Non-rigid Target Tracking in 2D Ultrasound Images Using Hierarchical Grid Interpolation

Lucas Royer, Marie Babel, Alexandre Krupa

Non-rigid Target Tracking in 2D Ultrasound Images Using Hierarchical Grid Interpolation

Lucas Royer *, Marie Babel † and Alexandre Krupa ‡

INRIA Rennes-Bretagne Atlantique, IRISA, INSA Rennes, France

ABSTRACT

In this paper, we present a new non-rigid target tracking method within 2D ultrasound (US) image sequence. Due to the poor quality of US images, the motion tracking of a tumor or cyst during needle insertion is considered as an open research issue. Our approach is based on well-known compression algorithm in order to make our method work in real-time which is a necessary condition for many clinical applications. Toward that end, we employed a dedicated hierarchical grid interpolation algorithm (HGI) which can represent a large variety of deformations compared to other motion estimation algorithms such as Overlapped Block Motion Compensation (OBMC), or Block Motion Algorithm (BMA). The sum of squared difference of image intensity is selected as similarity criterion because it provides a good trade-off between computation time and motion estimation quality. Contrary to the others methods proposed in the literature, our approach has the ability to distinguish both rigid and non-rigid motions which are observed in ultrasound image modality. Furthermore, this technique does not take into account any prior knowledge about the target, and limits the user interaction which usually complicates the medical validation process. Finally, a technique aiming at identifying the main phases of a periodic motion (e.g. breathing motion) is introduced. The new approach has been validated from 2D ultrasound images of real human tissues which undergo rigid and non-rigid deformations.

Keywords: Target tracking, ultrasound images, deformable, non-rigid motions, hierarchical grid interpolation

1 INTRODUCTION

In 2D ultrasound image sequence, targets may undergo both rigid and non rigid motions. Non rigid motion occurs when the target is deformed across the time which can be caused by patient breathing, or cardiac motions. These deformations complicate tasks which use the target tracking such as robotized needle insertion. In order to address these issues, several techniques of non-rigid target tracking within ultrasound images have been presented over the two last decades. Some of them rely on a physical model or on a prior knowledge such as Dydenko[1] which uses information related to the shape and the motion provided by a medical expert, or like Angelova[2] where prior dynamics of a contour are considered. These techniques are efficient, as they are dedicated to a particular target. Other techniques only estimate one of both motion types: Yeung[3] uses an adaptive deformable mesh for non-rigid tissue motion estimation, and Nadeau[4] only extracts the rigid motion. These techniques may fail due to the neglected motion part. To address this issue, Mikic[5] presented a non-rigid target tracking method based on the optical flow which allow tracking a deformable target undergoing rigid displacement. However, this technique does not take into account the speckle noise inherent to ultrasound image modality, or the non-linear displacement of the target. In this paper, we present a new target tracking method which allows to distinguish the rigid motion and the non-rigid motion. In addition, the ultrasound image modality is considered by using a hierarchical approach without any prior knowledge. In the following, we first describe a new non-rigid target tracking method, then we present our results obtained from a 2D ultrasound image sequence.

*Lucas Royer is with INSA Rennes - IRISA/INRIA Rennes Bretagne Atlantique, Campus de Beaulieu, 35042 Rennes, France. E-mail: First-name.Name@irisa.fr

†Marie Babel is with INSA Rennes - IRISA/INRIA Rennes Bretagne Atlantique, Campus de Beaulieu, 35042 Rennes, France. E-mail: First-name.Name@irisa.fr

‡Alexandre Krupa is with INRIA Rennes Bretagne Atlantique - IRISA , Campus de Beaulieu, 35042 Rennes, France. E-mail: First-name.Name@inria.fr

2 METHOD

2.1 Manual initialization

In order to propose a fully automatic method, the user interactions are limited to the selection of the target. This is done by simply extracting a rectangular area ("patch") around the desired anatomical structure thanks to a graphical interface. The reference patch corresponds to the area defined by the user at the beginning of the US sequence. The objective of the algorithm is to track the target by updating the position of the reference patch along the image sequence. The images are considered consecutive, which ensures that motion between two successive frames is small and continuous. We describe hereafter how the motion between the reference patch and the current patch is estimated.

2.2 Motion estimation

HGI algorithm

Once the user has selected the target, motions are estimated by using a dedicated HGI algorithm inspired by Huang[6] which provides a continuous motion representation by using a grid of control points. This algorithm allows describing highly localized deformation by interpolating the motion of any pixel regarding the displacement of the four surrounding control points. In this way, the large variety of deformation provided by HGI algorithm is well-suited for describing non-rigid motions compared to others methods such that BMA, or OBMC. However, this gain is not without shortcoming due to computational demanding of the motion estimation. In the following, we describe how we design the objective function, the model complexity, and the search strategy in order to tackle both the computational issue, and the speckle noise sensitivity.

Objective function

The basic principle of HGI algorithm consists in displacing the control points in order to minimize the cost function C_{SSD} expressed as:

$$C_{SSD} = \sum_{\mathbf{p}=0}^{N_{pixels}} (I_{ref}(\mathbf{p}) - I_{cur}(\mathbf{p}+\mathbf{m_p}))^2 \qquad \mathbf{m_p} = \sum_{p=1}^{4} w_i \mathbf{m_{c_i}} \tag{1}$$

where $\mathbf{p} = (x, y)$ and $\mathbf{m_p} = (m_x, m_y)$ respectively denote the pixel coordinates of the point and its interpolated motion vector. N_{pixels} represents the total number of pixels in the image. $\mathbf{m_{c_i}}$ are the motion vectors of the four surrounding control points and w_i expresses the control point weight which depends on the distance $|\mathbf{p} - \mathbf{c_i}|$. $I_{ref}(\mathbf{p})$ and $I_{cur}(\mathbf{p})$ are respectively the intensity of the point $\mathbf{p}$ in the reference and the current patch. We chose the sum of squared difference (SSD) as similarity criterion of image intensity because it provides a good trade-off between accuracy and computation time. However, the SSD is valid only if we consider that the US echo reflected by a physical point is time independent. In figure 1, an example of the HGI grid is illustrated where control points associated to motion vectors are represented by black circles.

Model complexity

The model complexity of the HGI algorithm depends on the number of control points. A fine grid can thus provides a detailed description of deformation, but is very sensitive to the noise. Due to the presence of speckle noise in ultrasound image, we employed a hierarchical approach which splits the patch with blocks of different sizes in order to find an optimal trade-off between robustness to noise and motion vector accuracy. The block are divided regarding a threshold based on the variance difference between the reference patch, and the current patch. Consequently, the number of block is increased when details are needed.

Search strategy

Our approach is based on local search which refines control point displacements one by one. In order to provide a faster convergence of the HGI algorithm, the motion vector of each control point is first initialized by using block matching algorithm with multi-resolution search which strongly reduces the computational time. The initialization consists in finding the motion of the block surrounding the control point. The optimal position of each control point is then refined by using a diamond search which drastically reduces the computational cost by selecting only nine candidates around the control point position instead of a block. The search is stopped when the central candidate is selected as the best. Once the displacement of each control point is computed, the algorithm is reiterated until the Peak Signal to Noise Ratio (PSNR) value converges to a maximum value. An high PSNR can be obtained when the reference patch and the current patch are similar, or when the motion vectors are well estimated.

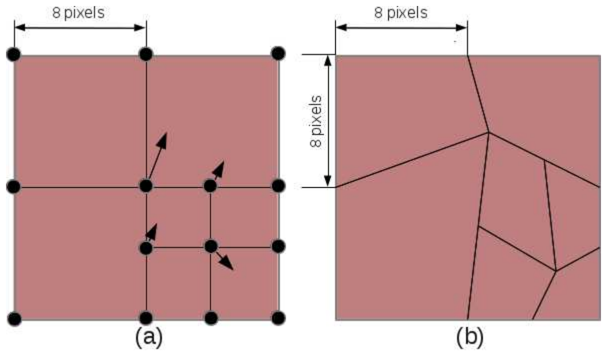


Figure 1: (a) Initial grid (b) Deformed grid

2.3 Extraction of rigid motion

As mentioned earlier, the main task of tracking target consists in finding the target rigid displacement and compensating the non-rigid deformation during the time. However, the HGI algorithm can not be used directly to track a target because it only provides a deformed grid which represents the control point displacements. In this section, we show how the rigid motion is extracted from these displacements. The rigid motion is described by the parameters t_x, t_y and θ_z which represent respectively the translation along the x-axis, the translation along the y-axis of the image, and the rotation along the z-axis. The previous parameters allow to compute the rotation matrix $\mathbf{R} = f(\theta_z)$, and the translation vector $\mathbf{t} = \begin{bmatrix} t_x & t_y \end{bmatrix}^T$ of the grid. The optimal parameters $\mathbf{R}$, $\mathbf{t}$ are found by minimizing the cost function defined by:

$$C(\mathbf{R}, \mathbf{t}) = \frac{1}{n} \sum_{i=1}^{n} ||\mathbf{c_i^{cur}} - (\mathbf{R}\mathbf{c_i^{ref}} + \mathbf{t})||^2 \tag{2}$$

where $\mathbf{c_i^{cur}}$ and $\mathbf{c_i^{ref}}$ are respectively the control points positions within the initial grid, and within the deformed grid. n represents the number of control points in the grid. Then, $\mathbf{R}$ and $\mathbf{t}$ are calculated by using a Singular Value Decomposition (SVD) inspired by Lee.[7] Each time $\mathbf{R}$ and $\mathbf{t}$ are computed, we can derive the corresponding homogeneous matrix $\mathbf{H}_{local}(k)$ which represents the target motion between the frame k and the frame $k-1$. It is given by the following expression:

$$\mathbf{H}_{local}(k) = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0_{1\times 2}} & 1 \end{pmatrix} = \begin{pmatrix} cos\ \theta_z & -sin\ \theta_z & t_x \\ sin\ \theta_z & cos\ \theta_z & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

In order to find the global motion of the target from the reference patch to the current patch, each local rigid motion is added to the global rigid motion through homogeneous matrix multiplication. It is well known that the homogeneous expression can be obtained from:

$$\underbrace{\mathbf{H}_{global}(k)}_{\substack{\text{Global motion} \\ \text{at time k}}} = \underbrace{\mathbf{H}_{local}(k)}_{\substack{\text{Local motion} \\ \text{at time k}}} \underbrace{\mathbf{H}_{global}(k-1)}_{\substack{\text{global motion} \\ \text{at time k-1}}} = \underbrace{\mathbf{H}_{local}(k) H_{local}(k-1) \mathbf{H}_{local}(k-2) \cdots \mathbf{H}_{local}(0)}_{\text{local motions}} \tag{4}$$

2.4 Reinitialization of the reference patch

As stated previously, the deformation caused by cardiac motions or patient breathing can deform the target in a periodic way. Thus, the reference patch and the current patch can be periodically similar during an amount of time. This results in a periodical evolution of the PSNR value which achieves both minimum and maximum values. However, it is well-known that a low PSNR value signifies that the current target undergoes an high deformation compared to its initial shape.

Consequently, the HGI algorithm may produce an unrealistic displacement field which can affect the tracking quality. In order to prevent this effect, a reinitialization technique is performed when a significant decrease of the PSNR value is detected. In this way, a new reference patch is created when the PSNR falls below a given threshold which ensures an high PSNR value and a good tracking. The old reference patches are kept in memory and could be reused if the PSNR value falls again below the threshold. In this way, the number of created reference patches are limited due to the periodicity of the motion, and the error due to reinitialization is minimized. This technique may allow identifying the main phases of the target shape during the breathing cycle by selecting the best reference patch.

2.5 Global Algorithm

In this section, an overview of the overall algorithm is given by using the block diagram presented in figure 2 which describes the different steps of our approach. The user first selects the target in the ultrasound image(section 2.1). This target corresponds to a particular anatomical structure which can be associated to either an hyperechogenic area (tumor), or hypoechogenic area (vessels). The motion vectors of the control points are then estimated from the HGI algorithm (section 2.2) between the current and the reference images. In order to prevent the failure of the tracking method due to high amount of deformation, the reference patch is thus reinitialized regarding the minimum threshold of PSNR value defined by the user (section 2.4). Finally, the local rigid motion of the target is extracted by using SVD decomposition (section 2.3), and is added to the global rigid motion through homogeneous matrices.

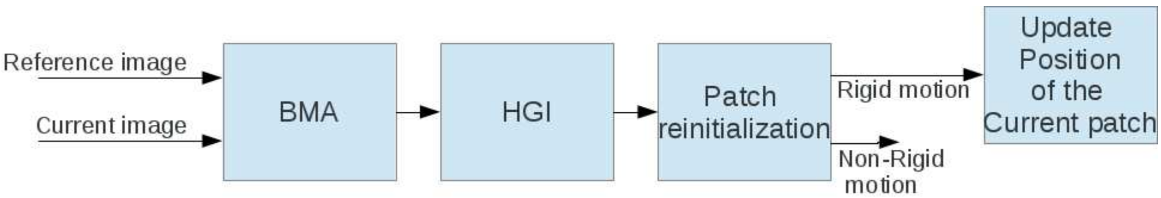


Figure 2: Block diagram of the non-rigid target tracking algorithm

3 RESULTS

The results of our target tracking method were obtained on an ultrasound sequence containing images of 640×480 pixels. The selected target represents an hyperechogenic area which undergoes both rigid and non-rigid motions. In order to test the robustness of our method, we also applied an extra simulated rigid motion to our image sequence. Toward that end, the original images are translated and rotated thanks to the software Blender. The results are presented for frame index 13, 51, 70 and 108 of the sequence in the figures 4, 5, 6, 7. The subimage (a) is the current image which changes each time a new image is acquired where the target is identified as the green square. The sub-images (b) and (c) respectively represents the reference target which can be reinitialized across the time depending on the PSNR evolution, and the current target extracted from the current image during the time. The subimage (d) describes the zoomed reference image in which we superimposed the grid of control points (red points) and its associated motions vectors (blue lines). The reference and the current patches are characterized by images of 96×96 pixels around the target.

As can be seen in the figure 3, the PSNR evolution may vary significantly from 33dB to 25dB due to the variation of the target deformation during the time. In order to prevent that the PSNR becomes unreasonably low, the re-initialization threshold value is set at 25dB. One can see that the reference patch change during the time, because the PSNR falls below 25 dB after the frame index 56, 72, 92, 114. In the figure 3, it is also shown that the first reference patch (patch 0) is used at the beginning and at the end of the sequence due to the motion periodicity which means that the target returns to its original shape due to periodical motion. The computation time of the algorithm varies between 100 ms and 600 ms for each frame depending on the similarity between the reference patch, and the current patch. These results are obtained with non-optimized C++ implementation on a 2,8 GHz dual core Intel pentium.

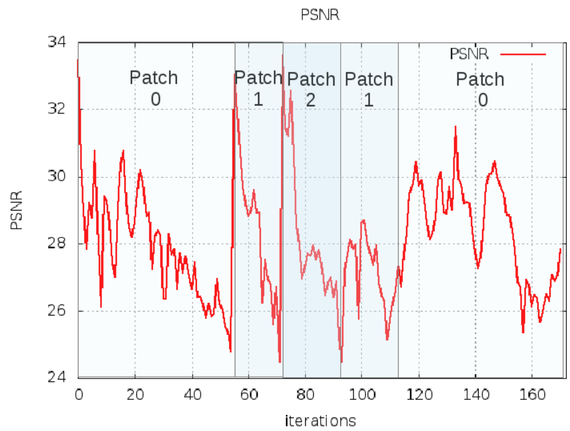
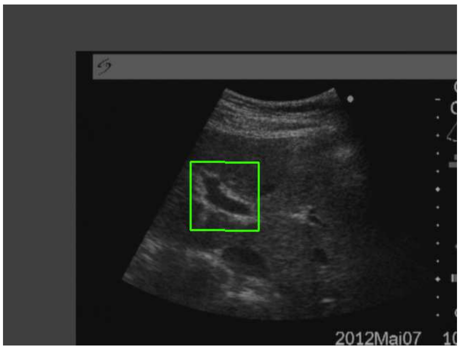
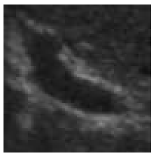
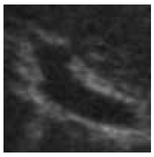


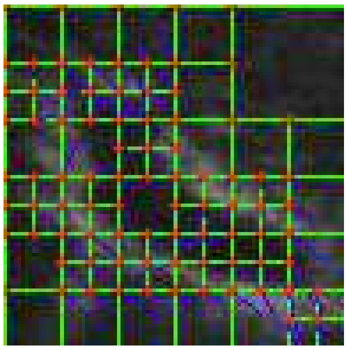
Figure 3: PSNR Evolution during 2D ultrasound sequence of 173 images (threshold=25dB)





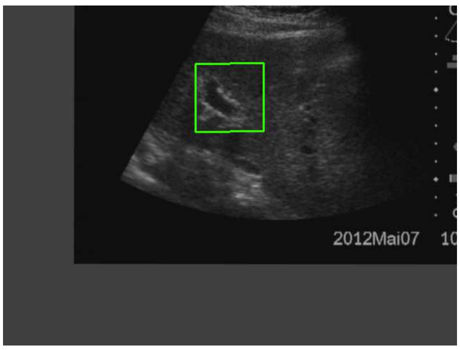
(b) Reference patch

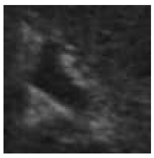




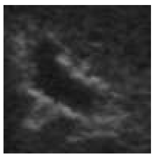
(a) Current frame (c) Current patch (d) HGI Grid

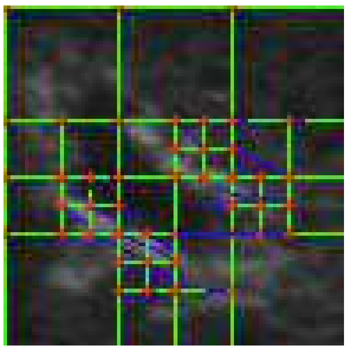
Figure 4: Target tracking for frame index 1: PSNR=33.48dB, t_x=0 pixels, t_y=0 pixels





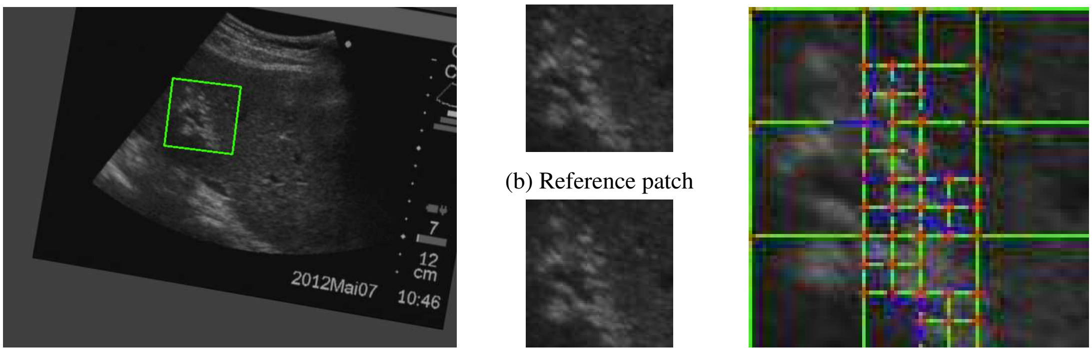
(b) Reference patch





(a) Current frame (c) Current patch (d) HGI Grid

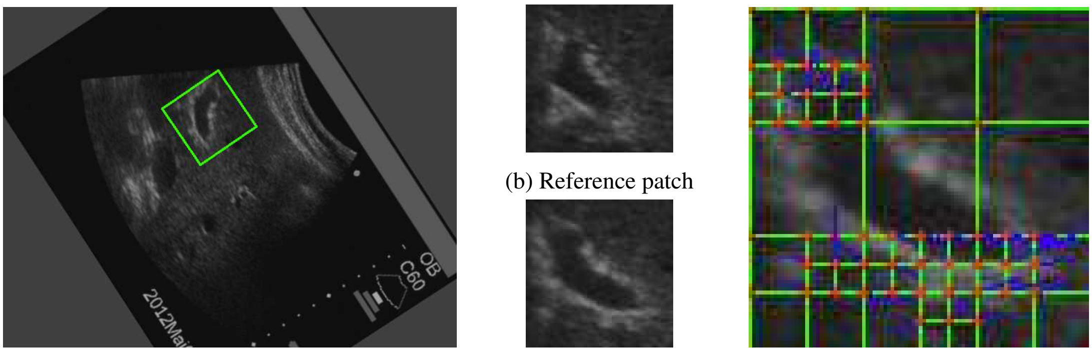
Figure 5: Target tracking for frame index 51: PSNR=26.17dB, t_x=79 pixels, t_y=-95 pixels, θ_z=1°



(b) Reference patch

(a) Current frame (c) Current patch (d) HGI Grid

Figure 6: Target tracking for frame index 70: PSNR=25.60dB, t_x=31 pixels, t_y=-76 pixels, θ_z=7°



(b) Reference patch

(a) Current frame (c) Current patch (d) HGI Grid

Figure 7: Target tracking for frame index 108: PSNR=26.46dB, t_x=-29 pixels, t_y=-80 pixels, θ_z=-55°

4 CONCLUSION

In this paper, a new approach which allows to track a deformable target within 2D ultrasound images has been proposed. The major contribution of our method consists in estimating rigid motion from the dedicated hierarchical grid interpolation algorithm which can represent various deformation types. Furthermore, the robustness is ensured by employing reinitialization technique summarized previously. As our method is automatic, the user interaction is limited to the target selection and no prior knowledge such as shape segmentation or elasticity information is needed. This method has been validated on 2D ultrasound sequences captured from human soft tissue. Several applications such as the robotized needle insertion procedures could benefit from the robustness of our approach regarding the deformation. However, several efforts have to be made in order to make this method work in real-time. Future works should not consist solely in improving computation time but also in exploiting this approach in order to track deformable target in 3D ultrasound volumes.

REFERENCES

[1] Dydenko, I., Jamal, F., Bernard, O., D'hooge, J., Magnin, I. E., and Friboulet, D., "A level set framework with a shape and motion prior for segmentation and region tracking in echocardiography," *Medical image analysis* **10**, 162–177 (Apr. 2006). 00035.

[2] Angelova, D. and Mihaylova, L., "Contour segmentation in 2D ultrasound medical images with particle filtering," *Machine Vision and Applications* **22**, 551–561 (May 2011). 00004.

[3] Yeung, F., Levinson, S., Fu, D., and Parker, K., "Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh," *IEEE Transactions on Medical Imaging* **17**(6), 945–956 (1998). 00119.

[4] Nadeau, C. and Krupa, A., "Intensity-based ultrasound visual servoing: Modeling and validation with 2-d and 3-d probes," *IEEE Transactions on Robotics* **29**(4), 1003–1015 (2013). 00000.

[5] Miki, I., Krucinski, S., Thomas, J. D., and Member, A., "Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates," *IEEE Trans. Medical Imaging* **17**, 274284 (1998). 00000.

[6] Huang, C.-L. and Hsu, C.-Y., "A new motion compensation method for image sequence coding using hierarchical grid interpolation," *IEEE Transactions on Circuits and Systems for Video Technology* **4**(1), 42–52 (1994). 00149.

[7] Lee, D. and Krupa, A., "Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4D ultrasound," in [*Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*], 28312836 (2011).