Model-based Segmentation and Image Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning

M. Bach Cuadra[1], S. Gorthi[1], F. I. Karahanoglu[1], F. Salvador[1],
A. Pica[2], H.P. Do[3], A. Balmer[4], F. Munier[4], J.-Ph. Thiran[1]
[1]*Signal Processing Laboratory (LTS5) Ecole Polytechnique Fédérale de Lausanne, Switzerland*
[2]*Radiation Oncology Department, Lausanne University Hospital (CHUV), Lausanne, Switzerland*
[3]*Institute of Applied Radiophysics, Lausanne, Switzerland*
[4]*Ophthalmic Hospital Jules Gonin, Lausanne, Switzerland*

For radiotherapy treatment planning of retinoblastoma in childhood, Computed Tomography (CT) represents the standard method for tumor volume delineation, despite some inherent limitations. CT scan is very useful in providing information on physical density for dose calculation and morphological volumetric information but presents a low sensitivity in assessing the tumor viability. On the other hand, 3D ultrasound (US) allows a high accurate definition of the tumor volume thanks to its high spatial resolution but it is not currently integrated in the treatment planning but used only for diagnosis and follow-up. Our ultimate goal is an automatic segmentation of gross tumor volume (GTV) in the 3D US, the segmentation of the organs at risk (OAR) in the CT and the registration of both. In this paper, we present some preliminary results in this direction. We present 3D active contour-based segmentation of the eye ball and the lens in CT images; the presented approach incorporates the prior knowledge of the anatomy by using a 3D geometrical eye model. The automated segmentation results are validated by comparing with manual segmentations. Then, for the fusion of 3D CT and US images, we present two approaches: (i) landmark-based transformation, and (ii) object-based transformation that makes use of eye ball contour information on CT and US images. **Keywords:** *Segmentation, Active contours, 3D Ultrasound imaging, Radiotherapy planning.*

1 INTRODUCTION

Retinoblastoma is one of the common primary ocular malignancy in childhood and it accounts for 5% of childhood blindness (Donaldson 1989). External beam radiation is now used to control advanced, often chemotherapy resistant, intra-ocular tumors in progression after chemotherapy and focal ophthalmologic therapy in children. Thus, primary endpoint of such radiation therapy is the eye and visual function preservation following intra-ocular progression. Therefore, delineating the tumor in order to optimize radiation doses, allowing minimization of dose to adjacent developing tissues is a crucial goal. To this end, it is necessary for patients to undergo a multi-modal imaging: Computed Tomography (CT) scan, which is very useful in providing information on physical density for dose calculation and morphological volumetric information, and 3D ultrasound (US), which allows a high accurate definition of the tumor volume thanks to its high spatial resolution. Thanks to its improving quality, the use of US imaging is increasing in ophthalmology (Fenster et al. 2001; Yannuzzi 2004). At the Ophthalmic Hospital Jules Gonin in Lausanne, Stwitzerland, 3D ultrasound imaging is routinely used for diagnosis and follow-up but it is not integrated in the treatment planning. In this paper we propose the combination of 3D US imaging, optimal for tumor volume delineation, and CT imaging, optimal for dose calculation and treatment planning. Such multimodal imaging framework has been already suggested for other organs therapy like the liver or kidney (Wein et al. 2008; Penney et al. 2004) but, as far as we know, this is the first attempt of combining these two modalities in the radiotherapy planning of the eye. Through an automatic segmentation of gross tumor volume (GTV) in the 3D US and simultaneous registration with the CT, the rather inaccurate and time-consuming manual volume definition could be shortened and the volume definition would become highly reproducible and comparable, overcoming the

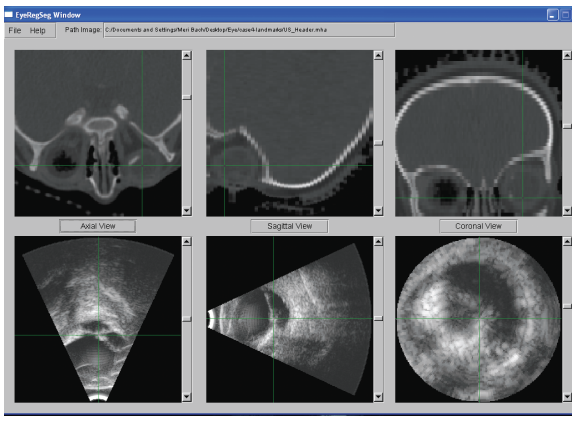

Figure 1: Data set: top row, 3D CT image of the head used for the treatment planning; bottom row, 3D ultrasound of the eye. Note that the green cross does not represent same spatial location.

intrinsic problems of inter- and intra-user variability, thus making radiation treatment techniques amenable to standardization.

In this paper we present the first steps that will allow us to work in this multimodal framework for the eye radiotherapy. We present the automatic segmentation of the eye exploiting the prior knowledge that we have on eye's anatomy by means of a 3D geometric eye model within the parametric active contour framework. We also present an initial fusion of these images using landmark-based transformation and object-based transformation that makes use of eye ball contours on both the modalities.

2 MULTIMODAL IMAGING

CT images are acquired in the Lausanne University Hospital, on a LightSpeed VCT General Electric Medical Imaging scanner. Images have a resolution of $0.7 \times 0.7 \times 2$ mm^3 (see Fig. 1 top row). The ophthalmic 3D ultrasound images used here are acquired with OTI Ophthalmic Technologies Inc. (OTI). The OTI-Scan 3D is one of the most advanced ophthalmic ultrasound system available today. The internal rotator assembly generates a 3D image in less than 2 seconds with a 0.1 mm resolution in each X, Y, and Z direction. An example of such acquisition is shown in Figure 1 bottom.

3 EYE SEGMENTATION IN THE CT

The importance of exploiting prior knowledge in the process of complex image segmentation has been largely proven. Thus, our aim in this model-based segmentation process is to combine information of 3D data (CT) with the anatomical information stored in a standard model. To this end we propose to use a 3D eye model based on ellipsoidal structures. This is not the first attempt of using a parametric eye model for segmentation. Two important works exists. In (Dobler and Bendl 2002), a very precise geometric model of the eye is used for proton therapy and they set most of the parameters manually via ultrasound measurements. Recently, in (Bekes et al. 2008), a simplified eye model has been present, where the setting of the parameters was almost completely automatized to provide minimal user interaction. The method presented here is inspired from (Bekes et al. 2008). Our segmentation model differs from the model of Bekes et al. mainly in the following aspects: The model of Bekes et al. is based on thresholding and soft classification function whereas our model uses parameterized active contours that takes advantage of characteristics of the edges and regions. Unlike Bekes et al., we are currently interested in the segmentation of eye ball and lens only, but not optic nerve and optic chiasm. Another difference is that Bekes et al. approximate the eye ball with a sphere while we use a more accurate ellipsoidal approximation.

The CT image properties and the very well known geometry of the eye lead to an adaptation of the active contours theory (Caselles et al. 1997) which aims at finding a curve C that minimizes the energy functional E, designed to be minimal when C has found the desired contour. The general expression of E is given by

$$E(C) = E_{image}(C) + E_{smooth}(C), \qquad (1)$$

Here, we take advantage of using a parameterized model, which means that the analytical expression of C is known everywhere on the domain of the image and is defined by a set of parameters θ. Thus, C is a parameterized model based on the parameters $\theta = \{x_c, y_c, z_c, r_x, r_y, r_z, \phi, \psi\}$ corresponding to the center coordinates, the length of the axes and the rotation angles.

Obviously, we will not impose any E_{smooth} term since our parameterized curve is already smooth. Then, the E_{image} term can be expressed in two subterms:

$$E_{image}(C(\theta)) = E_{boundary}(C) + E_{region}(C), \qquad (2)$$

which are computed from the image features.

$E_{boundary}$ attracts the curve towards the object boundaries using an edge detecting function f, and is given by:

$$E_{boundary} = \int_C f(C,s)ds, \qquad (3)$$

Once applied to the image I, $f(C,s)$ has to return an optimum value (minimum in our case) when its variables, which are the curve spatial coordinates $(x(\theta,s), y(\theta,s), z(\theta,s))$ mapped by the model on the image, are matching the edges of the object to be de-

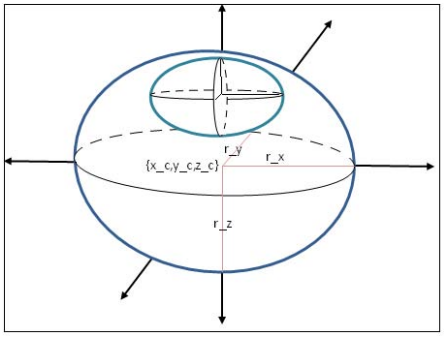

Figure 2: The proposed 3D parametric model of the eye consisting of ellipsoids for the eye ball and the lens. The model parameters for the eye lens are $\theta_{lens} = \{x_c, y_c, z_c, r_x, r_y, r_z, \phi, \psi\}$ corresponding to the center coordinates, the length of the axes and the rotation angles. Model parameters for the eye ball are $\theta_{ball} = \{x_c, y_c, z_c, r_x, r_y, r_z\}$ thus ϕ and ψ angles are neglected.

lineated. Therefore, f can be of the form:

$$f(C,s) = \begin{cases} -(I_{eq}(C,s) * G_\sigma) \text{ if } I_{eq}(C,s) > T_h, \\ 0 \text{ otherwise.} \end{cases} \quad (4)$$

where T_h is a threshold applied on the equalized image intensity and empirically set to 0.8, G is the Gaussian function with standard deviation σ and $I_{eq} * G_\sigma$ represents the smoothed version of the equalized image *I*.

E_{region} captures the regional statistics to drive *C* towards homogeneous regions. The idea here is to maximize the difference between two statistical descriptors (the mean value of intensities) related to two regions of the image, Ω_{in} and Ω_{out}, in our case, respectively the inside and the outside of an ellipsoid within a selected region of interest around the lens. Formally

$$E_{region} = -\text{difference}[\text{mean}(\Omega_{in}), \text{mean}(\Omega_{out})]. \quad (5)$$

The segmentation problem is now reduced to an energy optimization problem. Thus, the parameter-set θ that results in minimum energy provides the segmentation of the objects of interest.

$$\theta_{\text{object}} = \arg\min_{\theta} E(C(\theta)). \quad (6)$$

From our experiments, we found that boundary-based term ($E_{boundary}$) alone is sufficient for the accurate segmentation of the eye ball. Similarly for the lens, we found that region-based term (E_{region}) alone is sufficient. Hence, for the results that we present in the following Section, we use only $E_{boundary}$ term for the eye ball, and only E_{region} term for the lens segmentation. We also observed from our experiments for eye ball segmentation that the rotation angles for the eye ball are always varying around 0. Hence, we made a simplification for the eye ball segmentation by neglecting the rotation angles: $\{\phi, \psi\}$. As a result, the number of parameters to be optimized for the eye ball and the lens are 6 and 8 respectively. The proposed 3D parametric model of the eye is shown in Fig. 2.

3.1 Results: Eyeball and lens segmentation

The optimization is implemented with a Nelder-Mead simplex search included in the Matlab *fminsearch* function. Note that a set of parameters such as θ includes elements of different nature and therefore are their range very different. For example a semi-axis of the eye ball ranges from 5 to 20 *mm* while an angle ranges from 1 to 360 degrees. To overcome this problem we proceed to the optimization iteratively using 3 subsets of parameters, $[x_c, y_c, z_c]$, $[r_x, r_y, r_z]$, and $[\phi, \psi]$ (actually only the 2 first subsets for the eye ball). Moreover the algorithm is applied twice per object (first growing the ellipsoid and later decreasing it). Equalization of the image is performed as a preprocessing. Eye ball optimization is initialized by user clicks, (center coordinates x_c, y_c, z_c and the radius in z direction r_z). Lens optimization is restricted to one half of the eyeball only. The overall optimization algorithm is summarized as follows:

1. **Input data**
   Image with smoothed edges (lens) or threshold image (eyeball) (see Figure3)
2. **Ellipsoid initialization**
   Smaller (eyeball)/ Bigger (lens)
3. **Iterative loop**
   3 iterations loop optimizing 2 subsets of parameters for the eyeball, 3 subsets for the lens
4. **Security check**
   To avoid too much deviation from solution (axes length can be neither longer nor shorter than usual eye size)
5. **Stopping conditions**
   After $N = 15$ iterations or if the ellipsoid is outside normal anatomy size
6. **Go back** to 2
   Bigger than eyeball/lens, now ellipsoid shrinks instead of growing

Results for 3 patients are shown in Figure 4. Axial and sagittal views of the obtained segmentation show a very good segmentation for both lens and eyeball. 3D view reconstruction is also shown, note that the segmentation of the skull has been done simply by threshold for visualization purposes. Table 1 presents the radii and angles for the eyeball and the lens, obtained from the automated segmentation, for the three patients. Quantitative validation is also given in Table 2. Evaluation is done in comparison with manual segmentations done by an expert. Dice Similarity Measure (DSM) is presented showing a very high agreement between the manual and the automated segmentation for the eyeball and the lens.

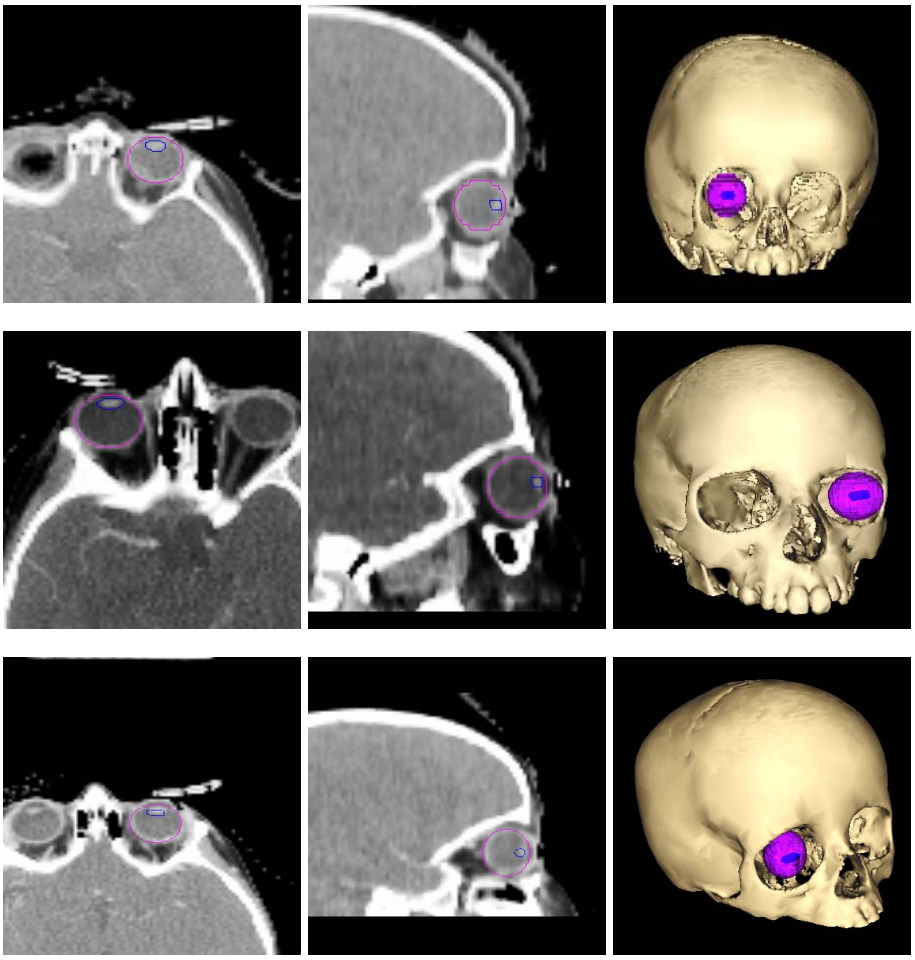

Figure 4: Axial, sagittal and 3D views of the automated segmentation results of the eye ball and the lens, for three patients.

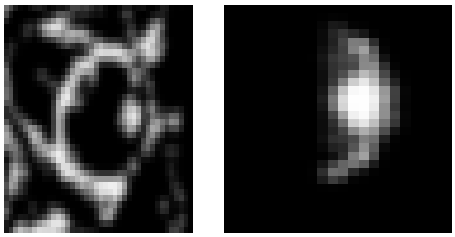

Figure 3: Input images for the eyeball (left) and lens (right) segmentation for one of the patients.

	Eye ball	Lens	
	(r_x, r_y, r_z) in mm	(r_x, r_y, r_z) in mm	(ϕ, ψ) in degrees
P-1	(10.0, 9.1, 10.0)	(3.6, 1.8, 2.0)	(186.6, 165.4)
P-2	(12.3, 10.6, 12.6)	(3.1, 1.5, 1.5)	(181.7, 181.1)
P-3	(10.5, 9.8, 11.0)	(4.9, 2.2, 2.4)	(185.3, 177.0)

Table 1: Radii and angles for the eye ball and the lens, obtained from the CT image segmentation.

	Volume (mm^3)			DSM
Eyeball	Manual	Automated	Error	
Patient 1	4764	4827	1.33%	89.9%
Patient 2	7758	7798	0.52%	92.6%
Patient 3	6114	5543	9.33%	90.2%
Lens				
Patient 1	141	102	-27.7%	78.4%
Patient 2	205	215	4.9%	76.9%
Patient 3	217	219	0.9%	76.8%

Table 2: Quantitative evaluation for the eyeball and lens.

4 IMAGE FUSION

4.1 Landmark-based transformation

Multimodal registration is a widely studied problem in medical imaging. There exists however less studies that consider the registration of 3D ultrasound with any other modality (Roche 2001; Wein et al. 2008; Penney et al. 2004). Most of these methods require a manual initialization of the transformation to ensure the convergence of the registration. This is even more necessary in our case, where both images do not contain the same information: note that while the CT contains the whole brain, the US image only presents

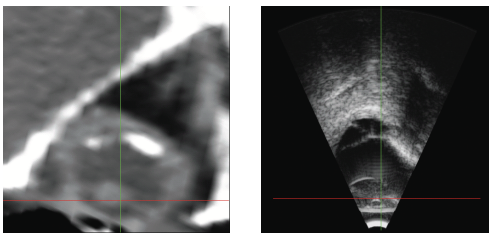

Figure 6: Resampled CT image to the US space and resolution, cross cursor represents the same spatial location.

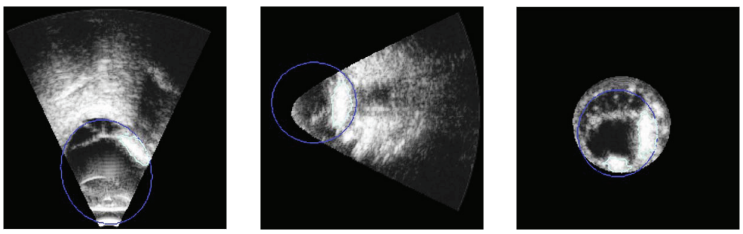

Figure 7: 3D US image with the superposed eye ball contours from the CT image, after applying the object-based transformation

partially one eye. Thus, in a preliminary step for the registration we will put the two images in correspondence by proceeding to a global rigid transformation (3 translations and 3 rotations). To this end we choose a landmark based matching, specifically the one in ITK library (ITK). Selection of the landmarks is not an easy task and must be done by an expert. Four landmarks per image have been selected (only 2 are necessary), in the presented pathological example (see Fig. 5: the lens (blue), head of the optic nerve (yellow) and two tumors (red and green). At this point we can proceed to the fitting of the rigid parameters that best approximate these 4 locations. Visualization of the resulted transformation is done by comparing with a cross at the same spatial position (see Figure 6). Rigid parameters are [0.560694, 0.29938, -0.0351702] angle rotations in radians, [-42.2236, 98.625, -26.3611] translations in mm. Landmark matching quality is: 0.951184 (blue), 0.88538 (red), 0.915436 (green), 0.915033 (yellow).

The main problem that we have faced with the current landmark-based approach is that identification of anatomical corresponding landmarks on CT and US images is a tedious and time consuming task. For example, it took around 30 minutes for an expert Doctor to make the 4 landmarks shown in Fig. 6. Further this always demands the intervention of an expert for identifying the landmarks. In order to avoid these inconveniences, we propose another type of initialization, an *object-based transformation*, and it is explained in detail in the following Section.

4.2 Object-based transformation

For finding out an object-based (eye ball in this case) transformation, we need to set the object in both image modalities. In Section 3, we have already obtained the segmentation of the eye ball on the 3D CT image. It is difficult to proceed similarly in the 3D US image, mainly because the field of view of US does not show the entire eye ball (neither the lens). However, we observe that: first, it is easy to identify at least the back portion the eye ball edges on the US and, second, we already know the eye ball dimension we are looking for. Thus, we suggest the following approach. First, we select few pixels along the visible edges of the eye ball in the US image. We have selected around 25 pixels for the results presented in this paper. This is done only on few slices in the sagittal and coronal directions (the eye ball edges are not clearly visible in the axial view). We then fit an ellipsoid to the selected pixels using least square error criteria. Note that the radii of the ellipsoid to be fit (r_x, r_y, r_z) are already known from the CT image segmentation; hence, we compute only the center (x_c, y_c, z_c) and angles (ϕ, ψ) of the ellipsoid on the US image, during the least square error procedure. Fig. 7 shows the contours of the fitted ellipsoid on the US image for one of the patients.

Now, from the values of the center and angles of the ellipsoid computed above, we can directly obtain the transformation (translation and rotation) between both segmented ellipsoids in the CT and US images. However, we noticed that since the object used for computing the transformation (ellipsoid) is symmetric about its axes of rotation, the rotational parameters obtained were not accurate. In order to compensate for this limitation, at present we have manually rotated the image towards the correct orientation. Hence, in the future work, in order to obtain a more accurate fusion, we need to either realign the images by selecting a landmark, or, perform a further image volume registration between the images. Figure 8 shows the spatially aligned CT and US images as well as fusion in the axial direction, by applying this object-based transformation. The translational parameters computed from the object-based approach are found to be very satisfactory.

5 DISCUSSION

In this paper, we have presented our preliminary work towards a multimodal framework for the radiotherapy planning of retinoblastoma. To our knowledge, this is the first attempt to combine 3D CT and 3D US for the eye radiotherapy. In the present work we have presented a model-based segmentation of the eye ball and the lens on the CT images. The main difference compared to the existing model-based segmentation methods is that here we proceed by using the well formulated active contour framework. Evaluation in comparison with manual segmentation has

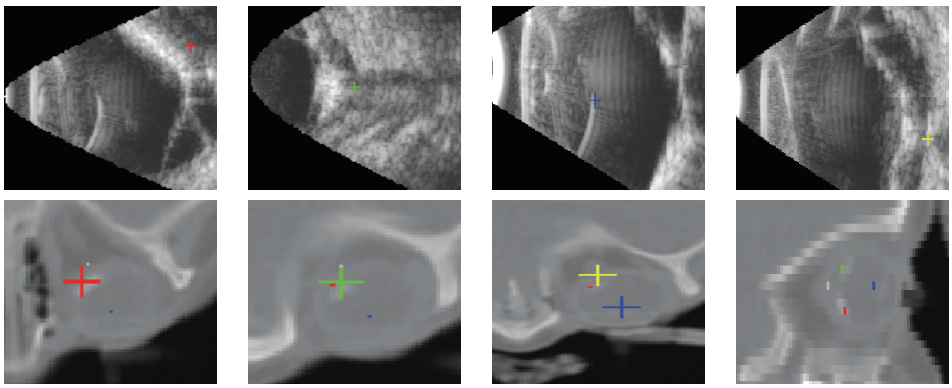

Figure 5: Selected landmarks: blue is in the lens, yellow is in the optic nerve, and green and red are in the back wall of the tumors. Note that zoom is shown for all US views.

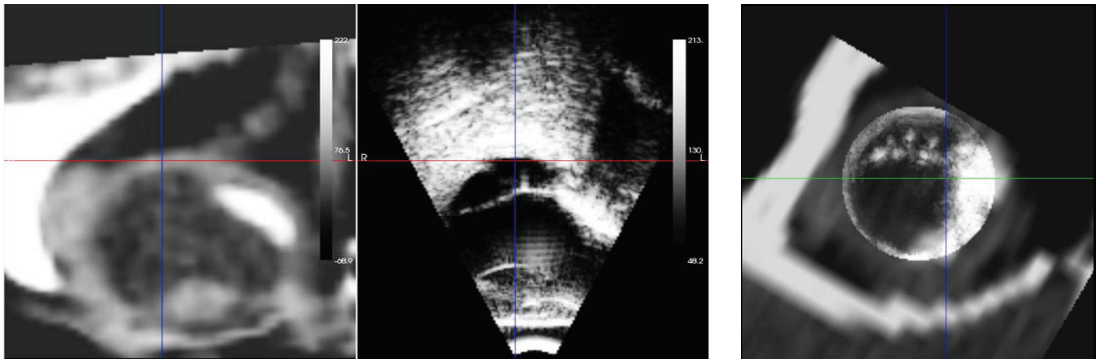

Figure 8: (a)Resampled CT image after applying the object-based transformation, (b) Corresponding US image with cross curser representing the same spacial location, (c) Fusion of CT and US in the axial direction.

shown a good accuracy of our segmentation model. Moreover, we have also computed a first required initialization of the rigid transformation that brings the CT, and thus the eye model, and the US into the same space. Two procedures have been suggested for the initial alignment of the CT and US: landmark-based and object-based. We prefer at this point the object-based transformation over the landmark-based transformation since the object-based approach is faster and does not require any expert intervention. Preliminary results of the fusion are shown. Further work consists of finding out a more accurate transformation between the modalities, testing the robustness of the our eye model on extended data, as well as proceeding to the tumor segmentation in the US images.

REFERENCES

ITK, www.itk.org.

OTI, http://www.oti-canada.com/.

Bekes, G., E. Máté, L. G. Nyúl, A. Kuba, and M. Fidrich (2008). Geometrical model-based segmentation of the organs of sight on CT images. *Medical Physics 35*, 735–743.

Caselles, V., R. Kimmel, and G. Sapiro (1997). Geodesic active contours. *International Journal of Computer Vision 22*(1), 61–79.

Dobler, B. and R. Bendl (2002). Precise modelling of the eye for proton therapy of intra-ocular tumours. *Phys. Med. Biol. 47*(4), 593–613.

Donaldson, SS; Smith, L. (1989). Retinoblastoma: biology, presentation, and current management. *Oncology 3*(10), 45–51.

Fenster, A., D. Downey, and H. Cardinal (2001). Three-dimensional ultrasound imaging. *Phys Med Biol 46*(5), 67–99.

Penney, G., J. Blackall, M. Hamady, T. Sabharwal, A. Adam, and D. Hawkes (2004). Registration of freehand 3D ultrasound and magnetic resonance liver images. *Med. Im. Anal. 8*, 81–91.

Roche, A.; Pennec, X. M. G. A. N. (2001). Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Trans. on Med. Imag. 20*(10), 1038–1049.

Wein, W., S. Brunke, A. Khamene, M. Callstrom, and N. Navab (2008). Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Med. Im. Anal. 12*, 577–585.

Yannuzzi, L.A., e. a. (2004). Ophthalmic fundus imaging: today and beyond. *Am J Ophthalmol 137*(3), 511–24.