GPU Accelerated Registration of a Statistical Shape Model of the Lumbar Spine to 3D Ultrasound Images

Siavash Khallaghi[a], Purang Abolmaesumi[b], Ren Hui Gong[c], Elvis Chen[c], Sean Gill[c], Jonathan Boisvert[e], David Pichora[d], Dan Borschneck[d], Gabor Fichtinger[c], Parvin Mousavi[c]

[a] Dept. of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada;
[b] Dept. of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada;
[c] School of Computing, Queen's University, Kingston, ON, Canada
[d] Kingston General Hospital, Kingston, ON, Canada
[e] National Research Council, Ottawa, ON, Canada

ABSTRACT

We present a parallel implementation of a statistical shape model registration to 3D ultrasound images of the lumbar vertebrae (L2-L4). Covariance Matrix Adaptation Evolution Strategy optimization technique, along with Linear Correlation of Linear Combination similarity metric have been used, to improve the robustness and capture range of the registration approach. Instantiation and ultrasound simulation have been implemented on a graphics processing unit for a faster registration. Phantom studies show a mean target registration error of 3.2 mm, while 80% of all the cases yield target registration error of below 3.5 mm.

Keywords: Graphics Processing Unit, Statistical Shape Model, Ultrasound, Registration, Spine

1. INTRODUCTION

Spinal needle injections, such as epidurals and facet joint injections, are performed on a regular basis in radiology clinics and hospitals. Traditionally these percutaneous procedures are performed under fluoro or CT guidance. This exposes both the patient and the physician to hazardous radiation. In order to remove the ionizing radiation, many papers suggest the registration of Statistical Shape Models (SSM) with intra-operative ultrasound (US) images.[1–3]

A statistical shape model is a mathematical entity, which captures the statistical information, including the mean and variations, of a group of objects. It not only represents the objects within the population of the constructed SSM, but can also produce new instances of the objects from a linear combination of the principal modes of variations. Existing SSMs can be divided into two categories: geometrical SSMs,[4–6] which describe the outline of images, and volumetric SSMs,[1,7,8] which model both the internal intensity and geometrical information. While geometrical SSMs are computationally less expensive, they are prone to bone surface segmentation errors.

Prior work proposes several approaches for SSM to CT,[9] SSM to MRI[10] and SSM to US registration.[1] The feasibility of US registration to SSMs of bony anatomy has been previously investigated.[3,11] Barratt *et al.*[1] and Foroughi *et al.*[2] built SSMs for the femur and pelvis, which are subsequently repositioned and deformed to fit a cloud of bone surface points extracted from a set of tracked US images. Schumann *et al.*[12] further improved their technique to compensate for variations in the speed of sound between calibration and intra-operative use.

Typically the most time-consuming components of a non-rigid registration algorithm are the transformation of the moving image and the calculation of the similarity metric. To alleviate this problem, several techniques base on parallel implementation of the registration technique have been proposed. Depending on the application, either of these blocks may be the bottleneck for the registration algorithm. Crookes *et al.*[13] implements an affine registration on the graphics processing unit. Ozcelik *et al.*[14] presents a parallel implementation of Thirion's Demons algorithm by applying four consecutive CUDA kernels: gradient, deformation, displacement smoothing and correlation. Rehman *et al.*[15] is another example of GPU-based deformable registration.

Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling,
edited by Kenneth H. Wong, David R. Holmes III, Proc. of SPIE Vol. 7964, 79642W
© 2011 SPIE · CCC code: 1605-7422/11/$18 · doi: 10.1117/12.878377

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

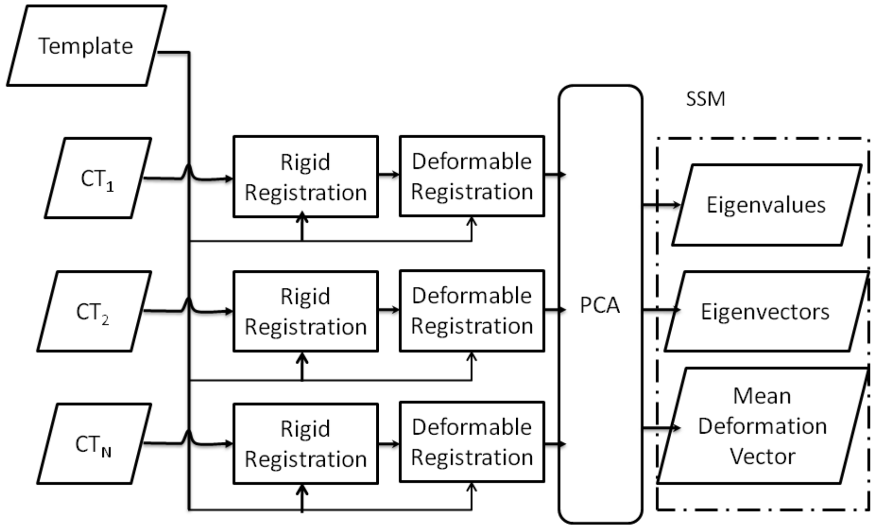

Figure 1. Outline of the SSM construction method from a set of CT images for an individual vertebra. The SSM of the ensemble of vertebra is created from three similar workflows for the L2, L3 and L4 respectively.

Multi-modality registration between CT and US images, due to to the different nature of input images, tends to be a time-consuming and complex task. Hence, the use of GPUs to accelerate the registration process is a natural choice. Kutter *et al.*[16] presents a modular design for the simulation of US images from CT data on the GPU. They further extend the approach to a multi-modality registration technique between US and CT data using a variation of the Linear Correlation of Linear Combination (LC2) similarity metric, originally proposed by Wein *et al.*[17] These works demonstrate a significant improvement in the registration runtime using the parallel implementation on the GPU, making it an affordable and available alternative for real-time applications.

In this paper we present a parallel implementation of an SSM to US registration method we proposed before,[3] for the alignment of an SSM of lumbar spine to 3D US data. Using the deformable B-spline transform, a statistical shape model for each of the L2-L4 vertebrae is constructed. For a faster registration, instantiation and ultrasound simulation have been implemented on the GPU. The registration is validated on three tissue mimicking phantoms with realistic spinal curvature.

2. METHODOLOGY

2.1 CT Data Collection

The CT data were collected at a local hospital under the approval of the Research Ethics Board. A set of CT images, acquired from 38 patients (19 male and 19 female), was used where the L2, L3 and L4 vertebrae were semi-automatically segmented using ITK-Snap and resampled into $120 \times 200 \times 100$ voxels with an isotropic spacing of 0.6 mm. The patient data was divided into two groups: 35 for constructing the SSM (hereafter referred to as training data), and three for validation (two male and one female).

2.2 Statistical Shape Model Construction

A SSM is contructed for each vertebra seperately as shown in Figure 1. For each SSM, a random patient CT volume is initially chosen as the template, I_{t}. Each training example, I_k , is registered to the template by a rigid registration followed by a B-spline deformable registration, such that $I_{\mathrm{t}} \approx T^k_{\mathrm{def}}(T^k_{\mathrm{rigid}}(I_k))$, where *T(.)* denotes a transformation. B-spline registration is performed in a $40 \times 30 \times 30$ grid using the mutual information metric.[7] To reduce the deformable registration time, a three stage multi-resolution approach is implemented. With the deformable transformation of all the training examples known with respect to the template, principal component analysis (PCA) is performed to construct the SSM.

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

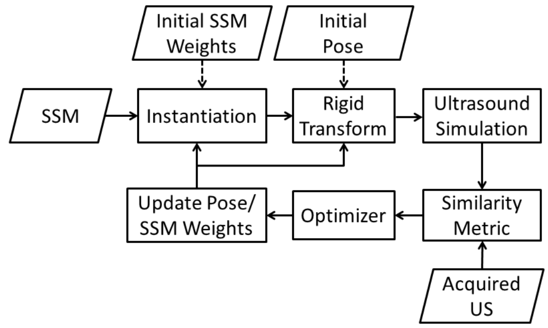

Figure 2. Registration workflow

After the SSM is constructed, it is used to generate new instances of the population. A new instance of the SSM, defined by the deformation vector D_{new} is produced by a linear combination of the mean deformation vector, $\bar{\phi}$, SSM weights, w_i, and the eigenvectors of the covariance matrix generated from all the deformation fields, v_i:

$$D_{new} = \bar{\phi} + \sum_{i=1}^{N} w_i v_i. \tag{1}$$

The SSM weights in equation (1) provide a compact description of the transformation needed to deform the mean shape into the shape of the new instance. In our case the first 12 eigenvectors cover 95% of variations in shape.

2.3 Registration Framework

This section describes the SSM to US registration method. The registration problem is solved in two steps: first rigid and then deformable. At the rigid stage, starting from six random initial rigid parameters (three for rotatation and three for translation, i.e. the pose) and zero SSM weights, the algorithm solves for the optimal rigid parameters. Subsequently, at the deformable stage, the algorithm finds 12 optimal deformable parameters (i.e. the SSM weights). Using the updated rigid and deformable parameters, the algorithm iterates through the rigid and deformable phases until convergence is achieved. For optimization, Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is used because it yields convergence in an irregular search space.[18] Decoupling the problem into rigid and deformable registration phases, not only decreases the complexity of the problem, but also allows for a faster convergence by decreasing the population size of the CMA-ES optimizer.

Figure 2 illustrates the general registration workflow. First, we set the initial SSM weights to zero, and generate an instance of each of the vertebra. Then, using a random initial rigid transformation, each generated instance is perturbed to an initial pose. Following that, a ray-casting based ultrasound simulation is applied to the reconstructed volume. Assuming that the Hounsfield units can be related to the acoustic impedance values used to calculate ultrasound transmission and reflection, each ultrasound beam is modeled as a ray passing down the columns of the image. Next, the LC2[17] metric is calculated. This similarity measure is fed to the CMA-ES optimizer until all the rigid parameters converge. A similar approach is employed for the deformable registration phase. The registration iterates between deformable and rigid phases until the correct pose and shape of the SSM are found.

Running the proposed registration algorithm on a desktop personal computer, will take approximately a day for a single registration. To speed the algorithm up, we have implemented a parallel version of our registration algorithm on an nVidia GTX 285 Graphical Processing Unit (GPU), details of which has been presented in this work.

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

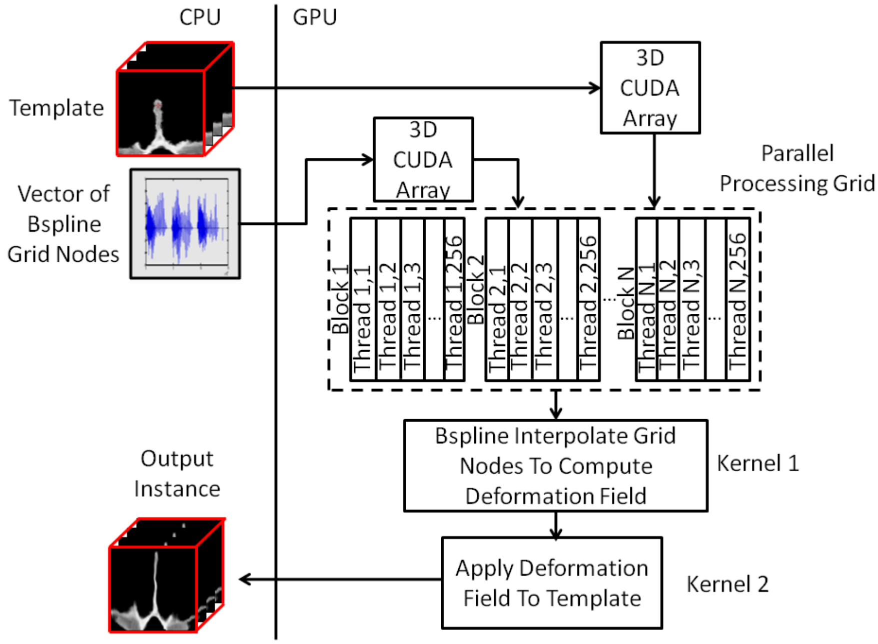

Figure 3. Parallel implementation of the instantiation block on the GPU.

2.3.1 GPU Accelerated Instance Generation

Figure 2 illustrates the parallel instance generation algorithm. Given the SSM weights, a linear combination of eigenvectors is set as the B-spline grid node vectors. The SSM template and grid node vectors are uploaded from the global memory to the device memory and stored in a CUDA-3D-arrays. A 3D-texture is also bound to each array, enabling the algorithm to read and manipulate data in the 3D-arrays. We have implemented the instantiation in two steps. The first kernel, interpolates the B-spline grid nodes to produce the deformation field at each voxel. This is done by placing a 2D grid that spans each slice and then going through the slices one at a time to produce the deformation vector. Subsequently, the second kernel transforms each voxel from the index representation to the physical coordinate system, applies the deformation vector and transforms it back to the index representation. In the case of unassigned voxels in the final volume, the empty voxel is filled with a default value. Finally the deformed volume is downloaded to global memory. Table 2 illustrates timing tests for instance generation on the CPU and the GPU.

2.3.2 GPU Accelerated Ultrasound Simulation

The ultrasound simulation is based on the work presented by Wein *et al.*[17] and Kutter *et al.*[16] Figure 4 illustrates the parallel implementation of ultrasound simulation on the GPU. First, the input CT volume is uploaded to the device memory and stored as a CUDA array. Then, a 2D grid is placed on top of the CT volume parallel to the coronal plane. the ultrasound is simulated in two steps: The first CUDA kernel computes the ultrasound reflection image, each ultrasound ray is assigned one CUDA-thread, traveling down the CT volume. The second kernel, computes the mapped CT image. The two images are fused together to produce the simulated ultrasound. The result of the ultrasound simulation is then downloaded to the global memory.

3. RESULTS AND DISCUSSION

3.1 Ultrasound Data Acquisition

As mentioned in subsection 2.1, we excluded three CT volumes from the patient data for validation. The three excluded CT volumes were used to construct 3D models of the entire lumbar spine, including L1 to L5. These

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

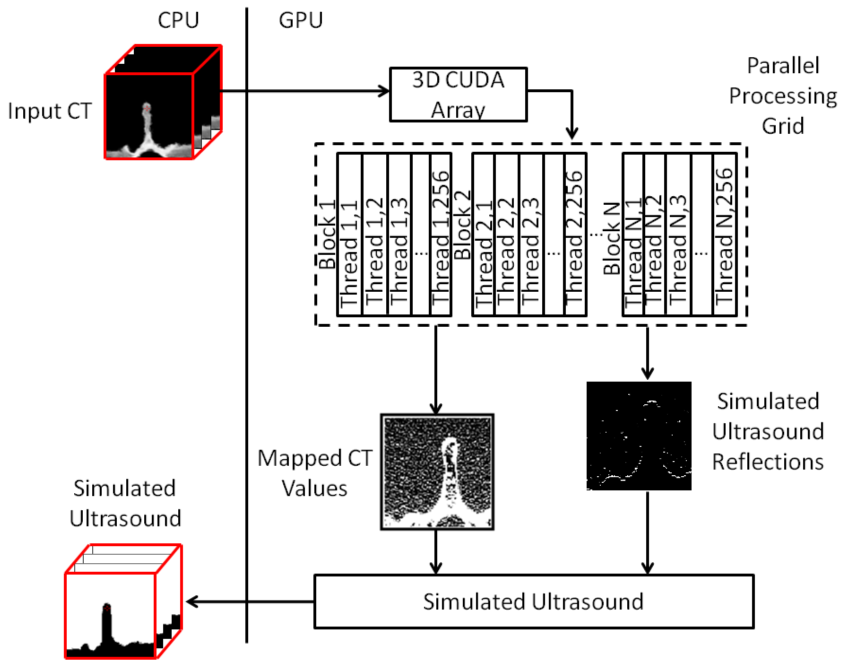

Figure 4. Parallel implementation of US simulation on the GPU.

virtual models were printed using a Cimetrix 3D shape printer (Cimetrix Solutions, Oshawa, ON, Canada). Three spine phantoms were constructed by submerging the printed vertebrae in an agar-gelatine-based gel. The resulting phantoms simulated the appearance of vertebrae and soft tissue in ultrasound. A high-resolution CT image ($0.46 \times 0.46 \times 0.625$ mm) and an ultrasound volume were acquired from each phantom from a freehand sweep with a linear-array transducer at 6.6 MHz with an imaging depth of 5.5 cm. Nine fiducial markers, mounted on the exterior of the phantom box, were localized with a tracked pointer and transformed to the US space. The CT and ultrasound volumes were aligned using these markers.

3.2 Results

The SSM mean shape and the ultrasound volumes were brought to an initial position by rigidly registering the mean shape to the corresponding phantom CT volume. For each phantom, thirty experiments were performed with perturbing the mean shape using a transformation generated from a uniform random distribution in the interval of [0,10] mm translation along each axis and [0-10]$^\circ$ rotation about each axis.

To evaluate the accuracy of the registration, an expert orthopedic surgeon was asked to identify five corresponding landmarks, three on the spinous process and two on the facet joints, on the registered SSM, the ultrasound volume and the corresponding CT. The average distance of these five landmarks is chosen as a measure of the final Target Registration Error (TRE). A registration is considered failed if the final TRE is more than 3.5 mm, as the clinically accepted error. Registration results are shown in Table 1 and an example of the initial misalignment and the registration result is illustrated in Figure 5. The registration results are sufficient for various spinal interventions, such as facet joint injections and epidural anesthesia, which are the primary focus of our research.

Results in Table 2 show a 350 speed up gain for the GPU implementation of the instance generation in comparison to the CPU implementation. The speed up gain is in agreement with the results presented by Gong *et al.*[19] Also by simulating the US on the device, the entire registration algorithm was implemented on the GPU which allowed for minimal interaction between the CPU and the GPU. This allowed the registration run-time to decrease from the order 8 hours to 20 minutes, a 24x improvement.

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

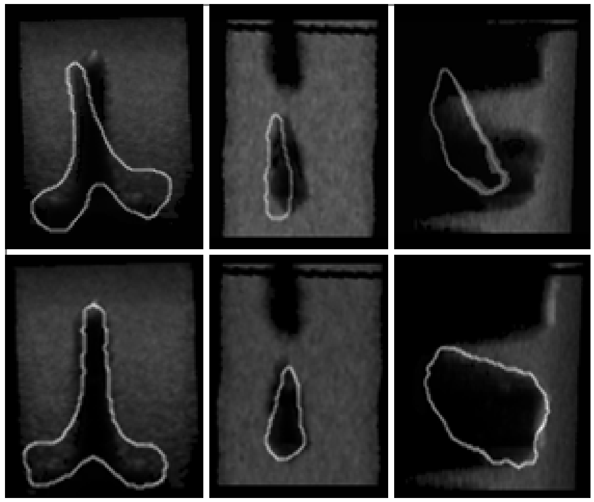

Figure 5. Transverse (left), coronal (center), and sagittal (right) slices of the original US volume overlaid with the bone contours of the misaligned (top) and registered (bottom) SSM volumes.

Table 1. Registration results for the SSM to US registration. SR (Success Rate) is defined as the ratio of the registrations where the overall TRE is less than 3.5 mm. SR is presented for each phantom with the maximum initial misalignment of 10 mm.

	L2			L3			L4		
Phantom	Mean	STD	SR	Mean	STD	SR	Mean	STD	SR
1	3.10	0.29	85%	3.38	0.42	81%	3.22	0.33	82%
2	3.05	0.43	80%	3.48	0.33	79%	3.41	0.45	78%
3	3.23	0.36	75%	3.25	0.45	82%	3.24	0.26	85%

Several factors limit the success rate of the registration. Due to partial volume occlusion, it is impossible to visualize the superior articular process in ultrasound images. Also, deforming the SSM to capture the sharp corners of the spinous and transverse processes is a challenging task. While the former is intrinsic to ultrasound imaging, the latter can be substantially improved by increasing the number of training images and using a denser grid in the construction of the SSM. With the deformation implemented on the GPU, denser B-spline grids are more accessible for real-time applications.

Table 2. CPU vs. GPU instantiation run-time and the performance gain computed as CPU/GPU. A total of 100 speed tests were performed on the CPU and the GPU. In each test a random combination of SSM weights was passed to the instantiation block.

SSM	Average CPU Time sec	Average GPU Time msec	Gain
L2	2.3	8.4	273
L3	2.4	7.2	342
L4	2.5	8.2	308

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

4. CONCLUSIONS

In this paper, we present a novel parallel implementation of the SSM to ultrasound registration of the lumbar vertebrae (L2-L4), which brings our previous report[3] closer to clinical applicability. To the best of our knowledge, this is the first report of a parallel implementation of instantiation of a volumetric SSM of the spinal vertebrae, which retains both the intensity and the geometry information. In addition, ray-casting is inherently parallel, making it possible to further decrease the registration time using GPU implementation. Data-intensive parts of the registration algorithm have been implemented on the graphics processing unit, and timing tests have been run to verify a high computational gain on the GPU vs. CPU. The high accuracy of the proposed GPU-accelerated registration technique makes it an ideal choice for various spinal intervention approaches, such as facet joint injections and spinal epidurals.

REFERENCES

[1] Barratt, D. C., Chan, C. S. K., Edwards, P. J., Penney, G. P., Slomczykowski, M., Carter, T. J., and Hawkes, D. J., "Instantiation and Registration of Statistical Shape Models of the Femur and Pelvis Using 3D Ultrasound Imaging," *Medical Image Analysis* **12**(3), 358–374 (2008).

[2] Foroughi, P., Song, D., Chintapani, G., Taylor, R. H., and Fichtinger, G., "Localization of Pelvic Anatomical Coordinate System Using US/Atlas Registration for Total Hip Replacement," in [*Medical Image Computing and Computer-Assisted Intervention*], 871–878 (2008).

[3] Khallaghi, S., Mousavi, P., Gong, R. H., Gill, S., Boisvert, J., Fichtinger, G., Pichora, D., Borschneck, D., and Abolmaesumi, P., "Registration of a Statistical Shape Model of the Lumbar Spine to 3D Ultrasound Images," in [*Medical Image Computing and Computer Assisted Interventions*], 68–75 (2010).

[4] Hu, Y., Ahmed, H. U., Allen, C., Pendse, D., Sahu, M., Emberton, M., Hawkes, D., and Barratt, D., "MR to Ultrasound Image Registration for Guiding Prostate Biopsy and Interventions," in [*Medical Image Computing and Computer-Assisted Intervention*], 787–794 (2009).

[5] Michopoulo, S. K., Costaridou, L., Panagiotopulous, E., Speller, R., Panayiotakis, G., and Todd-Pokropek, A., "Atlas-Based Segmentation of Degenerated Lumbar Interverbal Discs from MR Images of the Spine," *IEEE Transactions on Biomedical Engineering* **56**(9), 2225–2231 (2009).

[6] Sadowsky, O., Cohen, J. D., and Taylor, R. H., "Rendering Tetrahedral Meshes with Higher-Order Attenuation Functions for Digital Radiograph Reconstruction," in [*IEEE Visualization*], 303–310 (2005).

[7] Gong, R. H., Stewart, J., and Abolmaesumi, P., "Reduction of Multi-Fragment Fractures of the Distal Radius Using Atlas-based 2D/3D Registration," in [*Proceedings of the SPIE, Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*], **7261**, 726137–1 – 726137–9 (2009).

[8] Jurcak, V., Fripp, J., Engstrom, C., Walker, D., Salvado, O., Ourselin, S., and Crozier, S., "Atlas Based Automated Segmentation of the Quadratus Lumborum Muscle Using Non-Rigid Registration on Magnetic Resonance Images of the Thoracolumbar Region," in [*5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*], 113–116 (2008).

[9] Tang, T. S. and Ellis, R. E., "2D/3D Deformable Registration Using a Hybrid Atlas," in [*Medical Image Computing and Computer-Assisted Intervention*], 223–230 (2005).

[10] Leemput, K. V., "Encoding Probabilistic Brain Atlases Using Bayesian Inference," *IEEE Transactions on Medical Imaging* **28**(6), 822–837 (2009).

[11] Talib, H., Rajamani, K., Kowal, J., Nolte, L.-P., Styner, M., and Ballester, M. A. G., "A Comparison Study Assessing the Fesibility of Ultrasound-Initialized Deformable Bone Models," *Computer Aided Surgery* **10**(5/6), 293–299 (2005).

[12] Schumann, S., Puls, M., Ecker, T., Schwaegli, T., Stifter, J., Siebenrock, K.-A., and Zheng, G., "Determination of Pelvic Orientation from Ultrasound Images Using Patch-SSMs and a Hierarchical Speed of Sound Compensation Strategy," in [*Information Processing in Computer-Assisted Interventions*], Navab, N. and Jannin, P., eds., *Lecture Notes in Computer Science* **6135**, 157–167, Springer Berlin / Heidelberg (2010).

[13] Crookes, D., Boyle, K., Miller, P., and Gillan, C., "GPU Implementation of the Affine Transform for 3D Image Registration," in [*13th International Conference on Machine Vision and Image Processing*], 151–155 (2009).

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms

[14] Muyan-Ozcelik, P., Owens, J., Xia, J., and Samant, S., "Fast Deformable Registration on the GPU: A CUDA Implementation of Demons," in [*International Conference on Computational Sciences and Its Applications*], 223 –233 (june 2008).

[15] ur Rehman, T., Haber, E., Pryor, G., Melonakos, J., and Tannenbauma, A., "3D Nonrigid Registration via Optimal Mass Transport on the GPU," *Medical Image Analysis* **13**(6), 931–940 (2009).

[16] Kutter, O., Shams, R., and Navaba, N., "Visualization and GPU-accelerated Simulation of Medical Ultrasound from CT Images," *Computer Methods and Programs in Biomedicine* **94**(3), 250–266 (2009).

[17] Wein, W., Brunke, S., Khamene, A., Callstrom, M. R., and Navab, N., "Automatic CT-Ultrasound Registration for Diagnostic Imaging and Image-guided Intervention," *Medical Image Analysis* **12**(5), 577–585 (2008).

[18] Winter, S., Brendel, B., Pechlivanis, I., Schmieder, K., and Igel, C., "Registration of CT and Intraoperative 3-D Ultrasound Images of the Spine Using Evolutionary and Gradient-Based Methods," *IEEE Transactions on Evolutionary Computation* **12**(3), 284–296 (2008).

[19] Gong, R. H., Stewart, J., and Abolmaesumi, P., "A New Representation of Intensity Atlas for GPU-accelerated Instance Generation," in [*Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*], 4399 –4402 (2010).

Downloaded from SPIE Digital Library on 16 Aug 2011 to 137.82.117.28. Terms of Use: http://spiedl.org/terms