentation of the remaining images. This segmentation is then considered as training. Thus, the complexity of the labelling task can be reduced at least 70%. Moreover, it is possible to enrich and further validate this approach for other types of applications. For example, if we replace the temporal component by the third spatial axis (z), this method can be effectively used for volume segmentation that is based on an important a priori knowledge of shape. In this case, we must solve some additional issues such as correspondence between the slices of training volumes and the slices of the volume to be segmented as well as the automatic determination of the slices that contain the studied structure during the segmentation.

5. REFERENCES

[1] M. Kass, A. Witikin and D. Terzopoulos. “Snakes: Active contour models”. International Journal of Computer vision, pp. 321 – 331, 1988

[2] L. D. Cohen. “Active contour models and balloons”. CVGIP: Image Understanding, vol.53, pp. 211– 218, 1991

[3] S. Osher and J. A. Sethian. “Fronts propagating with Curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations”. J. Comput. Phys., vol.79, pp 12–49, 1988

[4] V. Caselles, R. Kimmel and G. Sapiro. “Geodesic Active Contours”. International Journal of Computer Vision, vol.22, pp 61–79, 1997

[5] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham. “Active Shape Models - Their Training and Application”. Computer Vision and Image Understanding, vol.61, pp 38–59, 1995

[6] T. F. Cootes, G. J. Edwards and C. J. Taylor. “Active Appearance Models”. IEEE Trans. Pattern Anal. Mach. Intell, vol.23, pp 681–685, 2001

[7] A. Singh, D. Goldgof and D. Terzopoulos. “Deformable Models in Medical Image Analysis”. IEEE Computer Society, Silver Spring, MD, 1998

[8] A. Fekir, N. Benamrane and A. Taleb-Ahmed. "Détection et Suivi d'Objets dans une Séquence d'Images par Contours Actifs". In Proc. CIIA, 2009

[9] F. Leymarie and M. Levine. “Tracking deformable objects in the plane using an active contour model”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, pp 617 – 634, 1993

[10] T. McInerney and D. Terzopoulos. “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, Vol.19, pp 69 – 83, 1995

[11] W. Niessen, J. Duncan, M. Viergever and B. Romeny. “Spatio-temporal Analysis of Left Ventricular Motion”. SPIE Medical Imaging, SPIE Press, pp 250 – 261, 1995,

[12] S. Sclaroff, and J. Isidoro. “Active blobs”. Proceedings of the Sixth IEEE International Conference on Computer Vision, pp 1146 – 1153, 1998

[13] G. Hamarneh and G. Tomas. “Deformable spatio- temporal shape models: extending active shape models to 2D+time”. Journal of Image Vision Computing, vol. 22, pp 461-470, 2004

[14] B. Lelieveldt, S. Mitchell, J. Bosch, R. van der Geest, M. Sonka, J. Reiber. “Time-continuous segmentation of cardiac image sequences using active appearance motion models”. Proceedings of the Information Processing in Medical Imaging, pp 446 – 452, 2001

[15] A. Signh, L. Von Kurowski and M. Chiu. “Cardiac MR image segmentation using deformable models”. SPIE Proceedings of the Biomedical Image Processing and Biomedical Visualization, pp 8 – 28, 1993

[16] G. Hamarneh. “Active Shape Models, Modeling Shape Variations and Gray Level Information and an Application to Image Search and Classification”. Technical report, Department of Signals and Systems, School of Electrical and Computer Engineering, Chalmers University of Technology, Goteborg, Sweden, 1998

[17] G. Strang. “Linear Algebra and Its Applications”. Saunders 1988

[18] J. Pladellorenst, J. Serraa, A. Castell and M. Yzuelll. “Using mathematical morphology to determine left ventricular contours”. Phys. Med. Biol., vol. 37, pp 3877-1894, 1992

[19] W. Gao and Y. Jin. “Auto threshold region-growing method for edge detection of nuclear medicine images”. Proceedings of SPIE, Medical Image Acquisition and Processing, pp. 126-130, 2001

[20] Damien, L. Itti, P. Egroizard and R. Itti. “Left ventricle detection in radionuclide ventriculograpy by a model of neural network”. Proc. 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society: IEEE-EMBS, pp 994-995, Paris, France, 1992

[21] N. Khlifa, N. Hraiech and K. Hamrouni. “Segmentation d’images par l’algorithme EM et les ondelettes. Premier Congrès International de Signaux, Circuits & systèmes : SCS’2004, pp 208-211, Monastir, Tunisie, 2004

[22] K. Hamrouni and N. Khlifa. “Two Methods for Analysis of Dynamic Scintigraphic Images of the Heart”. The International Arab Journal of Information Technology, vol. 3, pp 118-125. 2006

[23] A. Jouan, J. Verdenet and J. Cardot. “Extraction and analysis of left ventricular contours in cardiac radionuclide angiographies”. Journal of Optics, pp 239-246, 1991

[24] L. Sajn, M. Kukar, I. Kononenko and M. Milcinski. “Automatic segmentation of whole-body bone scintigrams as a preprocessing step for computer assisted diagnostics”. In lecture Notes in Computer Science, vol. 3581, pp 363-372, 2005

[25] N. Khlifa, T. Kraiem and K. Hamrouni. “Traitement et Analyse d’une Séquence d’Images Scintigraphiques du Cœur ”. La 2ème Conférence Internationale : JTEA’2002, pp 236-244, Sousse, Tunisie, 2002

[26] N. Khlifa, K. Hamrouni and T. Kraiem. “Etude de l’activité cardiaque par analyse d’images”. International Conference on Image and Signal Processing: ICISP’2003, pp 324-232, Agadir, Maroc, 2003

[27] N. Khlifa, A. Malek and K. Hamrouni. “Segmentation d’images par contours actifs : Application à la détection du ventricule gauche dans les images de scintigraphie cardiaque”. International Conference Sciences of Electronic, Technologies of Information and Telecommunication: SETIT’2005, pp 44, Sousse, Tunisie, 2005

[28] H. Najah, W. Daniel, K. Hamrouni. “An Active Contour Model Based on Splines and Separating Forces to Detect the Left Ventricle in Scintigraphic Images”. In 2nd International Conference on Machine Intelligence, Tozeur, Tunisie, 2005

[29] N. Khlifa, S. Ettaieb, Y. Wahabi and K. Hamrouni. “Left Ventricle Tracking in Isotopic Ventriculography Using Statistical Deformable Model”. The International Arab Journal of Information Technology, Vol. 7, pp 213–222, 2010

[30] D. Huttenlocher, D. Klanderman and A. Rucklige. “Comparing images using the Hausdorff distance”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, pp 850-863, 1993