

# Automatic Registration and Fusion of Ultrasound with CT for Radiotherapy

Wolfgang Wein<sup>1</sup>, Barbara Röper<sup>2</sup>, and Nassir Navab<sup>1</sup>

<sup>1</sup> Computer Aided Medical Procedures (CAMP) Group, TU Munich  
Boltzmannstr. 3, 85748 Garching, Germany  
{wein,navab}@cs.tum.edu

<sup>2</sup> Clinic and Policlinic of Radiation Oncology,  
Klinikum Rechts der Isar, TU Munich,  
Ismaninger Str. 22, 81675 Munich, Germany  
barbara.roeper@lrz.tum.de

**Abstract.** We present a framework for rigid registration of a set of B-mode ultrasound images to a CT scan in the context of Radiotherapy planning. Our main focus is on deriving an appropriate similarity measure based on the physical properties and artifacts of ultrasound. A combination of a weighted Mutual Information term, edge correlation, clamping to the skin surface and occlusion detection is able to assess the alignment of structures in ultrasound images and simulated slices generated from the CT data. Hence a set of ultrasound images, whose relative transformations are given by a magnetic tracking device, can be registered automatically to the CT scan. We validated our methods on neck data of patients with head and neck tumors and cervical lymph node metastases.

## 1 Introduction

*Overview.* Registration of ultrasound images to three-dimensional tomographic modalities such as CT and MRI is receiving a lot of attention in the past few years. On one hand, many intra-operative procedures, especially in neurology and orthopedics, can be guided with ultrasound while integrating pre-operative information from CT/MRI. On the other hand, data fusion for diagnosis and treatment planning can improve the outcome as well.

In the particular application of radiation treatment planning for inoperable head and neck cancer the identification of metastatic neck lymph nodes is mandatory for the correct target volume delineation. This can be achieved with a reported accuracy of 80-95% using high-frequency ultrasound [1]. However, the target volume definition is done on individual slices of a planning CT scan. In direct comparison with ultrasonography, diagnostic CT was equally predictive in revealing lymph node size, but performed worse in depicting internal nodal architecture, leading to a lower sensitivity and specificity than ultrasonography [2]. As in planning CTs for radiotherapy contrast medium is usually omitted, their diagnostic properties are particularly poor. Therefore, transferring the diagnostic information from ultrasound onto the CT data could yield a more precise

treatment.

In general, registration of multimodal data is especially desirable if they provide complementary information. At the same time, this complementary nature hampers image-based registration algorithms, which try to align structures present in both modalities. In our work, we will try to overcome some of these problems for ultrasound-CT registration.

**Related Work.** Due to the very different characteristics of ultrasound imaging with respect to CT / MRT, a lot of research has been carried out on using features extracted from the ultrasound images, in order to align them with corresponding structures in other modalities. Possible anatomical features comprise vessels [3, 4], bone surfaces [5], organ surfaces [6]. Pure intensity-based registration has been performed mainly for 3D ultrasonic data. Roche et al. [7] use an adapted correlation ratio similarity measure in order to register the ultrasonic data simultaneously to both the intensity and the gradient information of a MRI scan. A registration involving an automatic mapping of MR and Ultrasound data to Vessel probability values and successive registration of this information is proposed in [8]. Using a CT data of a kidney, where the intensity values are enhanced with strong edges from the gradient, a registration with freehand 3D ultrasound is performed in [9]. Voxel-based registration of MRA scans with Power Doppler ultrasound has been evaluated in [10].

## 2 Methods

### 2.1 Simulation from CT

Instead of a realistic simulation of ultrasound, we need an intelligent and efficient intermediate representation of the CT data at arbitrary cut-planes, such that an iterative registration can be performed in an acceptable time.

These slices have multiple components containing intensity, gradient and edge information, which are used to derive various parts of a similarity metric, so that the correspondence of anatomy contained therein with structures in 2D B-mode ultrasound images can be determined.

In our approach, first the three-dimensional gradient vector values are computed from the CT data set by convolution with a sobel filter cube. They are stored in a 4-channel volume together with the original voxel intensity. The interpolated slices contain four channels as well. For each pixel, the 4-vector is computed from the volume using trilinear interpolation. In the first channel of the slice, the original CT intensity is stored. The 3D gradient vector is scalar multiplied with each of the vectors indicating the horizontal and vertical slice plane directions, respectively. The resulting values, corresponding to the 2D gradient of the CT intensity within the slice, are stored in the second and third channel.

The 2D slice gradient values are then used to perform Canny edge-detection on the slice data, storing the result in the fourth channel. The most time-consuming steps within the Canny algorithm for 2D images are the computation of the 2D gradients, as well as filtering them with a sufficiently large Gaussian kernel for

smoothing. As we compute the 2D gradients directly from the precomputed 3D gradient values, we do not need to run a 2D filtering for gradient computation. In addition, those gradients are very smooth, as they originate from a three-dimensional Sobel filter using a 27-neighborhood. This makes further Gaussian filtering unnecessary. The two remaining steps for the Canny algorithm, non-maxima suppression and hysteresis thresholding, can be performed each in one traversal of the 2D slice. The horizontal gradient is weighted with a user-defined factor between 0 and 1, as the ultrasound data tends to show mainly vertical edges.

Thus we are able to construct intermediate slices from the CT data at estimated transformations of the US scan plane in very little time<sup>3</sup>. The individual components of the slice pixels are then used to compute a similarity metric with the ultrasound data.

## 2.2 Occlusion Handling

If an ultrasonic pulse hits bony structures, all image intensities in the ultrasound image further along the specific ray are occluded, and mainly determined by noise. Therefore, all ultrasound intensity values on a ray below such an occlusion should be disregarded in the registration method. In our implementation, we scan the US image from bottom to top, updating the variances for all ultrasonic pulse rays. Where they exceed a threshold (which is easily determined in the user interface), the first pixel to be considered is defined. Thus, our Region of interest  $\Omega$  is expressed by the following equations:

$$\Omega = \{(x, y) \mid (y < y_{top}) \wedge (y \geq b(x))\} \quad (1)$$

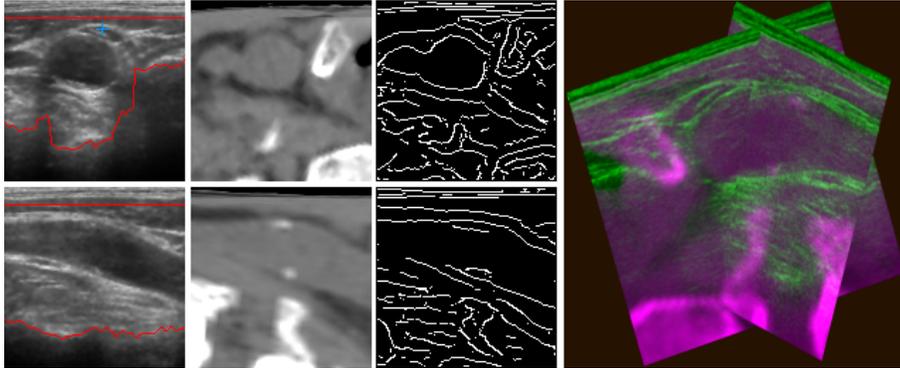
$$b(x) = \min_y \left| \frac{1}{y} \sum_{i=0}^{y-1} U(x, y)^2 - \left( \frac{1}{y} \sum_{i=0}^{y-1} U(x, y) \right)^2 \right| < \sigma_y^2 \quad (2)$$

By applying a median filter on the bottom function  $b(x)$ , discontinuities are removed before defining the ROI. In addition, we discard all pixels which are located above  $y_{top} = \frac{9}{10} size_y$ , as we observed that the anatomy is highly compressed there due to the probe pressure on the patients skin. This compressed region is very distinct from the remaining anatomical structures, its size ( $3.6mm$ ) being consistent on all data we obtained from patients (figure 1). This ROI definition is similar to the ones used in [8] and [9].

## 2.3 Similarity Measure

Based on both the physical properties of the imaging modalities, as well as the visible appearance of their images, we developed several components for a similarity measure, which can in turn be weighted to form a cost function value with respect to the transformation parameters.

<sup>3</sup> 1.1ms for a  $128^2$  pixel slice, interpolated from a  $512^2 \cdot 100$  CT/gradient volume, on an AMD Opteron 2.4 Ghz machine



**Fig. 1.** Two ultrasound images with ROI (red lines) and target, corresponding CT slices, edges from CT, and overlay in 3D. The physical image size is  $4 \times 4cm$ .

**Skin Surface Clamping.** In the compressed fraction of the ultrasound image, the interpolation from CT is done with 6 times the vertical scaling (figure 1 on top). As result, the interface between skin and air always has to be within that region, producing a large vertical gradient in the interpolated slice. When all vertical gradient pixels are summed to  $t$ , high and low thresholds  $t_h$ ,  $t_l$  can be defined in order to decide if the skin surface lies inside, outside or close to the compressed region:

$$f(t) = \begin{cases} 1 & \text{if } t > t_h \\ 0 & \text{if } t < t_l \\ (t - t_l)/(t_h - t_l) & \text{otherwise} \end{cases}; \quad S_0 = 3f(t)^2 - 2f(t)^3 \quad (3)$$

A cubic polynomial is used instead of the linear rise in order to avoid discontinuities. Used as a cost function component,  $S_0$  penalizes transformations which are physically impossible, as the patients skin is always on top of the ultrasound images.

**Edge Alignment.** As we have detected the edges in the simulated images, we would like to derive a similarity estimate based on the distance to edge structures in the ultrasound images. The straight-forward approach would be to 1) compute an edge-detection for the ultrasound images, 2) compute a 2D distance map for those edges and 3) sum over the distance map values at the locations indicated by the edges of the simulated data. Steps 1) and 2) need to be performed once for each ultrasound slice, while 3) establishes a similarity metric and thus has to be computed for each simulated slice during pose estimation. However, due to the very different nature of CT and ultrasound data, detected edges do not correspond in general. We therefore propose to skip the edge detection from ultrasound data, instead using the original ultrasound intensity just as indicator for edges.

Given a binary edge image, the distance of an image point  $\mathbf{x}$  to the edge structures  $Y = \{\mathbf{y}_i\}$  is  $d(\mathbf{x}) = \min_i |\mathbf{x} - \mathbf{y}_i|$ . Instead of the euclidian distance, we

can also express the proximity to edges by using a Gaussian expression, which allows us to adjust the sensitivity of the cost function value with respect to the distances, using  $\sigma^2$ :

$$d(\mathbf{x}) = \max_i \exp - \frac{(\mathbf{x} - \mathbf{y}_i)^2}{\sigma^2} \quad (4)$$

Taking into account that we do not have precise edge information, a proximity value can be defined as

$$d(\mathbf{x}) = \sum_i p_i \exp - \frac{(\mathbf{x} - \mathbf{y}_i)^2}{\sigma^2} \quad (5)$$

where  $p_i \in [0 \dots 1]$  is the probability for the image pixel  $\mathbf{y}_i$  being an edge. Assuming that the ultrasound image intensity directly scales with the edge probability, a two-dimensional proximity function  $p(\mathbf{x})$  can be computed by just convoluting the ultrasound image with a large gaussian kernel. The similarity measure component arises from this as  $S_1 = (\overline{p_e} - \overline{p})/\sigma_p$ , where  $\overline{p}$  is the mean of all values in the proximity image,  $\overline{p_e}$  the mean of just the pixels at locations where an edge is present in the simulated image, and  $\sigma_p$  the standard deviation of the proximity image values.

**Statistical Correspondence.** Different tissues in the anatomy cause different scattering characteristics for ultrasonic waves. Higher scattering in turn causes a larger portion of the ultrasound pulse to be reflected back to the transducer, resulting in higher intensities in the ultrasound image. It is therefore applicable to assess the statistical dependance of the CT intensities, which classify the tissue according to the X-Ray attenuation property, with the intensity in the ultrasound image. We therefore use Mutual Information on the CT and ultrasound intensities. The Normalized Mutual Information term uses the entropies of the combined and individual images, which are computed with the Shannon entropy from probability distributions of the image intensities:

$$\begin{aligned} NMI(U, S) &= 2 - 2H(U, S)/(H(U) + H(S)) \\ H(U) &= - \sum_j p_u(j) \log p_u(j); \quad H(S) = - \sum_i p_s(i) \log p_s(i) \\ H(U, S) &= - \sum_i \sum_j p(i, j) \log p(i, j) \end{aligned}$$

Here  $U$  denotes an ultrasound image, and  $S$  the corresponding simulated image, i.e. the slice interpolation of CT attenuation values. The probability distributions can be estimated using histogram information from the images:

$$p_u(i) = \frac{1}{n_\Omega} |\{(x, y) \in \Omega | U(x, y) = i\}| \quad (6)$$

$$p_s(j) = \frac{1}{n_\Omega} |\{(x, y) \in \Omega | S(x, y) = j\}| \quad (7)$$

$$p(i, j) = \frac{1}{n_\Omega} |\{(x, y) \in \Omega | U(x, y) = i \wedge S(x, y) = j\}| \quad (8)$$

Here we assume that each intensity value is mapped into one histogram bin, and  $n_\Omega = |\Omega|$  is the number of pixels in the region of interest. An equivalent formulation for constructing the probability distribution from a histogram can be written using a binary count function  $c_u$

$$p_u(i) = \frac{1}{n_\Omega} \sum_{(x,y) \in \Omega} c_u(x,y,i); \quad c_u(x,y,i) = \begin{cases} 1 & \text{if } U(x,y) = i \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

Due to the various physical effects in ultrasound imaging, both the chance that an image intensity reflects the anatomy, as well as the Signal to Noise Ratio (SNR), decrease with the distance from the ultrasound transducer. Thus we would like to give more emphasis on image pixels which are closer to the probe, i.e. with higher  $y$  values. In our approach, we introduce an integer weighting for assembling the distribution:

$$p'_u(i) = \frac{1}{n'_\Omega} \sum_{(x,y) \in \Omega} (y + c_0) c_u(x,y,i) \quad (10)$$

$$n'_\Omega = \sum_{x=0}^{n_x-1} \sum_{y=b(x)}^{y_{top}-1} (y + c_0) \quad (11)$$

Every intensity value is inserted  $y + c_0$  times into the histograms and the joint histogram. For  $c_0 \rightarrow \infty$  the original Mutual Information notation is obtained. Our weighted Mutual Information component  $NMI'$  of the similarity measure is assembled by inserting all used ultrasound slice images and the corresponding simulations into one histogram, as it increases the statistical significance of the derived entropy terms.

**Cost Function.** The final similarity measure from a set of  $n$  ultrasound slices  $\{U_i\}$  and their CT simulations  $\{S_i\}$  is

$$cf = w_0 \frac{1}{n} \sum_{i=1}^n S_0(U_i, S_i) + w_1 \frac{1}{n} \sum_{i=1}^n S_1(U_i, S_i) + w_2 NMI'(\{U_i\}, \{S_i\}) \quad (12)$$

## 2.4 Registration

In order to manually navigate the stack of ultrasound images to the desired position within the CT data, the user picks a reference slice  $k$ , whose position and orientation is changed by left-multiplication with a rigid transformation matrix. At the same time, all other transformations are updated in order for the relative locations to stay fixed, as they originate from the tracker.

For automatic registration, a non-linear optimization method maximizes the cost function  $cf$  iteratively with respect to the parameters of a rigid transformation (6 DOF, translation and Euler angles), which is initialized with zero and affects the location of all slices. We used three optimization schemes: simple hill climbing, Powell-Brent and an exhaustive hill climbing. The latter one evaluates all combinations of [forward, keep, backward] for all parameters, using the best result of all  $3^6 = 729$  evaluations as estimate for the next iteration.

### 3 Results

Three head and neck cancer patients with metastatic lymph node involvement were thoroughly examined with a 11 MHz linear array ultrasound probe. The images were recorded using a frame grabber card, while an Ascension microBIRD<sup>TM</sup> magnetic tracking sensor provided the spatial encoding. A set of 3-10 slices from the right carotid artery of each patient was picked for registration. Figure 1 depicts two slices from the first patient alongside the registered CT data.

A ground truth registration pose was established with manual registration by the physician. This could be done with an estimated precision of  $1mm$  in the first data set, as the calcifications in the carotid artery (figure 1) represented good anatomical landmarks.

In order to evaluate the robustness and accuracy of the automatic registration, 200 registrations were launched from initial transformations randomly displaced up to  $5mm/5^\circ$  in each parameter around the ground truth pose. The following table denotes both the root mean squared (RMS) error in the translational and rotational components, as well as the target registration error (TRE) for the lymph node (figure 1) picked as target. This evaluation was done for all three used optimization schemes on the data of patient 1.

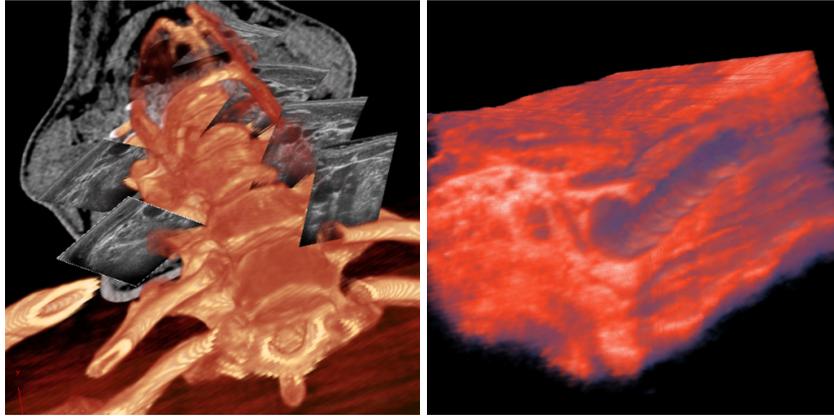
	trans.	rot.	TRE	iterations	time
Hill Climbing	$1.2mm$	$3.7^\circ$	$2.0mm$	242	3s
Powell-Brent	$1.0mm$	$2.8^\circ$	$1.8mm$	4	8s
Exhaustive H.C.	$0.8mm$	$2.5^\circ$	$1.2mm$	189	144s

All optimization methods are able to converge precisely to the ground truth registration, so that the registered data can be used reliably for therapy planning. To do so, the slices from the original CT data set, which are used to outline the target volume, are visualized together with the registered ultrasound slices and optionally volume rendering of both the CT and a 3D ultrasound volume spatially compounded from the tracking data (figure 2).

### 4 Conclusion

We developed methods which allow automatic registration of a set of ultrasound slices to a CT scan, despite the very difficult characteristics for registration of both modalities. The similarity metric is derived from the physical properties of ultrasound imaging, rather than from the particular anatomy used in our experiments. Therefore the algorithms are also applicable on any other part of the human body scanned with an external ultrasound probe. The registration is performed within a few seconds, and is therefore capable of supporting real-time applications, such as intra-operative navigation, as well. We evaluated our methods in the context of radiotherapy for head and neck cancer, where the use of registered data is beneficial for the treatment planning.

**Acknowledgements.** This research was funded by an academic grant from Siemens Corporate Research, Princeton, USA. The authors would like to thank in particular Frank Sauer and Ali Khamene for their continuous support.



**Fig. 2.** Left: Overlay of registered ultrasound images, a slice from the CT data set, and CT volume rendering. Right: Volume rendering of compounded 3D ultrasound.

## References

1. Yonetsu, K., Sumi, M., Izumi, M., Ohki, M., Eida, S., Nakamura, T.: Contribution of doppler sonography blood flow information to the diagnosis of metastatic cervical nodes in patients with head and neck cancer: assessment in relation to anatomic levels of the neck. *AJNR Am J Neuroradiol* **22** (2001) 163–169
2. Sumi, M., Ohki, M., Nakamura, T.: Comparison of sonography and CT for differentiating benign from malignant cervical lymph nodes in patients with squamous cell carcinoma of the head and neck. *AJR Am J Roentgenol* **176** (2001) 1019–1024
3. Penney, G.P., Blackall, J.M., Hayashi, D., Sabharwal, T., Adam, A., Hawkes, D.J.: Overview of an ultrasound to CT or MR registration system for use in thermal ablation of liver metastases. In: *Proc. Medical Image Understanding and Analysis*. (2001) 65–68
4. Porter, B., Rubens, D., Strang, J., Totterman, S., Parker, K.: Three-dimensional registration and fusion of ultrasound and MRI using major vessels as fiducial markers. *IEEE Transactions on Medical Imaging* **20** (2001) 354–359
5. Amin, D., Kanade, T., Gioia, A.M.D., Jaramaz, B.: Ultrasound registration of the bone surface for surgical navigation. *Computer Aided Surgery* **8** (2003) 1 – 16
6. Wu, R., Ling, K.V., Shao, W., Ng, W.S.: Registration of organ surface with intra-operative 3D ultrasound image using genetic algorithm. In: *MICCAI 2003 Proceedings*. (2003) 383–390
7. Roche, A., Pennec, X., Malandain, G., Ayache, N.: Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging* **20** (2001) 1038 – 1049
8. Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., Hawkes, D.: Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis* (2004) 81 – 91
9. Leroy, A., Mozer, P., Payan, Y., Troccaz, J.: Rigid registration of freehand 3D ultrasound and CT-Scan kidney images. In: *MICCAI Proceedings*. (2004) 837ff
10. Slomka, P., Mandel, J., Downey, D., Fenster, A.: Evaluation of voxel-based registration of 3-D power doppler ultrasound and 3-D magnetic resonance angiographic images of carotid arteries. *Ultrasound Med. Biol.* **27** (2001) 945–955