Technical University of Denmark

DTU

Cochlea Segmentation using Iterated Random Walks with Shape Prior

Ruiz Pujadas, Esmeralda; Kjer, Hans Martin; Vera, Sergio; Ceresa, Mario; González Ballester, Miguel Angel

DTU Library
Technical Information Center of Denmark

Cochlea Segmentation using Iterated Random Walks with Shape Prior

Esmeralda Ruiz Pujadas[a], Hans Martin Kjer[b], Sergio Vera[c], Mario Ceresa[a], and Miguel Ángel González Ballester[a,d]

[a]Universitat Pompeu Fabra, Barcelona, Spain
[b]Technical University of Denmark, Denmark
[c]Alma Medical Imaging, Barcelona, Spain
[d]ICREA Barcelona, Spain

ABSTRACT

Cochlear implants can restore hearing to deaf or partially deaf patients. In order to plan the intervention, a model from high resolution μCT images is to be built from accurate cochlea segmentations and then, adapted to a patient-specific model. Thus, a precise segmentation is required to build such a model. We propose a new framework for segmentation of μCT cochlear images using random walks where a region term is combined with a distance shape prior weighted by a confidence map to adjust its influence according to the strength of the image contour. Then, the region term can take advantage of the high contrast between the background and foreground and the distance prior guides the segmentation to the exterior of the cochlea as well as to less contrasted regions inside the cochlea. Finally, a refinement is performed preserving the topology using a topological method and an error control map to prevent boundary leakage. We tested the proposed approach with 10 datasets and compared it with the latest techniques with random walks and priors. The experiments suggest that this method gives promising results for cochlea segmentation.

Keywords: random walks, shape prior, distance map, segmentation, cochlea

1. INTRODUCTION

Hearing impairment or loss is among the most common reasons for disability. Worldwide, 27% of men and 24% of women above the age of 45 suffer from hearing loss of 26dB or more. The cochlear implant (CI) is a surgically placed device that converts sounds to electrical signals, bypassing the hair cells and directly stimulating the auditory nerve fibers. Even if cochlear implantation is able to restore hearing in patients with severe or complete functional loss, the level of restoration varies highly between subjects and depends on a variety of patient-specific factors.

The HEAR-EU[*] project aims at reducing the inter-patient variability in the outcomes of surgical electrode implantations by improving CI designs and surgical protocols using computational models.[1,2] The computational models are initially created using Statistical Shape Modeling (SSM) techniques[3] on a set of μCT images acquired from cadaveric samples. Once created, the models could then be propagated to clinical CT images using Active Shape Model (ASM) segmentation. This allowed us to provide detailed information of the internal structure of the cochlea in patient-specific scans, were the resolution is normally much lower.[4] In this context, we propose a method that enables an accurate segmentation of the cochlea from μCT. This will benefit model creation to generate accurate patient-specific models, implant design, insertion planning and selection of the best treatment strategy for each patient.

A few methods have been developed for high resolution cochlea segmentation. One work is based on 2D snakes to obtain the cochlea, where a high level of user interaction to locate the initial contour and adjustment of the parameters are required.[5] Another approach is based on active shape models[6] where the segmentation is

---

Further author information: (Send correspondence to Esmeralda Ruiz.)
Esmeralda Ruiz.: E-mail: esmeralda.ruiz@upf.edu, Telephone: (+34)935422173 x 1348
[*]http://www.hear-eu.eu/

performed by first building a high resolution statistical model from several training micro-CT images, which can then be used to predict the position of anatomical structures in CT images.

In this work, we will focus on the Random Walks algorithm. Random walks based image segmentation is a graph-based segmentation method proposed by Leo Grady[7] in 2006. This technique has become very popular because it is able to deal with weak boundaries efficiently, and the extension to 3d and multi-label segmentation is straightforward.[7] According to the author, random walks can outperform the well-known graph cuts[8] in the presence of weak boundaries and do not require any complex technique to be extended to multi-label segmentation unlike graph cuts which usually use complex alpha-beta techniques.

In addition, Grady also extended the method to include unary node information like the data term of graph cuts and added a non-parametric probability density model which allows localization of disconnected objects, and eliminates the requirement for user-specified labels.[9] The author only used the intensity profile to estimate the probability densities. In our paper, we will use this framework to incorporate a shape prior into the random walks formulation. The shape prior information and the region term will constitute the probability density model. The closest works to this paper, are based on graph cuts with prior knowledge[8] and analysis of confidence to decide how strongly priors should be enforced.[10]

The contribution of this paper is to implement a new framework to segment the cochlea in μCT using random walks and a new way to incorporate a prior term using a distance map of the aligned shape prior weighted by a confidence map to adjust its influence according to the strength of the image contour. This technique is more flexible than probabilistic maps or atlas-based methods. The distance prior has been widely used in level sets, watersheds and graph cuts and we would like to apply it in random walks due to its numerous advantages mentioned above. In addition, we add a refinement step preserving the topology of the cochlea to reduce segmentation leaks. In the remainder of this paper, we will explain the details of the proposed method and show the experimental results on μCT images of the cochlea.

2. RANDOM WALKS SEGMENTATION

An image can be represented as a graph where the nodes are the pixels of the image, and the weights represent the similarity between nodes. Nodes marked by the user as seeds are denoted by V_m and the rest by V_u. Given some seeds, $v_j \in V_m$, the random walker assigns to each node, $v_i \in V_u$, the probability, x_i^s, that a random walker starting from that node first reaches a marked node, $v_j \in V_m$ assigned to label g^s. The segmentation is then completed by assigning each free node to the label for which it has the highest probability. The formulation to obtain the probability for each label is as follows:[7]

$$E_{Total}^s = x^{sT} L x^s \tag{1}$$

where L is the combinatorial Laplacian matrix which can be defined as:

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where L_{ij} is indexed by vertices v_i and v_j and d_i is defined as $d_i = \sum_{j=1}^n w_{ij}$. The weight function can be computed as:

$$w_{ij} = \exp(-\beta(g_i - g_j)^2) \tag{3}$$

where g_i is the intensity at pixel i and β is a free parameter related to the bandwidth kernel. The weight range is between 0 and 1 and the higher the weight the larger the similarity between pixels.[11, 12]

An extension to random walks was proposed by incorporating a probability density model based on the gray-level intensity for each label.[9] Let λ_i^s be the probability that the intensity at node v_i belongs to the intensity distribution of label s. The modified random walks formulation is as follows:[9]

$$E_{Total}^s = x^{sT} L x^s + \gamma \left[\sum_{q=1, q \neq s}^{k} x^{sT} \Lambda^q x^s + (x^s - 1)^T \Lambda^s (x^s - 1) \right] \tag{4}$$

where $\Lambda = diag(\lambda^s)$, γ is a free parameter. The first term is the original random walk formulation. The minimization of the functional in Equation 4 is obtained by solving the following system:

$$\left(L + \gamma \sum_{r=1}^{k} \Lambda^r \right) x^s = \gamma \lambda^s \tag{5}$$

In this work, we will use this framework to perform image segmentation but instead of using an intensity-based distribution, we propose a more robust density estimation considering region information as well as shape prior knowledge. We will explain them in detail in the remaining part of the section.

2.1 Data Term Formulation

A histogram is built from the cochlea dataset as there is a good contrast between the cochlea and the surroundings. Then, two gaussian components representing the cochlea and the background are fitted to the histogram. A Gaussian Mixture Model (GMM) estimates a parametric probability density function as a weighted sum of gaussian densities to model background and foreground partitions in an image. The gaussian mixture is defined as:

$$P(x) = \sum_{k=1}^{K} \pi_k N(x | \mu_k, \Sigma_k) \tag{6}$$

where K are the number of gaussian components, μ_k is the mean, Σ_k the covariance and π_k are the mixture coefficients to ensure that a probability distribution holds $\sum_{k=1}^{K} \pi_k = 1$ with $\pi_k \geq 0$. Then, the parameters are estimated by Expectation Maximization (EM) and the region-based term can be defined as:

$$D_i(l_i) = \begin{cases} -\ln P(x_i | O) & \text{if } l_i = object \\ -\ln P(x_i | B) & \text{if } l_i = background \end{cases} \tag{7}$$

where x_i is the pixel indexed by i, l is the label and $P(x_i | O)$ and $P(x_i | B)$ are the probability of pixel at i belonging to object and background intensity, respectively.

2.2 Shape Prior Information

The shape prior will be obtained by applying a non-rigid registration from a random sample to the target object. The procedure follows a common registration work-flow and uses the formulation and framework of `elastix`.[13]

A non-rigid registration of the training dataset, I_M, towards the target dataset, I_F, is formulated as a (parametric) transformation, T_μ, where the vector $\hat{\mu}$ containing the p-parameters of the transformation model are found as an optimization of a cost function, C.

$$\hat{\mu} = \arg \min_{\mu} \mathcal{C}(T_\mu, I_F, I_M) \tag{8}$$

Afterwards the same transformation is applied to the segmentation of the training data, B_M, to obtain an estimate of the target dataset segmentation, B_F, i.e.:

$$B_F = B_M (T_\mu) \tag{9}$$

The transformation model used is the cubic B-spline transform in a multi-resolution setting. Image smoothing is applied with a gaussian kernel to both the fixed and moving image. For each level of resolution the spacing

between B-spline grid points and the width of the smoothing kernel decreases to gradually make the alignment finer.

The cost function used here is defined as:

$$\mathcal{C} = \alpha \cdot \mathcal{S}_{\text{Sim}}(\mu, I_F, I_M) + (1 - \alpha) \cdot \mathcal{P}_{\text{BE}}(\mu) \tag{10}$$

The metric for the image intensity similarity term, $\mathcal{S}_{\text{Sim}}$, is chosen as normalized correlation. Even though the data originate from the same modality the offset and scaling of the intensity-distribution was not the same for all datasets. Bending energy regularization, $\mathcal{P}_{\text{BE}}$, is applied to penalize strong changes and avoid fold-over in the transformations. The weighting parameter, α, between the two terms should be [0,1], and was by experimentation chosen to 0.75.

The optimization is solved using Adaptive Stochastic Gradient Descent. Once the non-rigid transformation is applied using Equation 9 and aligned to the target image, a distance map from the contour of the prior is built. The idea is that given an estimation of the location and shape of the object to segment, pixels close to the shape contour are more likely to be labelled as foreground and viceversa. The formulation can be defined as follows:[8]

$$S_i(\theta) = p(x_i = object|\Theta) = 1 - p(x_i = background|\Theta) = \frac{1}{1 + \exp(\mu \cdot (d(i, \Theta) - d_r))} \tag{11}$$

where $d(i, \Theta)$ is the distance of a pixel i from a shape Θ, being negative for inside the shape and positive for outside the shape. μ determines the ratio of the magnitude of the penalty of points outside the shape compared to the points inside the shape and d_r is the "width" of the influence of the shape.

Then, the distance prior term is

$$S_i(l_i, \Theta) = \begin{cases} p(x = object|\Theta) & \text{if } l_i = object \\ 1 - p(x = object|\Theta) & \text{if } l_i = background \end{cases} \tag{12}$$

2.3 Confidence Map

It may be interesting to adjust the influence of the prior according to the strength of the image contour by reducing the weight of the prior where strong contours are present. To do so, we compute a "confidence map" which represents strong contours with values close to 0 and homogeneous regions with values close to 1. In order to obtain the confidence map, we use the following formulation:[14]

$$c_i = \exp(-k_v \sigma_r^2(i)) \tag{13}$$

where σ_r^2 is the variance at pixel i computed on a patch with radius r and k_v is a free parameter.

2.4 Combining Region and Prior Knowledge Terms

We combine the region and prior terms using a weight for each term where the sum of weights must be one. The formulation is defined as:

$$E_{total}(l_i) = kS_i(l_i, \Theta) + (1 - k)D_i(l_i) \tag{14}$$

where k are the weights chosen empirically and in case the confidence map is required, Equation 14 becomes:

$$E_{total}(l_i) = kS_i(l_i, \Theta) \cdot c_i + (1 - k)D_i(l_i) \cdot \frac{1}{c_i} \tag{15}$$

This Equation 15 is used to obtain λ^s in the random walks formulation in Equation 5 for every label. In order to implement the confidence map efficiently, the map c_i is only considered in a band around the contour of

the current shape prior. This well-known technique is called narrow band.[15] Then, the values of $E_{total}(l_i) > 1$ in the narrow band might be set to 1.
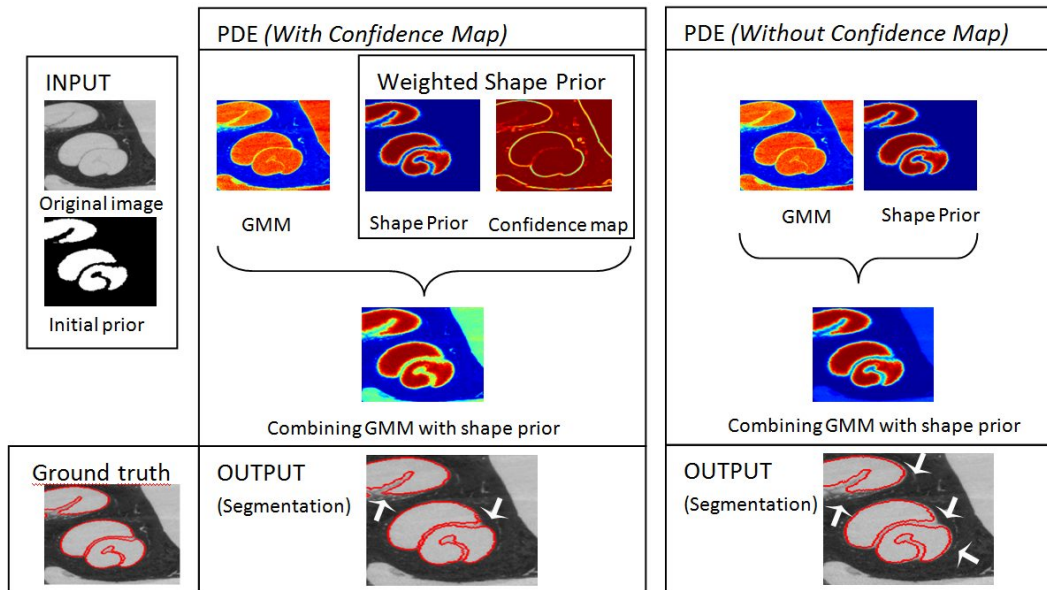

Figure 1. Comparison of using the confidence map (second column) versus not using it (third column).

In our experiments, the confidence map will be computed and integrated into the formulation. The main reason for using the confidence map is shown in Figure 1 where the intensities have been inverted and the image has been cropped around the cochlea for better visualization. In regions where the prior is not reliable around the high gradient areas of the image (edges), the confidence map adjusts the segmentation increasing the region term influence and reducing the prior knowledge strength. Doing this, we might cope with the artifacts of the prior in high gradient regions of the image. As we can see in the Figure 1 where we show the probability distribution estimation (PDE) using region and shape prior terms with and without confidence map, the segmentation adjusts better the borders of the image when the confidence map is used because it does not allow a high displacement from the information obtained by the region term due to a reduction of influence of the prior in these areas.

3. ITERATIVE SEGMENTATION ALGORITHM

To sum up, the proposed algorithm can be described as:

1. The non-rigid registration is performed as described in Section 2.2 to obtain the shape prior.

2. The parametric probability density function is estimated with the Equation 15.

3. The random walks segmentation is performed using the Equation 5.

4. The shape prior is aligned to the current segmentation using affine registration.

5. Go to 2 until convergence or maximum number of iterations is reached.

It is important to see that the weight k for the region term in Equation 15 must be higher than the prior term for the first iterations and the opposite for the last ones. The reason is that we want to give more importance to the region information first to align the shape properly and once aligned, the influence of the prior knowledge is increased. This value k is chosen empirically.

Figure 2. Leakage prevention. (a) Topology preservation where the red arrows indicate that the unit normal vectors of the corresponding points are pointing in opposite direction. Thus, the countour is not allowed to evolve along these points. (b) Error control map prevents leakage, for example the red arrow shows that the contour in this area will not evolve and merge with the other turn of the cochlea and in case there was leakage in the error map in this area, both turns would not merge either, due to the topology preservation method.

Then, a final step of refinement is performed in order to refine the segmentation. The k is reduced to increase the influence of the region but the less contrasted areas will not be affected due to the topology preservation method which is computed using the unit outward normal vector of the contour. When two vectors are pointing in opposite directions, we will not consider it a safe position and these points will not be updated in the narrow band. A smoothing filter might be used such as sgolay filter to reduce noise.[16] In addition, we prevent the leakage in most cases using an error control map which will be considered in order to update the segmentation in the narrow band. To build the error control map, the training data images will be non-rigidly registered to the target image, then a sum of these registrations will be computed and it will be interpreted as a voting map where the positions with small number of "votes" mean that it might be an error in the registration, so below a threshold, we will not consider it a safe position and these points will not be updated in the narrow band. An example of topology and error control map is shown in Figure 2. Summing up, the refinement step applies the weights locally for regions and can be described as:

6. Calculate the band around the contour of width w. The w is chosen empirically.

7. The parametric probability density function is estimated with the Equation 15.

8. The points in the narrow band are updated according to the Equation 15. The rest of points will have a weight of 0 or 1 depending if they are located outside or inside the shape prior, respectively for $E_{total}(l = foreground)$ and the opposite for $E_{total}(l = background)$.

9. The random walks segmentation is performed using the Equation 5.

10. Go to 8 until convergence or the maximum number of iterations for refinement is reached.

4. RESULTS

In this experiment, 10 μCT datasets from the cochlea are used to perform the segmentation in 3D using the proposed method. The original dataset was downsampled to a nominal isotropic resolution of 49μm for efficiency reasons. Every dataset contains around 213 slices with an average size of 413x275 pixels. The ground truth is manually annotated. The prior is obtained by applying a rigid followed by a non-rigid registration to a random sample in a training data. Some weighted multi-atlas method could be used to improve the prior but we choose a random sample to compute the non-rigid registration to show that the method is robust independently of the prior chosen.

The parameters to produce the results are the following: $d_r = 0$, $\mu = 1.0$ in Equation 11, $k_v = 0.5$ in Equation 13 and the total number of iterations are 4 with $k = 0.6$ in Equation 15 for the first 2 iterations and $k = 0.7$ for the rest. The parameters for the refinement step are $k = 0.6$ and the maximum number of iterations is 6.
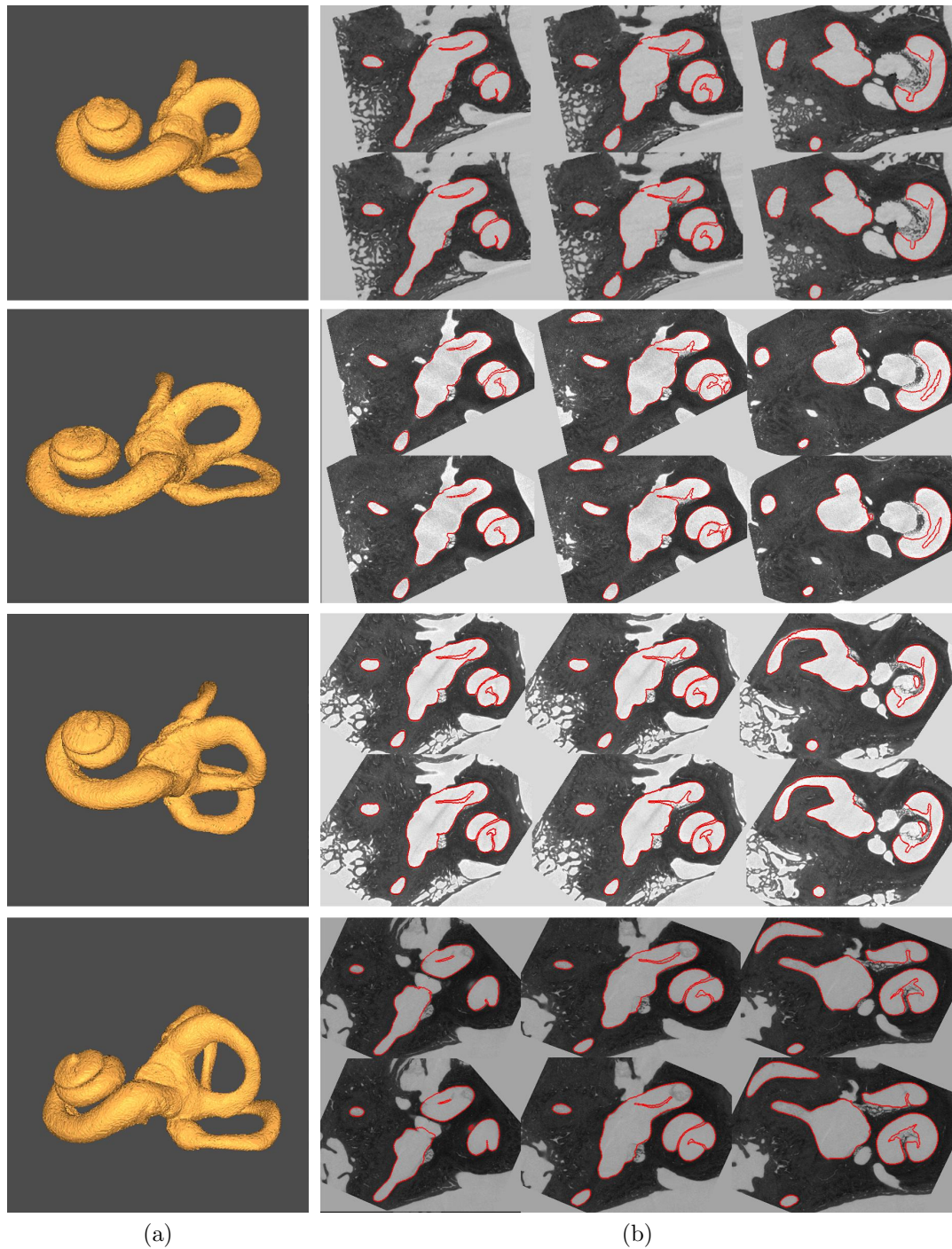
Figure 3. Cochlea segmentation in 4 datasets. (a) Segmentation in 3D. (b) Slices of the 3D segmentation. The top row is the segmentation using the proposed method and at the bottom their corresponding ground truth.

To quantify the segmentation quality assessment for the proposed method we compute the following measures: overlap, sensitivity, specificity and similarity *. Moreover, the proposed method is compared with the latest techniques in random walks for prior incorporation such as guided random walks[17] and constrained random walks[18] as well as the non-rigid registration described in Section 2.2 as a baseline. All parameters are optimized manually for all approaches. The prior for all techniques are obtained as follows. The guided RW use as a prior nine different datasets and their corresponding ground truth segmentations. The guided random walks is performed for every single dataset in the training set and the segmentation with the highest dice score is given. The constrained random walks use a probabilistic map built by averaging nine aligned datasets in the training set and then, registering rigidly to the target object.

	Cochlea segmentation			
	Proposed approach	Registration only	Guided RW	Constrained RW
Overlap	0.94 (0.0082)	0.88 (0.0134)	0.80 (0.0581)	0.65 (0.0402)
Sensitivity	0.95 (0.0125)	0.90 (0.0145)	0.99 (0.0016)	0.73 (0.0502)
Specificity	0.99 (0.0006)	0.99 (0.0003)	0.97 (0.0076)	0.98 (0.033)
Similarity	0.97 (0.0044)	0.93 (0.0074)	0.88 (0.0372)	0.77 (0.0304)

Table 1. Average performance measures for the cochlea, for the proposed approach, registration, guided random walks and constrained random walks. In parenthesis, standard deviation is shown.

Table 1 summarizes the average performance for the cochlea segmentations where we can observe an improvement in all metrics tested except for the sensitivity (fraction of pixels belonging to the cochlea correctly detected) in the guided random walks. This means that guided RW give more points belonging to the cochlea but also some of these points are not correctly detected as shown in the specificity (fraction of pixels not belonging to the cochlea correctly detected) which is lower. Therefore, they do not outperform the results of our method. Our approach gives better overlap and similarity as well as specificity than the other methods giving more accurate segmentations. This is due to the weighted combination of the distance map prior with the region term. In these images, the exterior of the cochlea can be efficiently separated as there is enough contrast between the cochlea and background and the small and invisible regions can be extracted with the guidance of the prior. It is clear that for internal regions, this method relies on the prior but for the exterior of the cochlea, the region term with the prior can provide promising results.

5. CONCLUSION

We presented a new framework for cochlea segmentation in μCT using random walks which is able to deal with weak boundaries efficiently. Some works have been developed for the cochlea segmentation based on snakes and active shape models that respectively require a high level of user interaction and a large amount of training examples. Therefore, we incorporate a distance map prior into the random walks segmentation. Distance map prior has been widely used in level sets and graph cuts but not in random walks. The combination of the distance map prior weighted by the confidence map with a region term into random walks provides accurate segmentations of the cochlea. Then, a refinement step is performed adding a topological preservation method and an error control map to prevent boundary leakage. Moreover, we conducted some experiments and compared with the latest techniques of random walks with prior knowledge. The experiments suggest that the proposed approach is robust and accurate for the cochlea segmentation.

ACKNOWLEDGMENTS

---

*The performance measures are defined as: $Overlap = \frac{TP}{TP+FN+FP}$, $Sensitivity = \frac{TP}{TP+FN}$, $Specificity = \frac{TN}{TN+FP}$ and $Similarity = \frac{2TP}{2TP+FN+FP}$ where TP and FP stand for true positive and false positive and TN and FN for true negative and false negative.

REFERENCES

[1] Ceresa, M., Mangado, N., Dejea, H., Carranza, N., Mistrik, P., Kjer, H. M., Vera, S., Paulsen, R. R., and González Ballester, M. A., "Patient-specific simulation of implant placement and function for cochlear implantation surgery planning," in [*Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*], 49–56, Springer International Publishing (2014).

[2] Ceresa, M., Mangado, N., Andrews, R. J., and González Ballester, M. A., "Computational models for predicting outcomes of neuroprosthesis implantation: the case of cochlear implants," *Molecular neurobiology* **52**(2), 934–941 (2015).

[3] Kjer, H. M., Fagertun, J., Vera, S., González Ballester, M. A., and Paulsen, R. R., "Shape modelling of the inner ear from micro-ct data," in [*Symposium on Statistical Shape Models and Applications (Shape 2014)*], (2014).

[4] Kjer, H. M., Vera, S., Fagertun, J., Perez, F., Herrero, J., González Ballester, M. A., and Paulsen, R. R., "Predicting detailed inner ear anatomy from pre-operational ct for cochlear implant surgery," *Int J Comput Assist Radiol Surg* **10**(Suppl 1), S98–S99 (2015).

[5] Poznyakovskiy, A. A., Zahnert, T., Kalaidzidis, Y., Lazurashvili, N., Schmidt, R., Hardtke, H.-J., Fischer, B., and Yarin, Y. M., "A segmentation method to obtain a complete geometry model of the hearing organ," *Hearing research* **282**(1), 25–34 (2011).

[6] Noble, J. H., Labadie, R. F., Majdani, O., and Dawant, B. M., "Automatic segmentation of intracochlear anatomy in conventional CT," *Biomedical Engineering, IEEE Transactions on* **58**(9), 2625–2632 (2011).

[7] Grady, L., "Random walks for image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **28**(11), 1768–1783 (2006).

[8] Kohli, P., Rihan, J., Bray, M., and Torr, P. H., "Simultaneous segmentation and pose estimation of humans using dynamic graph cuts," *International Journal of Computer Vision* **79**(3), 285–298 (2008).

[9] Grady, L., "Multilabel random walker image segmentation using prior models," in [*Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*], **1**, 763–770, IEEE (2005).

[10] Rao, J., Abugharbieh, R., and Hamarneh, G., "Adaptive regularization for image segmentation using local image curvature cues," in [*Computer Vision–ECCV 2010*], 651–665, Springer (2010).

[11] Pujadas, E. R. and Reisert, M., "Shape-based normalized cuts using spectral relaxation for biomedical segmentation," *Image Processing, IEEE Transactions on* **23**(1), 163–170 (2014).

[12] Ruiz, E. and Reisert, M., "Image segmentation using normalized cuts with multiple priors," in [*SPIE Medical Imaging*], 866937–866937, International Society for Optics and Photonics (2013).

[13] Klein, S., Staring, M., Murphy, K., Viergever, M. A., and Pluim, J. P., "elastix: a toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging* **29**, 196 – 205 (January 2010).

[14] Baudin, P.-Y., Azzabou, N., Carlier, P. G., and Paragios, N., "Prior knowledge, random walks and human skeletal muscle segmentation," in [*Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*], 569–576, Springer (2012).

[15] Xianfen, D., Siping, C., Changhong, L., and Yuanmei, W., "3d semi-automatic segmentation of the cochlea and inner ear," in [*Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*], 6285–6288, IEEE (2006).

[16] Press, W. H., Flannery, B., Teukolsky, S., and Vetterling, W., "14.8 savitzky-golay smoothing filters,," *Numerical Recipes in C: The Art of Scientific Computing (ISBN 0-521-43108-5), Cambridge University Press* **1992**, 650–655 (1988).

[17] Eslami, A., Karamalis, A., Katouzian, A., and Navab, N., "Segmentation by retrieval with guided random walks: Application to left ventricle segmentation in MRI," *Medical image analysis* **17**(2), 236–253 (2013).

[18] Li, K.-C., Su, H.-R., and Lai, S.-H., "Pedestrian image segmentation via shape-prior constrained random walks," in [*Advances in Image and Video Technology*], 215–226, Springer (2012).

[19] Andrews, S., McIntosh, C., and Hamarneh, G., "Convex multi-region probabilistic segmentation with shape prior in the isometric log-ratio transformation space," in [*Computer Vision (ICCV), 2011 IEEE International Conference on*], 2096–2103, IEEE (2011).