Real-time Tracking of Deformable Target in 3D Ultrasound Images

Lucas Royer, Maud Marchal, Anthony Le Bras, Guillaume Dardenne,
Alexandre Krupa

HAL Id: hal-01122026

https://inria.hal.science/hal-01122026

Submitted on 3 Mar 2015

Real-time Tracking of Deformable Target in 3D Ultrasound Images

Lucas Royer[1,2,4], Maud Marchal[1,2], Anthony Le Bras[1,3], Guillaume Dardenne[1], Alexandre Krupa[1,4]

Abstract— In this paper, we present a novel approach for tracking a deformable anatomical target within 3D ultrasound volumes. Our method is able to estimate deformations caused by the physiological motions of the patient. The displacements of moving structures are estimated from an intensity-based approach combined with a physically-based model and has therefore the advantage to be less sensitive to the image noise. Furthermore, our method does not use any fiducial marker and has real-time capabilities. The accuracy of our method is evaluated on real data acquired from an organic phantom. The validation is performed on different types of motions comprising rigid and non-rigid motions. Thus, our approach opens novel possibilities for computer-assisted interventions where deformable organs are involved.

I. INTRODUCTION

In intra-operative interventions such as biopsy, it is often desirable to follow the motions of deformable structures in order to adjust the needle position under image guidance. These applications are generally performed under ultrasound (US) guidance as it produces real-time visual feedback of soft tissue deformations. Furthermore, US imaging provides low-cost, portable and non-invasive imaging, since it is acquired from a non-cumbersome transducer and a non-ionizing energy. However, such systems suffer from different shortcomings, including the field of view often limited by small acoustic windows. Additionally, the quality of US images is often affected by speckle noise, artifacts, and shadows that makes soft-tissue tracking challenging. These limitations generally prevent from using classical tracking approaches proposed for other image modalities. To overcome the previous limitations, we propose a novel approach that combines a physically-based model with dense dissimilarity minimization in successive 3D ultrasound images. Moreover, we derive an adaption of our method that takes into account the US volume geometry. This allows to track only partially visible target and to propagate the deformation to structures that are not within the field of view. Our approach has the advantage of being real-time and robust to the noise. Furthermore, it provides an accurate estimation of target deformation and does not use any fiducial marker. The rest of the paper is organized as follows. In section II, related work combining visual tracking and physical model are presented. In section III, we detail our intensity-based approach combined with a physically-based model. In section IV, we present results on real data that demonstrate the performance of our method. Finally, section V concludes the paper.

II. RELATED WORK

Several techniques have been presented in order to extract the motions of a deformable target within ultrasound images. They can be classified in three types. Some of them are based on block-matching methods where local displacements are computed by minimizing image dissimilarity between consecutive frames. In the literature, typical examples using Sum of Squared Differences (SSD)[1] [2] or Cross-correlation [3] have been proposed. However, these methods assume that the displacement is the same within a local region block. Other approaches are based on the extraction of contours which represent the interfaces of soft tissues [4]. Nevertheless, these methods are generally based on a Bayesian framework in order to decrease the noise sensitivity. Angelova and Mihaylova [5] developed a method that detects contours over the time where prior dynamics are considered. These techniques assume that strong edges are always visible within US sequence. Other approaches rely on the estimation of the optical-flow where dense displacement field is obtained from spatio-temporal derivatives of the pixels intensities inspired from the work of Lucas and Kanade [6]. These methods are based on the hypothesis of the constancy of the US wave reflection by a given physical structure over the time. Lee [7] also provided a method based on thin-plate splines (TPS) for modeling the deformation. However, inaccurate results can be observed in practice due to the inability to fit high local deformation with TPS-based model. Mikic et al. [8] presented a non-rigid target tracking method based on the optical flow which allows tracking a deformable target undergoing rigid displacement. Nevertheless, these methods may produce inaccurate results due to the image noise sensitivity.

To cope with the previous issues, several works combining visual tracking and physical model have recently emerged. Some of them have been developed for others image modalities such as RGB-D camera [9], stereo-camera [10], or magnetic resonance (MR) imaging [11]. Few methods are based on principal component analysis (PCA) in order to generate a database of plausible shapes of the organ of interest as for example the prostate [12]. However, the PCA techniques are generally computationally expensive and are not suitable for real-time tracking. In contrast, several registration techniques driving directly the physically-based model without atlas generation have been developed. For example, Shen et al. [11] proposed a registration method

on MR imaging based on the mass-spring system in order to constrain the deformation. Zhang and al. [13] introduced another MRI registration based on finite element model where the extraction of the scale invariant features (SIFT), proposed in [14], is required to drive the model. However, the extraction of such points is not suitable in US images due to the speckle noise. Marami et al. [15] proposed very recently an elastic registration method applicable to multi-modality image registration. Their algorithm uses an estimation and a filtering process based on a linear elastic model of tissue deformation to automatically register the pre-operative 3D MR data to the intra-operative 2D US images. However, this approach does not focus on the 3D tracking of a specific target with appropriate elastic parameters and does not comply with the real-time constraint needed for robotic applications. To the best of our knowledge, no method combining real-time tracking and physically-based model within 3D US volume has been developed.

III. METHOD

A. Model Generation

The main objective of our method is to estimate the position of a deformable target within 3D ultrasound sequences. The target represents a structure of interest containing N_v voxels and is delimited by a visible surface in the US volume. The first step of our algorithm consists in extracting the target geometry by manually segmenting each slice of the first US volume of the sequence. The resulting surface is then smoothed in order to remove sharp edges, and a corresponding tetrahedral mesh is generated from it. The mesh is defined by a set of connected tetrahedral cells, each containing four vertices and a specific number of voxels. The number of vertices is defined as N_c. The position of a voxel $\mathbf{p} = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix}$ that lies inside a specific cell can be related to the positions of its four vertices $\mathbf{q}_j$ by using barycentric coordinates. This can be expressed in the following relation:

$$\mathbf{p} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \mathbf{q}_3 \\ \mathbf{q}_4 \end{pmatrix} \quad (1)$$

where $\mathbf{q}_i$ denotes the position of the vertice lying in the cell. $\mathbf{M}_i$ represents a diagonal matrix of 3×3 elements containing the barycentric coefficients that relates the voxel $\mathbf{p}$ to the positions of the vertices $\mathbf{q}_i$. From the equation (1), we can link all the voxel positions lying in the mesh model with all the vertices by defining a $(3 \cdot Nv) \times 1$ vector $\mathbf{p}_{im}$ containing all the voxel positions, and a $(3 \cdot Nc) \times 1$ vector $\mathbf{q}$ containing all the vertex positions.

$$\mathbf{p}_{im} = \mathbf{M}.\mathbf{q} \quad (2)$$

where $\mathbf{M}$ is a constant matrix containing $(3 \cdot N_v) \times (3 \cdot N_c)$ elements. Each 3-line of $\mathbf{M}$ defines the set of barycentric coefficients that relates a voxel position with the positions of the vertices. Thanks to equation (2), we can update the positions of the target when the vertices of the model are displaced. These positions will be obtained from the variation of the inter-volume intensity and the mechanical behavior of the mass-spring model as we will explain in the following.

B. Intensity-based Approach

In this work, we assume that the US wave reflected by a physical point is constant and time independent. This ensures that the intensity of the target is constant across the time. This assumption can be expressed as:

$$I_t(\mathbf{p}_{im}(t)) = I_{t_0}(\mathbf{p}_{im}(t_0)) \quad (3)$$

where t_0 and t represent respectively the acquisition times of the first volume and the current volume. $\mathbf{p}_{im}(t_i)$ denotes the voxel positions of the target at the time of acquisition of the volume i. I_t denotes the US intensity of the volume acquired at time t. Thus, $I_t(\mathbf{p}_{im}(t))$ is a column vector defining the intensity of the target at time index t. The objective is to find iteratively the optimal displacement $\Delta \mathbf{q}$ of the vertices in order to minimize the cost function C which computes the least square of the intensity difference between $I_t(\mathbf{p}_{im}(t))$ and $I_{t_0}(\mathbf{p}_{im}(t_0))$ such that:

$$C(\mathbf{q}^k(t)) = (I_t(\mathbf{p}_{im}(t)) - I_{t_0}(\mathbf{p}_{im}(t_0)))^2 \quad (4)$$

$$C(\mathbf{q}^k(t)) = (I_t(\mathbf{M}(\mathbf{q}^{k-1}(t) + \Delta\mathbf{q}) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0)))))^2 \quad (5)$$

where $\mathbf{q}^{k-1}(t)$ represents the estimation of the parameters at time t at iteration $k-1$ of the optimization algorithm. In order to compute the optimal displacement of the vertices, the cost function is linearized by using the first order Taylor approximation such that:

$$C(\mathbf{q}^k(t)) \approx (\mathbf{J}\Delta\mathbf{q} + I_t(\mathbf{M}(\mathbf{q}^{k-1}(t))) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0))))^2 \quad (6)$$

where $\mathbf{J}$ denotes the Jacobian matrix which links the variation of the intensity I of the voxels to the variation of the parameters $\Delta \mathbf{q}$. The intensity variation of the voxels can be linked to the variation of the voxel coordinates such that:

$$\dot{I} = \nabla I.\dot{\mathbf{p}}_{im} \quad (7)$$

where ∇I represents the spatial derivatives of the voxel intensities regarding the x, y, and z axis and $\dot{\mathbf{p}}_{im}$ the displacements of voxels. In our case, the image gradient of voxel $\mathbf{p}$ is computed from 6×1 spatial derivative filters. Combining equation (7), with the time derivative of equation (2), the Jacobian can be expressed as:

$$\mathbf{J} = \nabla I.\mathbf{M} \quad (8)$$

Let us recall that our objective is to minimize the difference between the intensity vector of the reference target $I_{t_0}(\mathbf{p}_{im}(t_0))$ and current target $I_t(\mathbf{p}_{im}(t))$. Several minimization strategies can be used in order to obtain the optimal displacements of the vertices. We chose to use the steepest gradient strategy instead of Gauss-Newton method as the pseudo-inverse of the Jacobian is highly computational demanding due to its size. We therefore obtain:

$$\Delta\mathbf{q} = -\alpha\mathbf{J}^t[I_t(\mathbf{M}(\mathbf{q}^{k-1}(t))) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0)))] \quad (9)$$

where $\alpha > 0$ denotes the step size of the steepest gradient method. This coefficient has an influence on the convergence rate of the difference between the current intensities $I_t(\mathbf{M}(\mathbf{q}^{k-1}(t)))$ and the initial ones $I_{t_0}(\mathbf{M}(\mathbf{q}(t_0)))$. $\mathbf{J^T}$ represents the transpose matrix of the Jacobian $\mathbf{J}$. However, straightforward minimization of the intensity variation yields to poor results in general due to possible local minima, in particular in presence of noise or partial object occlusion that are inherent to US images. In this work, we chose to address this problem by integrating a physically-based model of the expected deformation.

C. Deformable Component

In order to constrain the deformations of the target, a mechanical model is combined to the previous approach. It has the advantage to limit the noise sensitivity of the intensity-based approach. In this work, we assume that the deformation of the target is viscoelastic with a linear behavior. Our deformable model consists of a mass-spring-damper system where the vertices of the model are linked with springs. In order to tune the behavior of the model, a specific value of mass can be set to each vertex. Furthermore, specific elastic and damping parameters can be added to each link and can constrain the deformation in a variety of ways. Thus, the force $\mathbf{f}_{ij} = [f_{x_{ij}} \; f_{y_{ij}} \; f_{z_{ij}}]^T$ exerted on a vertex $\mathbf{q}_i$ from a neighbor vertex $\mathbf{q}_j$ is expressed as follows:

$$\mathbf{f}_{ij} = K_{ij}(d_{ij} - d_{ij}^{init})(\mathbf{q}_i - \mathbf{q}_j) + D_{ij}(\dot{\mathbf{q}}_i - \dot{\mathbf{q}}_j) \circ (\mathbf{q}_i - \mathbf{q}_j) \tag{10}$$

where d_{ij} is a scalar value representing the distance between the vertices $\mathbf{q}_i$ and $\mathbf{q}_j$ in their current positions. d_{ij}^{init} denotes the initial distance value at the rest state of the deformable mass-spring. The $\circ$ operator expresses the Hadamard product, K_{ij} is a scalar value denoting the stiffness of the spring that links the two vertices while D_{ij} is the damping coefficient value. The values of these coefficients can be different, depending on the homogeneities of the tissues. From the previous equation, we can compute the internal forces $\mathbf{f}_i$ applied on each vertex $\mathbf{q}_i$ of the deformable system such that:

$$\mathbf{f}_i = \sum_{n=0}^{N_i} \mathbf{f}_{in} + G_i \dot{\mathbf{q}}_i \tag{11}$$

where G_i is a scalar value representing the velocity damping coefficient associated to the vertex $\mathbf{q}_i$. N_i denotes the number of springs that are connected to the vertex $\mathbf{q}_i$. Once the internal forces are computed from the previous equation, we integrate them with a semi-implicit Euler integration scheme for simulating our system in order to obtain the associated displacements. Then, we compute the optimal positions of the vertices combining the displacement obtained from the intensity-based approach with the internal displacements. This can be expressed as follows:

$$\mathbf{q}(t) = \mathbf{q}(t-1) + \Delta\mathbf{q} + \Delta\mathbf{v} \tag{12}$$

where $\Delta\mathbf{v}$ represents the internal displacements obtained from the simulation of the model described above. The

external displacements $\Delta\mathbf{q}$ are obtained from the intensity-based approach presented in equation (9), where the α coefficient (step size of the steepest gradient method) also acts as the balance between the internal forces and the external displacements.

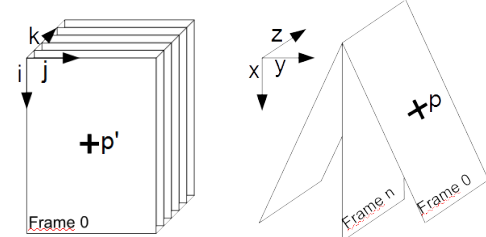## D. Prescan data adaptation

Fig. 1. (Left) Prescan Data. (Right) Rearranged Prescan Data

The original method described previously is well-suited if the target is completely visible within US volumes. However, it is likely that the target becomes partially invisible due to the limited field of view and large motions that can be applied to the probe. To cope with these issues, we adapted the intensity-based approach by taking into account the US volume geometry. To do so, we use the prescan volume instead of the 3D complete postscan volume. A prescan volume corresponds to the raw data acquired from the ultrasound probe, and is obtained by putting together each US scan-line. As it is illustrated in Figure 1, the prescan volumes are represented by a set of rectangular frames, and have therefore the advantage to present a simple geometry. However, since our transducer has a curved shape, prescan volumes are not geometrically correct as it does not show the exact position of the tissue being imaged. Thus, to adapt the intensity-based approach, a conversion scan step ψ, described in [7], is required in order to rearrange the prescan data. ψ is the transformation which maps a point $\mathbf{p}'$ of the prescan volume into the point $\mathbf{p}$ in the rearranged prescan volume. Equation (4) can now be expressed as:

$$C(\mathbf{q}_t^k) = (I_t(\psi(\mathbf{p}'_{im}(t))) - I_{t_0}(\psi(\mathbf{p}'_{im}(t_0))))^2 \tag{13}$$

where $\mathbf{p}'_{im}$ represents the voxel positions of the target in prescan data. I_t and I_{t_0} denote the rearranged prescan US intensity volume at time t and t_0. However, as we can see from Figure 1, several voxel positions do not have an associated intensity value in the prescan volume since the rearranged prescan data is not a continuous set of voxels. To overcome this limitation, we implemented a method that interpolates the unknown intensity of a point $\mathbf{p}$ by using the inverse distance weighting (IDW) interpolation of neighboring points. It can be expressed as follows:

$$\tilde{I}(\mathbf{p}) = \frac{\sum_{v=0}^{N_n} \gamma_v(\mathbf{p}) I(\mathbf{p}_i)}{\sum_{v=0}^{N_n} \gamma_v(\mathbf{p})} \tag{14}$$

where $\tilde{I}(\mathbf{p})$ represents the unknown intensity value associated to the point $\mathbf{p}$. $\mathbf{p}_i$ denotes the set of neighbors of $\mathbf{p}$ that have an associated intensity value, N_n expresses the number

of neighbors points. γ_v is the distance inverse between **p** and its neighbors $\mathbf{p}_i$. This distance is evaluated from rearranged prescan data which presents a correct geometry. To adapt the method described in the previous section, we need also to design a new spatial derivative filter adapted to prescan data. As previously stated, the image gradient of a voxel **p** is computed from 6×1 filters and takes into account the intensity value of nearest neighbors.Thus, by using the IDW interpolation technique, the image gradient of a voxel $\mathbf{p}'$ in prescan frame can be expressed as :

$$\Delta I(\psi(\mathbf{p}')) = \Delta I(\mathbf{p}) = \sum_{j=-3}^{3} w_j \tilde{I}(\mathbf{p}_j) \qquad (15)$$

where $\mathbf{p}'$ represents a point in the prescan volume. w_j depicts the filter coefficients, and $\mathbf{p}_j$ represents the nearest neighbors of **p** within rearranged prescan volume. When the target goes out of the field of view, our approach does not take into account the points $\mathbf{p}'$ that are outside of the prescan volume. This can be performed by setting their derivative value to zero. The rest of the method remains unchanged. This adaptation allows to take into account the US volume geometry. However, for this purpose, it should be mentioned that we can also use another method that consists in generating a mask on postscan data.
.

IV. RESULTS

The ultrasound volumes are obtained from an Ultrasonix station thanks to a motorized probe (4DC7-3/40) which allows to acquire volumes containing 44 frames per volume with an angular step of $1.4°$. Thus, the US field of view is equal to $64.4°$ and its associated volume rate is 0.8 vol/s. Since our US platform is designed for the purpose of research work, we have access to the prescan data as well as reconstructed postscan data. From this system, we can retrieve postscan and prescan volumes that consists respectively of $129 \times 189 \times 135$ voxels and $128 \times 120 \times 44$ voxels, and their associated depth is 15 cm. The size of a voxel represents a cube with 1.33 mm length sides. Our tracking system achieves real-time capabilities, since the computation time of 350 ms per 3D US frame is much lower than the acquisition time. It has been built with C++/GPU code by using Cuda [16] and VTK [17] libraries. The preprocessing steps described previously and including the segmentation and the model generation are performed respectively with the ITK-SNAP software [18], and the tetGen software [19]. In order to demonstrate the performance of our method, several experiments including both translation and deformation motions have been performed by using a home-made phantom which contains animal soft tissue (duck gizzard) and gelatin. As it can be seen in Fig. 2, the US probe is mounted on the end effector of a 6 Degrees of Freedom (DOF) robot (Viper 650, Adept). The probe is displaced on the surface of the phantom by applying different trajectories to the robot. We evaluated the performance of our tracking system by measuring the robot motion regarding the x, y and z axis with odometry. The tracking task is then performed on a target representing a convex surface visible within the US volume that can be delimited in the first US volume of the sequence with a manual segmentation as shown in Fig. 2. In the following experiments, we set empirically the elastic and damping parameters to $K_{ij} = 2.5$, $D_{ij} = 0.1$ for all the springs, and $G_i = 2.0$ for all vertices. An interesting approach could consists in evaluating these parameters from elastography. The step size of the steepest gradient method has been set to $\alpha = 2 \times 10^{-6}$.
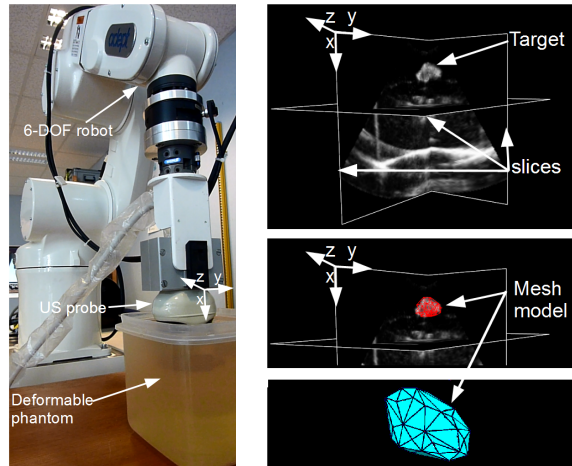

Fig. 2. (Left) Experiment setup. (Right) 3D slices of the ulltrasound volume representing the target and the tetrahedral mesh model (148 cells, 60 vertices).
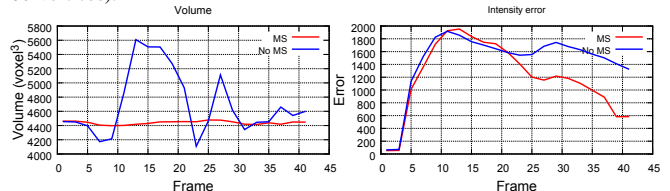

Fig. 3. Results obtained during lateral translation motion. (Left) Volume estimation of the target. (Right) Least square intensity error.

A. Experiments With Rigid Motion

At first, we evaluated the tracking task with a lateral translation motion where the robot makes the probe follow a square trajectory on the phantom surface. In this experiment, the probe is displaced along the y and z-axis and no deformation has been involved. Thus, the target volume is constant over the time and is not compressed. In order to validate our method, we compared our results to the approach without the mass-spring (MS) system by removing the $\Delta\mathbf{v}$ term of equation (12). Fig. 4 shows the comparison between the displacement estimation of the target obtained by computing the barycenter position of our model at each 3D US frame. Our results are compared to the displacement estimated from odometry. In Fig. 4, we observed that our method achieves a low tracking error of (1.23 mm±0.03 mm) compared to the method without mass-spring system (5.25 mm±0.42 mm). Thus, our approach provides subvoxel accuracy since this error is lower than the size of a voxel. Fig. 3 presents the evolution of the model volume in the US sequence. This is computed by summing the volume of each tetrahedral cell. We noticed that the estimated volume is constant with our method over each 3D US frame contrary to the method
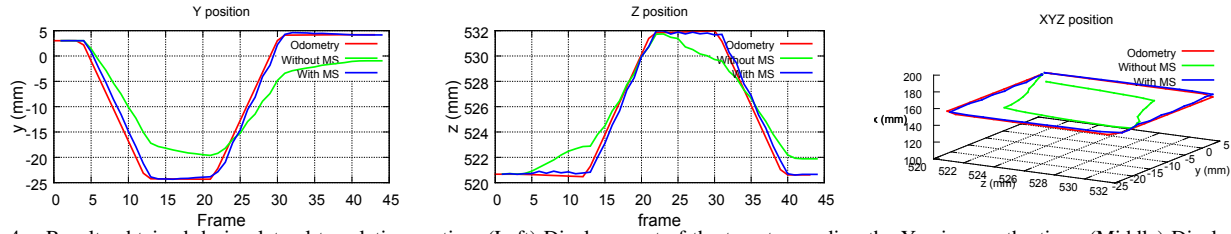
Fig. 4. Results obtained during lateral translation motion. (Left) Displacement of the target regarding the Y-axis over the time. (Middle) Displacement of the target regarding the Z-axis over the time. (Right) 3D Displacement of the target
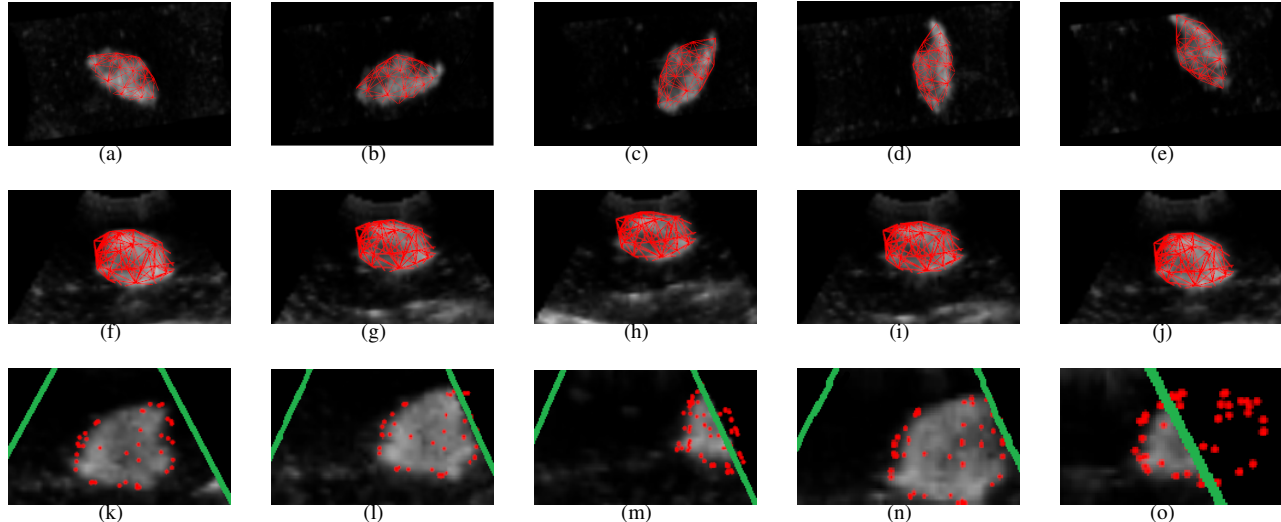

Fig. 5. Example of the tracking task. (Top) Mesh model during postscan tracking with rotation motion at frame 1 (a), at frame 16 (b), at frame 26 (c), at frame 36 (d), at frame 46 (e). (Middle) Mesh model during postscan tracking with deformation motion at frame 1 (f), at frame 16 (g), at frame 24 (h) , at frame 32 (i), at frame 40 (j). (Bottom) Partially visible target experiment showing a slice of the target regarding the Y-axis. Green lines are the borders of ultrasound images. Red points represent the intersections between the mesh model and the image slice. The green lines represent the borders of the US image slice. Template tracking with postscan method at frame 1 (k), at frame 21 (l), at frame 34 (m), at frame 42 (n). Template tracking with prescan method at frame 34 (o).

without MS system where the volume varies significantly due to the image noise. We can also see that our method produces a smaller final intensity error given by equation (4). An example of the tracking task during a probe rotation motion along the x-axis is presented in Fig. 5(a-e).

B. Experiment With Deformation

In the second experiment, we evaluated the response of the tracking task when the target is deformed. To do so, we repeatedly compressed and released the deformable phantom with the probe mounted on the robot. In this case, the deformation has been applied by moving the probe along the x-axis thanks to the robot. This produces significant and periodic variation of the target volume. The estimated displacement of the target is shown in the Fig. 6. Contrary to the previous experiment, we can observe that the barycenter displacement does not exactly match the displacement obtained from the robot odometry. This difference is due to the compression of both the target and neighborhood gelatin. This attenuates the barycenter displacement compared to the motion measured by odometry. The Fig. 6 shows that the estimation of the volume varies periodically over the time with our method. Due to the noise sensitivity, we can observe that without the combination of a physically-based model,

the volume varies unsteadily and even overcomes the initial volume of the target. An example of the tracking task with a deformation is presented in Fig. 5(f-j).
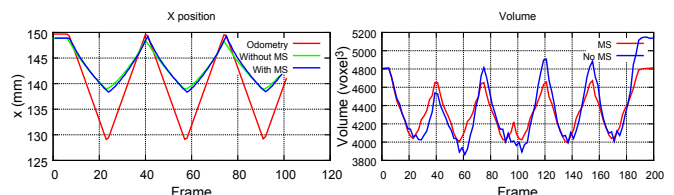

Fig. 6. Results obtained during deformation motion. (Left) Displacement of the target regarding the X-axis over each frame. (Right) Volume estimation of the model over each frame

C. Experiment With Partially Visible Target

We also evaluated the performance of our method when the target becomes only partially visible during the sequence. To do so, the robot followed a large square trajectory as well as an elevation motion which deformed the target. In this experiment, we compared the method applied on prescan data and the original method applied on postscan data. Fig. 7 shows the displacements of the model barycenter and the comparison with odometry measurement. We observed that the prescan adaption improves the tracking accuracy (1.94 mm±0.18 mm instead of 3.25 mm± 0.20 mm with postscan
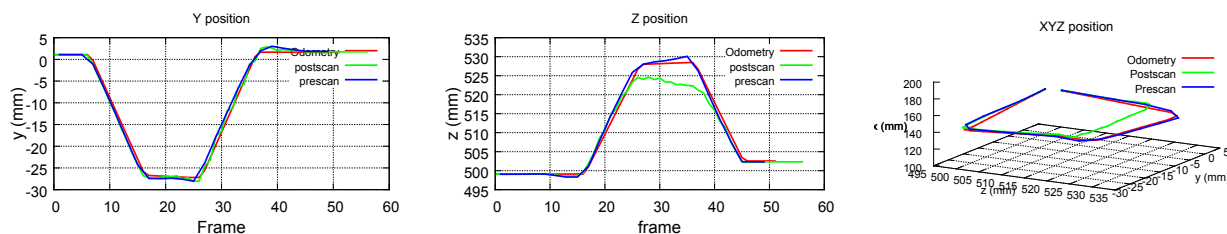
Fig. 7. Results obtained with partially visible target. (Left) Displacement of the target regarding the Y-axis over the time. (Middle) Displacement of the target regarding the Z-axis over the time. (Right) 3D Displacement of the target

data) when the target goes out of the field of view. Fig. 5(k-o) shows an example of the tracking task on several frames showing a slice of target which is displaced to the right and goes out of the field of view. We noticed that, with the postscan method, our model does not have a plausible shape when the target is not within the field of view (see Fig. 5(m). This can be also visualized in Fig. 8 from which the volume of the model is abnormally reduced when the target goes out of the field of view at frame 20. The postscan method remains however robust since the model returns to its initial shape when the target becomes visible again during its return to the initial position as shown in Fig. 5(n). We can also observe that the prescan method shows a stable evolution of the volume (see Figures 5(o)-8).
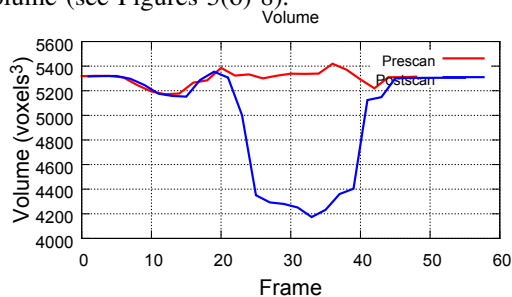

Fig. 8. Results obtained with partially visible target. Comparison of the volume estimation of the target with the prescan method and with the postscan method

V. CONCLUSION

In this paper, we have developed a real-time approach for tracking a deformable target combined with a physical-based model. This method has therefore the advantage to be robust to speckle noise. Additionally, our approach allows tracking of partially visible target within 3D ultrasound volumes. The performance of our algorithm has been validated on real-data acquired from organic phantom, and has been evaluated with visual assessment combined with robotic odometry ground truth. Thus, our method opens novel perspectives in computer-assisted interventions based on US imaging where deformable organs are involved, such as image-guided needle biopsy or ablation therapies. In future work, we will propose to enhance the physical model in order to improve the tracking accuracy. Furthermore, our approach will be validated on real data obtained from patients.

REFERENCES

[1] F. Yeung, S. F. Levinson, D. Fu, and K. J. Parker, "Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh," *IEEE Trans. on Medical Imaging*, vol. 17, no. 6, pp. 945–956, 1998.

[2] W. Yu, N. Lin, P. Yan, K. Purushothaman, A. Sinusas, K. Thiele, and J. S. Duncan, "Motion analysis of 3D ultrasound texture patterns," in *Functional Imaging and Modeling of the Heart.* Springer, 2003, pp. 253–261.

[3] C. B. Compas, B. A. Lin, S. Sampath, A. J. Sinusas, and J. S. Duncan, "Combining shape and speckle tracking for deformation analysis in echocardiography using radial basis functions," in *Proc. of IEEE International Symp. on Biomedical Imaging*, 2011, pp. 1322–1325.

[4] Q. Peng, R. C. Jones, and C. E. Constantinou, "2D Ultrasound Image Processing in Identifying Responses of Urogenital Structures to Pelvic Floor Muscle Activity," *Annals of Biomedical Engineering*, vol. 34, no. 3, pp. 477–493, Mar. 2006.

[5] D. Angelova and L. Mihaylova, "Contour segmentation in 2D ultrasound medical images with particle filtering," *Machine Vision and Applications*, vol. 22, no. 3, pp. 551–561, 2010.

[6] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision." in *Proc. of International Joint Conference on Artificial Intelligence*, vol. 81, 1981, pp. 674–679.

[7] D. Lee and A. Krupa, "Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4D ultrasound," in *Proc. of IEEE International Conference on Intelligent Robots and Systems*, 2011, pp. 2831–2836.

[8] I. Mikic, S. Krucinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates," *IEEE Trans. on Medical Imaging*, vol. 17, no. 2, pp. 274–284, 1998.

[9] I. Leizea, H. lvarez, I. Aguinaga, and D. Borro, "Real-time Deformation, Registration and Tracking of Solids Based on Physical Simulation," in *Proc. of IEEE International Symp. on Mixed and Augmented Reality*, 2014, pp. 165 – 170.

[10] N. Haouchine, J. Dequidt, I. Peterlik, E. Kerrien, M.-O. Berger, and S. Cotin, "Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery," in *Proc. of IEEE International Symp. on Mixed and Augmented Reality*, 2013, pp. 199–208.

[11] J.-K. Shen, B. J. Matuszewski, L.-K. Shark, A. Skalski, T. Zielinski, and C. J. Moore, "Deformable Image Registration - A Critical Evaluation: Demons, B-Spline FFD and Spring Mass System," in *Proc. of Fifth International Conference BioMedical Visualization*, 2008, pp. 77–82.

[12] Y. Hu, H. U. Ahmed, Z. Taylor, C. Allen, M. Emberton, D. Hawkes, and D. Barratt, "MR to ultrasound registration for image-guided prostate interventions," *Medical Image Analysis*, vol. 16, no. 3, pp. 687–703, Apr. 2012.

[13] J. Zhang and J. Wang, "Model-based nonrigid image registration using scale-invariant features," in *Proc. of IEEE International Conference on Signal Processing, Communication and Computing*, 2013, pp. 1–5.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[15] B. Marami, S. Sirouspour, A. Fenster, and D. W. Capson, "Dynamic Tracking of a Deformable Tissue Based on 3D-2D MR-US Image Registration," in *Proc. of SPIE Medical Imaging*, 2014.

[16] J. Nickolls, I. Buck, M. Garland, and K. Skadron, "Scalable parallel programming with CUDA," vol. 6, no. 2, pp. 40–53, 2008.

[17] W. Schroeder, K. W. Martin, and B. Lorensen, *The visualization toolkit: an object-oriented approach to 3D graphics.* Kitware, Inc., 2002.

[18] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, July 2006.

[19] H. Si, "TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator," *ACM Trans. on Mathematical Software*, vol. 41, no. 11, 2015.