Statistical Shape Model to 3D Ultrasound Registration for Spine Interventions Using Enhanced Local Phase Features

Ilker Hacihaliloglu[1], Abtin Rasoulian[1]
Robert N. Rohling[1,2], and Purang Abolmaesumi[1]

[1] Department of Electrical Engineering
[2] Department of Mechanical Engineering, University of British Columbia, Vancouver, B.C., Canada
{ilkerh,abtinr,rohling,purang}@ece.ubc.ca

Abstract. Accurate registration of ultrasound images to statistical shape models is a challenging problem in percutaneous spine injection procedures due to the typical imaging artifacts inherent to ultrasound. In this paper we propose a robust and accurate registration method that matches local phase bone features extracted from ultrasound images to a statistical shape model. The local phase information for enhancing the bone surfaces is obtained using a gradient energy tensor filter, which combines advantages of the monogenic scale-space and Gaussian scale-space filters, resulting in an improved simultaneous estimation of phase and orientation information. A novel statistical shape model was built by separating the pose statistics from the shape statistics. This model is then registered to the local phase bone surfaces using an iterative expectation maximization registration technique. Validation on 96 *in vivo* clinical scans obtained from eight patients resulted in a root mean square registration error of 2 mm (SD: 0.4 mm), which is below the clinically acceptable threshold of 3.5 mm. The improvement achieved in registration accuracy using the new features was also significant ($p < 0.05$) compared to state of the art local phase image processing methods.

Keywords: Ultrasound, local phase, spinal injection, gradient energy tensor, image registration, statistical shape model.

1 Introduction

Lower back pain is a common malady, affecting up to 90% of people at some point in their lifetime [1]. Minimally invasive epidural and facet joint injections have been employed in the treatment of back pain as an alternative to more invasive interventions. Epidural injection of anesthesia is also common in obstetrics, with 50-80% of women in labour electing to receive an epidural injection [2]. Although these injection procedures are common, accessing the lumbar epidural space and the facet joint can still present difficulties [2]. Image-guided systems based on ultrasound (US) have the potential for increasing the precision and

K. Mori et al. (Eds.): MICCAI 2013, Part II, LNCS 8150, pp. 361–368, 2013.

success rate of spinal injections [3]. US imaging is non-ionizing, fast, portable, inexpensive and capable of real-time imaging, but, unfortunately, US images typically contain significant speckle and other artifacts, which complicate image interpretation and automatic processing [4,5,6]. To improve the accuracy of guidance in spine injections US has been registered to pre-procedure data such as computed tomography (CT) or statistical shape models (SSM).

Rasoulian et al. [3] used a point-based registration approach of lumbar vertebrae for registering CT to US. Bone surfaces from two-dimensional (2D) US images were automatically segmented using a dynamic programming approach based on intensity/gradient information [4]. The authors reported a target registration error (TRE) of 2.47 mm for a sheep cadaver study. However, managing the sensitivity of intensity-based US bone segmentation techniques to US artifacts, machine settings and algorithm parameters remains a serious challenge. Also, the dependence of bone appearance on the US beam direction increases the number of false and missed surface segments. This is especially important when imaging the highly curved surfaces of the vertebrae. In other words, the complex vertebral shape produces high variability in the angle of reflection to the US beam, which results in low contrast, blurred and disconnected bone features. To overcome some of these challenges, intensity-invariant, local phase-based image processing methods, based on filtering the US data with 2D/3D Log-Gabor filters, have been proposed [5,6]. Although such methods are intensity invariant, their success highly depends on the optimization of the filter parameters [6]. To improve image guidance without segmentation, several intensity-based CT to US registration methods have been developed[7,8,9]. Gill et al. [7] simulated US images from CT data for intensity-based US registration. The method was validated by registering spine bone images obtained from sheep cadavers. The reported registration accuracy was 1.25 mm. Later, Khallaghi et al. [8] proposed a method where ultrasound images were simulated from SSM and registered to US images. A TRE of 3.48 mm was achieved on a phantom. The results suggested that a clear appearance of the bone surfaces is still a challenge in simulation-based approaches. Winter et al. [9] segmented bone surfaces from CT data using US transducer orientation information and registered these bone surfaces to the US scans obtained from patients. The success was influenced strongly by the quality of the US scans obtained. Datasets in which the bone surface was not clearly visualized were excluded from the registration [9].

In summary, enhancement of US image data and extraction of bone surfaces is still a major challenge affecting the overall success of many US-based, image-guided interventions. This paper addresses the specific challenge of guiding spine injection procedures, but the technique can be applied to other interventions requiring clear bone surfaces. We propose a new fully automatic image enhancement method that is based on the use of a gradient energy tensor *(GET)* filter to construct a new feature enhancement metric, which we call local phase tensor *(LPT)*, for enhancing bone features from US data. The tensor is constructed using an isotropic band-pass filtered image, its gradient and Hessian. We show that the response obtained from this new operator provides improved bone features

by combining the advantages of monogenic scale-space and Gaussian scale-space. In related work, we have recently proposed [10] a novel method for constructing a statistical multi-vertebrae shape+pose model where the model is registered to a CT dataset. The second contribution of the current paper is the registration of the proposed local phase features to the constructed statistical multi-vertebrae model using a Gaussian Mixture Model (GMM) based registration method. We validate the proposed method for improving the accuracy of SSM-US spine image registration of 96 *in vivo* 3D US scans obtained from eight patients. We also provide a comparison with previously proposed local phase-based filtering methods [6,11].

2 Methods

2.1 Gradient Energy Tensor-Based Local Phase Features

Phase-based operators are important tools for invariant processing and feature detection. Local phase information is obtained by convolving images with a pair of band-pass quadrature filters. A common choice for US image processing is the Log-Gabor filter. The output of this convolution is used to construct phase-based descriptors such as phase symmetry (PS) for bone [6] and vein [11] localization, or phase asymmetry (PA) for border enhancement in echocardiography US images [5]. Successful results were obtained, but these descriptors are usually only valid for a single feature type (e.g., a step edge for soft tissue or a ridge edge for bone) and give no or wrong responses at points where the underlying feature model is violated. This is especially problematic in US scans of the spine where different boundary feature types are present due to the complex shape of the vertebrae and ligaments. An improvement can be achieved by moving from scalar feature descriptors to tensor-based ones. For simultaneous estimation of orientation and phase information, the *GET* descriptor has shown better results compared to a structure tensor-based descriptor [12]. The *GET* response is given as:

$$\begin{aligned} GET(US_B(x,y))) &= T_{even} + T_{odd}, \text{ where} \\ T_{even} &= [\boldsymbol{H}US_B(x,y))]\,[\boldsymbol{H}US_B(x,y))]^T, \\ T_{odd} &= -0.5 \times ([\nabla US_B(x,y))]\,[\nabla \Delta US_B(x,y))]^T + \\ &\quad [\nabla \Delta US_B(x,y))]\,[\nabla US_B(x,y))]^T). \end{aligned} \tag{1}$$

The first term T_{even} represents symmetric features whereas the second term T_{odd} represents the asymmetric features. $\boldsymbol{H}$,∇ and Δ denote the Hessian, Laplacian and Gradient operations, respectively, and $US_B(x,y)$ represents the band-pass filtered US image. For band-pass filtering the B-mode US image, we have used a Log-Gabor filter with a transfer function of $G(\omega) = \exp(-log(\omega/\kappa)^2/2\log(\sigma_\omega^2))$, where κ is the centre frequency of the filter and σ_ω is related to the spread of the frequency spectrum in a logarithmic function. This filter is commonly applied in ultrasound image processing [5,6,11] of bone and soft tissue interfaces. The complex double angle orientation representation $\boldsymbol{z}$ and the orientation vector $\boldsymbol{o}$ are

given as $\boldsymbol{z}$=$GET_{11} - GET_{22} + i2GET_{12}$ and $\boldsymbol{o} = (real(\boldsymbol{z}^{0.5}), imag(\boldsymbol{z}^{0.5}))^T$. Here, the the superscript 'T' denotes the transpose operation. The instantaneous phase information is calculated as: $\varphi = arg(s_{even}\sqrt{trace(T_{even})} + is_{odd}\sqrt{trace(T_{odd})})$, where $s_{even} = -sign(\boldsymbol{o}^T[\mathbf{H}US_B(x,y)]\boldsymbol{o})$ and $s_{odd} = -sign(\boldsymbol{o}^T[\nabla US_B(x,y)])$. In monogenic signal analysis the image is modeled as $I(x,y) = A(x,y) \times cos(\theta)$ where *A(x,y)* represent the amplitude and θ represent the phase. Based on this, we present a new phase metric called the local phase tensor *(LPT)* as: $LPT(x,y) = \sqrt{T_{even}^2 + T_{odd}^2} \times cos(\varphi)$. The tensor trace, used during the construction of the instantaneous phase component (φ)of the *LPT*, indicates the local contrast independently of feature type and thus acts as a general boundary indicator providing improved local phase vertebrae features.

2.2 Statistical Shape Model Generation and Registration

For construction of the SSM, pose statistics are separated from the shape statistics since they are not necessarily correlated and do not belong to the same space. Pose, which is represented by a similarity (rigid+scale) transformation between the objects, form a Lie group which is a differentiable manifold and thus linear analysis is not applicable. However, analogous to principal components in the Euclidean space, Principal Geodesics (PG), are defined for Lie groups which provide the modes of variations for a set of similarity transformations [10]. The approximations of PGs are as follows: initially the transformations T_i are transformed to a tangent space at the average transformation, μ, by logarithmic mapping: $log_\mu T_i$. Here i represent the number of set of elements. Next Principal Component Analysis is applied and lastly principal components ν_l are transformed back to the original space by exponential mapping: $exp_\mu(\nu_l)$. The same approach is applied to shapes with the assumption that shapes form a Lie group as well [10]. Finding the pose and shape variations of a complex among a training set, a new instance of the model is generated by applying weights to pose and shape PGs: $s_l = \Phi(\omega^s, \omega^p) = \Phi_l^p(\Phi_l^s(\omega^s); \omega^p)$. Here $\Phi_l^p(.;\omega^p)$ and $\Phi_l^s(.)$ denote the similarity transformation and a shape, respectively, which are built by a combination of the pose and shape PGs with corresponding weights [10]. The model is then registered to an US image by finding proper weights. The registration is performed using a GMM-based technique proposed earlier [13]. In this iterative technique the previously generated model boundary points are defined as the centroids of the GMM. The target (local phase enhanced bone surfaces) is assumed to be an observation (data points) generated by the GMM. The registration is solved by an expectation maximization algorithm where the centroids of the model point set are transformed by a set of transformation parameters to fit onto the target. The transformation is based on the parameters obtained from the shape and pose variations of the model [13]. In each iteration, the mutual distance between the vertices of the statistical model and the points extracted from US, weighted by the output of local phase method, is optimized. Since all volumes are centred at inter-vertebral levels, and are 5 to 10 mm lateral to midline, an approximate location of the vertebrae is known *a priori*. This information is used to initialize the model during the registration. Note that the registration is performed based

on points of the model that are visible in US volumes, i.e. laminae, articular processes and transverse processes.

2.3 Data Acquisition and Experiments

After obtaining informed consent, 3D US volumes were captured from 8 patients by an expert sonographer using a Sonix Touch US machine (Ultrasonix, Medical Corp, Richmond, BC, Canada) with a curvilinear 3D transducer (4D C7-3/40), operating at 3.3 MHz with depth of 7 cm. 80 frames were captured for each volume over a 60^0 field of view. US volumes were acquired in the prone position. For each subject, the intervertebral levels were found by the sonographer and were marked on the skin. Four intervertebral levels (L1-L2, L2-L3, L3-L4, and L4-L5) were scanned. Twelve 3D US volumes were acquired for each patient. Bone surfaces were then marked on the US images manually by the expert sonographer to provide a gold standard for the registration. The manually selected anatomical landmarks were chosen from the spinous process, superior and inferior articular process of vertebra. Corresponding landmarks were also identified from the generated shape model. These landmarks were only used during validation and were not part of the registration. In order to evaluate the efficacy of the proposed US filtering methodology, we performed two experiments.

Experiment 1: For qualitative evaluation of the proposed LPT-based bone enhancement method, samples of processed *in vivo* 3D US scans were annotated. Furthermore, we also present qualitative results by graphically displaying examples of the registered surfaces.

Experiment 2: For quantitative evaluation we present registration results between the 96 *in vivo* US scans and the generated shape model. The US images were processed with either the proposed method or with the previously propsed PS approaches [6,11] before inputting them to the registration framework. Root mean square (RMS) distance error was calculated between the manually selected landmarks from the US scans and the statistical shape model after the registration. The proposed filtering, SSM generation and registration approaches were implemented in MATLAB and run on a 2.67 GHz Intel(R) $Core^{TM}$ i5 with 8 GB of RAM. The Log-Gabor filter was constructed using a filter centre frequency of $\kappa = 25$ pixels and frequency bandwidth value of $\sigma_\omega = 0.25$. The standard deviation for calculating the Hessian, Gradient and Laplacian was $\delta = 0.2$ pixels. These values were not changed throughout the experimental validation.

3 Results

Experiment 1: Figure 1 shows annotated examples of results obtained *in vivo*. The yellow arrows point to the blurred or low contrast lamina surfaces due to the previously mentioned difficulties encountered while imaging vertebrae. In contrast, the proposed method results in sharper and higher contrast lamina bone surfaces which could be seen for the bone surface locations corresponding to the yellow arrows. Investigating the last image in the fourth row, we can

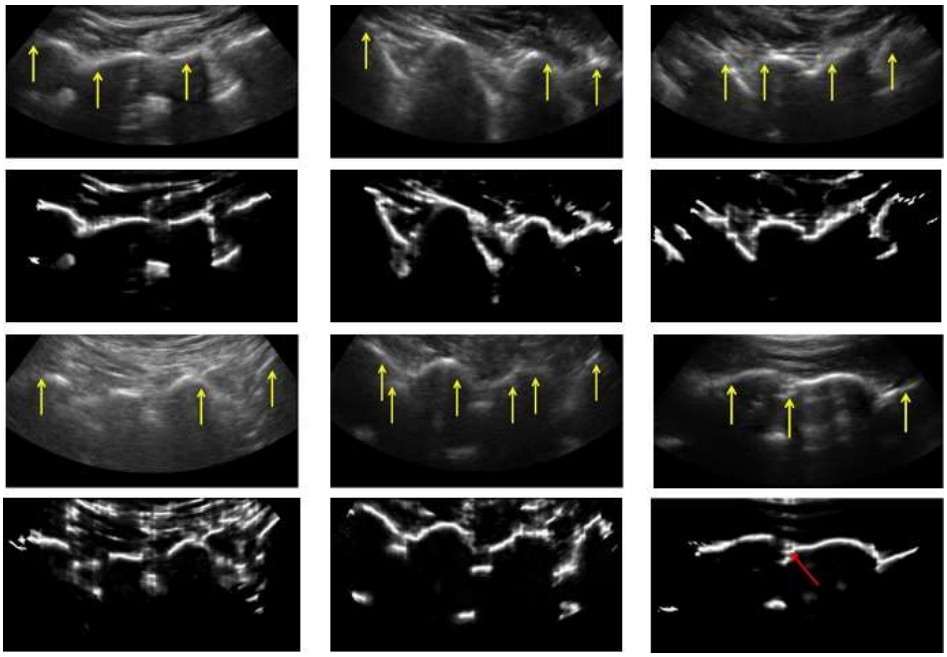

Fig. 1. Qualitative results obtained from clinical scans. First and third row: B-mode US images. Second and fourth row: Corresponding *(LPT)* images. Yellow arrows point to the locations where bone surface are low contrast and appear blurred. Bone surfaces appear sharper in LPT images. Red arrow in last row points to the enhanced LF.

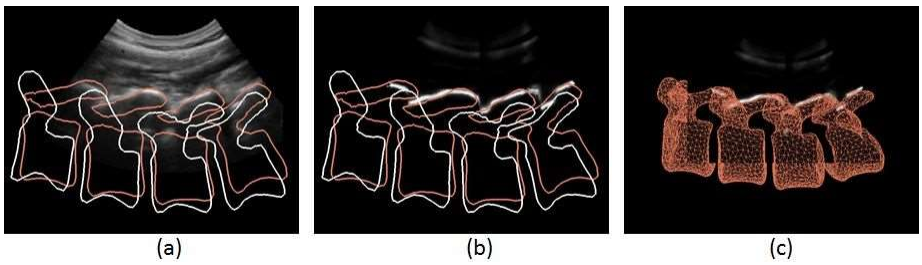

Fig. 2. Qualitative results for SSM to US registration. The initial position of the model (prior registration) is drawn in white and the registered model is in red. (a) Overlay of *in vivo* B-mode US slice with the outline extracted from the SSM. (b) Overlay of the LPT features with the outlines extracted from the SSM. (c) Overlay of theLPT features with the 3D SSM after registration.

also see the achieved enhancement for the LF compared to the corresponding B-mode image. The method appears unaffected by intensity variations present in different US scans. In Figure 2, we present results on registering SSM to the extracted surfaces. Investigating the registration result in the figure, we can see that the proposed method successfully aligns the two volumes.

Experiment 2: The processing time was 0.05 s for a 2D US image compared to 6 s for phase symmetry method [6]. Registration was achieved in 39 s. The registration results obtained with the proposed method were significantly better than the previously proposed PS methods [6,11] (paired t-test $p<0.05$). The

Table 1. Results of SSM to US registration using the proposed image enhancement, and previous local phase filtering approaches [6,11]

	Proposed Method		Monogenic Filter [11]		Phase Symmetry [6]	
	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)
L1-L2	2.0	0.5	4.6	3.4	12.3	9.2
L2-L3	2.3	0.7	5.9	4.5	14.4	7.6
L3-L4	1.7	0.3	1.8	0.5	3.7	2.4
L4-L5	2.0	1.1	4.7	5.5	14	11
Total	2.0	0.4	4.3	4.2	11.1	9.1

overall mean RMS error value calculated from the selected anatomical landmarks was 2.0 mm (SD 0.4 mm) for all the 96 scans with a maximum mean value of 2.3 mm for registering L2-L3 and minimum mean value of 1.7 mm for registering L3-L4. The previously proposed PS features [6] resulted in an overall mean RMS value of 11.1 mm (SD 9.1 mm). The Monogenic-filter approach [11] obtained a mean RMS error value of 4.3 mm (SD 4.2 mm)(Table 1).

4 Discussions and Conclusions

Obtaining a clear depiction of bone surfaces in US images continues to be a challenging problem affecting the success of many US guided interventions. Local phase image processing based on Log-Gabor filtering for bone enhancement has been reported in the past with successful results [6]. However, the phase metrics constructed previously are usually only valid for a single feature type and give no or wrong responses at points where the underlying feature model is violated. Furthermore, one is faced with the complexity of constructing oriented filter banks which makes the success of the proposed methods dependent on the selected orientations. Finally, orienting the constructed filter to cover the whole frequency spectrum increases the computational time of these methods as well. On the other hand, monogenic filter based approaches [11] overcome the filter bank construction problem by being isotropic filters, however, the descriptors constructed using this filter are again only valid for a single feature type and face similar problems as the Log-Gabor filter. In this paper, we provide solutions to these problems by proposing a new approach for enhancing spine bone surfaces from 3D US data. The proposed method is based on the use of the gradient energy tensor filter for constructing a novel phase metric. The tensor is constructed using an isotropic band-pass filtered image, its gradient and Hessian. Moving from scalar based descriptors to tensor-based ones allows the extraction of multiple feature types which provides improved results in the enhancement of spine surfaces where different boundary feature types are present due to the complex shape of the vertebrae and ligaments. The enhanced bone features are then registered to a SSM constructed using a novel method where pose statistics are separated from the shape statistics. The target area for a facet joint injection is approximately 1.5 cm in diameter. Given the size of the target, an accuracy of 3.5 mm should be sufficient for the registration approach to be acceptable in clinic. The proposed method enhances the blurred and low contrast vertebrae

surfaces from *in vivo* scans and achieves registration errors below the required clinical limit. We obtain significant improvements ($p < 0.05$) in accuracy compared to previous local phase methods. We also achieve improved registration results compared to state-of-the-art simulation based SSM-US registration approaches [8]. Future work will include the validation of the method on more clinical scans and different registration approaches for guiding spine injections.

Acknowledgments. This study is is jointly funded by grants from the NSERC, CIHR and the Institute for Computing,Information and Cognitive Systems(ICICS) at UBC.

References

1. Frymoyer, J.D.: Back pain and sciatica. N. England J. Med. 318, 291–300 (1988)
2. Osterman, M.J.K., Martin, J.A.: Epidural and spinal anesthesia use during labor: 27-state reporting area, 2008. Nat. Vital Statistics Reports, vol. 59(5). NCHS, Hyattsville (2011)
3. Rasoulian, A., Abolmaesumi, P., Mousavi, P.: Feature-based multi-body rigid registration of CT and US images of lumbar spine. Med. Phys. 39(6), 3154–3167 (2012)
4. Foroughi, P., Boctor, E., Swatrz, M.J., Taylor, R.H., Fichtinger, G.: US bone segmentation using dynamic programming. In: IEEE Ultrasonics Symp., pp. 2523–2252 (2007)
5. Mulet-Parada, M., Noble, J.A.: 2D+T boundary detection in echocardiography. Med. Im. Anal. 4(1), 21–30 (2000)
6. Hacihaliloglu, I., Abugharbieh, R., Hodgson, A., Rohling, R.N.: Automatic adaptive parameterization in local phase feature-based bone segmentation in US. Ultr. in Med. Biol. 37(10), 1689–1703 (2011)
7. Gill, S., Abolmaesumi, P., Fichtinger, G., Boisvert, J., Pichora, D., Borshneck, D., Mousavi, P.: Biomechanically constrained groupwise US to CT registration of the lumbar spine. Med. Im. Anal. 16(3), 662–674 (2010)
8. Khallaghi, S., Mousavi, P., Borschneck, D., Fichtinger, G., Abolmaesumi, P.: Biomechanically constrained groupwise statistical shape model to ultrasound registration of the lumbar spine. In: Taylor, R.H., Yang, G.-Z. (eds.) IPCAI 2011. LNCS, vol. 6689, pp. 47–54. Springer, Heidelberg (2011)
9. Winter, S., Pechlivanis, I., Dekomien, C., Igel, C., Schmieder, K.: Toward registration of 3D ultrasound and CT images of the spine in clinical praxis: design and evaluation of a data acquisition protocol. Ultr. in Med. Biol. 35(11), 1773–1782 (2009)
10. Rasoulian, A., Rohling, R.N., Abolmaesumi, P.: A statistical multi vertebrae shape+pose model for segmentation of CT images. SPIE Medical Imaging, 86710–86716 (2013)
11. Rahmatullah, B., Papageorghiou, A.T., Noble, J.A.: Integration of local and global features for anatomical object detection in US. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 402–409. Springer, Heidelberg (2012)
12. Felsberg, M.: The GET operator. LiTH-ISY-R, No. 2633 (2004)
13. Rasoulian, A., Rohling, R.N., Abolmaesumi, P.: Group-wise registration of point sets for statistical shape models. IEEE Trans. Med. Imag. 31(11), 2025–2034 (2012)