Ultrasound in Med. & Biol., Vol. 24, No. 3, pp. 427–441, 1998

0301-5629/98 $19.00 + .00

ELSEVIER

PII S0301-5629(97)00281-0

● *Original Contribution*

MULTILEVEL AND MOTION MODEL-BASED ULTRASONIC SPECKLE TRACKING ALGORITHMS

FAI YEUNG, STEPHEN F. LEVINSON and KEVIN J. PARKER
Departments of Electrical Engineering and Physical Medicine and Rehabilitation, University of Rochester, Rochester, NY

(*Received* 10 *October* 1996; *in final form* 11 *July* 1997)

Abstract—A multilevel motion model-based approach to ultrasonic speckle tracking has been developed that addresses the inherent trade-offs associated with traditional single-level block matching (SLBM) methods. The multilevel block matching (MLBM) algorithm uses variable matching block and search window sizes in a coarse-to-fine scheme, preserving the relative immunity to noise associated with the use of a large matching block while preserving the motion field detail associated with the use of a small matching block. To decrease further the sensitivity of the multilevel approach to noise, speckle decorrelation and false matches, a smooth motion model-based block matching (SMBM) algorithm has been implemented that takes into account the spatial inertia of soft tissue elements. The new algorithms were compared to SLBM through a series of experiments involving manual translation of soft tissue phantoms, motion field computer simulations of rotation, compression and shear deformation, and an experiment involving contraction of human forearm muscles. Measures of tracking accuracy included mean squared tracking error, peak signal-to-noise ratio (PSNR) and blinded observations of optical flow. Measures of tracking efficiency included the number of sum squared difference calculations and the computation time. In the phantom translation experiments, the SMBM algorithm successfully matched the accuracy of SLBM using both large and small matching blocks while significantly reducing the number of computations and computation time when a large matching block was used. For the computer simulations, SMBM yielded better tracking accuracies and spatial resolution when compared with SLBM using a large matching block. For the muscle experiment, SMBM outperformed SLBM both in terms of PSNR and observations of optical flow. We believe that the smooth motion model-based MLBM approach represents a meaningful development in ultrasonic soft tissue motion measurement. © 1998 World Federation for Ultrasound in Medicine & Biology.

Key Words: **Speckle tracking, Motion estimation, Acoustics, Ultrasonics, Elasticity, Biomechanics, Doppler ultrasound, Image processing, Computer simulation, Muscles.**

INTRODUCTION

The application of tissue motion estimation to ultrasonic tissue characterization and blood flow imaging is one of the most promising areas of research in medical imaging today (Bohs et al. 1993; Jensen 1996; Krouskop et al. 1987; Trahey et al. 1988; Tristam et al. 1986). This is particularly true of sonoelastography, which derives the elastic properties of soft tissues from measurements of displacements resulting from an applied perturbation (Gao et al. 1996). The potential applications of sonoelastography include tumor detection (Gao et al. 1995; Huang et al. 1992), studies of muscle elasticity (Levinson et al. 1995), the investigation of vascular tissue properties (Ryan et al. 1992), as well as other areas of clinical investigation (Céspedes et al. 1993; O'Donnell et al. 1994; Ophir et al. 1996; Yamakoshi et al. 1990).

In contrast to frequency domain (Doppler) methods, time domain tissue motion estimation techniques use quantitative methods to track the movement of speckle patterns in B-mode ultrasound scans (Hein and O'Brien 1993). Time domain techniques have been used in a variety of biomedical applications, including the use of an optical flow technique to assess local myocardial deformation (Mailloux et al. 1989), the use of a block matching algorithm for blood flow assessment (Trahey et al. 1988) and to derive tissue elasticity information (Levinson et al. 1994).

Two-dimensional block matching time domain approaches to speckle tracking have found widespread ap-

Address correspondence to: Dr. Stephen F. Levinson, Department of Physical Medicine and Rehabilitation, University of Rochester, Rochester, NY 14642, USA.

plication because of their inherent simplicity and relative immunity to noise. In these methods, speckle patterns are tracked by matching a target block within a search window in successive image frames. Traditionally, a fixed size matching block is employed using a single match at each image point. This is referred to as single-level block matching (SLBM) to distinguish it from the multilevel approach proposed here.

One of the most significant drawbacks inherent in all methods of speckle tracking is the loss of motion information due to speckle decorrelation. Ultrasonic speckle patterns arise from the constructive and destructive interference of coherent echo signals (Wagner et al. 1983). Speckle pattern decorrelation can be induced by out-of-plane motion, nonuniform motion of subresolution scatterers, nonuniformity of the ultrasound field and nonrigid tissue deformation (Trahey et al. 1988; Wagner et al. 1988; Wear et al. 1987). Since speckle tracking algorithms depend on the stability of the speckle pattern, they often yield inaccurate motion estimates in regions where speckle decorrelation is high.

In a previous experiment attempting to quantify the decorrelation of speckle patterns, the rate of speckle decorrelation was determined from the cross-covariance curves between ultrasound B-scans of a soft tissue-mimicking phantom with translational movement of the transducer (Ramamurthy and Trahey 1991). It was reported that the correlation coefficient fell to 0.82 for lateral translations of 2.0 mm and 0.94 for axial translations of 2.0 mm. The correlation curve fell even more sharply with increasing translations. These results, however, were for rigid body motion. Sonoelastography, on the other hand, relies on the nonrigid deformation of soft tissues (Ophir et al. 1996). It is anticipated that speckle decorrelation would be even greater for the case of nonrigid tissue motion.

To obtain good estimates of the motion field, a large matching block is desirable to increase the uniqueness of each data block and to align the local data of weak or homogeneous speckle patterns to those of strong or structural speckle patterns. The gain in accuracy, however, is usually at the expense of spatial resolution. For this reason, block matching algorithms using a large matching block often fail to resolve nonrigid or highly varying motion fields. Indeed, one of the underlying assumptions of block matching is that all tissue elements within the matching block have identical motion vectors, an assumption that is often violated in using a large matching block when nonrigid or rotational motion is involved. By using a smaller size matching block, block matching algorithms perform better in resolving highly varying motion fields and avoid the problem of nonuniform motion vectors, often at the expense of increased susceptibility to noise and speckle decorrelation-related tracking errors.

One of the goals of this article is to address the inherent trade-off between speckle tracking accuracy and spatial resolution based on the selection of block size by adapting a multilevel approach that incorporates the advantages of both large and small matching blocks. In contrast to single-level block matching, the multilevel algorithms use a matching block and search window of variable size. A large block size initially is used to provide a coarse-resolution estimate of the overall motion field. Each subsequent level uses a smaller block size and search window so as to increase the spatial resolution without sacrificing the noise immunity of the prior levels. The multilevel method is similar to hierarchical schemes used in block matching algorithms for video coding, in that both provide a global-to-local framework for motion estimation (Tekalp 1996). They differ, however, in the use of different schemes of subsampling.

In the multilevel scheme presented here, ultrasound images are not decimated to construct a multiresolution pyramid of images because ultrasound images are characterized by discrete speckle patterns rather than by smooth intensity functions. Excessive filtering and decimation of ultrasound images could transform the speckle statistics from a Rayleigh distribution to one that approaches a Gaussian distribution (Ashton 1996). Although the effect of these changes is worthy of further investigation, that is beyond the scope of this article. In the multilevel algorithms, only the motion field is estimated based on a multiresolution grid of motion vectors. This coarse-to-fine approach also improves computational efficiency through the use of significantly smaller matching blocks and search windows than would be the case if large matching blocks and search windows were used at the finest motion field resolution level. The improvement in computational efficiency (Chen 1995; Wang and Shung 1996) is important to applications in which a dense motion field is required.

Another goal of this work was to devise speckle tracking algorithms that are robust in the presence of noise and relatively insensitive to speckle decorrelation. As speckle decorrelation is signal and motion dependent, its effects cannot be compensated by simple averaging or spatial or temporal filtering. The problem of speckle decorrelation is addressed by using *a priori* knowledge based on the physical properties of tissue motion. The spatial motion model-based block matching algorithm combines the multilevel scheme with a smoothness constraint based on the assumption that the tissue motion field is continuous. This assumption is generally valid when the motion vectors are contained within a single moving or deforming object. Even at the boundaries of

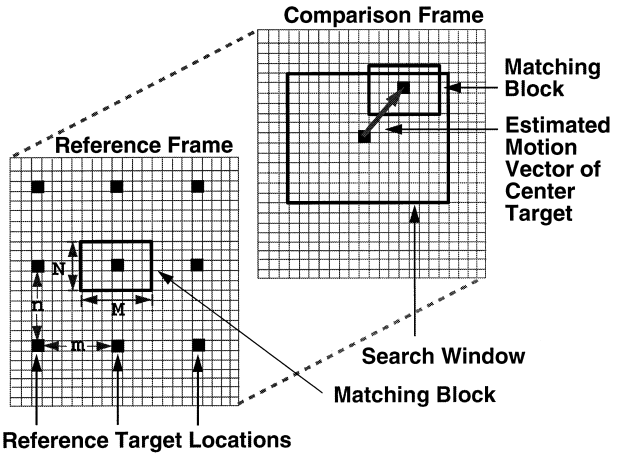

Fig. 1. Schematic representation of the general block matching algorithm.

tissues with differing properties, motion continuity is still valid provided that stress also is continuous (nonslip condition). This smoothness constraint can be imposed on local neighborhoods of motion estimates to exploit the connectivity among neighboring motion vectors. The speckle tracking problem is formulated as a minimization of a weighted sum of block matching estimates and the motion smoothness constraint. Conceptually, this motion model-based method is similar to regularization approaches (Poggio et al. 1985) based on variational principles (Horn and Schunck 1981) or statistical formulations (Konrad and Dubois 1992), in that admissible solutions of motion estimates are restricted to smooth functions and that data with weak estimates depend on data from reliable neighboring sites.

BLOCK MATCHING ALGORITHMS

Single-level block matching

In general block matching algorithms traditionally used for SLBM, as illustrated in Fig. 1, each target location in a reference frame is matched to a corresponding location in a comparison frame. The displacement between the two locations is an estimate of the motion vector for a given point over a time interval represented by the two frames. The target locations in the reference frame are distributed on a grid with m pixels separating the locations horizontally and n pixels vertically. The estimated motion vector at each location is defined as the relative position of the ''best-matched'' $M \times N$ block within a predefined search window centered at the same grid point of the comparison frame relative to the reference frame. The algorithm searches for all possible locations within the search window, selecting the best-matched one based on specific criteria. The grid spacing $M \times N$ is ideally chosen to equal the spatial resolution of the estimated motion field, which is strongly related to, but not necessarily the same as, $M \times N$. We define a *moxel* as being a single motion resolution element, equivalent to $m \times n$ pixels in the original frame. Since the correlation cell size of fully developed speckle under the condition of diffuse scattering is strongly related to the point spread function of the ultrasound system, the ratio of M and N should ideally be roughly equal to that of the system point spread function to achieve a symmetrical tracking precision in both axial and lateral directions (Trahey et al. 1988).

Matching criteria

The probability that a motion estimate represents the tissue motion in a two-dimensional image plane is maximized when the cross-correlation between the matching blocks in the reference and comparison frames also is maximized. Since cross-correlation is relatively computationally intensive, we have chosen to use an essentially equivalent approach, the sum of squared difference (SSD) criteria, which is based on minimizing the sum of the squared difference between pixels in the two images (Tekalp 1996). The sum of absolute difference (SAD) criteria used by others is simpler still but, unlike SSD that penalizes large pixel differences more heavily, SAD weighs both small and large discrepancies equally (Bohs et al. 1993; Chen 1995).

The SSD between a block in the reference frame, k, and a block displaced by a motion vector $d(x)$ in the comparison frame, assumed to be the next frame in a sequence, $k + 1$, is formulated as:

$$\mathrm{SSD} = \sum_{x \epsilon B} [f_k(x) - f_{k+1}(x + d(x))]^2, \qquad (1)$$

where f denotes the pixel intensity and x the position vector within the matching block B. The estimated motion vector $\hat{d}(x)$ is then given as $d(x)|_{\mathrm{SSD}_{\min}}$ for all possible locations within the search window.

Multilevel block matching

In basic multilevel block matching (MLBM), a coarse-to-fine approach is used in the selection of the sizes of matching blocks, search windows and moxels. For a multilevel scheme consisting of $L_{\max}$ levels, if L denotes the number of the level, $L = L_{\max} - 1$ represents the top level and $L = 0$, the bottom level. At each level, the algorithm searches target locations spaced $2^L m_0 \times 2^L n_0$ pixels apart using a $2^L M_0 \times 2^L N_0$ pixel matching block, where m_0, n_0, M_0, N_0, represent the values of m, n, M, N, at $L = 0$, respectively. Starting from the top level, in each subsequent level the target location spacing,

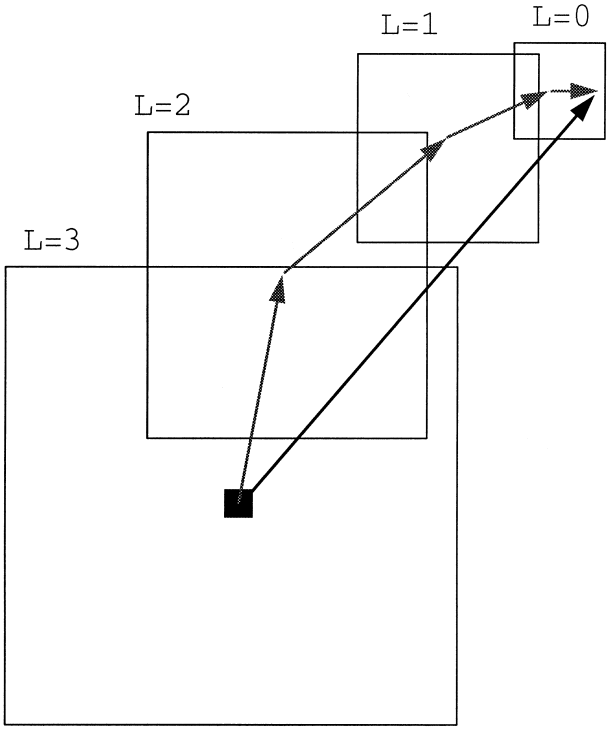

Fig. 2. Final motion vector estimation in the multilevel algorithm.

block size and search window are reduced by half. The grid size of estimated motion vectors in the final stage is represented by m_0 and n_0. Note that the images are not subsampled and the image size used in every level is the same as the original image size.

The center of the search window at each subsequent level is displaced by the motion vector estimated in the previous level. Because the moxel grid at a finer level has a finer resolution than that of the coarser level, the coarser-resolution motion estimates passed from the previous level are upsampled to the same resolution as the current moxel grid by interpolation. This procedure is repeated until the bottom level is reached. Since the motion estimates from coarser levels are passed to the finer levels as offsets of the center of the search windows, the final motion estimate is then the sum of the vectors found in all stages of the hierarchy, as shown in Fig. 2.

Spatial motion model-based block matching

Cost function. In spatial motion model-based block matching (SMBM), a cost function is generated based on physical criteria. Displacement vectors that violate physically realistic motion are associated with a correspondingly high cost. The displacement field then is optimized by minimizing the corresponding overall cost function, which consists of a weighted sum of SSD and smoothness constraints as defined in the following equation:

$$H(x, \hat{d}(x)) = \sum_{x \in \Omega} \mathrm{SSD}(x, \hat{d}(x)) + \beta_L \times \sum_{c_L \in C_L} \|\hat{d}(x_i) - \hat{d}(x_j)\|^2, x_i, x_j \in c_d, x_i \neq x_j, \quad (2)$$

where Ω is the set of all target locations on the moxel grid, c_L denotes a local neighborhood of moxels and C_L is the full set of all neighborhoods of moxels at resolution level L. In this equation, c_d represents one possible realization of c_L. The first term is the SSD error estimate and the second term is a smoothness constraint that accounts for the connectivity between neighboring moxels.

Figure 3 illustrates the effective local neighborhood at two different levels. The cost function represents an *a*

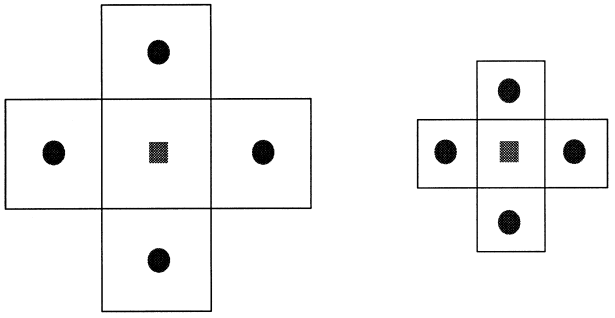

Fig. 3. Demonstration of 4-moxel neighborhood at different motion resolution levels.

priori motion model because a smooth motion field is assumed within the tissue region. The parameter β_L is a non-negative function that controls the relative weight of motion smoothness in the cost function. The weighting function β_L is level dependent because the connectivity between neighboring moxels varies with the physical size of the local neighborhood.

Binning. The cost function in eqn (2) is iterative in nature because each local motion estimate affects the cost function of its neighbors. We have devised a process of ''binning'' to accelerate the convergence of solutions by selecting a subset, or ''bin,'' of highly matched locations within the search window. For example, for an 11×11 search window, there are 121 possible locations for the estimated motion vector. If block matching is used to select the top 10 (8%) best-matched locations for the bin and the bin population $\tilde{d}(x)$ is used in the optimization of the cost function $H(x,\hat{d}(x)|\hat{d}(x) \in \tilde{d}(x))$, the number of calculations is reduced significantly. This process has three major advantages. By limiting the number of possible motion vector candidates, optimization converges more rapidly. Since only a percentage of the best-matched locations are used in the cost function, the effect of the smoothness constraint will be limited, reducing the blurring effect seen at motion boundaries. Finally, by using only the moxels with the highest degree of correlation, the effect of poorly correlated moxels on the overall cost function will be minimized.

Optimization. Minimization of the cost function can be accomplished using a deterministic optimization method similar to the iterated conditional models (ICM) optimization technique (Geman and Geman 1984). The cost function is optimized using the following procedure:

(1) Initialize the iteration counter g.
(2) Initialize the bin population $\tilde{d}(x)$.
(3) Initialize the update counter $u(x)$ to one for all target locations where motion vectors are to be estimated, $x \in \Omega$.
(4) For each target location, if $u(x) \neq 0$ or $u(x_j) \neq 0$, where $x, x_j \in c_d$ and $x \neq x_j$, compute the cost function $H(x, d'(x)|d'(x) \in \tilde{d}(x))$ for each possible candidates in $\tilde{d}(x)$.
(5) If $\hat{d}(x) \neq d'(x)|_{H_{min}}$, set the motion estimate $\hat{d}(x)$ to $d'(x)|_{H_{min}}$ and increment the update counter $u(x)$, otherwise reset $u(x)$ to zero.
(6) Increment g and terminate if $g \geq g_{\max}$.
(7) Terminate if $\Sigma_{x\in\Omega}\, u(x) = 0$.
(8) Go to the next iteration starting at step 4.

The same procedure is applied for all the moxels at each level. After the first pass of cost function minimization, the cost function of a newly updated motion estimate ($u(x) = 1$) or motion estimate with a newly updated neighbor ($u(x) = 1$) must be reminimized. Minimization is repeated until no new motion estimates ($\Sigma_{x\in\Omega}\, u(x) = 0$) are obtained or until the number of iterations has exceeded an arbitrary limit ($g_{\max} = 5$ in our work). Because of the deterministic nature of the optimization process, iteration converges rapidly. Furthermore, because the binning process limits the total number of motion vectors estimated in the optimization process, the additional computational overhead associated with the SMBM algorithm is insignificant. A flow diagram illustrating the operation of SMBM implemented using the multilevel scheme is shown in Fig. 4.

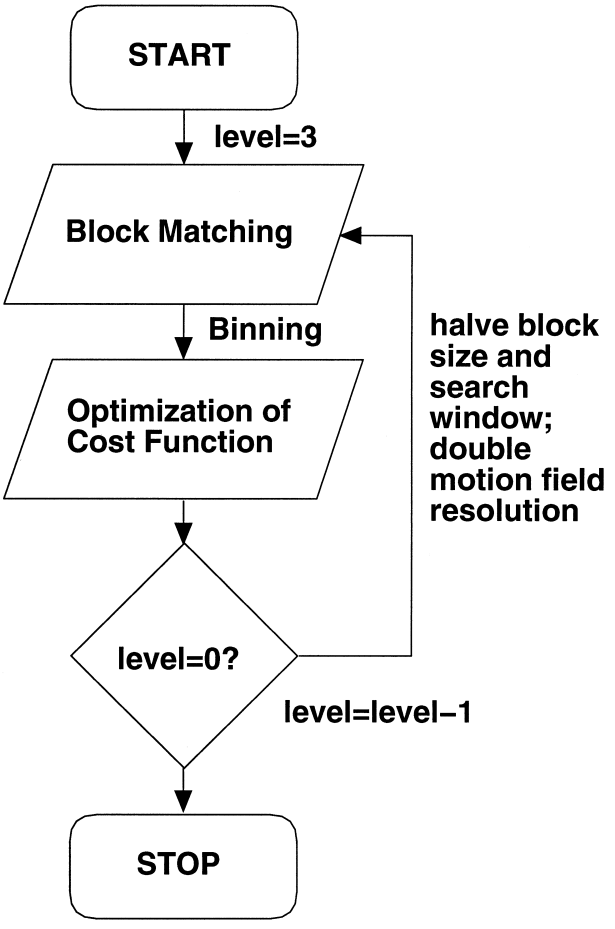

Fig. 4. Flow diagram of spatial motion block matching algorithm.

EXPERIMENTS AND RESULTS

Evaluation of tracking algorithms

All speckle tracking algorithms were tested using SUN SPARC 20 (Sun Microsystems, Mountain View, CA, USA) workstations and the computation time, in minutes, was recorded as a secondary indicator of the efficiency of the software implementation of the algorithms. Since this measure depends on the code architec-

ture in software implementation, we used the total number of SSD computations *per image* as the primary indicator of algorithm efficiency. Calculation of total SSD computations, however, excludes the filtering and interpolation processes. Still, the ratio of total SSD computations for the different algorithms represents an upper limit of the gain or loss in efficiency among the algorithms.

Three different measures were used to evaluate the performance of each of the speckle tracking algorithms. Mean squared tracking error was used when the actual motion was known, as was the case for the simulated motion experiments and the experiments involving translation in phantoms. In the case of muscle motion in humans, however, the actual motion field could not be determined and, hence, peak signal-to-noise ratio (PSNR) and visual perception of optical flow were used as indirect measures.

Mean squared tracking error. Mean squared tracking error is defined as the mean squared magnitude of the difference between the estimated motion vectors $\hat{d}(x)$ and actual motion vectors $d(x)$:

$$\varepsilon_t = \frac{1}{Q} \sum_{x \in \Lambda} \| d(x) - \hat{d}(x) \|^2, \qquad (3)$$

where Q is the total number of pixels within the image and Λ denotes a set of target locations in the image plane.

Peak signal-to-noise ratio. PSNR is a measure of how closely the corresponding points described by the motion vectors between two frames are matched, based on the mean squared displaced pixel difference (dpd). The PSNR, in dB, is formulated as:

$$\mathrm{PSNR} = 10 \log_{10} \frac{255 \times 255}{\frac{1}{Q} \sum_{x \in \Lambda} [\mathrm{dpd}(x, \hat{d}(x))]^2}. \qquad (4)$$

The displaced pixel difference is expressed as:

$$\mathrm{dpd}(x, \hat{d}(x)) = f_{k+1}(x + d(x)) - f_k(x), \qquad (5)$$

where $\hat{d}(x)$ is the estimated motion vector. The dpd is essentially the difference between pixel intensities at a target and its ''matched'' location in the next frame.

Visual perception of optical flow

Visual perception of optical flow, though subjective, is often more reliable than mathematical models in selecting the best algorithm for a given set of images. To minimize bias, we asked two blinded observers to choose the vector field image that best matched the observed motion. The observers each viewed continuous cineloops of the time-varying speckle patterns in B-mode images and compared these to unlabeled motion vector field images representing the SLBM, MLBM and SMBM algorithms.

Simple phantom translation experiment

To demonstrate the performance of the algorithms on actual ultrasonic B-mode scans in a controlled manner, we devised a simple translation experiment using a gelatin-agar tissue-mimicking phantom. A 7.5-MHz linear array ultrasound transducer (Quantum, Issaquash, WA, USA) was clamped rigidly to a 3D precision manipulator (Velmex, Holcomb, NY, USA) and used to image the phantom, which was submerged in a water tank. Movement of the Velmex slider was controlled by a PC so that the transducer could be translated continuously and precisely in 2 mm ($\pm$0.1 mm) lateral increments. The video output of the ultrasound scanner was connected to another PC with a frame grabber that was used to capture and store the B-scan sequences. Sequences of eight B-scan frames then were transferred to a workstation for speckle tracking using each of the three algorithms.

A 100 $\times$ 100 pixel rectangular region of interest (ROI) of the B-scan frame was used for speckle tracking. A low-pass filter with a 5 $\times$ 5 Gaussian-shaped (variance = 1.0) kernel was used to remove system noise. This filter was applied to the image data used in testing

Table 1. Parameters used in motion estimation algorithms.

		MLBM and SMBM			
Level	SLBM	3	2	1	0
Block size	41 × 25	41 × 25	21 × 13	11 × 7	5 × 3
Search window	31 × 31	31 × 31	15 × 15	7 × 7	3 × 3
Motion field resolution	2:1	16:1	8:1	4:1	2:1
Bin size (SMBM)		13	11	10	5
β_L (SMBM)		16	64	256	1024

MLBM = multilevel block matching; SLBM = single-level block matching; SMBM = spatial motion block matching.

Table 2. Performance of speckle tracking algorithms for linear translation.

	SLBM	MLBM	SMBM
Computation time (min)	144	5	5
SSD computations ($\times 10^6$)	2500	5.7	5.7
Tracking error (pixels)	0.23	2.28	0.97
PSNR (dB)	66.72	68.47	66.87

MLBM = multilevel block matching; PSNR = peak signal-to-noise ratio; SLBM = single level block matching; SMBM = spatial motion block matching.

all of the algorithms to avoid selection bias. Table 1 lists the parameters used in the speckle tracking algorithms. Each moxel at the finest level corresponds to two pixels in each direction, for a ratio of four pixels per moxel. The smoothness coefficient was increased with decreasing hierarchical levels to account for increased connectivity of the motion vectors with decreasing moxel size. Note that, even though the transducer displacement is in the lateral direction only, the error of motion estimation was computed in both the vertical and lateral directions.

Results. Table 2 lists the speckle tracking results for the simple translation experiment. The SLBM algorithm using a large matching block performed better than the MLBM and SMBM algorithms, with a mean squared tracking error of 0.23 (pixels) compared to 2.28 and 0.97 for MLBM and SMBM, respectively. Since only rigid motion was involved in this experiment, the assumption of a uniform local motion field within the matching block was valid. By using more data in computing each motion vector, SLBM had a more robust performance than MLBM, which suffered from inaccurate initial motion estimates propagated from previous levels and from sensitivity to local variations. Figure 5 shows the performance of SLBM with different sizes of matching blocks. Mean squared tracking errors increased monotonously with reduced matching block size. The performance of the motion-model algorithm was very close to that of the single-level algorithm with a large matching block, because the motion constraint helped to align motion vectors in regions with poor block correlation from frame to frame to neighboring sites with more reliable estimates.

The computation time was dramatically different among the single-level and multilevel algorithms. The single-level algorithm using the largest matching block required 144 min while the multilevel and motion model-based algorithms needed just 5 min each. The ratio between the computation times is 29. From the total SSD computations per image, SLBM using a large matching block required 2.5×10^9 computations per image whereas MLBM required 5.7×10^6 computations per image. The ratio between these values is 439. Although the ratio decreased significantly as the SLBM matching block size was reduced, the tracking error increased similarly (Fig. 5), demonstrating the inherent trade-off between the efficiency of computation and the resulting accuracy with SLBM.

Rotational, compressional and shearing motion fields

Because of the difficulty of accurately controlling displacements involving deformation of the phantoms, we instead used computer simulations of rotational, compressional and shearing motion fields. Speckle image pairs were obtained from envelope-detected data from radio frequency signals generated by convolving random scatterers with an ultrasound transducer point spread function. The speckle pattern satisfied a Rayleigh distribution (Wagner et al. 1983).

Rotational motion was generated by rotating the speckle images by 5° in a counterclockwise direction. By rotating the envelope-detected image instead of the scatterers, we avoided the speckle motion artifact under tissue rotation as described by Kallel et al. (1994). For the compressional motion field, the second image scatterers were compressed vertically by 10% and expanded horizontally by 10%. The shearing motion field was in the form of pure shear in the horizontal direction, with a shearing angle of 5°. The ROI of the rotational motion field was 195×195 pixels and that of the compressional and shearing motion fields was 120×120 pixels. The speckle tracking algorithms used the same parameters as shown in the Table 1, with the exception that the block size in the final level of the multilevel scheme was set to 1×1 pixels rather than 5×3 pixels to allow for the greater spatial variation expected as compared to simple

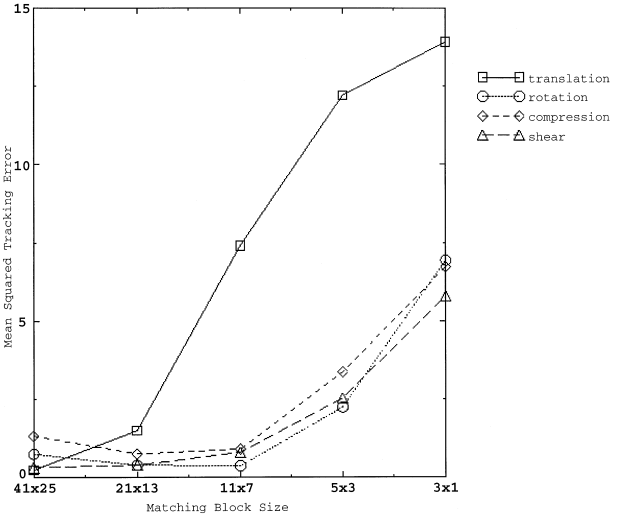

Fig. 5. Performance of one-level block matching algorithm (SLBM) with different matching block size.

Table 3. Performance of speckle tracking algorithms for simulated rotation, compression and shear.

	Rotation			Compression			Shear		
	SLBM	MLBM	SMBM	SLBM	MLBM	SMBM	SLBM	MLBM	SMBM
Computation time (min)	447	16	17	281	10	11	281	10	11
SSD computations ($\times 10^6$)	9300	21.6	21.6	3500	7.9	7.9	3500	7.9	7.9
Tracking error (pixels)	0.74	0.88	0.72	1.32	1.86	0.70	0.32	0.33	0.17
PSNR (dB)	69.53	77.21	72.56	63.14	75.50	70.78	71.12	79.68	75.33

MLBM = multilevel block matching; PSNR = peak signal-to-noise ratio; SLBM = single-level block matching; SMBM = spatial motion block matching; SSD = sum of squared difference.

translation. Other parameters were kept the same as the parameters used in the experiment using the phantom.

Results. The results are listed numerically in Table 3 and shown graphically by the vector diagrams in Figs. 6 and 7, and in the plot in Fig. 8. For the shear motion field, a one-dimensional plot was used because the alignment of motion vectors in a single direction made it difficult to observe the results. For the rotational motion field, all three algorithms essentially performed equally well by objective criteria (the ''Rotation'' column of Table 3). Our blinded observers, however, both noted that SLBM failed to track the motion correctly in the center region, where the speckle pattern rotated more rapidly (Fig. 6). With its limited spatial resolution, SLBM fails to resolve rapid changes in the motion field at the center. Both observers agreed that SMBM most accurately represented the motion observed. A similar observation was made for the simulated compressional motion field (Fig. 7). In this case, however, the tracking error associated with SMBM was less than half that with either SLBM or MLBM (the ''Compression'' column of Table 3), as was the case for simulated shear (the ''Shear'' column of Table 3). The nature of the tracking error is particularly evident for simulated shear as shown in Fig. 8, in which the motion along a single line has been plotted. For SLBM, a stairstep pattern is evident, indicating sharp discontinuities related to the matching block size. MLBM failed to improve on tracking error, whereas SMBM provided the smallest degree of mean squared tracking error.

The effect of decreasing matching block size on the performance of SLBM is illustrated in Fig. 5. The curves of mean squared tracking error drop slightly with decreasing matching block size in all cases of rotational, compressional and shearing motion due to the increased resolving power of smaller matching blocks, but rise rapidly as the matching block size is further reduced due to the progressive effect of false matches.

Forearm flexor muscle contraction experiment

To evaluate the three algorithms in living tissue, a complex muscle motion experiment was performed in a single subject. Although it was not possible to measure the true motion of the muscles independently of speckle tracking, we felt that the data obtained could provide important insights into the ability of the algorithms to handle motion fields with a high degree of spatial variation.

Informed consent was obtained in accordance with a protocol approved by the research subjects review board at our institution. A 7-MHz linear array ultrasound scanner (Acuson, Mountain View, CA, USA) was used to image the forearm from a single subject in cross-section at a frame rate of 16 Hz. Speckle tracking was performed between the first and fourth frames. A 200 $\times$ 300 pixel ROI was used, as shown in Fig. 9. The brightest region at the bottom of the ROI is the ulna. The region above the bone contains the forearm flexor muscles, blood vessels and nerves. Two groups of muscles, the flexor *digitorum superficialis* (FDS) at the top right and the *flexor digitorum profundus* (FDP) in the middle of ROI are of interest in this experiment. With the wrist supported in partial flexion so as to relax the wrist flexors, the subject was asked to bend his middle finger against resistance so as to contract these two muscles to a greater degree than the surrounding musculature.

Observation of optical flow. From the optical flow of speckle patterns in the image sequence, it was observed that the FDP expands outward toward the lower left and the FDS raised up toward the upper left. The muscle in the middle left region just above the bone also seemed to be pushed slightly to the right. There is no significant motion observed in the muscles in the top left region. The ''false'' motion below the bone was due to the reverberation of echo signals.

Results. The motion fields were estimated according to the parameters in Table 1. The same low-pass filter used in the phantom experiments was applied to the image data before the motion estimates were obtained. As seen in Fig. 10 and noted by both blinded observers, SLBM failed to detect accurately the sophisticated motion of the muscle tissues. It seems likely that the large

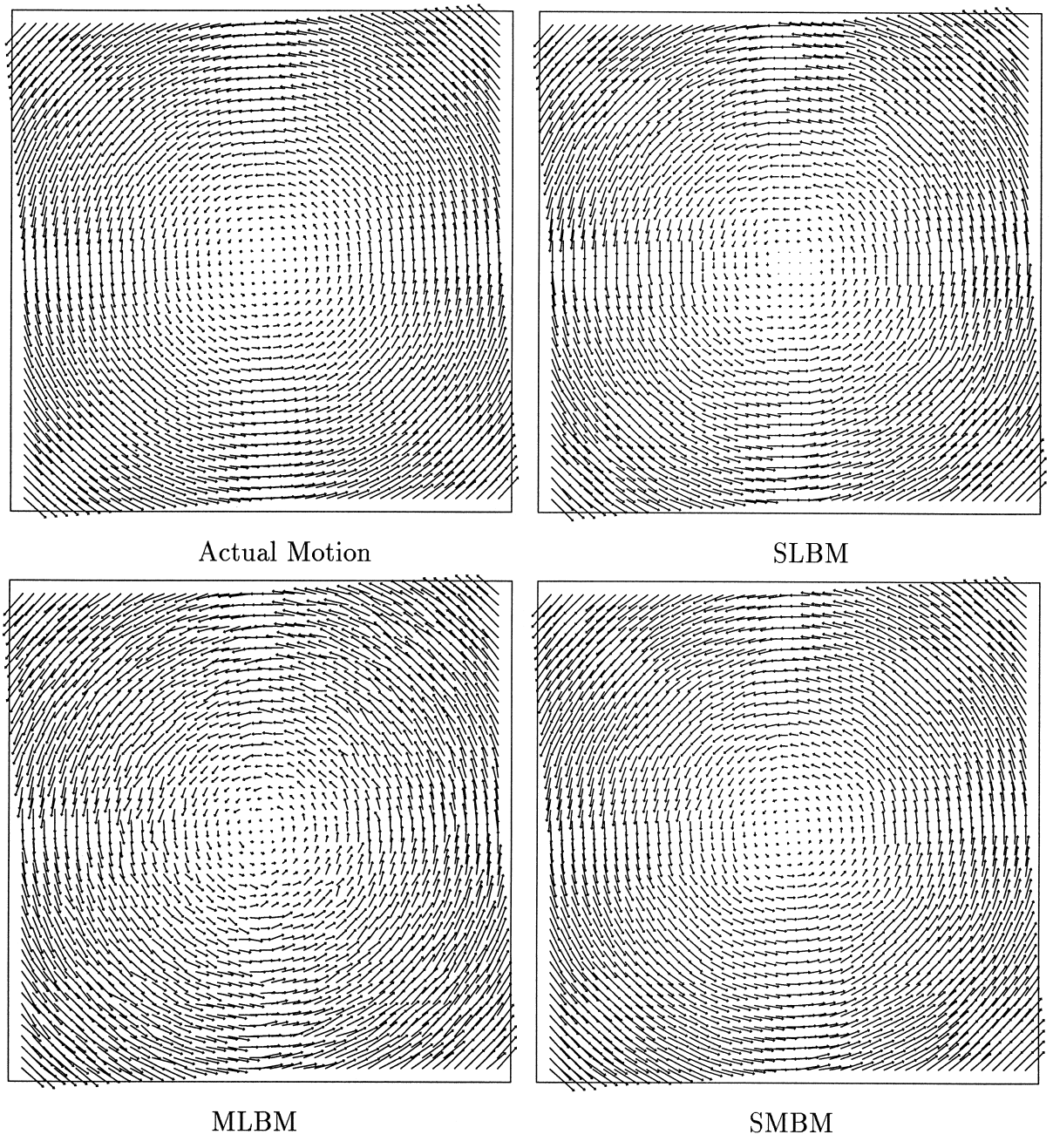

Fig. 6. Performance of single-level (SLBM), multilevel (MLBM) and spatial motion (SMBM) algorithms in tracking simulated rotation.

matching block was not able to differentiate multiple motions within the same block. To illustrate this point, we halved the matching block size to 21×13. The column under the heading SMBM* in Table 4 shows the results. Less computation time (324 min) and SSD computations (3.9×10^9) were required. Sophisticated muscle movement could be differentiated more easily by using the smaller matching block, but at the expense of numerous erroneous motion estimates as noted by both blinded observers.

Both observers felt that MLBM detected the muscle tissue motion reasonably well, as seen in Fig. 10, although the motion estimates appeared to be less orderly. From the motion vector field, it was relatively easy to differentiate among the muscle groups involved in finger contraction. The observers chose the SMBM vector plots

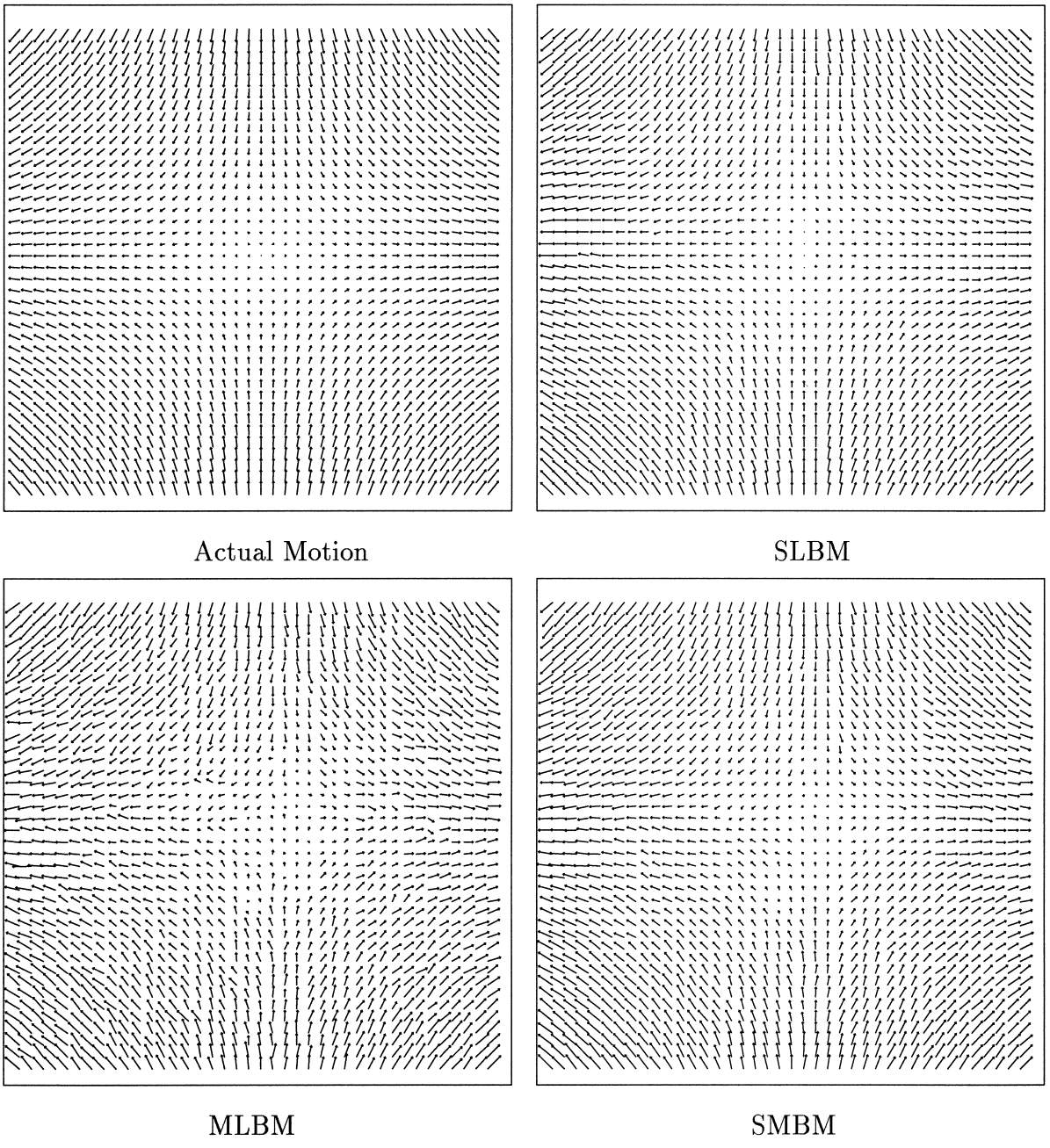

Fig. 7. Performance of single-level (SLBM), multilevel (MLBM) and spatial motion (SMBM) algorithms in tracking simulated compression.

as the most representative of the actual motion observed. The motion field was more physically realistic, with smooth motion within each muscle group and motion boundaries between different muscle groups being well defined.

DISCUSSION

Recovery of motion fields from ultrasonic speckle images is an ill-posed problem because of speckle decorrelation and noise. Although an incremental tracking strategy (Chen 1995) has been verified experimentally as a way to reduce speckle decorrelation, this technique is limited ultimately by the frame rate of the ultrasound scanner. The strategy of using a large matching block with SLBM (Ramamurthy and Trahey 1991) may help to reduce the number of false matches due to speckle pattern decorrelation but, unfortunately, at the expense of spatial resolution. We have proposed a different ap-

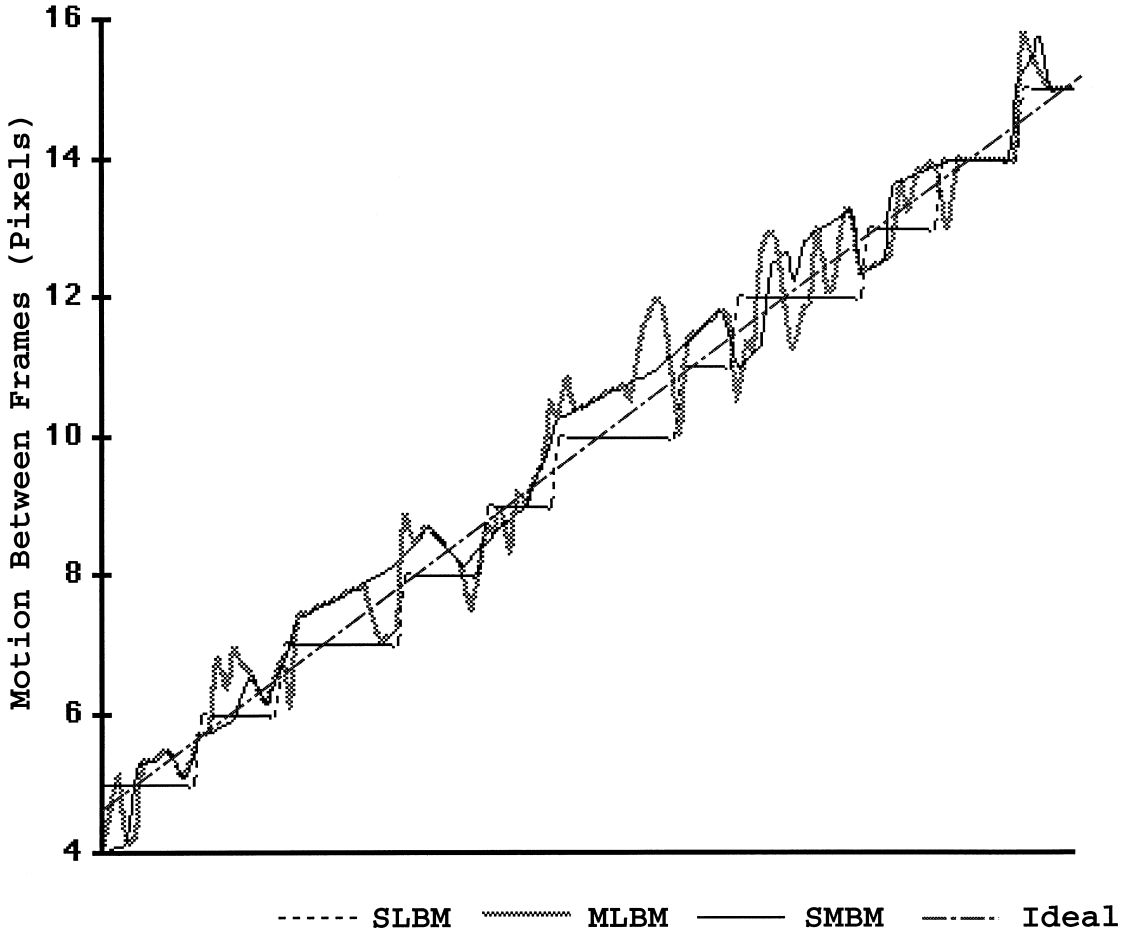

Fig. 8. Performance of single-level (SLBM), multilevel (MLBM) and spatial motion (SMBM) algorithms in tracking simulated tissue shear. The *x*-axis quantity is in pixels.

proach to the problem of speckle decorrelation. The multilevel scheme combines the relative immunity to noise of SLBM using a large matching block with the high spatial resolution associated with the use of a smaller matching block. The addition of a spatial smoothness constraint in the SMBM model correlates the motion estimates to their neighboring motion vectors. Because of this additional constraint, SMBM achieves an accuracy comparable to that of SLBM with a large matching block, but without sacrificing spatial resolution.

The selection of matching block and search window size is central in block-based speckle tracking algorithms, as these parameters determine the accuracy, resolution and computational load of motion estimation. The trade-off associated with the use of large *versus* small blocks can be illustrated by the plotting of SSD values at a given location as a function of all possible motion vectors within the search window. Figure 11 shows one such plot of SSD space based on experimental data from simple phantom translation. Figures 11a and 11b are surface plots of SSD space based on a search window of 41 × 41 using matching block sizes of 31 × 31 and 11 × 11, respectively. The troughs in the SSD space plots represent local minima and the location of the deepest trough corresponds to the optimum solution to $d(x)|_{\mathrm{SSD}_{\min}}$.

In Fig. 11a, it can be noted that there are relatively few local minima, although the location of the minimum in the deepest trough is not well defined. This corresponds to the limited spatial resolution associated with the use of a large matching block. On the other hand, Fig. 11b shows a larger number of better-defined local minima. To limit the number of local minima within the search area, the search window size can be decreased, but with a loss in the dynamic range of the motion estimates. MLBM combines the large dynamic range associated with the use of a large matching block with the higher spatial resolution associated with the use of a smaller matching block. Without the addition of the smoothness

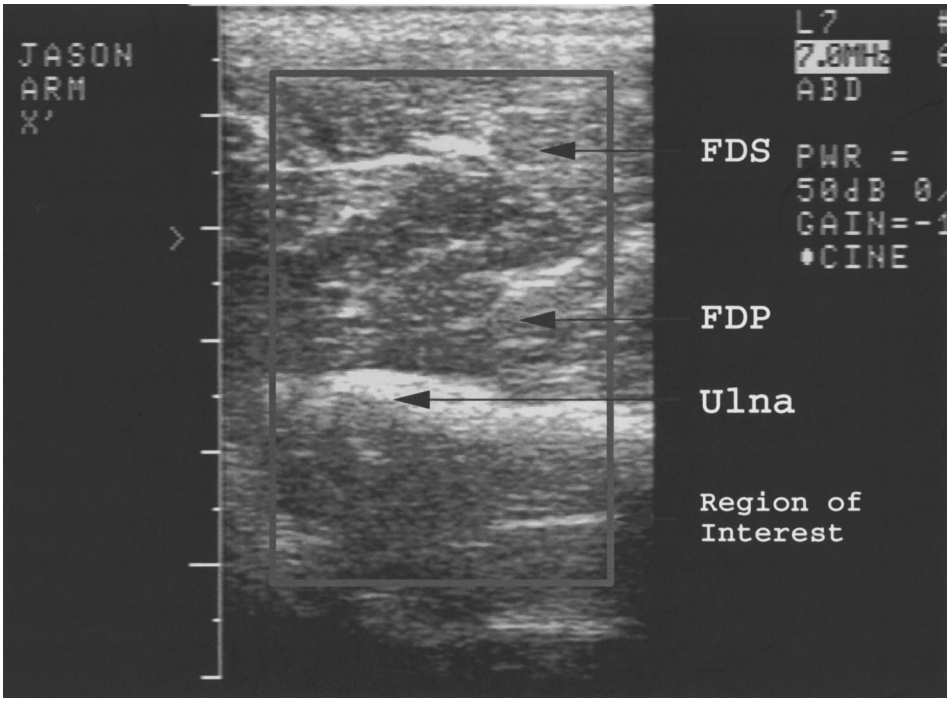

Fig. 9. Region-of-interest (ROI) in a B-scan of forearm flexors. The rectangular box shows a 200 × 300 ROI. The flexor digitorum superficialis (FDS), flexor digitorum profundus (FDP) and ulna are marked.

constraint, however, the estimate using a small matching block is often trapped in local minima of the SSD curve (false matches).

The effects of noise and nonuniform motion are well illustrated in Fig. 5. The curve from the phantom translation experiment represents the effect of noise on tracking error as a function of block size. As the underlying motion within the matching block is uniform regardless of the block size, the increase in tracking error with decreasing block size is representative of increased sensitivity to noise. With the computer-simulated motions, on the other hand, tracking error results in spite of an ideal noise-free environment. With large matching blocks, false tracking results from the inability to resolve varying motions within the matching block. With smaller matching blocks, on the other hand, false matches result from trapping within local minima in the SSD curve, as illustrated in Fig. 11b. It could be argued that an ideal algorithm would search for the optimum block size based on tracking error curves. As in the muscle experiment, however, the underlying motion is generally not known so that it is not possible to generate true tracking error curves with clinical images. The multilevel scheme addresses this problem by being adaptive, using a variable matching block size.

The measure of PSNR [eqn (4)] provides an indirect measure of the accuracy of motion estimates that does not require *a priori* knowledge of the actual motion field. Because it is pixel based rather than moxel based, however, it is susceptible to speckle decorrelation and other system noise. The motion estimates of an image pair with a high degree of speckle decorrelation or a high noise level usually show a low PSNR value. This most likely explains why our motion estimates from the computer simulations had the highest PSNR values, whereas estimates based on B-scans of tissue-mimicking phantoms had lower PSNR values but still higher than those from muscle. PSNR also measures the sensitivity of an algorithm to local variations in image intensity, whether they are the result of motion or noise. In our experiments, the PSNR values associated with SMBM were lower than those of MLBM but higher than those of SLBM, suggesting moderate spatial resolution.

One potential disadvantage in using a smoothness constraint is a blurring effect at motion boundaries. Konrad and Dubois (1992) used line fields to account for sharp transitions in motion vectors at motion boundaries, but with a dramatic increase in computational complexity. Simultaneous motion estimation and segmentation (Chang et al. 1994) is another approach, although initial segmentation classes from soft tissue speckle tracking generally are unavailable. We have handled the problem of motion blurring by limiting the number of motion vectors used in the opti-

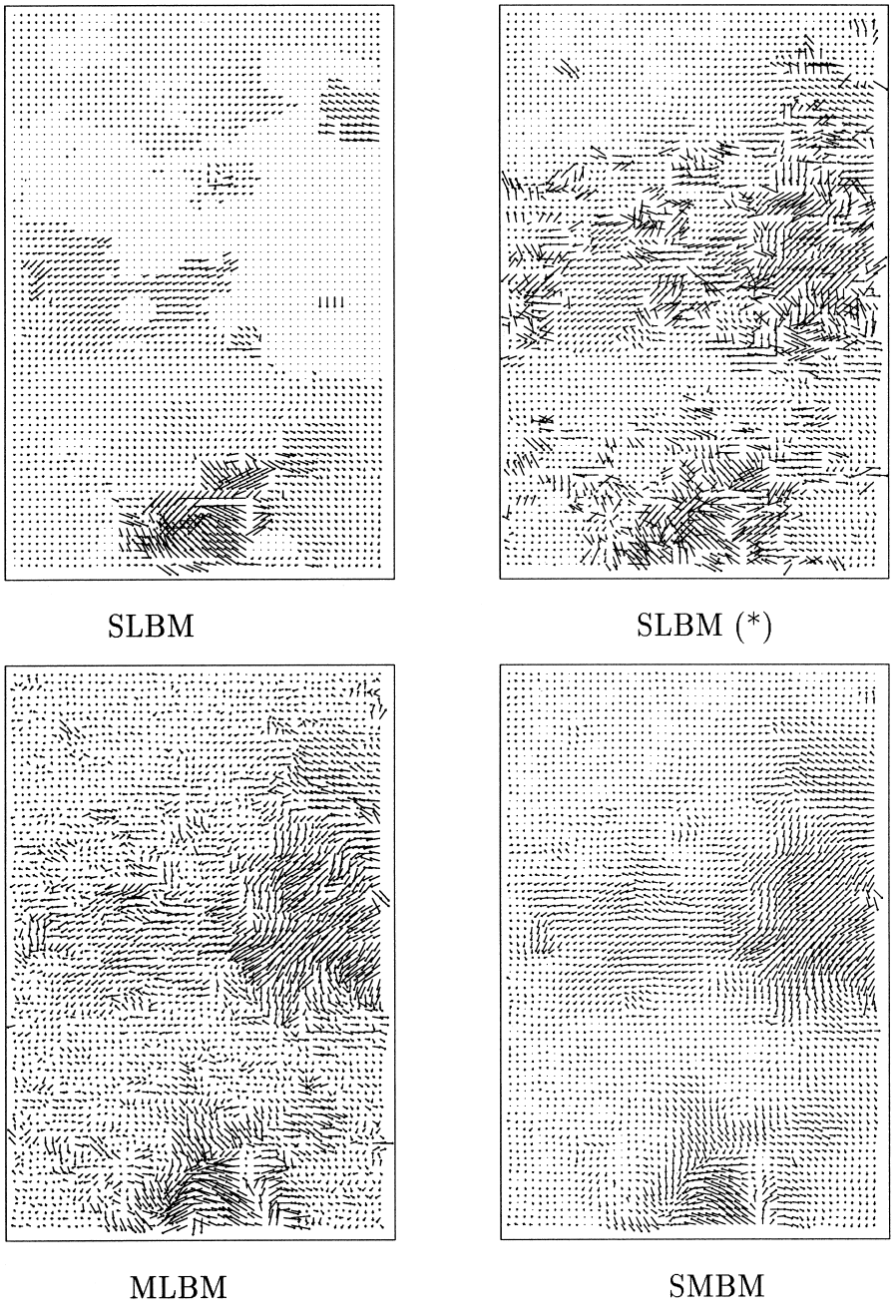

Fig. 10. Performance of single-level (SLBM), multilevel (MLBM) and spatial motion (SMBM) algorithms in tracking muscle motion in the forearm.

mization process. Although the binning process is neither as sophisticated nor as robust as the line field and segmentation techniques, it can be implemented very easily and adds little additional computational complexity, and may even reduce the overall computation time because the extra computations used in ranking motion candidates can be offset by the savings associated with using a limited subset of the available motion vectors. Our experiments were not designed to test sensitivity to motion discontinuities, however, and further investigation will be needed to test our hypothesis.

The choice of the spatial smoothness coefficients, β_L, in eqn (2), is currently a matter of empirical trials for different hierarchical levels and is likely related to the particular tissue investigated and the equipment. Basi-

Table 4. Performance of speckle tracking algorithms in tracking contraction of flexor muscles of the forearm.

	SLBM	SLBM*	MLBM	SMBM
Computation time (min)	983	324	65	72
SSD computations ($\times 10^6$)	15000	3900	304	304
PSNR (dB)	40.95	44.67	51.34	49.41

MLBM = multilevel block matching; PSNR = peak signal-to-noise ratio; SLBM = single-level block matching; SMBM = spatial motion block matching; SSD = sum of squared difference.

SLBM* = single-level block-matching with reduced matching block size.

cally, a large value of the smoothness coefficient should be used if a higher degree of speckle decorrelation or noise level is expected from the image data. As the physical size of the motion field neighborhood shrinks at finer hierarchical levels, the connectivity within the neighborhood is expected to increase and, therefore, the value of the smoothness coefficient should also be increased. In this article, the values of β_L in the four levels were chosen initially in increments of 4, 8, 16 and 32. Since the smoothness constraint represents a summation of the squared differences of neighboring moxel values, the values of β_L also should be squared: 16, 64, 256 and 1024. In addition to the values of the smoothness coefficient, the bin size also is important. We have found that a bin size of 8% of the total number of search locations works well, as compared to sizes larger and smaller than this number. We hope to develop more objective criteria for the selection of smoothness coefficients and bin size in the near future.

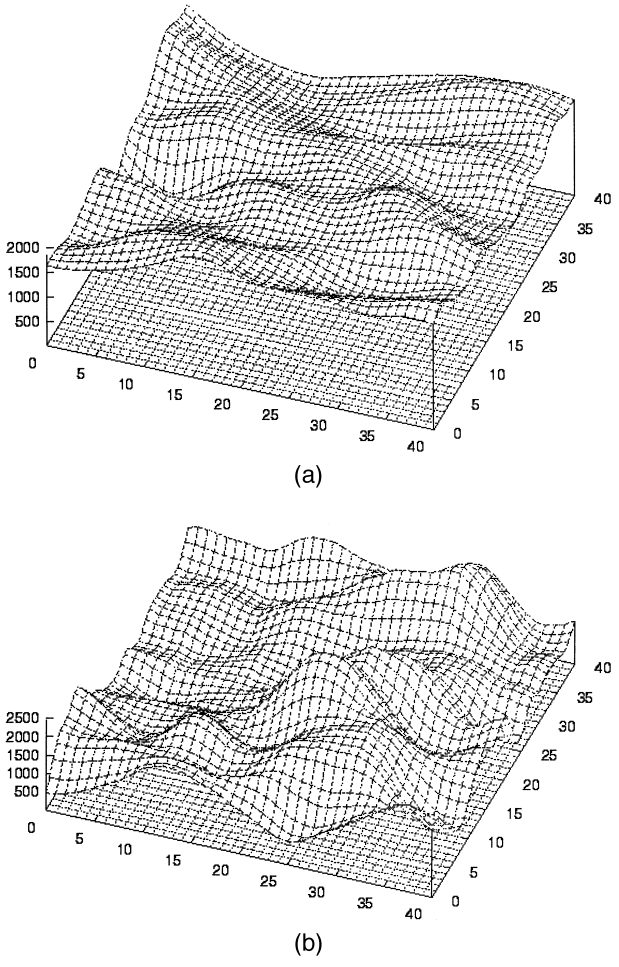

Fig. 11. Representation of SSD space with (a) large and (b) small matching blocks.

CONCLUSION

We demonstrated that 1) a traditional single-level block matching speckle tracking algorithm with a large matching block size provides good accuracy for rigid motion but fails to resolve highly varying nonrigid motion fields, and 2) block matching with a reduced matching block size provides good spatial resolution but at the expense of tracking accuracy. The motion model-based speckle tracking algorithm presented here combines the tracking accuracy of a large matching block with the spatial resolution associated with a smaller matching block through the use of a hierarchical coarse-to-fine scheme in conjunction with a physically based smoothness constraint. Our experimental results demonstrate that the motion model-based algorithm with a multilevel scheme has a comparable tracking accuracy and better spatial resolution than block matching using a large matching block, while saving substantial computation time.

REFERENCES

Ashton EA. Segmentation and feature extraction techniques, with applications to biomedical images. Ph.D. thesis. Rochester: University of Rochester, 1996.

Bohs LN, Friemel BH, McDermott BA, Trahey GE. A real time system for quantifying and displaying two-dimensional velocities using ultrasound. Ultrasound Med Biol 1993;19:751–761.

Céspedes I, Ophir J, Ponnekanti H, Maklad N. Elastography: Elasticity imaging using ultrasound with application to muscle and breast in vivo. Ultrason Imaging 1993;15:73–88.

Chang MM, Sezan MI, Tekalp AM. An algorithm for simultaneous motion estimation and scene segmentation. Proc. Int. Conf. ASSP, Adelaide, Australia, April 1994.

Chen EJ. Ultrasound tissue displacement and tissue elasticity imaging. Ph.D. thesis. Urbana-Champaign: University of Illinois at Urbana-Champaign, 1995.

Gao L, Alam SK, Parker KJ, Lerner RM. Sonoelasticity imaging: Theory and experimental verification. Ultrasound Med Biol 1995; 97:3875–3886.

Gao L, Parker KJ, Levinson SF. Imaging of the elastic properties of tissue—A review. Ultrasound Med Biol 1996;22:959–977.

Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of image sequences. IEEE Trans Pattern Anal Machine Intell 1984;6:721–741.

Huang SH, Lerner RM, Parker KJ. Time domain Doppler estimators of

the amplitude of vibrating targets. J Acoust Soc Am 1992;91:965–974.
Hein IA, O'Brien WD. Current time-domain methods for assessing tissue motion by analysis from reflected ultrasound echoes—A review. IEEE Trans Ultrason Ferroelec Freq Contr 1993;40:84–102.
Horn BK, Schunck BG. Determining optical flow. Artif Intell 1981; 17:185–203.
Jensen JA. Estimation of blood velocities using ultrasound: A signal processing approach. Boston: Cambridge University Press, 1996.
Kallel F, Betrand M, Meunier J. Speckle motion artifact under tissue rotation. IEEE Trans Ultrason Ferroelec Freq Contr 1994;41:105–122.
Konrad J, Dubois E. Bayesian estimation of motion vector fields. IEEE Trans Pattern Anal Machine Intell 1992;14:910–927.
Krouskop TA, Dougherty DR, Levinson SF. A pulsed Doppler ultrasonic system for making noninvasive measurements of the mechanical properties of soft tissue. J Rehab Res Dev 1987;24:1–8.
Levinson SF, Shinagawa M, Sato T. Sonoelastic determination of human skeletal muscle elasticity. J Biomech 1995;28:1145–1154.
Levinson SF, Yeung F, Walker W, Trahey G. A sonoelasticity imaging display for 2-d speckle tracking. Ultrason Imaging 1994;16:38.
Mailloux GE, Langlois F, Simard PY, Bertrand M. Restoration of the velocity field of the heart from two-dimensional echocardiograms. IEEE Trans Med Imaging 1989;8:143–153.
O'Donnell M, Skovoroda AR, Shapo BM, Emelianov SY. Internal displacement and strain imaging using ultrasonic speckle tracking. IEEE Trans Ultrason Ferroelec Freq Contr 1994;41:314–324.
Ophir J, Céspedes I, Garra B, et al. Elastography: Ultrasonic imaging of tissue strain and elastic modulus in vivo. Eur J Ultrasound 1996;3:49–70.
Poggio T, Torre V, Koch C. Computational vision and regularization theory. Nature 1985;317:314–319.
Ramamurthy BS, Trahey GE. Potential and limitations of angle-independent flow detection algorithms using radio-frequency and detected echo signals. Ultrason Imaging 1991;13:252–268.
Ryan LK, Lockwood GR, Starkoski BG, et al. A high frequency intravascular imaging system for investigation of vessel wall properties. Proc IEEE Ultrason Symp 1992;2:1101–1105.
Tekalp AM. Digital video processing. Englewood Cliffs, NJ: Prentice-Hall, 1996.
Trahey GE, Hubbard SM, von Ramm OT. Angle independent ultrasonic blood flow detection by frame-to-frame correlation of B-mode images. Ultrasonics 1988;26:271–276.
Tristam M, Barbosa DC, Cosgrove DO, et al. Ultrasonic study of *in vivo* kinetic characteristics of human tissues. Ultrasound Med Biol 1986;12:927–937.
Wagner RF, Insana MF, Smith SW. Fundamental correlation lengths of coherent speckle in medical ultrasonic images. IEEE Trans Ultrason Ferroelec Freq Contr 1988;35:34–44.
Wagner RF, Smith SW, Sandrik JM, Lopez H. Statistics of speckle in ultrasound B-scans. IEEE Trans Son Ultrason 1983;30:156–163.
Wang LM, Shung KK. Adaptive pattern correlation for two-dimensional blood flow measurements. IEEE Trans Ultrason Ferroelec Freq Contr 1996;43:881–887.
Wear KA, Popp RL. Theoretical analysis of a technique for the characterization of myocardium contraction based upon temporal correlation of ultrasonic echoes. IEEE Trans Ultrason Ferroelec Freq Contr 1987;34:368–375.
Yamakoshi Y, Sato J, Sato T. Ultrasonic imaging of internal vibration of soft tissue under forced vibration. IEEE Trans Ultrason Ferroelec Freq Contr 1990;37:45–53.