An AM-FM model for Motion Estimation in Atherosclerotic Plaque Videos

V. Murray*, S.E. Murillo*, M.S. Pattichis*, C.P. Loizou†, C.S. Pattichis‡, E. Kyriacou§ and A. Nicolaides¶

*Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM-87131
e-mails: vmurray@ieee.org, {smurillo, pattichis}@ece.unm.edu
†Department of Computer Science
Intercollege Limassol Campus, CY-3507 Limassol, Cyprus
e-mail: christosl@lim.Intercollege.ac.cy
‡Department of Computer Science
University of Cyprus, CY1678 Nicosia, Cyprus
e-mail: pattichi@ucy.ac.cy
§Cyprus Institute of Neurology and Genetics, CY1683, Nicosia, Cyprus
e-mail: ekyriac@ucy.ac.cy
¶Imperial College, London, England
University of Cyprus and Vascular Screening and Diagnostic Center
Nicosia, Cyprus
e-mail: anicolai@cytanet.com.cy

***Abstract*—We present new multidimensional Amplitude-Modulation Frequency-Modulation (AM-FM) methods for motion estimation. For a single AM-FM component we show that the optical flow constraint leads to separate equations for amplitude modulation (AM) and frequency modulation (FM). We compare our approach with phase-based estimation developed by Fleet and Jepson and also the original optical flow method by Horn and Schunck. An advantage of the proposed method is that it provides for dense estimates that remain accurate over the entire video. We also present preliminary results on atherosclerotic plaques videos where the AM method appears to work best.**

I. Introduction

In this paper, we present a new method for pixel-level based motion estimation using an Amplitude-Modulation Frequency-Modulation (AM-FM) model for digital video. We also consider an application of this model in analyzing atherosclerotic plaque motion.

Our study is motivated from a desire to extend traditional motion estimation methods into the development of reliable methods for video trajectory estimation. To accomplish this for atherosclerosis, we are interested in developing realistic plaque motion models that are motivated from clinical experience. We expect that accurate motion estimation will help us develop more accurate models that can predict plaque rupture.

In related work [1], the authors computed optical flow estimates from 45 patients and reported a significant increase in the maximal discrepant surface velocity for the symptomatic cases, as compared to the asymptomatic cases. In [2], the authors used 3D intravascular ultrasound to provide a computational analysis of stress distribution. Clearly, the development of accurate motion estimation methods can also benefit this related research.

Using realistic motion models, we propose a new, amplitude and phase based motion estimation method based on robust, multidimensional Amplitude-Modulation Frequency-Modulation (AM-FM) methods. Extending out prior work reported in [3], we investigate the limits of the new method as compared against traditional motion estimation methods.

We provide background information about related work in section II. We describe our new method in section III. Results are shown in section IV and finally, the conclusions are given in section V.

II. Background

A. Phase-Based Methods for Motion Estimation

Fleet and Jepson [4] proposed the use of an AM-FM model for modeling digital video based on

$$\mathbf{R}(\mathbf{x},t) = \rho(\mathbf{x},t)\, e^{j\varphi(\mathbf{x},t)} \tag{1}$$

where $\mathbf{x}$ denotes the spatial variables $\mathbf{x} = (x, y)$, t denotes time, $\rho(\mathbf{x},t)$ denotes the amplitude and $\varphi(\mathbf{x},t)$ denotes the phase component. For estimating the AM-FM components, Fleet and Jepson used a set of band-pass complex value Gabor filters.

For pixel velocity estimation, the basic model of (1) can only be used to estimate the projected component velocities that are in the direction of the instantaneous frequency. We express this using $\mathbf{v}_n = \alpha\, \mathbf{n}(\mathbf{x},t)$ where $\varphi_{\mathbf{x}}(\mathbf{x},t) = (\varphi_x, \varphi_y)^T$, $\mathbf{n}(\mathbf{x},t) = \varphi_{\mathbf{x}}(\mathbf{x},t) / \|\varphi_{\mathbf{x}}(\mathbf{x},t)\|$ and $\alpha = -\varphi_t(\mathbf{x},t) / \|\varphi_{\mathbf{x}}(\mathbf{x},t)\|$.

978-1-4244-2110-7/08/$25.00 ©2007 IEEE

For estimating the instantaneous frequency, they proposed to use:

$$\nabla\varphi(\mathbf{x},t) = \frac{Im\left[\mathbf{R}^*(\mathbf{x},t)\,\nabla\mathbf{R}(\mathbf{x},t)\right]}{\rho^2(\mathbf{x},t)} \tag{2}$$

where $\mathbf{R}^*$ denotes the complex conjugate of $\mathbf{R}$, $Im[\mathrm{z}] \equiv (Im[z_1], Im[z_2], Im[z_3])$.

Note that in order to recover the velocity components that are orthogonal to the instantaneous frequency vectors we will need to apply some type of smoothing over the estimated velocities. Fleet and Jepson accomplish this by fitting a local linear model over 5×5 neighborhoods:

$$\mathbf{v}(\mathbf{x},t) = (\alpha_0 + \alpha_1 x + \alpha_2 y, \beta_0 + \beta_1 x + \beta_2 y)$$

Then, from (II-A), the estimated 2D velocity is then taken to be (α_0, β_0).

To provide for a method of identifying accurate measures, Fleet and Jepson require that the estimated instantaneous frequency is within the range of the estimating filter. This is expressed as

$$\|\nabla\varphi(\mathbf{x},t) - (k_i, w_i)\| \leq \tau\sigma_k \tag{3}$$

where (k_i, w_i) is the peak tuning frequency of the i-th filter, σ_k is the standard deviation of the filter's amplitude spectrum, and τ is a threshold used to reject unreliable estimates of instantaneous frequencies. Similarly, for the amplitude, they require that the local signal amplitude must be as large as the average local amplitude, and at least 5% of the largest response amplitude across all the filters at that frame. When either one of these two conditions is not met, the method does not provide velocity estimates.

B. Discrete-Space Demodulation Using the Quasi Eigenfunction Approximation

The proposed method is based on a 3D extension of the separable 2D method presented in [5]. Instead of a 3D Gabor filterbank, we propose the use of a multiscale, separable 3D filterbank. First, we take a 3D FFT of N frames of the input video. If we let $I(x,y,t)$ denote the input video, we then use $\tilde{I}(\omega_x, \omega_y, \omega_t)$ to denote its DFT. Similar to the 2D case, we zero-out all the frequency components that have negative ω_t. In other words, we set $\tilde{I}(\omega_x, \omega_y, \omega_t < 0) = 0$. Let $\hat{I}_{AS}$ denote the resulting video.

III. Methods

A. An AM-FM model for motion estimation

In what follows, consider a single AM-FM component approximation to the input video:

$$I(x,y,t) = a(x,y,t)\exp(j\varphi(x,y,t)). \tag{4}$$

Recall the optical flow constraint equation:

$$I_x u + I_v v + I_t = 0. \tag{5}$$

We apply (5) to (4) and separate out the real from the imaginary parts to get the amplitude constraint equation

$$a_x u + a_y v + a_t = 0, \tag{6}$$

and the frequency modulation constraint equation

$$\varphi_x u + \varphi_y v + \varphi_t = 0. \tag{7}$$

The advantage of (6)-(7) is that they provide us with two equations per pixel. We also add a smoothness constrain to add a third equation. Collectively, the AM, FM and continuity constraints give:

$$E_s = \iint \left[u_x^2 + u_y^2 + v_x^2 + v_y^2\right] dxdy. \tag{8}$$

$$E_{AM} = \iint \left[a_x u + a_y v + a_t\right]^2 dxdy \tag{9}$$

$$E_{FM} = \iint \left[\varphi_x u + \varphi_y v + \varphi_t\right]^2 dxdy. \tag{10}$$

We combine all constraints together to get

$$E = E_s + \lambda E_{FM} + \beta E_{AM}. \tag{11}$$

B. Discrete Optimization

For minimizing (11) we consider finite-difference approximations:

$$\begin{aligned} u_x &\approx [u(i+1,j) - u(i,j)]/2 \\ u_y &\approx [u(i,j+1) - u(i,j)]/2 \\ v_x &\approx [v(i+1,j) - v(i,j)]/2 \\ v_y &\approx [v(i,j+1) - v(i,j)]/2. \end{aligned}$$

We then convert the sums to integrals and take derivatives with respect to $u(m,n)$ and $v(m,n)$ to get

$$\begin{aligned} \frac{\partial E}{\partial u(m,n)} &= 2\left[u(m,n) - u_{ave}(m,n)\right] \\ &\quad + 2\lambda\left[\varphi_x u + \varphi_y v + \varphi_t\right]\cdot\varphi_x \\ &\quad + 2\beta\left[a_x u + a_y v + a_t\right]\cdot a_x = 0 \\ \frac{\partial E}{\partial v(m,n)} &= 2\left[v(m,n) - v_{ave}(m,n)\right] \\ &\quad + 2\lambda\left[\varphi_x u + \varphi_y v + \varphi_t\right]\cdot\varphi_y \\ &\quad + 2\beta\left[a_x u + a_y v + a_t\right]\cdot a_y = 0. \end{aligned}$$

where

$$\begin{aligned} u_{ave}(m,n) &= \frac{1}{4}[u(m+1,n) + u(m-1,n) + u(m,n+1) \\ &\quad + u(m,n-1)] \quad \text{and} \\ v_{ave}(m,n) &= \frac{1}{4}[v(m+1,n) \\ &\quad + v(m-1,n) + v(m,n+1) + v(m,n-1)] \end{aligned}$$

For optimal values, we require that the gradient should be zero

$$\frac{\partial E}{\partial u(m,n)} = \frac{\partial E}{\partial v(m,n)} = 0. \tag{12}$$

We re-write (12) as

$$\mathsf{A}\mathsf{x} = \mathsf{b},$$

where:

$$\mathsf{A} = \begin{bmatrix} (1+\lambda\varphi_x^2+\beta a_x^2) & (\lambda\varphi_x\varphi_t-\beta a_x a_y) \\ (\lambda\varphi_x\varphi_t-\beta a_x a_y) & (1+\lambda\varphi_y^2+\beta a_y^2) \end{bmatrix},$$
$$\mathsf{x} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \text{and}$$
$$\mathsf{b} = \begin{bmatrix} (u_{ave}-\lambda\varphi_x\varphi_t-\beta a_x a_t) \\ (v_{ave}-\lambda\varphi_y\varphi_t-\beta a_y a_t) \end{bmatrix}.$$

Solving for $u(m,n)$ and $v(m,n)$ yields:

$$u = u_{ave} - \lambda\frac{u_{N1}}{D}\varphi_x - \beta\frac{u_{N2}}{D}a_x$$
$$v = v_{ave} - \lambda\frac{v_{N1}}{D}\varphi_y - \beta\frac{v_{N2}}{D}a_y$$

where:

$$\begin{aligned} u_{N1} &= \left(\varphi_x+\beta a_y^2\varphi_x-\beta a_y a_x\varphi_y\right)u_{ave}+\varphi_y v_{ave} \\ &\quad +\left(\beta a_y^2+1\right)\varphi_t-\beta a_y\varphi_y a_t \\ u_{N2} &= \left[\lambda\left(a_x\varphi_y-a_y\varphi_x\right)\varphi_y+a_x\right]u_{ave}+a_y v_{ave} \\ &\quad -\lambda a_y\varphi_y\varphi_t+\left(\lambda\varphi_y^2+1\right)a_t \\ D &= 1+\lambda\left(\varphi_y^2+\varphi_x^2\right)+\lambda\beta\left(\varphi_x a_y-\varphi_y a_x\right)^2 \\ &\quad +\beta\left(a_x^2+a_y^2\right), \end{aligned}$$

and similarly for v_{N1}, v_{N2}, after exchanging x with y derivatives.

The instantaneous frequency components are estimated using the AM-FM demodulation process described in section II. For each video, we compute the dominant AM-FM components over a three-scale 3D filterbank. To estimate the amplitude derivatives a_x, a_y, a_t, we take differences of Gaussian averages over 3×3 neighborhoods ($\sigma = 1$).

Using the derivative estimates, we iteratively compute velocity estimates at each iteration p:

$$u^{(p+1)} = u_{ave}^{(p)} - \lambda\frac{u_{N1}}{D}\varphi_x - \beta\frac{u_{N2}}{D}a_x \tag{13}$$
$$v^{(p+1)} = v_{ave}^{(p)} - \lambda\frac{v_{N1}}{D}\varphi_y - \beta\frac{v_{N2}}{D}a_y. \tag{14}$$

We initialize the estimation by using zero-velocity estimates.

Using the estimated motion vectors, we apply a Kalman filter to track the trajectories throughout the video. In summary, for the clinical videos, we establish the validity of the estimated trajectories based on: (i) the density of the estimated velocities, (ii) we require that trajectories remain valid throughout the video, (iii) estimation consistency and (iv) agreement with clinical expectations.

IV. Results

We apply the developed AM-FM method for motion estimation using separable 3D filter-banks. We use a dyadic decomposition with three-levels of decomposition (see [5] for a 2D example). Furthermore, at each pixel, we compute motion estimates over the 3D channel that gives the maximum response (dominant component analysis).

For the examples that we present we set the constraint parameters in (11) using:

- AM only: $\lambda = 0$, $\beta = 10$.
- FM only: $\lambda = 10$, $\beta = 0$.
- AM-FM: $\lambda = 10$, $\beta = 10$.

Here, we note that in (11), the continuity constraint is weighted by 1. Thus, in all of our experiments, we have placed a significantly more weight on the new AM-FM methods.

A. Results on a synthetic example

To test the implementations, we first present results using motion generated using a chirp image. This is shown in Fig. 1. The chirp image is given by

$$I(n_1,n_2) = \cos\varphi(n_1,\ n_2)$$

where

$$\varphi(n_1,n_2) = 2\pi\left[\alpha_1 n_1+\beta_1\frac{n_1^2}{2}+\alpha_2 n_2+\beta_2\frac{n_2^2}{2}\right] \tag{15}$$
$$\nabla\varphi(n_1,n_2) = 2\pi\left(\alpha_1+\beta_1 n_1,\ \alpha_2+\beta_2 n_2\right) \tag{16}$$

with parameter values described in the caption of Fig. 1. Here, we note that the chirp image covers all the possible discrete values of the instantaneous frequency.

As documented in [6], we are interested in simulating periodic motion that closely resembles atherosclerotic plaque motion. This is accomplished by using:

$$\begin{aligned} x(t) &= A_h\sin(\frac{2\pi}{N}f_h t)+\frac{A_h}{2}\sin(\frac{2\pi}{N}(2f_h)) \\ &\quad +\frac{A_h}{3}\sin(\frac{2\pi}{N}(3f_h)) \end{aligned} \tag{17}$$

and

$$\begin{aligned} y(t) &= A_v\sin(\frac{2\pi}{N}f_v t)+\frac{A_v}{2}\sin(\frac{2\pi}{N}(2f_v)t) \\ &\quad +\frac{A_v}{3}\sin(\frac{2\pi}{N}(3f_v)t). \end{aligned} \tag{18}$$

As noted in [6], as measured in atherosclerotic videos, the harmonic amplitudes decay at a rate that is inversely proportional to the harmonic frequency. This is consistent with a discontinuity in the motion and has also been observed in the power spectra of the estimated trajectories.

We compare the AM-FM results against optimal results obtained with Horn's method and also the phase-based method developed by Fleet and Jepson. We report the results in Table I. For Horn, we compared against two different runs using a Lagrange multiplier of $\alpha = 10$, 20 for cases 1 and 2. In both cases, for the Gaussian smoothing parameter, we used $\sigma = 1.00$. For the Fleet and Jepson approach, to obtain a reasonable density, for the valid constraint: $\|\nabla\varphi-(k_i,w_i)\| \le \tau\sigma_k$ and the amplitude constraint (see subsection II-A), we used

- Case 1: $\sigma = 0.75$, $\text{Amp}_{threshold}$ = 1%, τ = 2,
- Case 2: $\sigma = 1.00$, $\text{Amp}_{threshold}$ = 1%, τ = 2,

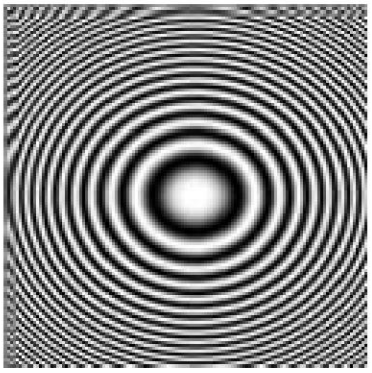

Fig. 1. Chirp used for our first test in motion estimation. For the simulation, from (15), we use square images of size $N = 128$, $\alpha_1 = \alpha_2 = \frac{1}{N-1}\frac{9}{40}$, and $\beta_1 = \beta_2 = \frac{2}{N-1}\frac{9}{40}$. This gives an instantaneous frequency magnitude range interval of $[-\frac{9\pi}{20}, \frac{9\pi}{20}]$ for both directions. For the motion, we set $A_h = A_v = 2$, and $f_h = f_v = 1$.

TABLE I
MSE AND DENSITY IN THE VELOCITY ESTIMATION OF THE SYNTHETIC CHIRP SIGNAL.

Method	Motion Magnitude	Density
Fleet (1)	Did not work	0%
Fleet (2)	0.0075	10.9851%
Fleet (3)	0.0140	33.4311%
Fleet (4)	0.0073	9.6752%
Horn (1)	0.0141	28.9235%
Horn (2)	0.0153	28.9235%
FM	0.0081	100%
AM	0.0090	100%
AM-FM	0.0085	100%

- Case 3: $\sigma = 1.00$, $\text{Amp}_{threshold} = 5\%$, $\tau = 10$,
- Case 4: $\sigma = 1.25$, $\text{Amp}_{threshold} = 1\%$, $\tau = 2$.

In Table I, we report the mean-square-error (MSE) for the magnitude of the estimated velocity vectors and the density of the estimates.

In comparing the results we note that the AM-FM approach provides reliable motion estimates at every pixel. In Table I, this is demonstrated by the fact that AM, FM and AM-FM estimates have a density of 100%. On the other hand, the methods by Horn and Fleet and Jepson do not produce densities over 30%. This is attributed to the fact that the filter-banks used in the proposed approach cover the entire frequency plane.

In terms of accuracy, we note that the phase-based methods gave the best results. At 100% density, the proposed method achieved an accuracy that is comparable to what Fleet and Jepson achieved at 10%. It is also interesting to note that the AM method gave very good results and that the combined AM-FM method gave an accuracy that ranges between the AM and the FM methods. Here, we note that since the chirp has constant amplitude, the generated AM comes from processing the motion through the dyadic filter-bank.

B. Results on atherosclerotic plaque videos

We present motion estimation results on four atherosclerotic plaque videos. We present the first video frames in Fig. 2 and the tracked video frames in Fig. 3. For estimating the motion, we applied the AM, the FM, and the AM-FM method.

To verify tracking, we magnified the plaque regions that we wanted to analyze and carefully examined how each pixel is tracked through time. We also note that transducer motion can cause artificial plaque motion and this is something that we need to guard against. To avoid this problem we carefully examined the videos to confirm that: (i) we do not have significant intensity changes from frame to frame and (ii) there is a clear periodicity in the video frames, showing that there is no detectable drift between consecutive cardiac cycles. Now, having confirmed that there is no drift, we are led to believe that the motion estimates tended to be unbiased and thus any small, noisy variations were correctly filtered out by the Kalman filter trackers.

The results indicated that the best results were obtained by the AM method, closely followed by the full AM-FM and then the FM method. Apparently, ultrasound speckle dominated the FM while the AM captured the essential brightness variations without the noise. As a result, the Fleet and Jepson phased based method (similar to FM) was not able to track the pixel motion through the video. For Fleet and Jepson, the periodic motion resulted in stationary estimates and the motion was not tracked throughout the periods. On the other hand, the optimized Horn's method appeared to track the pixels within the period, similar to the newly developed AM method. The correct pixel tracking is clearly visible in the examples in Fig. 3.

V. CONCLUSIONS

We presented new AM-FM methods for motion estimation, allowing us to estimate pixel motion (requiring two displacements vectors (u, v)) with up to three equations per pixel. The new methods use a multi-scale filterbank (with three scales) to cover the entire frequency plane, allowing us to provide dense motion estimates. From the results on atherosclerotic plaque videos, we show that the Amplitude Modulation method can provide visibly correct tracking over the plaques.

We are currently in the process of collecting more videos and also evaluating the motion as a function of the cardiac cycle. We are also researching methods for extending the method for estimating strain on the plaques.

VI. ACKNOWLEDGEMENTS

The authors would like to thank Maura Griffin (Ph.D.) and Ms. Niki Georgiou (S.R.N.) for selecting the case studies and carefully collecting the ultrasound videos that have made this research possible.

REFERENCES

[1] S. Meairs and M. Hennerici, "Four-dimensional ultrasonographic characterization of plaque surface motion in patients with symptomatic and asymptomatic carotid artery stenosis," *Stroke*, vol. 30, no. 9, pp. 1807–1813, Sep. 1999.

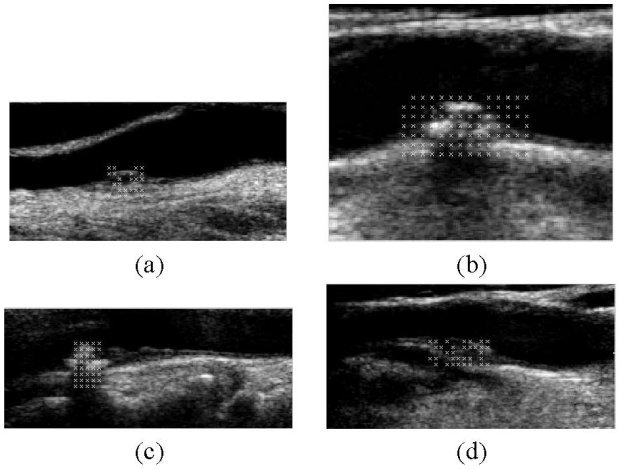

Fig. 2. First frame of atherosclerotic plaque videos that were used. We display a 'x' over pixels where the Kalman filter was able to track throughout the video. The individual video characteristics are as follows: (a) video of size 125 rows $\times 250$ columns $\times 67$ frames, (b) video of size 125 rows $\times 150$ columns $\times 111$ frames, (c) video of size 100 rows $\times 240$ columns $\times 87$ frames, and (d) video of size 125 rows $\times 250$ columns $\times 67$ frames.

[2] K. Imoto, T. Hiro, T. Fujii, A. Murashige, Y. Fukumoto, G. Hashimoto, T. Okamura, J. Yamada, K. Mori, and M. Matsuzaki, "Longitudinal structural determinants of atherosclerotic plaque vulnerability a computational analysis of stress distribution using vessel models and three-dimensional intravascular ultrasound imaging," *J. Am. Coll. Cardiol.*, vol. 46, pp. 1507–1515, Sept. 2005.

[3] S. Murillo, M. Pattichis, C. Loizou, C. Pattichis, E. Kyriacou, A. Constantinides, and A. Nicolaides, "Atherosclerotic plaque motion analysis from ultrasound videos," *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pp. 836–840, Oct.-Nov. 2006.

[4] D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 77–104, 1990.

[5] V. Murray, P. Rodriguez V., and M. S. Pattichis, "Robust multiscale am-fm demodulation of digital images," *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 1, pp. I –465–I –468, Sept. 16 2007-Oct. 19 2007.

[6] S. Murillo, M. Pattichis, C. Loizou, C. Pattichis, and E. Kyriakou, "Atherosclerotic plaque motion trajectory analysis from ultrasound videos," in *IEEE Int. Special Topic Conf. on Inform. Technology in Biomedicine*, Ioannina, Greece, Oct. 2006, pp. 1–5.

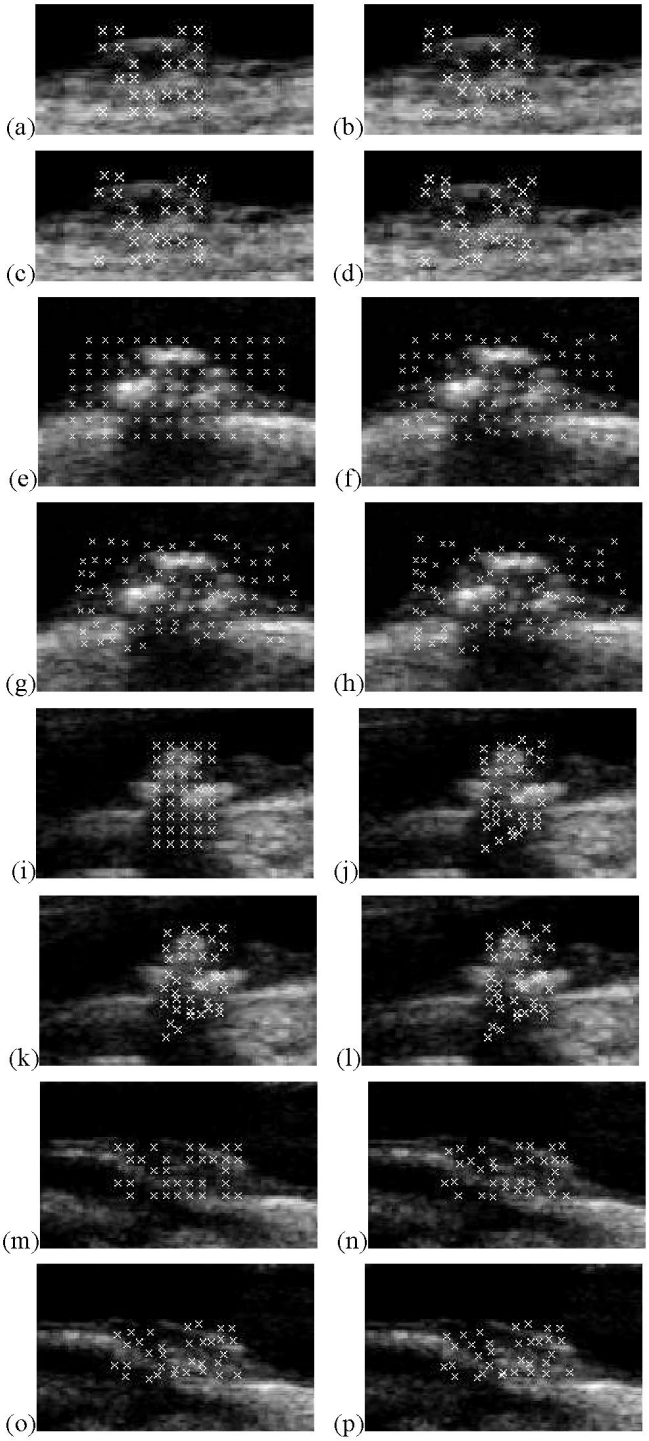

Fig. 3. Pixel tracking results from frames 1, 21, 41 and 61 for each video. For the video shown in Fig. 2(a), we have extracted a Region of Interest (ROI) of $30-140 \times 30-210$. We then show frame 1 in (a), frame 21 in (b), frame 41 in (c), frame 61 in (d). Similarly, we show the same frames for the videos shown in (ii) ROI: $90-160 \times 120-270$ from Fig. 2(b) in (e)-(h), (iii) ROI: $100-260 \times 80-315$ from Fig. 2(c) in (i)-(l), and (iv) ROI: $40-160 \times 100-300$ from Fig. 2(d) in (m)-(p).