



## Random walks with statistical shape prior for cochlea and inner ear segmentation in micro-CT images

Ruiz Pujadas, Esmeralda; Piella, Gemma; Kjer, Hans Martin; González Ballester, Miguel Angel

*Published in:*  
Machine Vision & Applications

*Link to article, DOI:*  
[10.1007/s00138-017-0891-x](https://doi.org/10.1007/s00138-017-0891-x)

*Publication date:*  
2017

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
Ruiz Pujadas, E., Piella, G., Kjer, H. M., & González Ballester, M. A. (2017). Random walks with statistical shape prior for cochlea and inner ear segmentation in micro-CT images. *Machine Vision & Applications*, 1-10. DOI: 10.1007/s00138-017-0891-x

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# Random walks with statistical shape prior for cochlea and inner ear segmentation in micro-CT images

Esmeralda Ruiz Pujadas<sup>1</sup> · Gemma Piella<sup>1</sup> · Hans Martin Kjer<sup>2</sup> · Miguel Angel González Ballester<sup>1,3</sup>

Received: 30 March 2017 / Revised: 12 September 2017 / Accepted: 26 October 2017  
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

## Abstract

A cochlear implant is an electronic device which can restore sound to completely or partially deaf patients. For surgical planning, a patient-specific model of the inner ear must be built using high-resolution images accurately segmented. We propose a new framework for segmentation of micro-CT cochlear images using random walks, where a region term estimated by a Gaussian mixture model is combined with a shape prior initially obtained by a statistical shape model (SSM). The region term can then take advantage of the high contrast between the background and foreground, while the shape prior guides the segmentation to the exterior of the cochlea and to less contrasted regions inside the cochlea. The prior is obtained via a non-rigid registration regularized by a statistical shape model. The SSM constrains the inner parts of the cochlea and ensures valid output shapes of the inner ear.

**Keywords** Inner ear segmentation for micro-CT images · Statistical shape prior · Statistical non-rigid registration · Random walks segmentation

## 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 26 dB or more [1]. The cochlear implant (CI) is a surgically placed device that generates hearing by direct electrical stimulation of 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 [1,2]. The HEAR-EU<sup>1</sup> project aims at reducing the inter-patient variability in the outcomes of surgical electrode implantation by improving CI designs and surgical proto-

cols using computational models. These models are generally built from the segmentations of high-resolution images, in which intra-cochlear structures are visible. In this context, we propose a method that enables an accurate segmentation of the inner ear in micro-CT images which contain the cochlea (in charge of translating vibrations induced by sound into electrical nerve signals sent to the brain) and the semi-circular canals, which are part of the balance system (see Fig. 1).

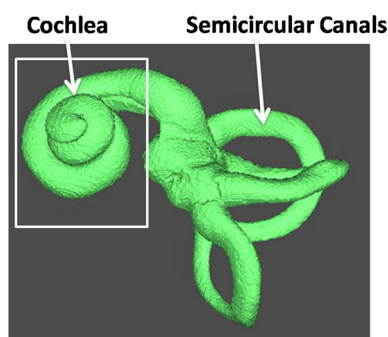
The segmentations play a crucial part in the generation of accurate patient-specific computational models, which can guide implant design, insertion planning and selection of the best treatment strategy for each patient [2,3].

Most high-resolution segmentations of the inner ear are aimed toward the generation of tetrahedral models for simulations. There are a few studies on semi- or fully automatic inner ear segmentation from micro-CT data. However, due to the complexity of the anatomical structure, it is generally a manual procedure. One approach is based on 2D snakes to obtain the cochlea [4], but it requires a high degree of user interaction to locate the initial contour and adjustment of the parameters. Another approach is based on statistical shape models (SSMs) [5], where the high-resolution segmentations are used to build a statistical model and assist the segmentation of low-resolution cochlear images. Using such

<sup>1</sup> <http://www.hear-eu.eu/>.

✉ Esmeralda Ruiz Pujadas  
esmeralda.ruizpujadas@gmail.com

<sup>1</sup> Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain  
<sup>2</sup> Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark  
<sup>3</sup> ICREA, Barcelona, Spain



**Fig. 1** Inner ear scheme showing the cochlea and the semicircular canals which are the auditory and balance parts of the inner ear, respectively

a technique to segment other high-resolution images would require a large amount of processed training data sets to learn the correct anatomical variability and to obtain sufficiently accurate results.

In order to alleviate these issues, Ruiz Pujadas et al. [6] proposed a new algorithm using random walks with a distance-based shape prior, which is robust independently of the chosen prior and which requires no user interaction. Random walks-based image segmentation has become very popular, since it is able to deal with weak boundaries efficiently, and its extension to 3D and to multi-label segmentation is straightforward [7]. There are some techniques to include prior knowledge into random walks. For example, in [8], random walks are constrained by a spatial shape prior model for pedestrian segmentation. A similar work applied to the skeletal muscles was proposed by Baudin et al. in [9]. The segmentations in both works may fail when the average model is too different from the target image or the registration is not very accurate. Therefore, Baudin et al. proposed a new technique by introducing principal components into the random walks [10]. However, the method only yields an approximate solution [11]. In order not to be constrained to the average shape, the guided random walks were proposed [12], where the closest subject in the database is retrieved to guide the segmentation. The weak points of this method are that all the samples of the training data must be considered.

Grady extended the standard random walks segmentation method to include unary node information and added a nonparametric probability density model which allows localizing disconnected objects, and eliminates the requirement for user-specified labels [13]. In this work, we use this framework to incorporate the shape prior into the random walks formulation and segment the cochlea. The main contribution of this paper is the incorporation of the statistical shape model (SSM) via non-rigid registration into the random walks formulation. Additionally, it also employs a confidence map to adapt the influence of the prior according to the gradient of the image. A topology preservation method is also

proposed to avoid leakage in the interior and turns of the cochlea [14]. This paper extends a previous conference communication [15] by presenting a more detailed description of the method, as well as a novel way to regulate the influence of the confidence map in the proposed formulation. We also present novel results, including a thorough analysis of the method as well as a comparison with other similar methods. In the remainder of this paper, we explain the details of the proposed method and show the experimental results on micro-CT images of the inner ear.

## 2 Materials and methods

Random walks algorithm is used to perform the inner ear 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. Vertices 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  $g^s$  as opposed to first reaching a node,  $v_j \in V_m$  with label  $g^{q \neq s}$ . The random walks segmentation is then completed by assigning each free node to the label for which it has the highest probability [7].

An extension was proposed in [13] by incorporating a probability density model based on the gray-level intensity for each label. Let  $\lambda_i^s$  be the probability density that the intensity at node  $v_i$  belongs to the intensity distribution of label  $s$ . The modified random walks formulation is as follows [13]:

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

where the first term is the original random walks formulation,  $\Lambda^s$  is the diagonal matrix with the values of  $\lambda^s$  on the diagonal,  $p$  is the number of labels,  $\gamma$  is a free parameter and  $L$  is the Laplacian matrix defined as:

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

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

$$w_{ij} = \exp(-\beta(I_i - I_j)^2) \quad (3)$$

where  $I_i$  is the intensity at pixel  $i$  and  $\beta$  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 [16,17].

The minimization of the functional in Eq. 1 is obtained by solving the following system:

$$\left( L + \gamma \sum_{r=1}^p \Lambda^r \right) x^s = \gamma \lambda^s. \tag{4}$$

For more details, we refer to [13]. In this work, we 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 given by a SSM. The region term, based on intensities, allows detecting the contours of the cochlea and the statistical shape model allows discarding the areas with the same intensity profile which do not belong to the cochlea. We explain them in detail in the remaining part of the section.

### 2.1 Region term

The region term partitions the image with regard to intensities (bright against dark). One of the partitions contains the inner ear among other non-cochlear structures which have a similar intensity profile (and will be discarded with the shape prior). The procedure to estimate the region term is as follows. Since there is a good contrast between the cochlea and the surroundings, we can represent them by two Gaussian components fitted to the histogram obtained from one slice of the image. 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 the image. The Gaussian mixture is defined as:

$$P(X) = \sum_{k=1}^K \pi_k N(X|\mu_k, \Sigma_k) \tag{5}$$

where  $K$  is the number of Gaussian components,  $\mu_k$  is the mean,  $\Sigma_k$  the covariance and  $\pi_k$  are the mixture coefficients which can be interpreted as prior probabilities, and hence,  $\sum_{k=1}^K \pi_k = 1$  with  $\pi_k \geq 0$ . Then, the parameters are estimated by expectation maximization (EM) [18] and the region-based term can be defined as:

$$D(l_i) = \begin{cases} -\ln p(X_i|O) & \text{if } l_i = \text{object} \\ -\ln p(X_i|B) & \text{if } l_i = \text{background} \end{cases} \tag{6}$$

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 probabilities of pixel at  $i$  belonging to object and background intensity, respectively.

## 2.2 Shape prior knowledge

The shape prior is computed to discard areas that do not spatially belong to the cochlea, but have a similar intensity. The initial shape prior is obtained by performing a non-rigid registration between the target and reference images. We refer to this procedure as a statistical non-rigid registration. The transformation is then applied to the segmented reference image that results in the initial shape prior. From its contour, the distance-based prior is built. The details of this procedure are described in detail in this section.

### 2.2.1 Initial shape prior

The use of a SSM can provide a realistic first prior to initialize the segmentation process, and further be a source of plausible shape regularization.

From a training data set of 17 micro-CT images with corresponding manual ground-truth inner ear segmentations, one data set was chosen as the reference. The 16 remaining data sets are in turn aligned to the reference using the registration model described in [19]. The output of each registration is a vector of deformation parameters (more specifically, a description of a B-Spline deformation field). A PCA on the deformation fields provides a statistical description of the deformation variability in a reduced parameter space.

This SSM models how the reference image,  $I_R$ , is allowed to deform to match another target image,  $I_T$ . In practice, this alignment to the target object regularized by the SSM is performed as a non-rigid image registration between the data sets [20]. The transformation that best aligns the two images,  $T_\eta : I_R \rightarrow I_T$ , is described by the vector  $\eta$  containing  $q$ -parameters, which is found by optimization of a similarity function,  $\mathcal{C}$ :

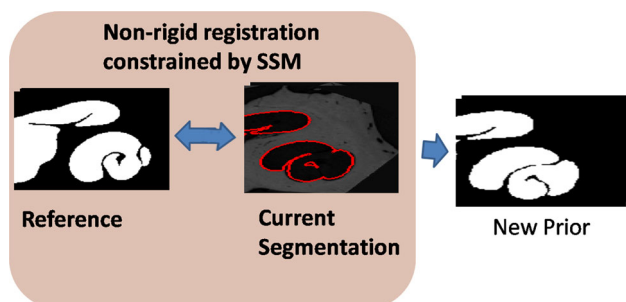
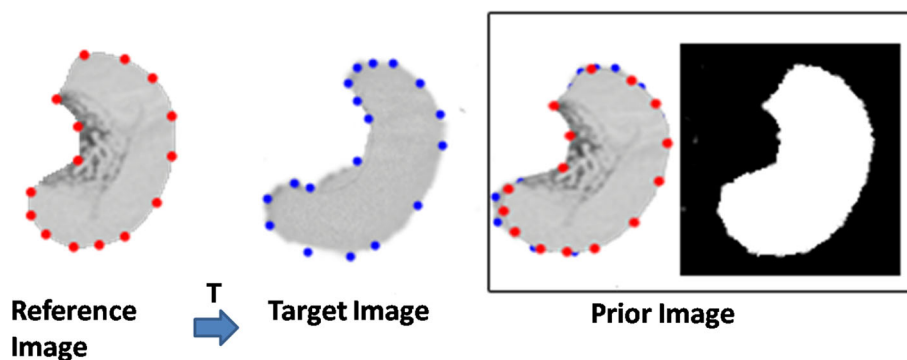
$$\hat{\eta} = \arg \max_{\eta} \mathcal{C} \left( T_\eta^{\text{SSM}}, I_R, I_T \right). \tag{7}$$

The chosen transformation model is the learned SSM—i.e., a B-Spline transformation regularized to be within the allowed deformation variability. The similarity function,  $\mathcal{C}$ , is an image similarity measurement using normalized correlation coefficient. No further regularization terms are needed in the similarity function, since the SSM inherently enforces regularization.

From the statistical non-rigid registration, the deformation between the grayscale reference and target images is applied to the segmentation of the reference data set to obtain the initial shape prior. An example of this procedure is shown in Fig. 2.

In addition to obtaining the initial prior from the statistical non-rigid registration between the grayscale target and reference images, we can use the same procedure to constrain the

**Fig. 2** Construction of the initial shape prior. A non-rigid registration constrained by the SSM finds the transformation that aligns the reference to the target image. This deformation applied to the binary segmentation of the reference image results in the initial prior



**Fig. 3** Updating the shape prior after random walks segmentation. Non-rigid registration constrained by the SSM between the segmented reference and the current segmentation provides the new shape prior

random walks segmentation in every iteration. In this case, the registration is performed between binary segmentation obtained by the random walks and the segmented reference image. In this way, a new prior is obtained, which is constrained to have a valid cochlea shape as shown in Fig. 3.

### 2.2.2 Distance-based shape prior

Once the shape prior regularized by the SSM is obtained, 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 labeled as foreground and vice versa. The formulation can be defined as follows [21]:

$$\begin{aligned} S(l_i = \text{object}, \Theta) &= p(X_i = \text{object}|\Theta) \\ &= 1 - p(X_i = \text{background}|\Theta) \\ &= \frac{1}{1 + \exp(\mu \cdot (d(i, \Theta) - d_r))} \end{aligned} \quad (8)$$

where  $d(i, \Theta)$  is the distance of a pixel  $i$  from a shape  $\Theta$ , being negative inside the shape and positive outside the shape. Here,  $\mu$  is the ratio of points that lie outside the shape to the points inside the shape and  $d_r$  is the “width” of the influence of the shape.

Then, the shape prior term is:

$$S_i(l_i, \Theta) = \begin{cases} p(X_i = \text{object}|\Theta) & \text{if } l_i = \text{object} \\ 1 - p(X_i = \text{object}|\Theta) & \text{if } l_i = \text{background.} \end{cases} \quad (9)$$

### 2.3 Confidence map

It may be interesting to adjust the influence of the prior according to the strength of the image contour (i.e., reducing the weight of the prior where strong contours are present). To do so, we compute a “confidence map” that represents strong contours with values close to 0 and homogeneous regions with values close to 1. The confidence map is obtained using the formulation:

$$c_i = \exp(-k_v \sigma_r^2(i)) \quad (10)$$

where  $\sigma_r^2(i)$  is the variance at pixel  $i$  computed on a patch with radius  $r$  and  $k_v$  is a free parameter. Using this formulation, we obtain a map with a small weight close to 0 in high-gradient areas of the image (i.e., strong edges) and the opposite in flat regions. This map is used to weight the contribution of the region and shape prior as explained in the next section. The influence of the confidence map is regularized by a weighted band around the contour of the prior minimizing any leakage or artifacts in the segmentation.

### 2.4 Combining region and prior knowledge terms

The shape prior and region terms are combined as:

$$E_{\text{total}}(l_i) = kS(l_i, \Theta) \cdot c_i + (1 - k)D(l_i) \cdot \frac{1}{c_i} k_c \quad (11)$$

where  $k$  balances the contribution of the shape and region term and  $k_c$  regulates the influence of the confidence map with respect to the shape prior. This weight  $k_c$  is defined only in a band around the contour of the prior (narrow band). The higher the value, the higher the region strength on edges is.

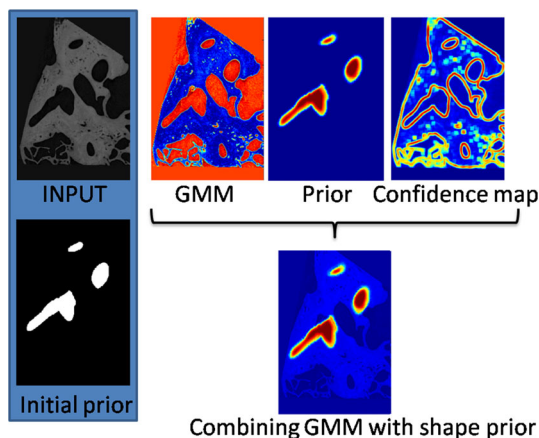


Fig. 4 Combination of the region term given by a GMM and the distance-based shape prior weighted by the confidence map

Adding this narrow band with a corresponding weight  $k_c$ , the influence of the confidence map can be better adjusted around the prior and the rest of the image is not affected minimizing leakage and other artifacts. This energy function combines the region and shape information, which are incorporated into the random walks formulation (using  $E_{total}$  normalized between 0 and 1) in the  $\lambda^s$  in Eq. 1.

An example of the combination of both terms is illustrated in Fig. 4. Given the original image and the initial prior provided by the SSM, the Gaussian mixture model and distance-based shape prior are computed weighted by a confidence map, and then, both terms are combined using Eq. 11. For simplicity, the values of the region term close to 0 will be directly discarded in the segmentation, having an  $E_{total}(l_i = object) = 0$  as this is always satisfied.

### 2.5 Iterative segmentation algorithm

To sum up, the pipeline of the algorithm, which is illustrated in Fig. 5, is as follows:

1. A statistical non-rigid registration of the target image with the grayscale reference data set is performed as described in Sect. 2.2.1 and the distance-based prior is computed with Eq. 8.
2. The region term is estimated with Eq. 6 and combined with the distance-based prior in Eq. 11.
3. The random walks segmentation is performed using Eq. 4.
4. The current segmentation is statistically non-rigidly registered with the segmented reference data set to obtain a new prior constrained by the SSM.
5. Go to step 2 until convergence or until the maximum number of iterations is reached.

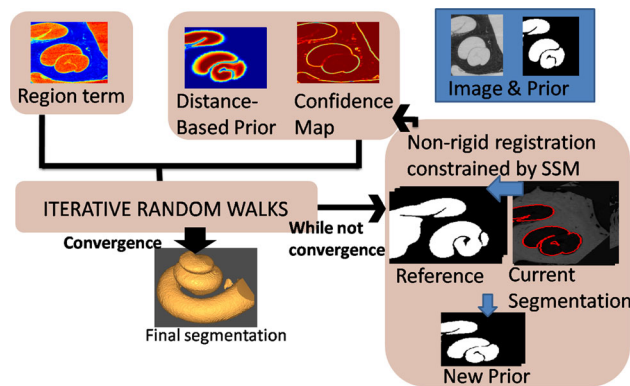


Fig. 5 Pipeline showing the main steps of the proposed algorithm. The SSM recovers the missing parts of the segmented image obtaining a valid shape of the cochlea

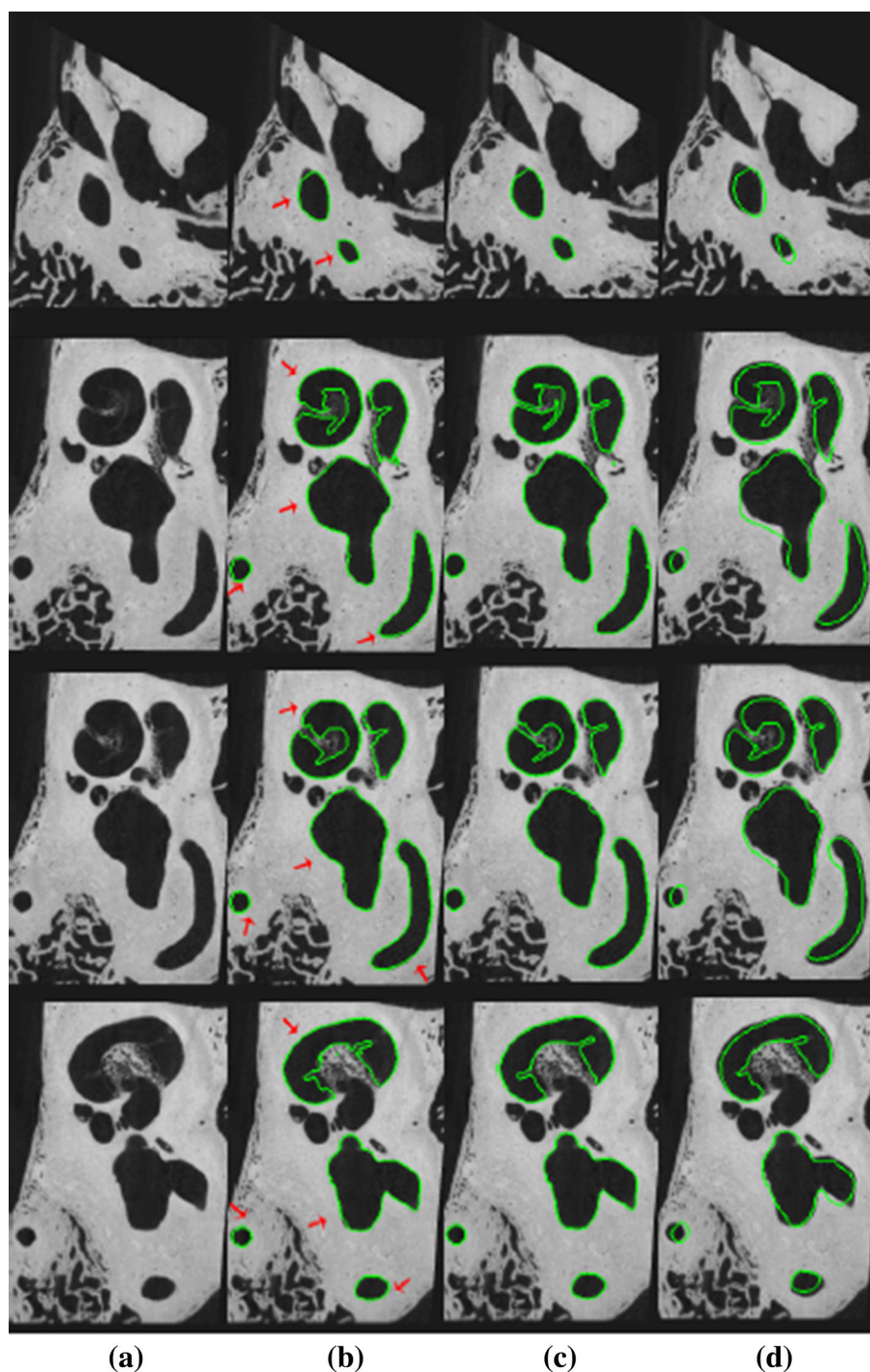
In order to avoid merging in the non-contrasted areas of the inner ear (e.g., turns of the cochlea and the semicircular canals), a topology preserving method is proposed using the unit outward normal vector of the contour. When two vectors within certain distance are pointing in opposite directions, the contour around those points (windows) is not allowed to evolve along the segmentation contour to prevent leakage. This approach detects the interior and different turns of the cochlea and prevents them from merging, whereas the exterior of the cochlea can evolve around the prior until it converges. We refer to [14] for more details on the topology preservation method for the cochlea.

### 3 Experiment and results

The proposed approach was tested in 10 micro-CT data sets of the inner ear where each ear belongs to a unique subject. Every data set contains around 213 slices with an average size of  $413 \times 275$  pixels. The original 3D data set was downsampled from a nominal isotropic resolution of  $24.5\text{--}49 \mu\text{m}$  for computational efficiency reasons. We use 17 labeled data sets to build the SSM. The execution time for the statistical non-rigid registration is on average around 133 s. However, the time can be reduced by reducing the number of iterations performed during the registration. The computational time of the random walks segmentation with the calculations of the weights of the graph as well as the region and shape prior terms is around 41 s in average including the execution time of solving the random walks system given by Eq. 1 which has a mean running time of 12 s.

Figure 6 illustrates some inner ear segmentation results using our approach. It is shown that the segmentation corrects inaccuracies of the prior adjusting the contour to the high-gradient areas of the image (edges) using the confi-

**Fig. 6** Inner ear segmentation. **a** Original image. **b** Final segmentation. Arrows show the improvement w.r.t. the baseline segmentation (column **d**). **c** Ground truth. **d** Baseline segmentation (i.e., using only the initial prior without the distance-based prior)



dence map. The interior of the cochlea and the non-contrasted regions are preserved thanks to the shape prior and the topology preservation method which does not allow leakage in these areas. The arrows in Fig. 6 show areas where the segmentation is more adjusted to the contour of the image than the initial prior, giving satisfactory segmentation results. The segmentation of three different data sets is also shown in Fig. 7. In this case, the 3D segmentation as well as the

corresponding slices are illustrated. In this example, we can observe from the 3D volume that the topology of the cochlea shape is preserved and that the contour of the segmentation is adjusted to the edges of the image, whereas the interior of the cochlea and less contrasted areas are conserved.

To quantify the segmentation quality for the proposed method, we compute the following measures: overlap, sensi-

**Fig. 7** Inner ear segmentation. **a** Segmentation in 3D. **b** Slices of the 3D segmentation. **c** Ground truth



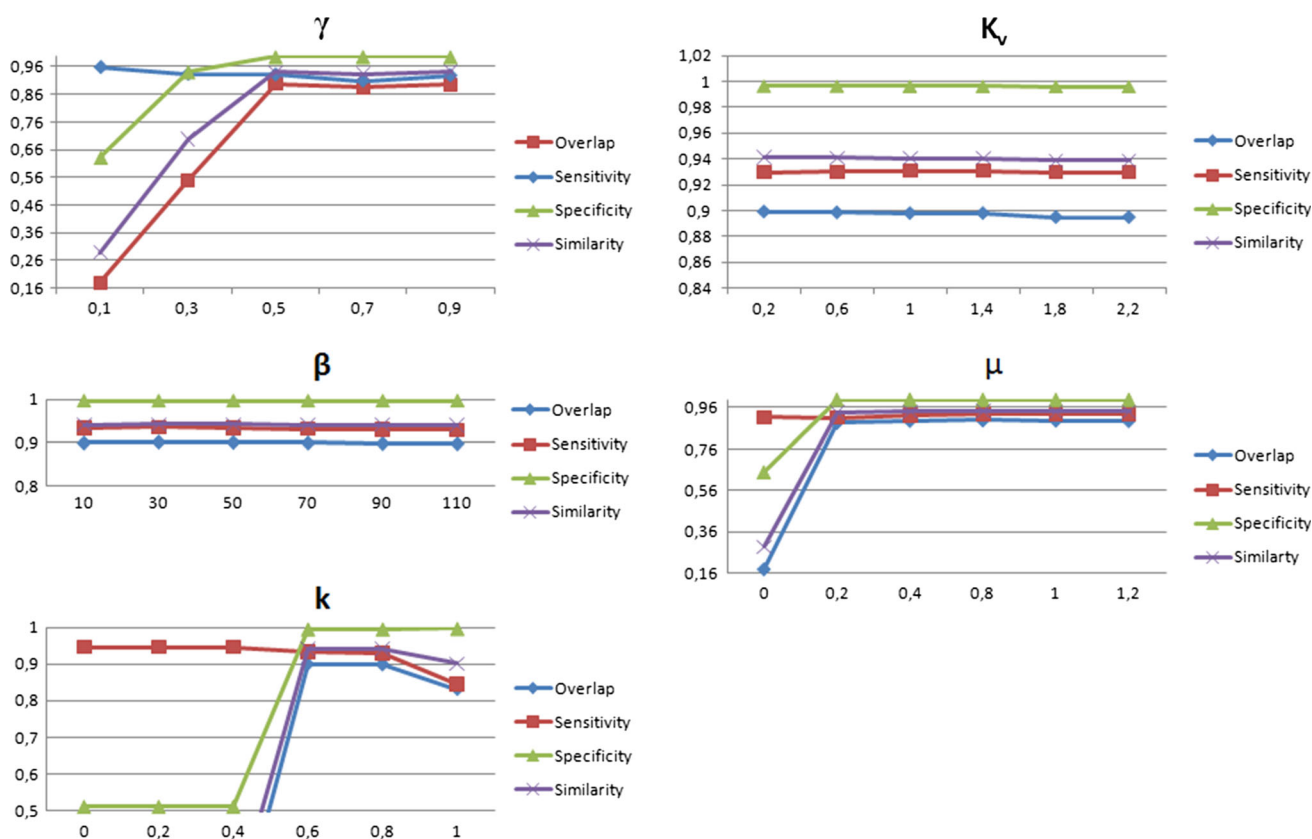
tivity, specificity and similarity.<sup>1</sup> The following parameters were used for the proposed approach to produce the results:  $\beta = 90$  in Eq. 3,  $\gamma = 0.8$  in Eq. 4,  $d_r = 0$  and  $\mu = 1.0$  in Eq. 8,  $k_v = 0.5$  in Eq. 10 and the total number of iterations are 4 with  $k = 0.8$  in Eq. 11.

To find some guidance for setting the parameters, the sensitivity of the method regarding these parameters was

<sup>1</sup> 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.

investigated. We found that the results were quite stable throughout a wide range of values and that the same behavior was obtained for all data sets. This shows that this method is not very sensitive to how they are chosen. The study of parameters shown in Fig. 8 indicates the segmentation performance for all data sets as a function of each parameter. Each plot is obtained by varying one parameter, while the remaining were fixed with the empirically chosen values. According to this study, the possible ranges of values for the algorithm to perform well could be:  $\beta \in [10 \dots 110]$ ,  $\gamma \in [0.5, \dots, 0.9]$ ,  $k_v \in [0.2, \dots, 2.2]$  and  $\mu \in [0.2, \dots, 1.2]$ . The improvement in the method when  $\gamma$  increases shows the improvement in





**Fig. 8** Robustness of the proposed segmentation method with respect to the parameters  $\gamma$ ,  $k_v$ ,  $\beta$ ,  $\mu$  and  $k$

using the probability distribution with respect to the standard random walks. When  $\gamma$  parameter takes a small value, the random walks behavior of segmenting areas with similar intensity profiles accentuates creating leakage in areas that do not belong to the cochlea. When the  $\gamma$  value increases, the influence of the prior also increases (provided that the influence of the prior is high with  $k \geq 0.7$ ) discarding the areas which do not belong to the inner ear. The most sensitive parameter found was  $k$  which is very simple to select by choosing a higher influence of the prior than the region term (e.g.,  $k \geq 0.7$ ) in order to avoid leakage in areas with the same intensity as the cochlea but avoiding the highest values (e.g.,  $k \geq 0.9$ ) as the result will be more similar to the prior shape as expected, including possible artifacts that the prior can have.

The proposed method is compared with the latest techniques in random walks for prior incorporation such as guided random walks [12] and constrained random walks [8]. It is also compared with the segmentations obtained from the statistical shape model algorithm (initial prior) and an atlas-based non-rigid registration, implemented in [6], as baselines. All parameters are optimized manually for all approaches. The prior for all techniques is obtained as follows. The proposed approach obtains the initial prior by deforming the segmentation of the reference data with the

deformation field obtained by the statistical non-rigid registration. The guided random walks use as a prior nine different data sets and their corresponding ground-truth segmentations. The guided random walks are performed for every single data set 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 data sets in the training set and then registered rigidly to the target object.

Table 1 summarizes the average performance for the inner ear segmentations. We can observe that the SSM algorithm attains a lower performance than the atlas-based method due to the limited number of training data sets. However, the results of the SSM are constrained to a valid cochlear shape, an important feature for further analysis. Our approach gives better overlap, similarity and specificity than the other methods while maintaining the advantages of using a statistical model. This is due to the weighted combination of the distance map prior with the region term. The table also shows the benefit of using the confidence map. This map increases the influence of the energy term in high-gradient areas of the image around the prior and helps the adjustment of the contour to the exterior of the cochlea. As this influence on the prior is constrained within a narrow band, it only affects areas close to the prior mini-

**Table 1** Average performance measures for the inner ear using the following methods: our approach with confidence map (CM), our approach without CM, an atlas-based non-rigid registration, SSM used to obtain

the initial prior, guided random walks (guided RW) and constrained random walks (constrained RW)

	Inner ear segmentation					
	Proposed with CM	Proposed without CM	Atlas	SSM	Guided RW	Constrained RW
Overlap	<b>0.90</b> (0.0094)	0.86 (0.021)	0.88 (0.0134)	0.77 (0.0216)	0.80 (0.0581)	0.65 (0.0402)
Sensitivity	0.93 (0.0105)	0.88 (0.0252)	0.90 (0.0145)	0.86 (0.0388)	<b>0.99</b> (0.0016)	0.73 (0.0502)
Specificity	<b>0.99</b> (0.0014)	0.99 (0.0007)	0.99 (0.0003)	<b>0.99</b> (0.0024)	0.97 (0.0076)	0.98 (0.033)
Similarity	<b>0.94</b> (0.0052)	0.92 (0.0117)	0.93 (0.0074)	0.87 (0.0137)	0.88 (0.0372)	0.77 (0.0304)

The standard deviation is shown in parentheses. The best value for each measure is highlighted in bold

mizing or avoiding any possible artifact or leakage in other structures. The interior part of the cochlea, which is composed of the non-contrasted areas, is recovered by the prior through the SSM, which is updated iteratively enforcing a valid shape of the cochlea. Thus, it is clear that for internal regions, this method relies on the statistical prior which constrains the shape of these areas, and for the exterior of the cochlea, the region term with the prior provides promising results even when the prior shape has some artifacts.

## 4 Discussion

We presented a new framework for the inner ear segmentation in micro-CT using random walks which is able to deal with weak boundaries efficiently. The region term can then take advantage of the high contrast between the background and foreground, while the shape prior guides the segmentation to the exterior of the cochlea and to less contrasted regions inside the cochlea. In this work, the statistical shape model segmentation is performed using a statistical non-rigid registration, which allows constraining the inner ear to a valid shape.

We evaluated the accuracy of our segmentation results given by the proposed approach and compared it with the latest techniques of random walks with prior knowledge. The experiments suggest that the proposed approach is robust and accurate for the inner ear segmentation in micro-CT.

Despite the limited sample size, the resulting segmentations are accurate. Moreover, the same data have been used for further analysis [1–3] where it has shown to be representative of the different patient variations. Thus, we believe that the proposed method will generalize well to other data sets. Note that this type of ex vivo data are scarce, difficult to get and very time-consuming to segment manually.

As future work, it would be interesting to investigate the use of this method on non-regular grids such as proposed in

[22]. This would allow reducing the computational complexity of the proposed methods.

**Acknowledgements** The research leading to these results received funding from the European Union Seventh Frame Programme (FP7/2007-2013) under Grant agreement 304857.

## Compliance with ethical standards

**Conflict of interest** E.Ruiz, H.M Kjer, G.Piella and M.A González declare that they have no conflict of interest.

**Human and animal rights** All human and animal studies have been approved.

**Informed consent** All patients/volunteers gave their informed consent.

## References

1. Ceresa, M., Lopez, N.M., Velardo, H.D., Herrezuelo, N.C., Mistrik, P., Kjer, H.M., Vera, S., Paulsen, R.R., Ballester, M.A.G.: Patient-specific simulation of implant placement and function for cochlear implantation surgery planning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, vol. 8675, pp. 49–56. Lecture Notes in Computer Science. Springer, Berlin (2014)
2. Ceresa, M., Mangado, N., Andrews, R.J., Ballester, M.A.G.: Computational models for predicting outcomes of neuroprosthesis implantation: the case of cochlear implants. *Mol. Neurobiol.* **52**(2), 934–941 (2015)
3. Mangado, N., Ceresa, M., Duchateau, N., Kjer, H.M., Vera, Sergio, Velardo, Hector Dejea, Mistrik, Pavel, Paulsen, R.R., Fagertun, J., Noailly, J., Piella, G., González, M.A.: Automatic model generation framework for computational simulation of cochlear implantation. *Ann. Biomed. Eng.* **44**, 2453–2463 (2016)
4. Poznyakovskiy, A.A., Zahnert, T., Kalaidzidis, Y., Lazurashvili, N., Schmidt, Rolf, Hardtke, Hans-Jürgen, Fischer, Björn, Yarin, Yury M.: A segmentation method to obtain a complete geometry model of the hearing organ. *Hear. Res.* **282**(1), 25–34 (2011)
5. Noble, J.H., Labadie, R.F., Majdani, O., Dawant, B.M.: Automatic segmentation of intracochlear anatomy in conventional CT. *IEEE Trans. Biomed. Eng.* **58**(9), 2625–2632 (2011)
6. Pujadas, E.R., Kjer, H.M., Piella, G., Ceresa, M., Ballester, M.A.G.: Random walks with shape prior for cochlea segmentation in ex vivo  $\mu$ CT. *Int. J. Comput. Assist. Radiol. Surg.* **11**(9), 1647–1659 (2016)

7. Grady, Leo: Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(11), 1768–1783 (2006)
8. Li, K.-C., Su, H.-R., Lai, S.-H.: Pedestrian image segmentation via shape-prior constrained random walks. In: *Advances in Image and Video Technology*, pp. 215–226. Springer, Berlin (2012)
9. Baudin, P.-Y., Azzabou, N., Carlier, P.G., Paragios, N.: Prior knowledge, random walks and human skeletal muscle segmentation. In: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2012*, pp 569–576. Springer, Berlin (2012a)
10. Baudin, P.-Y., Azzabou, N., Carlier, P.G., Paragios, N.: Manifold-enhanced segmentation through random walks on linear subspace priors. In: *Proceedings of the British Machine Vision Conference (2012b)*
11. Baudin, P.-Y.: De la segmentation au moyen de graphes d'images de muscles striés squelettiques acquises par RMN. Ph.D. thesis, Ecole Centrale Paris (2013)
12. Eslami, A., Karamalis, A., Katouzian, A., Navab, N.: Segmentation by retrieval with guided random walks: application to left ventricle segmentation in MRI. *Med. Image Anal.* **17**(2), 236–253 (2013)
13. Grady, L.: Multilabel random walker image segmentation using prior models. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 1, pp. 763–770. IEEE (2005)
14. Pujadas, E.R., Kjer, H.M., Vera, S., Ceresa, M., Ballester, M.A.G.: Cochlea segmentation using iterated random walks with shape prior. In: *SPIE Medical Imaging*, pp. 97842U–97842U. International Society for Optics and Photonics (2016b)
15. Pujadas, E.R., Kjer, H.M., Piella, G., Ballester, M.A.G.: Statistical shape model with random walks for inner ear segmentation. In: *SESAMI, International Conference on Medical Image Computing and Computer-Assisted Intervention. Lecture Notes in Computer Science*. Springer, Berlin (2016a)
16. Ruiz, E., Reisert, M.: Image segmentation using normalized cuts with multiple priors. In: *SPIE Medical Imaging*, pp. 866937–866937. International Society for Optics and Photonics (2013)
17. Pujadas, E., Reisert, M.: Shape-based normalized cuts using spectral relaxation for biomedical segmentation. *IEEE Trans. Image Process.* **23**(1), 163–170 (2014)
18. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. soc. Ser. B Methodol.* **39**, 1–38 (1977)
19. Kjer, H.M., Fagertun, J., Vera, S., Gil, D., González Ballester, M.Á., Paulsen, R.R.: Free-form image registration of human cochlear  $\mu$ CT data using skeleton similarity as anatomical prior. *Pattern Recognit. Lett.* **76**, 76–82 (2016)
20. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, Josien P.W.: elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* **29**(1), 196–205 (2010)
21. Kohli, P., Rihan, J., Bray, M., Torr, P.H.S.: Simultaneous segmentation and pose estimation of humans using dynamic graph cuts. *Int. J. Comput. Vis.* **79**(3), 285–298 (2008)
22. Bampis, C.G., Maragos, P., Bovik, A.C.: Graph-driven diffusion and random walk schemes for image segmentation. *IEEE Trans. Image Process.* **26**(1), 35–50 (2017)