

Integrating Statistical Shape Models into a Graph Cut Framework for Tooth Segmentation

Johannes Keustermans¹, Dirk Vandermeulen¹, and Paul Suetens²

¹ KU Leuven, Faculty of Engineering, ESAT - PSI, Leuven, Belgium

² IBBT-KU Leuven Future Health department, Leuven, Belgium

johannes.keustermans@esat.kuleuven.be

Abstract. The segmentation of teeth is of great importance for the computer aided planning of dental implants, orthodontic treatment, and orthognathic surgery. However, it is hampered by metallic streak artifacts present in Computed Tomography (CT) images in general, and the lack of contrast between the teeth and bone in Cone-Beam CT (CBCT) images particularly. Therefore, we propose a novel graph cut based algorithm that effectively integrates a statistical shape model based on a probabilistic shape representation. The statistical shape model is obtained from a set of training samples and imposes a Gaussian distribution on the shape space. The presented algorithm minimises an energy function that is formulated according to a maximum a posteriori criterion and consists of three terms: an image likelihood term, a segmentation likelihood term integrating the shape model into the graph cut framework, and a shape model term favoring shapes that are more likely according to the statistical shape model.

1 Introduction

Recently, three-dimensional orthognathic surgery and dental implant planning software systems became available, enabling visualization, quantification, non-invasive diagnosis, treatment planning, and evaluation of treatment outcome in an unprecedented way. The introduction of Cone-Beam Computed Tomography (CBCT) has instigated a breakthrough towards the routine use of these three-dimensional treatment planning software systems, due to the low radiation dose, unique accessibility and low cost. The segmentation of teeth from CBCT images is of particular interest for these software systems as it may significantly broaden or help their applicability, e.g. virtual tooth extraction, dental implant planning, and orthodontic treatment planning and evaluation. Yet, the segmentation of teeth from CBCT images is challenging problem, as it is hampered by various factors. First, the presence of metal streak artifacts in the CBCT images, caused by orthodontic braces or dental fillings. Next, since the teeth are anchored in the jaw bone there is only little contrast between the bone and the teeth, predominantly at the level of the apex. Subsequently, the signal to noise ratio of CBCT images is in general lower compared to CT images. Finally, the shape of teeth shows a significant variability over individuals. In the literature a number of methods for the segmentation of teeth have been proposed [1,2]. To our knowledge, however, none of these methods are capable of coping with the mentioned problems.

Since a manual segmentation is very time-consuming and subjective, hindered even more by the dimensionality and size of the CBCT images, as well as the number of teeth, there is a strong need for an automated or semi-automated approach. In the last decades the computer vision community has produced a large variety of image segmentation algorithms. Earlier approaches to the segmentation problem are based on heuristic rules. Despite being not robust, these methods are still widely known and used in practice due to their simplicity, predictability, and speed. More recently, optimization methods have become established as being more powerful and mathematically sound. These algorithms minimize an appropriate energy function, leading to an optimal image segmentation in some sense. The Bayesian formulation of the energy function allows to introduce prior shape knowledge into the image segmentation framework. To make the algorithm as generic as possible, a statistical learning approach can be used, in which the prior shape knowledge can be obtained from training data.

This paper presents an algorithm that integrates a statistical shape model into the graph cut framework. The graph cut algorithm defines a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ consisting of a set of nodes \mathcal{V} , representing image voxels, and a set of edges \mathcal{E} connecting neighboring nodes. Two extra terminal nodes s (*source*) and t (*sink*) are added, representing the object and background label. All edges $e \in \mathcal{E}$ are assigned some non-negative weight w_e . The goal of the graph cut algorithm is the optimal separation of source and sink by slicing a set of edges that no remaining path exists between the source and the sink. The cost of a cut \mathcal{C} is defined by the sum of the weights of the edges sliced by the cut. This cut can be computed efficiently in low order polynomial time using the max-flow/min-cut algorithm [3].

Integrating statistical shape models into the graph cut framework is not straightforward. Malcolm et al. [4] present an interactive graph cut framework employing statistical shape models based on an implicit shape representation. Freedman et al. [5] integrate a deformable template based on an implicit representation into the interactive graph cut framework. Although not based on graph cuts, Schoenemann et al. [6] present a model-based segmentation algorithm providing global minima. First, however, only a deformable template is used compared to the statistical shape models used in this work. Second, the algorithm is inherently two dimensional, so extending the method to higher dimensions is not straightforward.

Compared to previous work we have chosen to integrate a statistical shape model based on a probabilistic shape representation reflecting the probability of a point to belong to the object boundary, originally presented by Hufnagel et al. [7], into the graph cut framework. As such, we retain advantages of both explicit and implicit shape representations combined with an effective optimization algorithm. The segmentation algorithm presented in this paper is applied to the segmentation of teeth from CBCT images. The organization of this paper is as follows. Section 2 presents the statistical shape model. Section 3 further details the segmentation algorithm itself. Subsequently, section 4 discusses some experiments and results. Finally, section 5 formulates a conclusion and some ideas for future work.

2 Statistical Model Building

This section presents the procedure for constructing the statistical shape models. This procedure was originally presented by Hufnagel et al. [7], and is extended here. A clear distinction is made between the observation parameters and the model parameters. Since the proposed statistical shape model imposes a Gaussian distribution on the shape space, the set of model parameters Θ consists of the mean shape $\bar{M} = \{\bar{\mathbf{m}}_j\}_{j=1}^{N_m}$, the eigenmodes $v_p = \{\mathbf{v}_{pj}\}_{j=1}^{N_m}$, the eigenvalues λ_p and the number of eigenmodes n . The observation parameters $Q = \{Q_k\}_{k=1}^N$ are the bandwidth parameters σ_k , the rigid transformations $T_k = \{\mathbf{R}_k, \mathbf{t}_k\}$, and the deformation coefficients $W_k = \{w_{kp}\}_{p=1}^n$ with respect to the eigenmodes that fit the statistical shape model to each of the training samples $\mathcal{S}_k = \{\mathbf{s}_{ki}\}_{i=1}^{N_k}$. Unlike Hufnagel et al., the bandwidth parameters are training sample specific, and included with the observation parameters. As such, their values are estimated in an optimal and automated manner. In order to optimize both model and observation parameters a maximum a posteriori criterion is formulated

$$p(Q, \Theta | \mathcal{S}) = p(\Theta) \prod_{k=1}^N \frac{p(\mathcal{S}_k | Q_k, \Theta) p(Q_k | \Theta)}{p(\mathcal{S}_k)}. \quad (1)$$

The term $p(Q_k | \Theta)$ can be further expressed according to the Gaussian shape model and by assuming a constant prior for the rigid transformations. Furthermore, the term $p(\mathcal{S}_k | Q_k, \Theta)$ can be further expressed by imposing the probabilistic object representation based on a kernel density estimator with Gaussian kernels

$$p(\mathcal{S}_k | Q_k, \Theta) = \prod_{i=1}^{N_k} p(\mathbf{s}_{ki} | Q_k, \Theta) = \prod_{i=1}^{N_k} \sum_{j=1}^{N_m+1} p(j) p(\mathbf{s}_{ki} | j, Q_k, \Theta), \quad (2)$$

where $p(j)$ are equal membership probabilities, and $p(\mathbf{x} | N_m + 1, Q_k, \Theta) = \frac{1}{N_m}$ is a uniform (pseudo) outlier distribution. Unlike Hufnagel et al., an outlier distribution is added to improve the robustness against outliers. The term $p(\mathbf{x} | Q_k, \Theta)$ can be further expanded as,

$$p(\mathbf{x} | Q_k, \Theta) = \frac{1 - \omega}{N_m} \sum_{j=1}^{N_m} \frac{\exp\left(-\frac{\|\mathbf{s}_{ki} - \mathbf{R}_k \mathbf{m}_{kj} - \mathbf{t}_k\|^2}{2\sigma_k^2}\right)}{(2\pi\sigma_k^2)^{d/2}} + \frac{\omega}{N_m}, \quad (3)$$

where $0 \leq \omega \leq 1$ is parameter balancing the relative importance of the (pseudo) outlier distribution. Furthermore, \mathbf{m}_{kj} can be expanded as a linear combination of the principal components and the mean, $\mathbf{m}_{kj} = \bar{\mathbf{m}}_j + \sum_{p=1}^n w_{kp} \mathbf{v}_{pj}$.

Assuming $p(\Theta)$ to be uniform and taking the negative logarithm yields the energy function to be minimized. Unlike Hufnagel et al., the energy function is optimized using the EM algorithm [8]. In the E-step an optimal upper bound to the energy function is formulated, by applying Jensen's inequality and introducing specific probability distributions $q_k(i, j)$, representing probabilistic correspondences. The probabilistic correspondences $q_k(i, j)$ are determined as to optimize the upper bound that it touches the energy function using the current model Θ^t and observation parameters Q^t and adding

the constraints $\sum_{i=1}^{N_k} q_k(i, j) = 1$ and $\sum_{j=1}^{N_m+1} q_k(i, j) = 1$. As such, the E-step consists of iterating the following equations until convergence

$$\hat{q}_k^{t+\frac{1}{2}}(i, j) = \frac{\hat{q}_k^t(i, j)}{\sum_{j=1}^{N_m} \hat{q}_k^t(i, j) + \frac{\omega(2\pi\sigma_k^2)^{d/2}}{1-\omega}} \quad \text{and} \quad \hat{q}_k^{t+1}(i, j) = \frac{\hat{q}_k^{t+\frac{1}{2}}(i, j)}{\sum_{j=1}^{N_m} \hat{q}_k^{t+\frac{1}{2}}(i, j)}, \quad (4)$$

where $\hat{q}_k^0(i, j) = \exp\left(-\frac{\|\mathbf{s}_{ki} - \mathbf{m}_{kj}\|^2}{2\sigma_k^2}\right)$. This approach is similar to the softassign approach of Chui et al. [9], and leads to more accurate probabilistic correspondences compared to Hufnagel et al. [7]. In the M-step the obtained upper bound is minimized with respect to the observation parameters Q and the model parameters Θ while keeping, respectively, Θ and Q fixed. This upper bound can be formulated as

$$C_k^M(Q_k, \Theta) \simeq \frac{1}{2\sigma_k^2} \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} q_k(i, j) \|\mathbf{s}_{ki} - \mathbf{R}_k \mathbf{m}_{kj} - \mathbf{t}_k\|^2 + \frac{N_{qk}d}{2} \log(\sigma_k^2) + \sum_{p=1}^n \left(\log(\lambda_p) + \frac{w_{kp}^2}{2\lambda_p^2} \right), \quad (5)$$

where $N_{qk} = \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} q_k(i, j)$. Closed-form solutions for almost all parameters can be obtained. For the rotation matrix \mathbf{R}_k and translation vector \mathbf{t}_k , the expressions are similar to the rigid coherent point drift algorithm [10]. For the bandwidth parameter σ_k the following expression can be derived

$$\sigma_k = \sqrt{\frac{1}{N_{qk}d} \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} q_{ki}(j) \|\mathbf{s}_{ki} - \mathbf{m}_{kj}\|^2}. \quad (6)$$

For the deformation coefficients, as well as all model parameters we refer to Hufnagel et al. [7], since similar expressions are obtained, differing only in the probabilistic correspondences $q_k(i, j)$.

3 Image Segmentation

This section will provide details of the energy function and the optimization thereof. In the following paragraphs we will formulate this energy function and derive expressions for the edge weights in the graph cut framework. The objective of the presented algorithm is the optimal segmentation of teeth from CBCT images according to some criterion. This criterion is formulated through a maximum a posteriori probability formulation of the segmentation $\phi: \Omega \mapsto \{0, 1\}$ and the observation parameters Q given the image $\mathcal{I}: \Omega \mapsto \mathbb{R}$ and the shape model Θ . The segmentation ϕ is evolved such that $p(\phi, Q|\mathcal{I}, \Theta)$ is maximal which, according to the rule of Bayes, can be stated as follows:

$$\arg \max_{\phi, Q} \frac{p(\mathcal{I}|\phi, Q, \Theta) p(\phi|Q, \Theta) p(Q|\Theta)}{p(\mathcal{I})}. \quad (7)$$

Since the observation parameters Q and the shape model Θ do not add any information when the segmentation ϕ is known, \mathcal{I} is conditionally independent with respect to Q and Θ . Therefore, and since $p(\mathcal{I})$ is constant, $p(\phi, Q|\mathcal{I}, \Theta)$ can be expressed as

$$p(\phi, Q|\mathcal{I}, \Theta) \propto p(\mathcal{I}|\phi) p(\phi|Q, \Theta) p(Q|\Theta) . \quad (8)$$

Maximization of the probability is converted to energy minimization by taking the negative logarithm of equation 8. The following paragraphs formulate expressions for the different terms of this energy function.

3.1 Image Likelihood Term

The term $p(\mathcal{I}|\phi)$ of equation 8 is the image likelihood. Since the main contribution of this paper lies in the integration of statistical shape models into the graph cut framework, we restrict ourselves to a simple term. Therefore we exploit the knowledge that object boundaries typically coincide with edges in the image and is based on an edge detector function $g(|\nabla I|)$.

$$E(I, \phi) = \int_{\partial\Omega} g(|\nabla I|) dx , \quad (9)$$

where $\partial\Omega$ denotes the segmentation boundary. To translate this image likelihood term into edge weights for the n -links we can use the geo-cut method of Boykov et al. [11]. It should be noted that the image likelihood supports a variety of different energy terms, such as regional ones. This is application specific and can greatly improve the segmentation outcome.

3.2 Segmentation Likelihood Term

The second term $p(\phi|Q, \Theta)$ of equation 8 is the segmentation likelihood, and connects the segmentation and the statistical shape model. Due to the probabilistic shape representation reflecting the probability of a point to belong to the object boundary, the shape model nicely integrates in the graph cut framework. As such we can formulate

$$p(\phi|Q, \Theta) = \prod_{e \in \mathcal{C}_n} p(e|Q, \Theta) , \quad (10)$$

where \mathcal{C}_n is the subset of n -links belonging to the cut corresponding to segmentation ϕ and $p(e|Q, \Theta)$ is the probability of cutting edge $e \in \mathcal{E}$ given the observation parameters Q and the shape model Θ . This probability is given by the sum of the probabilities for each point along the edge to belong to the object boundary.

$$p(e|Q, \Theta) = \int_0^1 p(u(\mathbf{h}_e - \mathbf{t}_e) + \mathbf{t}_e|Q, \Theta) du . \quad (11)$$

where u is the parametrization variable and \mathbf{t}_e and \mathbf{h}_e , respectively, are the tail and head node of edge e . This can be further expressed as

$$p(e|Q, \Theta) = \sum_{j=1}^{N_m} \frac{|\mathbf{a}_e| e^{\left(-\frac{|\mathbf{a}_e \times \mathbf{b}_e^j|^2}{\sigma^2}\right)}}{N_m 2\pi\sigma^2} \left(\operatorname{erf} \left(\frac{|\mathbf{a}_e|^2 + \mathbf{a}_e \cdot \mathbf{b}_e^j}{\sigma} \right) - \operatorname{erf} \left(\frac{\mathbf{a}_e \cdot \mathbf{b}_e^j}{\sigma} \right) \right) , \quad (12)$$

where $\mathbf{a}_e = \mathbf{h}_e - \mathbf{t}_e$, $\mathbf{b}_e^j = \mathbf{t}_e - \mathbf{m}_j$, $\sigma_e = \sqrt{2}\sigma|\mathbf{a}_e|$, $|\mathbf{x}|$ is the norm of vector \mathbf{x} , $\mathbf{x} \cdot \mathbf{y}$ is the dot product and $\mathbf{x} \times \mathbf{y}$ is the cross product between vectors \mathbf{x} and \mathbf{y} .

Taking the negative logarithm of equation 10 converts the probability of a cut given the observation and model parameters into the energy of the cut. From this energy function, the weights assigned to the n -links e can be defined in a straightforward manner as $w_e = -\log(p(e|Q, \Theta))$. As such, an effective means for the integration of a statistical shape model into the graph cut framework is achieved. However, a bias is introduced favoring shorter object contours, as can be seen from equation (10). This is also reflected in the work of Hufnagel et al. [12], as can be seen from the curvature term in equation 7.

3.3 Shape Model Prior Term

The term $p(Q|\Theta)$ favors objects that are more probable according to the statistical shape model. Again assuming a constant prior for transformations leads to the following equation

$$\log(p(Q|\Theta)) \simeq \sum_{p=1}^n \left(\log(\lambda_p) + \frac{w_p^2}{2\lambda_p^2} \right), \quad (13)$$

with $Q = \{T, W\}$ and $W = \{w_p\}_{p=1}^n$, being the observation parameters. This term does not influence the edge weights of the graph.

3.4 Optimization

The statistical shape model term in the energy function enforces an iterative optimization procedure. In each iteration, starting from an initial segmentation, alternatively the observation parameters Q and the segmentation ϕ are optimized while keeping, respectively, ϕ and Q fixed. The observation parameters are optimized in a similar manner as explained in section 2. Optimization of the segmentation is performed by the max-flow/min-cut algorithm [3]. Although the graph cut algorithm provides global minima, since an iterative optimization is pursued, depending upon the initialization, only local minima can be proven to be obtained.

The statistical shape model in this framework causes the next segmentation boundary after a complete iteration to be located close to the current segmentation boundary. Therefore, computing the edge weights throughout the entire graph and applying the max-flow/min-cut algorithm to the full graph is not needed. In order to enforce a significant speedup of the algorithm, a narrow-band approach is pursued.

3.5 Initialization

As stated above, the outcome of the optimization is highly dependent upon the initialization. Therefore the initialization procedure is an important aspect of the algorithm. Different approaches can be used in order to provide a fast, yet accurate initial segmentation. Here we used the interactive graph cut segmentation algorithm of Boykov et al. [3]. Hereby, based on the edge-consistency prior explained in paragraph 3.1 and manually indicated seed points an initial segmentation is obtained. These manually indicated seed points are included in the graph cut framework as hard constraints, by setting the weights of the corresponding t -links to $+\infty$.

4 Experiments and Results

The segmentation algorithm presented in this paper is applied to the segmentation of teeth from Cone-Beam Computed Tomography (CBCT) images. A training data set of 22 patients is used of which the upper and lower left incisors, canines, premolars and molars are manually segmented. Since a manual segmentation of all left teeth is available, statistical shape models of the right teeth can be obtained as well, by a simple mirroring operation. Training the algorithm on all except one of the training samples and testing on the remaining data sample produces the results shown in figure 1. The leave-one-out validation results for all left teeth are provided in table 1. Here the Dice coefficient is used to measure the overlap between the ground truth (manual) segmentation and the segmentation provided by the algorithm. The results for the lower teeth are slightly worse compared to the upper teeth. The main reason for this is the lower contrast between the bone and teeth in the lower jaw compared to the upper jaw. Comparison to the results reported by Gao et al. [1] reveals a seemingly inferior performance of the algorithm presented in this paper. However, comparison of the results is not straightforward. At first, the algorithm presented in this paper is validated on a different data set. Second, the data set used by Gao et al. [1] does not contain orthodontic braces. Third, no range of results is reported by Gao et al. Fourth, a more advanced image likelihood term will most likely further improve the results, but was not the main scope of this paper.

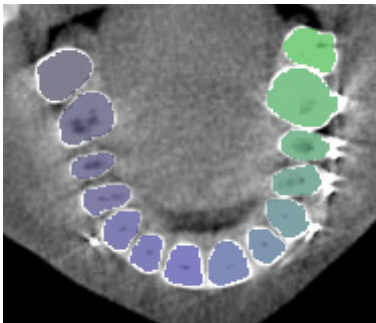


Fig. 1. Segmentation of teeth in a CBCT image. A single slice is shown with a different color for each tooth. The metallic streak artefacts caused by the orthodontic braces are visible.

Table 1. Validation results for the algorithm presented in this paper obtained from a leave-one-out approach applied to the training data. For each tooth the Dice coefficient is given as a measure of overlap between the ground truth (manual) segmentation and the segmentation provided by the algorithm.

Tooth	Root number	Dice coefficient
incisor 1 (lower)	1	0.8535 ± 0.0302
incisor 2 (lower)	1	0.8047 ± 0.1057
canine (lower)	1	0.8625 ± 0.1004
premolar (lower)	1	0.8369 ± 0.2164
molar (lower)	2	0.8522 ± 0.0766
incisor 1 (upper)	1	0.8820 ± 0.0332
incisor 2 (upper)	1	0.8734 ± 0.0641
canine (upper)	1	0.8882 ± 0.0431
premolar (upper)	1 or 2	0.8780 ± 0.0331
molar (upper)	2 or 3	0.8517 ± 0.0588

5 Discussion

In this paper a novel graph cut based segmentation algorithm is presented for the segmentation of teeth from CBCT images. Using a probabilistic shape representation a statistical shape model is constructed that integrates into the graph cut framework. The

algorithm optimizes an energy function, formulated according to a maximum a posteriori criterion. This energy function consists of three components: an image likelihood, a shape model term, and a segmentation likelihood term efficiently integrating the shape model into the graph cut framework. The energy function is optimized in an iterative manner. In future work we would like to extend this framework. At first, the linear shape model is not perfectly suited to segment objects that arise from a complex underlying distribution. We therefore would like to extend the probabilistic object representation framework to nonlinear shape models. Next, the image likelihood term is very general and can easily be replaced by a variant more tailored to a specific application. Finally, since the segmentation likelihood term introduces a shrinking bias, favoring objects with shorter contours, we would like to circumvent this. However, care must be taken in order to ensure that the derived metric is graph-representable.

References

1. Gao, H., Chae, O.: Individual tooth segmentation from ct images using level set method with shape and intensity prior. *Pattern Recognition* 43(7), 2406–2417 (2010)
2. Hosntalab, M., Zoroofi, R.A., Tehrani-Fard, A.A., Shirani, G.: Segmentation of teeth in ct volumetric dataset by panoramic projection and variational level set. *International Journal on Computer Assisted Radiology and Surgery* 3(3-4), 257–265 (2008)
3. Boykov, Y., Funka-Leah, G.: Graph cuts and efficient n-d image segmentation. *International Journal of Computer Vision* 70(2), 109–131 (2006)
4. Malcolm, J., Rathi, Y., Tannenbaum, A.: Graph cut segmentation with nonlinear shape priors. In: *IEEE International Conference on Image Processing*, vol. 4, pp. 365–368 (2007)
5. Freedman, D., Zhang, T.: Interactive Graph Cut Based Segmentation with Shape Priors. In: *CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, vol. 1, pp. 755–762. IEEE Computer Society (2005)
6. Schoenemann, T., Cremers, D.: A combinatorial solution for model-based image segmentation and real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(7), 1153–1164 (2010)
7. Hufnagel, H., Pennec, X., Ehrhardt, J., Handels, H., Ayache, N.: Shape Analysis Using a Point-Based Statistical Shape Model Built on Correspondence Probabilities. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 959–967. Springer, Heidelberg (2007)
8. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B* 39(1), 1–38 (1977)
9. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* 89(2-3), 114–141 (2003)
10. Myronenko, A., Song, X.: Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(12), 2262–2275 (2010)
11. Boykov, Y., Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: *Proceedings of International Conference on Computer Vision*, vol. I, pp. 26–33 (2003)
12. Hufnagel, H., Erhardt, J., Pennec, X., Schmidt-Richberg, A., Handels, H.: Level set segmentation using a point-based statistical shape model relying on correspondence probabilities. In: *Proc. of MICCAI Workshop Probabilistic Model for Medical Image Analysis, PMMIA 2009* (2009)