A FREEHAND ULTRASOUND ELASTOGRAPHY SYSTEM WITH TRACKING FOR IN VIVO APPLICATIONS

PEZHMAN FOROUGHI,* HYUN-JAE KANG,* DANIEL A. CARNegie,† MARK G. VAN VLEDDER,† MICHAEL A. CHOTI,‡ GREGORY D. HAGER,* and EMAD M. BOCTOR‡

*Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA; †Department of Surgery & Oncology, Johns Hopkins Medical Institutions, Baltimore, MD, USA; and ‡Department of Radiology, Johns Hopkins Medical Institutions, Baltimore, MD, USA

(Received 1 December 2011; revised 15 June 2012; in final form 10 September 2012)

Abstract—Ultrasound transducers are commonly tracked in modern ultrasound navigation/guidance systems. In this article, we demonstrate the advantages of incorporating tracking information into ultrasound elastography for clinical applications. First, we address a common limitation of freehand palpation: speckle decorrelation due to out-of-plane probe motion. We show that by automatically selecting pairs of radio frequency frames with minimal lateral and out-of-plane motions, combined with a fast and robust displacement estimation technique, greatly improves in vivo elastography results. We also use tracking information and image-quality measures to fuse multiple images with similar strains that are taken from roughly the same location so as to obtain a high-quality elastography image. Finally, we show that tracking information can be used to give the user partial control over the rate of compression. Our methods are tested on a tissue-mimicking phantom, and experiments have been conducted on intraoperative data acquired during animal and human experiments involving liver ablation. Our results suggest that in challenging clinical conditions, our proposed method produces reliable strain images and eliminates the need for a manual search through the ultrasound data in order to find radio frequency pairs suitable for elastography. (E-mail: pezhman@cs.jhu.edu and eboctor1@jhmi.edu) © 2013 World Federation for Ultrasound in Medicine & Biology.

Key Words: Elastography, Elasticity imaging, Tracking, Strain, Electromagnetic tracker, In vivo imaging.

INTRODUCTION

Ultrasound elastography is enabled by applying a mechanical stimulus and estimating the disturbance created by this stimulus. Stimulus can be static mechanical load (Ophir et al. 1991) or can be generated by external vibration (Parker et al. 1990) or by an acoustic push signal (Bercoff et al. 2004). Estimation can be achieved by either tracking displacement or measuring the speed of the propagating wave. A common flavor of this technology is quasistatic elastography in which the elastography image is generated by comparing the pre- and postcompression images to form a displacement map. The displacement can be then differentiated axially to find the strain map. When compression is applied to the tissue by hand via the ultrasound transducer, the elastography is referred to as freehand elastography. Many ultrasound systems are shipped with freehand elastography modules already installed. Note that this differs from freehand ultrasound, in which the transducer is swept over the region of interest to create a volumetric (three-dimensional [3D]) image.

Physicians find elastography with freehand palpation natural because it resembles the centuries-old practice of examination by hand palpation. Moreover, ultrasound elastography does not necessarily require special hardware or major alterations in the equipment, so it can be an integral part of any ultrasound system from high-end cart-based to pocket-sized systems.

Despite these advantages, freehand elastography introduces several challenges that have prevented the routine use of elastography in the diagnoses and treatments of patients. This technique is highly qualitative and user dependent. The best result is achieved when the examiner compresses and decompresses the tissue uniformly in the axial direction with the speed that produces proper strain rate (Ophir et al. 1999; Hall et al.
Small lateral or out-of-plane motions can result in decorrelation, which reduces signal-to-noise ratio (SNR) (Hall et al. 2003; Chandrasekhar et al. 2006). However, it is difficult to induce pure axial motion with freehand compression, especially for slippery or oblique surfaces or for intraoperative imaging where the access to the tissue may be limited. The widespread clinical use of elastography could benefit from the development of elastography techniques that are less influenced by factors such as expertise of the user. It would also enable the longitudinal study of tumors in which the strain images acquired during different time periods are compared.

Mechanical attachments designed to control the motion of the probe can reduce role of the user. In Hiltawsky et al. (2001), a mechanical compression applicator that applies uniform axial motion via compression plates was developed. The mechanical applicator reduced the amount of out-of-plane motion and consequently the dependency of the experience of the examiner. A 50% reduction in out-of-plane motion was reported by Kadour and Noble (2009) using an assistive hand-held device. In Rivaz and Rohling (2007), a low-frequency vibration was induced into the tissue while reaction forces were compensated for. In a recent work, robotic palpation was studied in laparoscopic elastography (Billings et al. 2012). The main issues with mechanical solutions are that they require extra hardware, limit elastography to specific applications, make the probe bulky, and reduce its maneuverability.

Sophisticated algorithms have been developed that extend the search for displacement to the lateral direction and tolerate certain amounts of decorrelation (Pellot-Barakat et al. 2004; Rivaz et al. 2008; Jiang and Hall, 2009). These algorithms can only partially address the problem by compensating for in-plane motions and enforcing smoothness constraints. However, they remain vulnerable to out-of-plane and large in-plane motions.

Another approach is to define a high-quality metric to evaluate the performance of strain images. Jiang et al. (2006) defined this metric as the multiplication of the normalized cross-correlation (NCC) of the motion compensated RF fields (i.e., motion tracking accuracy) and the NCC of the motion compensated strain fields (i.e., consistency of strain images). This quality metric may be provided to the user as a feedback or used to adaptively selected for different frame step sizes.

Lindop et al. (2008) developed a framework in which the streams of strain images are normalized for the displacement estimates by a nonlinear function to compensate for uneven distribution of force and varying strain rates. The generated “pseudo-strain” images were then averaged over time on a per-pixel basis, which depends on the precision of displacement estimation. The weights were also presented to the user as a feedback on the quality.

Normally the gap between frame pairs in elastography is fixed, and it is assumed that poor elastography results are detectable. Even when this gap is dynamically selected, the search range remains very small. The other limitation of these approaches is that the strain has to be estimated prior to the calculation of the quality metric. In typical settings, the ultrasound frame rate can reach well above 30 Hz. For consecutive frames, an efficient implementation of these image-based metrics (probably with graphics processing unit [GPU] acceleration) might accommodate this frame rate. However, the task will be extremely difficult when the aim is to evaluate all plausible combinations of frame pairs in a series of images.

It is common to localize the ultrasound transducer by attaching a position sensor to it. Some of the main applications of tracked ultrasound include volume reconstruction and extended field of view (Fenster et al. 2001); multimodality registration (Gobbi et al. 1999; Foroughi et al. 2008); spatial compounding (Rohling et al. 1998); and navigation/guidance (Banovac et al. 2002; Sjolie et al. 2003). The growing interest in tracked ultrasound has led to the emergence of ultrasound systems from companies such as Ultrasonix, GE, and Siemens, which have electromagnetic sensors already embedded within their transducers.

The aim of this study was to show that by tracking the ultrasound transducer, the information from the tracking device can be exploited to select pairs of RF frames for elastography that produce images with high SNR, contrast-to-noise ratio (CNR), and consistency. The frame selection happens before the calculation of displacement and strain and requires negligible computational power. The selected frame pairs contain minimal lateral and out-of-plane motion and a predefined compression range with respect to each other. Based on tracking information, the strain images taken from roughly the same location are fused to obtain a high-quality elastography image. The method is validated by using tissue mimicking phantom as well as intraoperative experiments on a porcine subject and on human-patient data.

METHODS

The block diagram of our algorithm summarizing the major steps are shown in Figure 1. As shown in this
Distributions of phase of the signal reflected by the scatterers randomly distributed within the resolution cell (Wagner et al. 1983) of the collective response of the ultrasound (Wagner et al. 1983; Shankar, 2000): the decorrelation model is normally employed to estimate the displacement, whereas here, the correlation is estimated knowing the displacement. Extending eqn (2) to all 3 directions of displacement, a pseudocorrelation function, $C_{rr}$, is defined as follows:

$$ C_{rr}(D_x, D_y, D_z) = \exp \left( -K_x D_x^2 - K_y D_y^2 - K_z D_z^2 \right), $$

(3)

where $D_x$, $D_y$, and $D_z$ represent the displacement in out-of-plane, axial, and lateral directions, respectively. $K_x$, $K_y$, and $K_z$ determine the sensitivity to motion in their corresponding directions.

**Distance metric**

Given the relative homogeneous transformation between 2 frames, eqn (3) can be used directly only when the relative rotation is ignored. However, even a small rotation may cause large decorrelation at the bottom of the image and degrade the quality of the strain. At the same time, a component-wise metric is needed because the axial motion has to be isolated as the desired type of motion. Suppose $a = [a_x, a_y, a_z]^T$ is the axis-angle representation of the relative rotation, and $t = [t_x, t_y, t_z]^T$ is the relative translation. Assuming a small rotation, the relative displacement of a point, $P = [x, y, 0]^T$, will be $d = a \times P + t$. We then define the distance vector of 2 frames, $D = [D_x, D_y, D_z]^T$, as the root mean square of the components of $d$ for all the points in the region of interest:

$$ D_x = \sqrt{a_x^2 + (a_y + t_y)^2}, $$

$$ D_y = \sqrt{a_x^2 + (a_y + t_y)^2}, $$

$$ D_z = \sqrt{a_x^2 + (a_y + t_y)^2}, $$

where $a_x$, $a_y$, and $a_z$ represent the random amplitude and the phase of the signal reflected by the $i$th scatterer. The distributions of $a_i$ and $\theta_i$ are often assumed to be uniform.
where \( \sqrt{\cdot} \) returns the root. The region of interest is assumed to be rectangular and determined by \( x_1, x_2, y_1, \) and \( y_2 \). Symbols \( x_{21} \) and \( y_{21} \) are equal to \( x_2-x_1 \) and \( y_2-y_1 \), respectively. The vector \( D \) provides a measure of distance for each direction separately.

**Frame pair selection**

Given a sequence of RF frames, all possible combinations of frame pairs are evaluated using a slightly modified version of \( C_{rr} \). For \( M \) frames, there will be \( \binom{M}{2} = M(M-1)/2 \) pair combinations which provide more choices than consecutive frames. Because the pairs are compared directly, it suffices to minimize the negated exponent of eqn (3) in order to maximize \( C_{rr} \). Axial motion is desired, so the term for axial motion is also modified to penalize compressions that are higher or lower than the optimum compression value, \( t_{opt} \). This forms a cost function, \( Cost \) being defined as follows:

\[
Cost(D) = K_\delta D_x^2 + K_y \left( \frac{D_x - t_{opt}}{D_x + c} \right)^3 + K_z D_z^2, \tag{5}
\]

\( t_{opt} \) is determined based on the desired strain rate, which is set by the user considering the amount of decorrelation that the elasticity estimation algorithm can handle. \( c \) is a small number that limits the cost of zero compression. With \( D_\delta \) much larger than \( t_{opt} \), the second term of the cost function grows quadratically, similar to that of lateral and out-of-plane motion. A \( D_\delta \) value close to zero also increases the cost, preventing the selection of pairs with no axial motion. The sensitivity to the exact value of \( t_{opt} \) is adjusted by \( K_y \) (Fig. 2).

Using this function, multiple frame pairs with the lowest cost values are identified. In our implementation, a maximum of 8 frame pairs are selected to ensure an upper bound for the required processing time. A pair is rejected if the value of the cost function is above \( ln(2) \) (or equivalently when the pseudocorrelation drops to lower than 0.5).

**Elastography and fusion**

To compute the displacement map for a pair of RF frames we have employed the 2D AM technique (Rivaz et al. 2011), which is shown to be robust and fast. It computes a subpixel estimation of the motion field in both axial and lateral directions. The regularizing terms makes it tolerant to outliers and noise. Furthermore, it can handle small internal motions induced by heartbeat or respiration. Because localization information is assumed to be available to us, we can automatically adjust the search range for the recovery of the motion field and increase the speed of displacement estimation.

Given the displacement map, the strain is computed by applying least-squares fitting. The strain image is calculated for all selected pairs. The final image is found by merging the motion-compensated strain images. Motion compensation is aimed at reducing the effect blurring caused by mixing strain images. Assume the selected frame pairs are \( \{ (I_1^1, I_1^2), (I_2^1, I_2^2), \ldots, (I_n^1, I_n^2) \} \), where \( n \) is the number of selected frame pairs resulting in the strain set \( \{ S_1, S_2, \ldots, S_n \} \). \( I_1^1 \) is chosen as the reference frame, and \( I_2^1 \) to \( I_n^1 \) are registered back to the reference frame to find their relative global compression and lateral motion. This registration is applied to a subset of RF intensities forming a grid, \( G \). The tracking information is used to initialize the registration. \( G \) is composed of a constant number (20) of rows and columns. This approach speeds up the optimization and makes it independent of the number of samples in the RF frames. For a pair of frames, \( I_1^1 \) and \( I_2^2 \), the optimization function will be:

\[
\arg \max_{a,b} \left\{ \sum_{x,y \in G} \{ I_1^1(x,y) \cdot I_2^2(x+a,y) + I_1^1(x-a,y) \cdot I_2^2(x,y) \} \right\} \tag{6}
\]

The optimization variables \( a \) and \( b \) determine the lateral motion and axial compression, respectively. The symmetric form of the objective function ensures that the optimization does not depend on the order of frames, and exchanging the frames affects only the sign of the variables. The search spans of \( a \) and \( b \) are initialized from the tracking information. In the case of \( b \), this range will be from 0 to the value reported by the tracker because the actual observed compression might be less than that value. For both \( a \) and \( b \), the search span is also extended on both sides, with an assumed maximum error value for
the tracker readings (0.2 mm). Note that the relative motion error is expected to be much lower than the absolute error.

The motion-compensated strain images, $S'_i$, are constructed by applying the global deformation to the strain images ($S_i$). We are interested only in global motion, so the registration is fast and unlikely to fail. These images are merged by weighted averaging on the pixel level to increase the SNR forming the output image, $S_{\text{final}}$:

$$S_{\text{final}} = \frac{\sum_{i=1}^{n} w_i(p) \cdot S'_i(p)}{\sum_{i=1}^{n} w_i(p)}$$

$$w_i(p) = \rho_{i}^{\text{cost}} \cdot \rho_{i}^{\text{cc}} \cdot \rho_{i}(p).$$

The weight factor, $w_i(p)$, is found by multiplying 3 values, $\rho_{i}^{\text{cost}}$, $\rho_{i}^{\text{cc}}$, and $\rho_{i}(p)$. $\rho_{i}^{\text{cost}}$ is the estimated correlation coefficient from the cost function of eqn (5) ($\rho_{i}^{\text{cost}} = \exp\{-\text{Cost}_{i}\}$). $\rho_{i}^{\text{cc}}$ and $\rho_{i}(p)$ represent the overall and pixel-level correlation coefficient of the RF frame pair after applying the estimated displacement. $\rho_{i}^{\text{cost}}$ reduces the bias toward lower compression created by the other 2 factors. This bias is due to the fact that in general pairs with lower compression yield higher correlation coefficients.

Whenever the overall correlation coefficient is low for a pair of images, the whole strain estimation cannot be trusted, even if the pixel-level correlation for a specific area is high. The inclusion of $\rho_{i}^{\text{cc}}$ reduces the share of strain images with low overall correlation coefficients.

### Cross-section test

It is meaningful to fuse only images that are taken from roughly the same cross-section of tissue. The tracking data may be utilized to detect which frames are from 1 cross-section of the tissue, with minimal lateral and out-of-plane motion. This is a step taken prior to the selection of frame pairs and provides the pool of frames from which the pairs are selected.

Axial motion is not of interest in this case, so $K_y$ is set to 0 in eqn (3), providing a closeness measure. Given a single frame, the closeness measure of this frame with respect to all other frames in the sequence is computed. These values are sorted in descending order, and the frames with closeness measures greater than 0.5 are chosen. A threshold on the maximum number of frames, $M$, is also applied, which ensures that the number of frames from that cross-section is equal to or less than $M$. This process is carried out for all frames in the sequence and the set of frames with the largest sum of closeness measures is chosen.

One can also extract sets of frames from multiple cross-sections and generate a strain image for each cross-section. The availability of localization information means that a 3D strain volume could be automatically reconstructed.

### DATA ACQUISITION

#### System specifications

Two different ultrasound systems were used to collect RF data. The data from phantom and pig experiments was collected using the SonixSP ultrasound system (Ultrasonix Medical Corporation, Richmond, BC, Canada) and the L12-5 transducer with central frequency of 10 MHz. The RF data was obtained via our interface software based on Ultrasonix research software development kits. This software controls RF acquisition and sends the data over a local network to a laptop computer where the RF frames are saved along with the tracking information. This is our preferred setup because it provides an excellent research interface, and we have full control of the parameters of the ultrasound machine and the flow of the data.

In the case of patient experiments, the raw ultrasound data were obtained from ACUSON Antares (Siemens Medical Solutions USA, Malvern, PA, USA), a high-end ultrasound system with an intraoperative ultrasound transducer (VF13-5SP) at a center frequency of 8.89 MHz. The Axius Direct Ultrasound Research Interface was employed to enable RF acquisition and trigger the ultrasound system. For synchronization purposes, our software signaled both the ultrasound system and the external tracker at the start of data collection. The duration of each sequence of RF frames was limited by the buffer size on the machine.

The localization information was gathered via a medSAFE electromagnetic (EM) tracker (Ascension Technology, Milton, VT, USA) with model 180 sensors. The midrange cubic transmitter supplied the magnetic field for all experiments except for the patient experiment where a flatbed transmitter was placed under the mattress of the surgical bed before the arrival of the patient. The data from the flatbed transmitter was filtered by a low-pass filter to reduce the inherent jitter in the readings. The ultrasound transducer was calibrated using the standard cross-wire calibration (Prager et al. 1998).

#### Synchronization

It is important for the success of this method to have correspondence between the RF frames and the tracking data. Here, we relied on the timestamp of the tracker readings and the RF frames. The task of synchronization would be to find the constant delay between the 2 timestamps and compensate for it. This goal was achieved by exploiting the same freehand palpation used for elastography, which implies that no special procedure is required for synchronization. This also enables dynamic synchronization, where not only the constant delay between RF and tracker data is estimated as the data are collected, but also small jitters in delay are compensated for.
In this work, a constant delay was assumed because time-stamps are available for the acquired data. To find this delay, the global axial compression was first recovered by correlating the first frame and the stretched version of the next frame. The same optimization procedure described by eqn (6) was employed for this purpose. The compression values were integrated over a period of time to get the compression with respect to the first frame in that period. The resulting curve was matched with the axial motion reading from the EM tracker using NCC. For both the Ultrasonix and the Antares machines, the tracker and RF data were synchronized using this technique. Figure 3 shows the 2 signals after the delay was compensated for. To find the exact transformation for each RF frame, the tracking data interpolated. The translations and rotations were separately interpolated using spline interpolation for the translations and spherical linear interpolation (Shoemake, 1985) for the rotations.

**EXPERIMENTS**

The performance of our algorithms is demonstrated in 3 sets of experiments: phantom, animal, and patient experiments. The phantom experiment provides a controlled environment for which the shape, location, and elasticity of the target is known. The pig experiment replicates the challenging environment of the operation room and surgery but at the same time, control of the duration of the experiment and of the ultrasound system parameters is maintained. The gross pathology is also available because the pig is sacrificed at the end of the experiment. The patient experiment imposes a less controlled environment in which the time for data collection is limited, and minimal interference is introduced to the routine flow of the surgery.

**Phantom experiment**

The data were taken from the CIRS (Norfolk, VA, USA) elasticity phantom model 049A. Out-of-plane and lateral motion are less problematic in this experiment because the surface of the phantom is flat and there is more control over hand motion. However, low compression rates can affect the quality of the strain image. Figure 4 shows 4 strain images with increasing compression rates taken from a 1.58 mm-diameter lesion located at a depth of roughly 3 cm. The strain images are automatically generated from a single sequence of RF frames using our frame-selection technique. This experiment demonstrated the capability of our frame selection of providing some level of control over the strain rate. It also showed that higher strain rates are more effective for scanning small lesions. This lesion is almost 3 times softer than the background and does not show up in the B-mode image (Fig. 4b). In Figure 4 (c–e), as the strain rate increases, enough SNR and contrast are achieved to make the lesion easily detectable. The diameter of the lesion in the image seems to be slightly larger than the original size. This could be due to smoothing constraints applied in the elastography algorithm and the least-square fitting. The strain rates chosen by the user in Figure 4 (c–f) are given in the figure’s caption. The algorithm tries to find the strain rate closest to these values. However, the strain rate may be lower than the selected value when not enough compression is applied by hand. In this case, the phantom was not compressed more than about 4% so as to prevent damaging it. The frame rate for RF data was about 30 frames per s which means, for a slow palpation, a large gap between the frames (12 frames for 4f) was necessary to achieve the required strain.

**Animal experiment**

Generating strain images intraoperatively is difficult because factors, such as experimental time constraint, safety, and the introduction of surgical tools, have to be taken into account. The access to the imaging area may be limited, and internal motions can be disruptive. In these conditions, axial hand motion does not usually translate into simple up-and-down palpation. The force has to be exerted at an oblique angle, making it difficult to avoid lateral and out-of-plane motions. Moreover, the slippery surface of the tissue or the ultrasound gel applied to the surface of the tissue causes the transducer to slide over the tissue. The aim of this experiment was to show the benefit of the proposed technique for intraoperative imaging.

Our protocol for animal experiment was approved by the Johns Hopkins University Institutional Care and Animal Use Committee. For this experiment, the cubic electromagnetic transmitter was secured on the bedside prior to the arrival of the pig. The pig was then prepared.
for the surgery, and the liver was exposed for ablation. Next, the ablation needle was inserted inside the liver under ultrasound guidance. The liver was then ablated to create an ablation zone of about 2 to 3 cm. After the ablation, the needle was removed and a sequence of tracked RF data were collected from the ablated region. The data collection was carried out a few minutes after the ablation to allow for the dispersion of the bubbles caused by the ablation procedure. The placement of the probe and the direction of the needle were marked on

Fig. 4. Four strain images with increasing compression rates are automatically generated from a single sequence. The structure of the phantom is shown in (a). The small lesion is not visible in the B-mode image displayed in (b). The resulting strain image for the selected strain rate is shown in (c) through (f). The lesion is clearly visible in (f).
the surface of the liver by 2 small burns on each side of the transducer and a third arrow-shaped burn mark. Three zones were ablated, creating 3 hard lesions (a fourth lesion was also ablated for a separate experiment). Finally, the pig was sacrificed and the liver was harvested. Guided by the burn marks, the liver was cut on approximately the same plane on which the ultrasound data were gathered.

**Figure 5** shows the B-mode, strain, translational motion from tracking, and gross pathology of the ablated zone. The core of the ablation appears as a slightly hypoechoic region in the B-mode image. Compared to the whole ablated region, this core is smaller and contains the dehydrated tissue for which the acoustic properties are mildly modified. The needle imprint is also visible as a hyperechoic line inside the ablation. The strain image shown in **Figure 5** (b) is the output of our algorithm. The ablation is harder than the surrounding tissue and appears darker in the strain image. It is larger than the hypoechoic region in the B-mode image, and its size matches the size of ablation from gross pathology (Fig. 5d).

The graph in **Figure 5** (c) displays the translational motion of the ultrasound plane with respect to the first frame in the sequence. The effect of respiration is captured in the lateral and elevational components of the tracking information as a periodic cycle. In this case, the breathing pushed the liver which, in turn, caused the motion of the transducer placed on the top of it. Only the frames from the resting portion of the breathing cycle are suitable for computation of strain because the breathing motion may cause the transducer to slide over the tissue. The presented algorithm can automatically compute the strain from only the resting phase of respiration as long as the breathing cycle is captured in the tracking data. This observation was possible in the animal experiment because the data collection system permits the obtaining of long sequences of images containing several breathing cycles.

**Fig. 5.** The figure shows an in vivo ablation experiment on pig liver. The B-mode image (a) shows only the core of the ablation, composed of dehydrated tissue. The strain in (b) is generated by our method and depicts the ablated zone in shaded areas. The size of a hard lesion in the strain image matches the gross pathology shown in (d). The relative translational motion of the frames with respect to the first frame is presented in (c). The breathing motion is captured in the lateral and elevational components of the tracking data.
The protocol for patient experiment was approved by the Johns Hopkins Institutional Review Board, and informed consent was acquired from all patients. In this experiment, raw ultrasound data were obtained from a cancer patient undergoing multiple liver wedge resections and radiofrequency ablations of metastatic lesions. In order to introduce minimal deviation into the normal flow of the surgery, only 1 lesion was imaged before and after ablation. The ultrasound data acquisition was quick and added only a few minutes to the total length of the surgery.

To collect the ultrasound data, the surgeon had to slide the ultrasound probe inside the abdominal cavity through the surgical incision, making an out-of-plane motion and probe sliding nearly unavoidable. This can be seen in Figure 6, which shows the translational component of the hand motion of the surgeon during data collection. The frames selected by the cross-section test are marked by small squares on the top of the motion curve. From Figure 6 it can be noted that these frames have lower lateral and elevational (out-of-plane) motion with respect to each other.

The pre- and postoperation CT scans of the patient were taken as a routine part of the operation. In the preoperative CT scan, the contrast between the tumor and the background was very low, and an accurate comparison between CT and ultrasound could not be established. The postoperation CT scan of the patient is presented in Figure 7, where (a) contains the ablated region and (b) is the resliced cross-section that roughly matches the plane of ultrasound imaging. The vascular structure in the CT scan taken during the venous phase appears brighter, and the ablated region is darker compared to the normal liver tissue. To compare strain and CT, the plane in CT volume corresponding to the strain should be found. This could be achieved relying on anatomic features, as proposed by Clements et al. (2008). In this work, we relied on 3 clues to reslice the computed tomography (CT) volume. The first was the shape of the ablation, which was extended toward the direction of the needle. From the shape, we were able to approximate the position of the needle in the CT scan. We also collected a series of tracked ultrasound images containing the track of the needle. The position of the needle in CT and ultrasound was then matched. The elastography data were obtained from a plane almost perpendicular to the plane of the needle, avoiding the distortion by the needle. Assuming minimal motion during ultrasound data acquisition, the spatial relationship of the image with the needle and the elastography image could be determined. Although the rotation of the slice could not be determined by this method, the size of the ablated zone was measurable. In Figure 7 (a), 2 ablated regions are visible of which only the ablation on the right side was imaged. The pixel spacing (x and y directions) is 0.7 mm and the slice thickness is 3 mm. The diameter of the ablated zone is about 19 mm.

The elastography results of the tumor before and after ablation are depicted in Figures 8 and 9. Prior to ablation, the tumor appeared as a hypoechoic region in the B-mode ultrasound image (Fig. 8a). Because the
tumor is harder than the surrounding tissue, it is darker in the strain image. The border of the tumor and its shape are more easily distinguishable in the elastography image. The B-mode and strain images of the tumor after ablation shown in Figure 9 have characteristics similar to those in the animal experiment. A slightly hypoechoic core is visible in the B-mode image, which is created by the dehydrated tissue, similar to the animal experiment. This region is smaller than what is detectable in the CT scan. The size of the ablation in the elastography image is approximately 17 mm, which is closer to the size of the ablation in Figure 7 (b).

**QUANTITATIVE RESULTS**

SNR and CNR are the most common measures used for quantifying the quality of a strain image. The SNR and CNR values are computed from Chaturvedi et al. (1998):

\[
\text{CNR} = \sqrt{\frac{2(\bar{I}_b - \bar{I}_t)^2}{\sigma_b^2 + \sigma_t^2}}, \quad \text{SNR} = \frac{\bar{I}}{\sigma}
\]

where \(\bar{I}\) and \(\sigma\) denote the mean and standard deviation of intensities. The \(t\) or \(b\) subscripts mean that the
computation is only for the target or the background region, respectively.

Figure 10 shows the SNR and CNR on the CIRS phantom with a hard inclusion that is twice as stiff as the background. The results are presented for strain images generated from the automatically selected frames, consecutive frames, and a constant gap of 2 frames. Our method consistently selected frame pairs that produce strain images with higher SNR and CNR values. The drop in SNR is more prominent in consecutive frames where there is not enough compression between the frames. Increasing the gap improves the SNR and CNR.
in some frames but reduces these values in other frames due to decorrelation (see the bottom row of Fig. 10). The periodic behavior of SNR for the constant gap roughly matches cycles of the axial motion by hand. The reason is that the speed of hand motion is variable and the step size is not adjusted accordingly.

Averaging a large number of strain images increases the SNR of the signal as shown by Lubinski et al. (1999). However, it blurs the strain image, as demonstrated in Figure 8 (c) and (d). The blurring is caused mainly by the displacement of the tumor in the strain sequence, which includes out-of-plane motion. Table 1 shows the surge in SNR and CNR as a result of averaging, especially for successive frames of the pre-ablated tumor in our patient experiment. These values are still lower than those of our method, although the number of frame pairs used is tenfold larger. At the same time, the quality of the final image is severely degraded due to the displacement of the tumor during data collection.

In Table 2, our frame selection technique is compared to the selection of consecutive frames. The SNR, correlation coefficient, and strain consistency of the selected frames are being compared for all our experiments. The number of frames used in both methods are fixed (5 frame pairs). For consecutive frame selection, the average over the whole sequence (100 frames) is presented. The standard deviation of each value is shown in parenthesis.

The correlation coefficient is the same as $r_{cc}$ in eqn (7). The 2 methods produce comparable correlation coefficients, but for consecutive frames this number is slightly higher. Selecting consecutive frames normally yields the highest correlation coefficient. This does not translate into a higher quality of elastography. The reverse is true, however. If the correlation coefficient is low, the output image is not reliable. In all cases, the value of the correlation coefficient remained above 0.80, which means that the displacement estimation did not fail.

Consistency among strain images has been suggested as a good measure of the quality of the images. Jiang et al. (2006) defined a performance descriptor based on the consistency of 2 consecutive stain images. They showed that this descriptor correlates with the CNR and human ranking scores. The consistency is defined as the NCC of the 2 strain images. Here, the consistency of more than 2 images is examined. For this purpose, the NCC for all possible pairs of strain images is computed. The average and standard deviation of this value is reported in Table 2. The range of this value is from 0 to 1. For both consecutive frames and selected pairs, the number of strain images is set to 5, resulting in 10 possible combinations. In all cases of our experiments, the average consistency was significantly better for the frame pairs selected using the tracking information while the standard deviation is lower.

**DISCUSSION**

Choosing the right frame pairs is essential for the success of elastography in clinical conditions. Frame selection using tracking reduces the dependency of elastography on the quality of hand motion and provides

<table>
<thead>
<tr>
<th>N</th>
<th>Consistency Mean (SD)</th>
<th>Correlation coefficient</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FS</td>
<td>Consecutive</td>
<td>FS</td>
</tr>
<tr>
<td>1</td>
<td>0.75 (0.11)</td>
<td>0.21 (0.15)</td>
<td>0.92 (0.04)</td>
</tr>
<tr>
<td>2</td>
<td>0.93 (0.05)</td>
<td>0.56 (0.25)</td>
<td>0.92 (0.01)</td>
</tr>
<tr>
<td>3</td>
<td>0.85 (0.08)</td>
<td>0.57 (0.26)</td>
<td>0.84 (0.04)</td>
</tr>
<tr>
<td>4</td>
<td>0.72 (0.14)</td>
<td>0.56 (0.21)</td>
<td>0.86 (0.05)</td>
</tr>
<tr>
<td>5</td>
<td>0.80 (0.15)</td>
<td>0.54 (0.18)</td>
<td>0.98 (0.01)</td>
</tr>
<tr>
<td>6</td>
<td>0.83 (0.10)</td>
<td>0.49 (0.20)</td>
<td>0.98 (0.01)</td>
</tr>
</tbody>
</table>

FS, our frame-selection method; N, experiment numbers that correspond to the following experiments: 1, phantom; 2, pig liver lesion; 3, pig liver lesion; 4, pig liver lesion; 5, pre-ablated tumor in patient; 6, post-ablated tumor in patient.
control over the rate of compression. It is an effective tool, especially because the rate of ultrasound data is very high. The computational cost of evaluating frame transformations is practically nil compared to the cost of strain estimation, and that makes it possible to analyze hundreds of frame pairs prior to strain calculation. The combination of this technique with strain motion compensation and the pixel-wise image fusion significantly elevates the quality and reliability of elastography.

Although an additional tracking device is required for this method, localization data are already available to many ultrasound navigation systems. The arrival of new ultrasound systems that embed the localizing sensor within the probe suggests a rising trend for such systems. We believe that elastography can become an essential element of these systems and will help in diagnosis, monitoring, and targeting in clinical applications. In our experiments, the source of tracking was an EM tracker. Nonetheless, other types of tracking devices, such as optical trackers or accelerometers, could also be employed for this purpose.

We have showed that blindly averaging strain images could introduce significant blurring. Much better results can be achieved with a smaller number of images that are carefully selected using the tracking information. For the last step of this technique, a similar approach to that suggested by Jiang et al. (2007) and Lindop et al. (2008) could also be adopted to further refine the results of elastography.

The frame-selection method generally produces results with higher SNR. Both the average SNR of the selected frames and the SNR of the fused image are reported in Table 2. As expected, the combined SNR value is higher than that of individual images. Although the SNR value seems to be a better indicator of the quality of the image, there are some cases where this value might be misleading. Specifically in the second pig ablation, the consecutive frames produced better SNR. However, this is because the image in that case contained a large hard lesion that affected the mean of the strain and therefore reduced SNR (Fig. 11). As shown in Figure 11 (c), the noise in the individual strain images has caused the lesion appear softer than its actual stiffness in the averaged image. This is also partly the result of the regularization term in displacement estimation when the compression is very low. In this case,
the relatively higher value of SNR (1.14) is only the result of the increase in the mean value of the strain. In the frame selection method, low compression and unwanted motions are avoided, resulting in an image that is in agreement with gross pathology. The 3 cm size of ablation in Figure 11 (d) matches the measurement in gross pathology (Fig. 11b).

Consistency among frames seems to be the best indicator of failure. With multiple strain images from 1 cross-section available, this indicator could be employed to detect failed images and eliminate them from the final step. However, this will introduce additional computational costs to the system.

Our data acquisition interface, developed in “C++” programming language, undertakes the tasks of connecting to the ultrasound machine and tracking device, synchronizing the acquisition, and recording the data. Currently, our implementation is in MATLAB (MathWorks, Natick, MA, USA) and all the processing is performed off-line. However, there are no computational limitations for processing the data in real time as the data are being collected. The 2D AM method (Rivaz et al. 2011) appears to be both fast and robust, and the image fusion could be implemented in real time as it involves only 2D-image mapping and weighted averaging.

**CONCLUSION**

In this article, we have presented a method of incorporating the information from a tracking device in elastography to tackle the challenges of clinical applications. In our experiments, the consistency and the SNR of strain images was increased by 67% and 97%, respectively. This was a pilot study aimed at demonstrating the potential of this approach to produce reliable, high-quality strain images. Further clinical trials are necessary to evaluate fully the diagnostic impact of our frame selection. This system could be particularly effective in the intraoperative procedures in which traditional manual search through the frames is needed after data collection in order to find the pair that generate an acceptable image. In conjunction with the incorporation of tracking, our method benefits from efficacious displacement estimation and image-fusion techniques, making it resilient to failure.

There are numerous possibilities of improving this elastography system. In the presence of tracking data, an immediate improvement could be the addition of hand-motion feedback to the user. This will involve developing easy-to-understand interfaces that will provide real-time feedback during hand palpation. 3D strain imaging from the 2D images will also be a natural extension of this system. This will be similar to the volume reconstruction from 2D B-mode images, except that the localization information will be used not only for placing the strain images inside the volume but also for frame selection. Finally, the full integration of this system into a clinical platform will allow for clinical trials for specific applications.

**Acknowledgments**—Pezhan Foroughi is supported by the U.S. Department of Defense predoctoral fellowship program. This work was also supported by National Institute of Health (NIH) award CA134169. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. We thank Hassan Rivaz for providing the displacement estimation code and Linda Wang for her help with the experiments.

**REFERENCES**


