3D TRANSLATION ESTIMATION USING THE MONOGENIC ORIENTATION VECTOR

Rémi Abbal, Adrian Basarab, Denis Kouamé

Université de Toulouse
IRIT UMR CNRS 5505
118 Route de Narbonne, F-310162 Toulouse cedex 9, France

ABSTRACT

Three dimensional motion estimation is an active research field in ultrasound imaging, motivated by the recent progress in 3D acquisition. Out-of-plane motion and pattern decorrelation induced by azimuthal displacement yield bias in 2D estimation methods. However, 3D estimation may easily become time-consuming because of the large amount of data. Authors proposed different ways to estimate 3D displacement, for the most part using Normalized Cross-Correlation (NCC) combined with original refinement methods. In this paper, we propose an alternative which uses the 3D local orientation, obtained using the monogenic signal, in order to estimate 3D local translations. We show that the use of local orientation provides better results than the NCC and the classical optical flow approach. Results on a 3D simulated ultrasound volume show that the proposed estimation is more robust to noise than classical methods. Giving a signal to noise ratio of 25 dB, results show that the mean absolute error of our orientation-based optical flow estimator is respectively (47.8%, 62.1%, 84.9%) lower than the one generated when using intensity-based optical flow in lateral, azimuthal and axial direction.

***Index Terms*—** Motion Estimation, 3D Speckle Tracking, Correlation, Phase, Optical Flow, Monogenic Signal

1. INTRODUCTION

Motion estimation is used in ultrasonic imaging for several applications such as tissue motion estimation (elastography, cardiovascular disease detection) or flow estimation. Even if several 2D methods have been developed, it has been shown that they suffer from bias due to azimuthal displacement which induces out-of-plane motion and thus speckle decorrelation [1, 2]. The reference method for 2D motion estimation in ultrasound imaging is speckle tracking (ST). It is based on a local similarity criterion measure, such as sum of absolute differences (SAD) or Normalized Cross-Correlation (NCC). The lags corresponding to the maximum or minimum value of this measure correspond to the estimated displacement. In number of applications, it is necessary to interpolate the similarity measure and get subpixelic estimation to accurately estimate motion. Optical flow and phase-based approaches have also been proposed and provide subpixelic accuracy without interpolation [3, 4]. The nature of these methods restricts the estimation to only small displacements. Therefore, they are usually preceded by an initialisation using a coarse estimation based on classical NCC without interpolation [5]. Concerning the 2D phase-based methods, they use the local spatial phase extracted from 1D or 2D analytical signals [6], or more recently from the monogenic signal [7]. Thus, several studies showed that using this type of structural information gives more robust results than using amplitude-based methods such optical flow or NCC.

For 3D motion estimation in ultrasound imaging, the first proposed methods were straightforward extensions of 2D approaches such as 3D Speckle Tracking (3DST). For subvoxelic estimation, the authors proposed different techniques, running from classical interpolation to refinements based on a 1D or 2D phase-based estimation [2, 8]. For registration purpose, a new 3D cost function based on features extracted from 3D monogenic signal has also been proposed [9]. However, the main drawback of these methods, especially of 3DST, still remains the lateral and mainly azimuthal low resolutions which implies less accurate estimations in these directions.

In this paper, we propose a method based on features extracted from 3D monogenic signal. The method proposed in this paper is based on a local optical flow method [10] applied to the local 3D orientation extracted from the monogenic signal. The aim is to evaluate monogenic signal potential and the performances of its local phase and orientation information, for a local motion model corresponding to 3D translations. We compare our estimations with the results returned by intensity-based optical flow and classical 3D NCC. Since we need to estimate subvoxelic displacements, it is necessary to interpolate the NCC, unlike in local optical flow and phase-based methods.

2. BACKGROUND

2.1. 3D Normalized Cross-Correlation

This first method against which we compare the estimator proposed in the paragraph 3 is based on the 3D NCC.

$$NCC_{3D}(\mathbf{I}) = \frac{\sum_{\mathbf{x}} V(\mathbf{x},t)V^*(\mathbf{x}+\mathbf{I},t+1)}{\sqrt{\sum_{\mathbf{x}} |V(\mathbf{x},t)|^2 \sum_{\mathbf{x}} |V((\mathbf{x}+\mathbf{I},t+1)|^2}} \quad (1)$$

Where $\mathbf{I} = [i, j, k]^T$ represents the lags where the NCC is calculate, $V(\mathbf{x}, t)$ is the voxel intensity at position $\mathbf{x} = (x, y, z)$ at time t and $V^*(\mathbf{x}, t)$ is its conjugate. Translation estimation vector noted by $\hat{\mathbf{\Delta}}$ is obtained by maximizing this similarity function:

$$\hat{\mathbf{\Delta}} = \underset{\mathbf{I}}{\operatorname{argmax}}(NCC_{3D}(\mathbf{I})) \tag{2}$$

For subvoxelic estimation, 3D spline interpolation is used in this paper.

2.2. Local 3D Optical Flow

The second method against which the performance of the proposed estimator is compared is a local 3D optical flow. For one voxel, the optical flow equation can be written as follows:

$$\frac{\partial V}{\partial x}\Delta_1 + \frac{\partial V}{\partial y}\Delta_2 + \frac{\partial V}{\partial z}\Delta_3 + \frac{\partial V}{\partial t} = 0 \tag{3}$$

Where $\left[\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right]$ are the spatial derivatives of the volume V, $\frac{\partial V}{\partial t}$ is the temporal derivative and $\mathbf{\Delta} = [\Delta_1, \Delta_2, \Delta_3]^T$ is the translation vector to estimate. For a local volume of N voxels, the equation (3) can be written as:

$$\begin{bmatrix} \frac{\partial V_1}{\partial x} & \frac{\partial V_1}{\partial y} & \frac{\partial V_1}{\partial z} \\ \vdots & \vdots & \vdots \\ \frac{\partial V_N}{\partial x} & \frac{\partial V_N}{\partial y} & \frac{\partial V_N}{\partial z} \end{bmatrix} \times \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix} = - \begin{bmatrix} \frac{\partial V_1}{\partial t} \\ \vdots \\ \frac{\partial V_N}{\partial t} \end{bmatrix} \tag{4}$$

(4) can be written as $M\mathbf{\Delta} = b$, which yields the translation vector noted by $\hat{\mathbf{\Delta}}$, using the least squares method:

$$\hat{\mathbf{\Delta}} = \left(M^T M\right)^{-1} M^T b \tag{5}$$

3. PROPOSED APPROACH

3.1. 3D Monogenic Signal

In this paper, we propose a 3D translation estimator based on local features extracted from the monogenic signal. The monogenic signal is a recent generalisation of 1D analytical signal in two dimensions, proposed by Felsberg et al. in [7]. Its extension to three dimensions has the following form:

$$V_{mono}(\mathbf{x}) = p(\mathbf{x}) + iq_1(\mathbf{x}) + jq_2(\mathbf{x}) + kq_3(\mathbf{x}) \tag{6}$$

Where $\mathbf{x} = [x, y, z]^T$ is the spatial variable, p corresponds to the native data V, filtered by a passband filter f_{BP} ($p = f_{BP} \otimes V$, $\otimes$ being the convolution operator). In our case, a difference of Gaussian 3D filter was employed. q_1, q_2, q_3 are defined by $q_i = p \otimes h_i$, $i = 1, 2, 3$ and h_i are anisotropic quadrature filters expressed in the frequency domain as:

$$H_i(\mathbf{u}) = \frac{u_i}{\sqrt{u_1^2 + u_2^2 + u_3^2}}, i = 1, 2, 3 \tag{7}$$

Where $\mathbf{u} = [u_1, u_2, u_3]^T$ is the 3D frequency variable. From (6), it is possible to extract local features as amplitude A, phase φ and orientation vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$ [7, 9].

$$A(\mathbf{x}) = \sqrt{p(\mathbf{x})^2 + q_1(\mathbf{x})^2 + q_2(\mathbf{x})^2 + q_3(\mathbf{x})^2} \tag{8}$$

$$\varphi(\mathbf{x}) = arg\left(p(\mathbf{x}) + i\sqrt{q_1(\mathbf{x})^2 + q_2(\mathbf{x})^2 + q_3(\mathbf{x})^2}\right) \tag{9}$$

$$\theta_i(\mathbf{x}) = \frac{q_i(\mathbf{x})}{\sqrt{q_1(\mathbf{x})^2 + q_2(\mathbf{x})^2 + q_3(\mathbf{x})^2}}, i = 1, 2, 3 \tag{10}$$

Moreover, a phase vector $\mathbf{r} = [r_1, r_2, r_3]^T$ can be obtained by projecting the phase on the orientation $r_i(\mathbf{x}) = \varphi(\mathbf{x}) \times \theta_i(\mathbf{x})$.

3.2. Proposed Orientation-based Estimator

In the previous sections, we presented the 3D extension of monogenic signal and a 3D local optical flow estimator. In this section, we explain how we used monogenic orientation vector components to estimate local 3D translations.

For a voxel, the orientation vector $\boldsymbol{\theta}$ has three components θ_1, θ_2, θ_3. Each θ_i is obtained using the i^{th} imaginary part of the monogenic signal q_i, as shown in (10). Therefore, θ_i depends on the local orientation of the i^{th} anisotropic quadrature filter h_i. In other words, θ_i contains signal information for the i^{th} direction. Local phase φ, orientation and phase vectors, $\boldsymbol{\theta}$ and $\mathbf{r}$, all contain structure information on the data. As a consequence, it is obvious that they contain also the motion information between the two volumes data. Thus, we propose in this paper to use them instead of the amplitude V, in order to estimate motion in a local optical flow manner. Moreover, the originality of this paper is that we show that among all these local features, the local orientation gives the best result, when the i^{th} θ is used to estimate the motion in the i^{th} direction. With this assumption, the three optical flow equations employed are:

$$\begin{bmatrix} \frac{\partial \theta_{i_1}}{\partial x} & \frac{\partial \theta_{i_1}}{\partial y} & \frac{\partial \theta_{i_1}}{\partial z} \\ \vdots & \vdots & \vdots \\ \frac{\partial \theta_{i_N}}{\partial x} & \frac{\partial \theta_{i_N}}{\partial y} & \frac{\partial \theta_{i_N}}{\partial z} \end{bmatrix} \times \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix} = - \begin{bmatrix} \frac{\partial \theta_{i_1}}{\partial t} \\ \vdots \\ \frac{\partial \theta_{i_N}}{\partial t} \end{bmatrix} \tag{11}$$

The equation (11) can be written as $M_i\Delta = b_i$, for i running from 1 to 3. Taking into account the explanations above, the retained estimator of Δ_i is:

$$\hat{\Delta}_i = \left\{\left(M_i^T M_i\right)^{-1} M_i^T b_i\right\}_i \tag{12}$$

Where $\{\mathbf{.}\}_i$ is the i^{th} component of the vector $\mathbf{.}$.

4. SIMULATION AND RESULTS

4.1. 3D ultrasound simulation

The motion estimation was performed on a 3D B-mode volume simulated with Fusk (Fast Ultrasound imaging Simulation in K-space) [11]. The dimension of the simulated

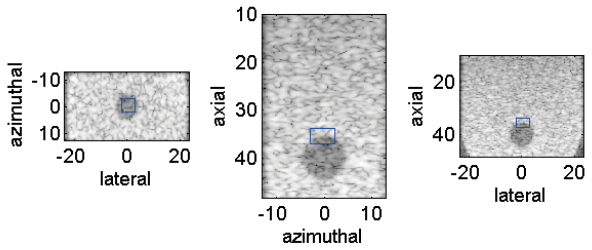


Fig. 1. Slices of a 3D volume containing a sphere simulated with Fusk.

medium was (50, 30, 50) mm in the lateral, azimuthal and axial dimensions, respectively. As shown on Fig.1, the simulated tissue represented an homogenous medium with a spherical inclusion in the center. To simulate it, scatterers were generated at random positions, with a density of 200 scatterers per mm^3. Their amplitudes were generated randomly using a normal distribution, with different standard deviations for the inclusion and for the surrounding medium. The focal depth was fixed at 40 mm. The transmit frequency was set at 2.5 MHz, with a bandwidth of 1 MHz. The estimated ultrasound volume was sampled so that the dimension of one voxel was 0.15x0.15x0.1 mm^3.

Two such volumes were generated. Between the two volumes, a constant displacement was imposed (on the scatterers), of (0.075, 0.0325, 0.075) mm in the lateral, azimuthal and axial directions. In voxels, the imposed displacement was (0.5, 0.25, 0.75). The proposed estimator and the classical ones were performed on a small region extracted from these volumes, of 4.5x4.5x3.4 mm^3, as shown by the rectangles superimposed to the images in Fig.1.

In order to evaluate the performances of the different estimators, additive Gaussian noise was added to the volumes, corresponding to Signal to Noise Ratios (SNR) running from 0 dB to 50 dB. For the monogenic signal, as a band pass filtering is processed, the noise was not added to the native volumes, but on each component of the monogenic signal (p, q_1, q_2 and q_3). This way, each method is evaluated with the same SNR. For each SNR level, 128 tries were performed, and the mean and standard deviation errors were compared.

4.2. Results

First, the results obtained using optical flows on features extracted from the monogenic signal are shown. On Fig.2, the results obtained for each direction are given for all the three orientation maps. As predicted in paragraph 3.2 the best result for each Δ_i is given by the corresponding θ_i. Indeed, the bias is largely reduced compared to the other estimations. We can note that the standard deviations are slightly the same and very small. This is due to the way that the optical flow is processed, using a least square method. On Fig.3, the same type of result is obtained for the phase vectors r_i. Results on figures 2 and 3 let us conclude that the retained estimators, when orientation and phase vectors are used, are, for each direction,

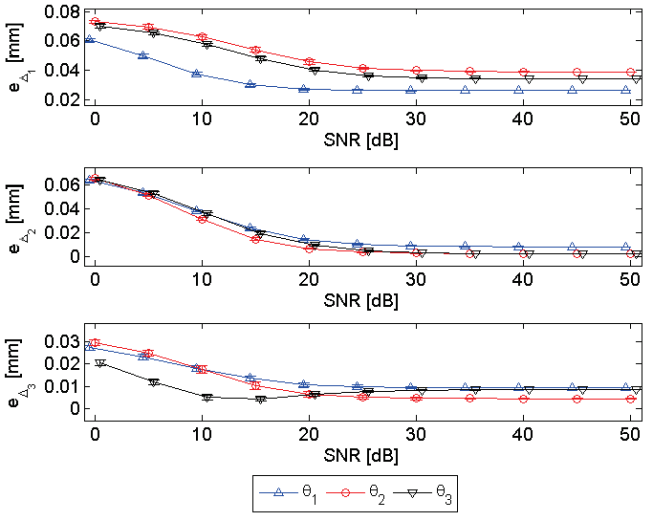


Fig. 2. Comparison of mean and standard deviation values of the absolute error between true and estimated displacement vectors using orientation vector components. Absolute errors for $\Delta_1, \Delta_2, \Delta_3$ are displayed respectively from top to bottom.

those provided by the corresponding component.

Further, the proposed estimators are compared to classical optical flow, optical flow applied to monogenic phase φ and NCC. As the imposed displacement is subvoxelic, the NCC was interpolated using splines. The interpolation factor was set at 4 in each direction. The choice was made so that the final sampling of the correlation function corresponds to the imposed displacement. Despite this optimal choice, the NCC was largely biased for our US simulation. This bias is certainly caused by the large size of the speckle, caused by the characteristics of the chosen Point Spread Function (PSF). However, the choice of these parameters was made so that they correspond to a realistic ultrasound scanner.

Figure 4 shows that globally, the proposed orientation based estimator has the smallest bias among all the methods: averaging for all SNR estimations, the absolute error of this estimator is 33.3% and 54.8% lower than intensity-based optical flow estimation, and 15.7%, 50.3% lower than the phase-based estimator, respectively for lateral and azimuthal directions. Axially, the error is 62.5% higher than intensity-based approach and 76.8% higher than the phase-based method. The intensity and phase based approaches give better results in axial direction, for SNR larger than 35 dB, due to the sphere contour. This edge is orthogonal to the axial motion component and improves the axial optical flow. On the other hand, the error is (12.4%, 32.4%, 27.7%) lower than the method using phase vector components, respectively for lateral, azimuthal and axial directions. The phase vectors give less accurate results than the proposed orientation-based estimator. The main reason is that in the case of our simulated volume, combining phase and orientation yields a very smooth feature, which biases optical flow method.

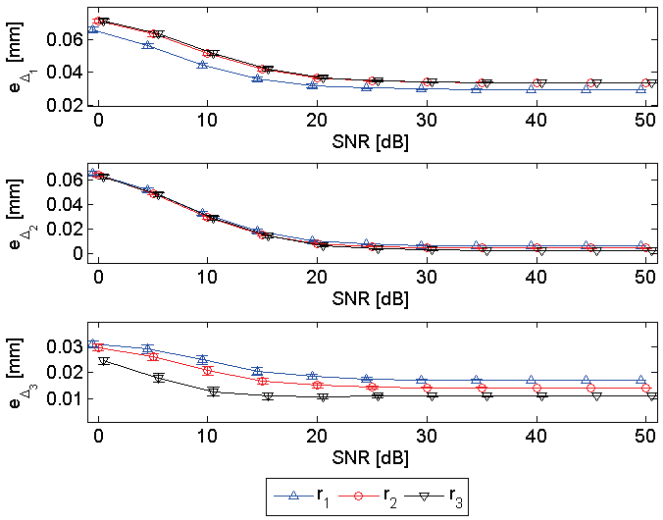


Fig. 3. Comparison of mean and standard deviation values of the absolute error between true and estimated displacement vectors using phase vector components. Absolute errors for $\Delta_1, \Delta_2, \Delta_3$ are displayed respectively from top to bottom.

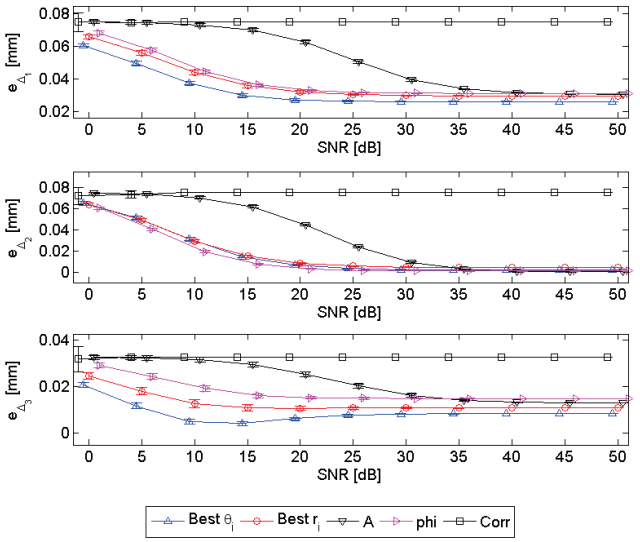


Fig. 4. Comparison of mean and standard deviation values of the absolute error between true and estimated displacement of correlation, optical flow using amplitude, phase, phase vector and orientation vector. Best θ_i and best r_i are orientation and phase components giving the lowest bias in figures 2 and 3. Absolute errors for $\Delta_1, \Delta_2, \Delta_3$ are displayed respectively from top to bottom.

5. DISCUSSION AND CONCLUSION

We have shown in this preliminary work that using angles information, such as phase and orientation, is an interesting alternative for motion estimation. The main purpose was to evaluate, in a controlled simulated case, the contribution of using features issued from the monogenic signal. The main result shows that the best choice of estimating 3D translations, with an optical flow approach, is the orientation vector. Moreover, we showed that each component of the orientation vector was the most appropriate choice for estimating the motion in its direction. The proposed estimator outperformed classical optical flow using the amplitude of the data and 3D normalized cross-correlation. Moreover, displacement models more complex than rigid translations, such as the affine transformation, will be studied in order to show the gain of accuracy using angular feature optical flows.

6. REFERENCES

[1] J.A. Hossack, "Influence of elevational motion on the degradation of 2d image frame matching," in *Ultrasonics Symposium, 2000 IEEE*, oct 2000, vol. 2, pp. 1713 –1716 vol.2.

[2] X Chen, H Xie, R Erkamp, K Kim, C Jia, J M Rubin, and M O'Donnell, "3-d correlation-based speckle tracking.," *Ultrasonic Imaging*, vol. 27, no. 1, pp. 21–36, 2005.

[3] P. Baraldi, A. Sarti, C. Lamberti, A. Prandini, and F. Sgallari, "Evaluation of differential optical flow techniques on synthesized echo images," *Biomedical Engineering, IEEE Transactions on*, vol. 43, no. 3, pp. 259 –272, march 1996.

[4] Adrian Basarab, Hervé Liebgott, and Philippe Delachartre, "Analytic estimation of subsample spatial shift using the phases of multidimensional analytic signals," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 440–447, january 2009.

[5] Adrian Basarab, Pierre Gueth, Hervé Liebgott, and Philippe Delachartre, "Phase-based block matching applied to motion estimation with unconventional beamforming strategies," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 56, no. 5, pp. 945–957, may 2009.

[6] G. Sommer T. Bulow, "Hypercomplex signals - a novel extension of the analytic signal to the multidimensional case," *IEEE Trans. Signal Processing*, vol. 49, pp. 2844, 2001.

[7] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Processing*, vol. 49, pp. 3136, 2001.

[8] Andriy Myronenko, Xubo B. Song, and David J. Sahn, "Lv motion tracking from 3d echocardiography using textural and structural information.," in *MICCAI (2)'07*, 2007, pp. 428–435.

[9] Vicente Grau, Harald Becher, and J. Alison Noble, "Phase-based registration of multi-view real-time three-dimensional echocardiographic sequences," in *MICCAI (1)*, 2006, pp. 612–619.

[10] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in *IJCAI81*, 1981, pp. 674–679.

[11] T. Hergum, S. Langeland, E.W. Remme, and H. Torp, "Fast ultrasound imaging simulation in k-space," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 56, no. 6, pp. 1159 –1167, june 2009.