

Automatic Segmentation of Jaw Tissues in CT Using Active Appearance Models and Semi-automatic Landmarking*

Sylvia Rueda¹, José Antonio Gil¹, Raphaël Pichery², and Mariano Alcañiz¹

¹ Medical Image Computing Laboratory (MedICLab), Universidad Politécnica de Valencia, UPV/ETSIA, Camino de Vera s/n, 46022 Valencia, Spain
silruelo@degi.upv.es, jgil@dsic.upv.es, malcaniz@degi.upv.es

² Ecole Supérieure d'Ingénieurs de Luminy, Marseille, France
rpichery@esil.univ-mrs.fr

Abstract. Preoperative planning systems are commonly used for oral implant surgery. One of the objectives is to determine if the quantity and quality of bone is sufficient to sustain an implant while avoiding critical anatomic structures. We aim to automate the segmentation of jaw tissues on CT images: cortical bone, trabecular core and especially the mandibular canal containing the dental nerve. This nerve must be avoided during implant surgery to prevent lip numbness. Previous work in this field used thresholds or filters and needed manual initialization. An automated system based on the use of Active Appearance Models (AAMs) is proposed. Our contribution is a completely automated segmentation of tissues and a semi-automatic landmarking process necessary to create the AAM model. The AAM is trained using 215 images and tested with a leave-4-out scheme. Results obtained show an initialization error of 3.25% and a mean error of 1.63mm for the cortical bone, 2.90mm for the trabecular core, 4.76mm for the mandibular canal and 3.40mm for the dental nerve.

1 Introduction

Dental implants are titanium roots placed in the bone of the jaw to support replacement teeth. Preoperative planning systems are commonly used for oral implant surgery. The dentist or surgeon placing the implant needs to determine if the quality and quantity of bone is sufficient for long-term function and stability.

The body of the mandible is formed by a hard exterior, the cortical bone, and a soft spongy inside, the trabecular core. Teeth are fixed to the jaw by their roots and nerved to the mandibular canal. The canal contains the mandibular nerve which has to be avoided to prevent permanent or temporary lip numbness. This nerve runs from the area behind the wisdom teeth, passes under the molars and emerges under the skin of the face in the region where the premolar teeth are or used to be (Fig. 1).

* This research has been partially supported by 3DENT Sistemas Dentales, Spain.

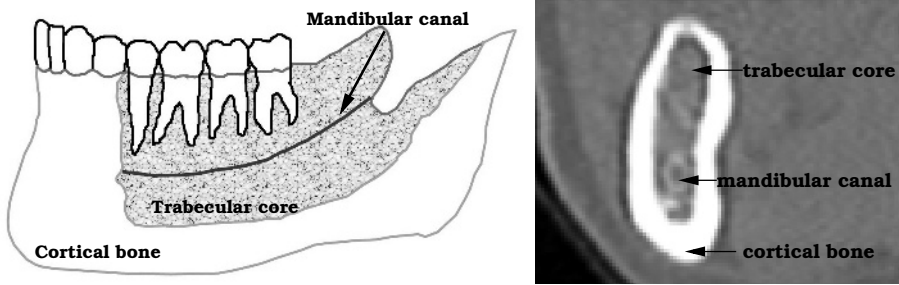


Fig. 1. Mandible structure: diagram and CT sectional view

Oral implant surgery uses different views on CT images to plan the intervention. Our work aims to segment cortical bone, trabecular core and the mandibular canal in sectional views (Fig. 1, right) because of their simplicity and the coherence existing between nearby slices. Depending on the patient, the shape of the bone can be very different and present mandible bone reabsorption. Moreover, the boundary of the trabecular core is not always well defined. We attempt to detect this region roughly because the mandibular canal is located inside. The canal is not always distinguishable in the trabecular zone because of the similarity of intensities in this region. Furthermore, these images present artifacts and sometimes teeth which make segmentation difficult.

2 Previous Work

We aim to automate the segmentation of cortical bone, trabecular core and the mandibular canal in CT images. Most of the studies carried out in the area of implantology has focused on 3D mandible reconstruction from projections [1] [2] avoiding the segmentation of tissues. Few studies have been found about mandible tissues segmentation. In most cases, the method was based on threshold and needed the intervention of an expert. For example, Fütterling et al. [3] developed an approach to automate the segmentation process of jaw tissues. They concluded that it was not possible to reliably detect any border between the cortical bone and the trabecular core without massive interaction. Stein et al. [4] proposed an approach to the detection of the mandibular canal using Dijkstra's algorithm to find the path and a deformable model to build the tubular structure. The process required the initialization of an expert.

De Bruijne et al. [5] obtained good results in the segmentation of tubular structures using Active Shape Models (ASMs). ASMs adapt the shape of the object to be segmented according to the statistical information of a training set previously annotated by an expert.

Previous work has shown that the problem to solve is complex and often requires expert interaction. Threshold techniques have proved to be inefficient to separate cortical bone and trabecular core. Nonetheless, the use of ASM to segment tubular structures gave good results. Therefore, if we want more accuracy

we have to consider texture information. Active Appearance Models is an extension of ASMs that takes into account texture and shape variability. We will use this method to segment mandibular tissues automatically.

3 Method

Active Appearance Models (AAMs), recently proposed by Cootes et al. [6][7], can model both shape and texture variability existing in a training set. To create a model we need a set of labeled images representative of the real variability of the object we want to segment. AAMs have proved to be powerful tools for medical image segmentation and understanding [8]. The whole process followed is summarized in Fig. 2.



Fig. 2. Methodology

3.1 Selection of the Training Set

The images forming the training set must be representative of the variability existing in real cases. The data and the source should reflect the target population for the algorithm that is being developed.

3.2 Semi-automatic Annotation of the Training Set

Because of the high variability of shape and the few anatomic landmarks existing in these images, manual landmarking, with high correspondence of points between images, will be difficult, tedious and error prone. The best solution is to automatize the process as much as possible. We use threshold techniques to find the contour of each structure. First, we find the external contour of the cortical bone (Fig. 3.a). Second, we define five mathematical landmarks (points of high curvature) on the cortical contour (Fig. 3.b). Third, we find the contour of the trabecular core (Fig. 3.c). Then, we locate the dental nerve, which is the only anatomical landmark, in the center of the mandibular canal (Fig. 3.d). Finally, we select the radius of the canal (Fig. 3.e).

Once this is done, we automatically define pseudo-landmarks equally spaced between the anatomical and mathematical landmarks previously located. To describe each structure we use a large number of landmarks. As stated in [9], duplicating the structure gives more specificity to the model and more accurate segmentation. The annotation for each structure can be seen in black in Fig. 4. Each black point is a landmark.

The cortical bone (Fig. 4.a) is described with 30 landmarks and a double contour of 28 points. The structure is open to avoid the presence of teeth. The

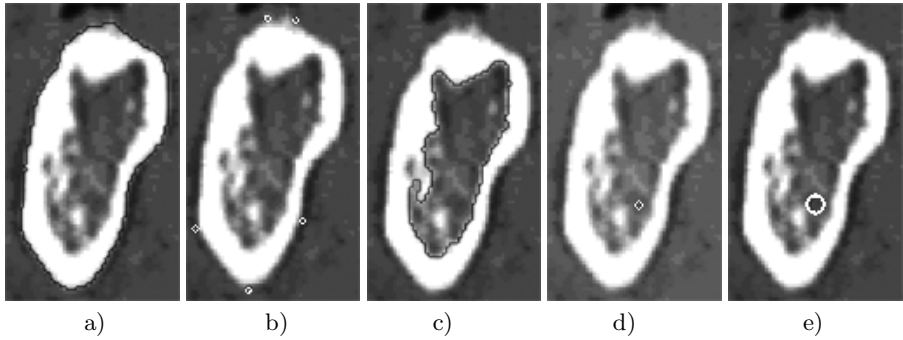


Fig. 3. Semi-automatic landmarking steps

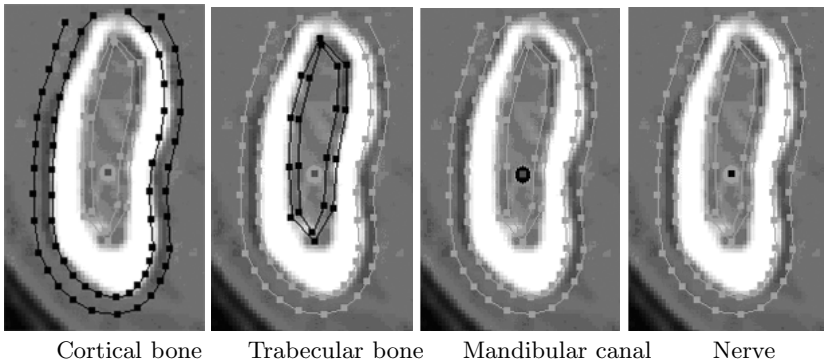


Fig. 4. Model landmarking

trabecular core is described with 10 landmarks and a double contour of 10 points. This structure has high variability and we use few points to avoid the adaptation of the shape to details. The mandibular canal uses 8 landmarks and the dental nerve only one.

3.3 Training the Model

To generate a statistical model of shape and texture variation it is necessary to describe the shape and the texture of each training example. Therefore, we represent n landmark points, (x_i, y_i) , for each image as a $2n$ vector, \mathbf{x} , where $\mathbf{x} = (x_1, \dots, x_n, y_1, \dots, y_n)^T$ describes the shape of an object. The annotated training set is then aligned in a common co-ordinate frame using a Procrustes Analysis. Hence, we obtain the Point Distribution Model (PDM) for all the images of the training set. The mean shape is extracted and the appearance variation collected by establishing a piece-wise affine warp (based on the Delaunay triangulation) between each image of the training set and the mean shape. Next, the intensity is sampled from the shape-normalized images over the region covered by the mean shape. The resulting samples are then normalized to minimize the effect of global

lighting variation and the texture (grey-levels) vector \mathbf{g} is obtained. Finally, for both shape and texture, we perform a Principal Component Analysis (PCA) on the aligned training set to describe the shape and appearance variations of the object. Often, shape and texture are correlated. Therefore, we can deduce a combined model of shape and texture with parameters \mathbf{c} , controlling the shape and texture at the same time. We obtain:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c} \quad (1)$$

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c} \quad (2)$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{g}}$ are the mean shape vector and the mean normalized grey-level vector, \mathbf{Q}_s and \mathbf{Q}_g are matrices describing the modes of variation derived from the training set and \mathbf{c} the combined PCA parameters. New images of the object can be synthetically constructed from \mathbf{c} . More details about AAMs can be found in [6][7]. The whole process is summarized in Fig. 5.

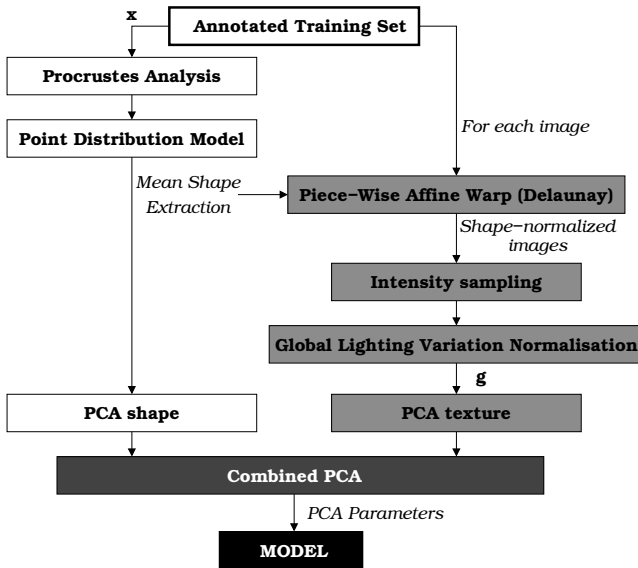


Fig. 5. Building an Active Appearance Model

3.4 Segmenting New Images

We first place an initial template model over the unseen image. We then use a principal component multivariate linear regression model to generate new images to fit the unseen image in the best way. Once the process converges, a match can be declared. Finally, an optimization scheme accomplishes further refinement of the match.

3.5 Model Evaluation

The final step of the whole process consists of testing the performance of the algorithm. Care should be taken to test the algorithm in data that has not been used in the training.

In this case, we will use a common methodology called leave-N-out. In this approach, a set of M images (ground truth) is split in N different ways into a training set of size M-N and a test set of size N. For each of the N splits, training is done on the M-N images and then testing is done on the remaining N. Performance is then estimated as the average of the N tests.

For each image segmented, we compare the result with the corresponding ground truth image. We calculate the distance between the structures segmented and those previously annotated, which gives the average error for each image and for each structure. There are two ways of measuring this distance: point-to-point (pt.pt) or point-to-curve (pt.crv). The pt.pt distance measures the Euclidean distance between each corresponding landmark, whereas the pt.crv distance measures the shortest distance to the curve in the neighborhood of the corresponding landmark. Pt.crv distance is more representative of the global segmentation because it does not evaluate each isolated point. We therefore based our evaluation on the pt.crv distance.

4 Results

The method described to segment cortical bone, trabecular core and the mandibular canal was applied to 215 CT images which were selected from 62 patients coming from different sources in order to cover the variability existing in real cases. The entire model was described by 87 landmarks annotated on each image of the training set. Semi-automatic landmarking was undertaken to give more accuracy to the system and assure that all the intermediate points were really equally spaced. The open C++ source code AAM-API[10] was partially used in this study. We created a tool with OpenCV libraries in C++ to annotate the images semi-automatically and to build AAM models automatically. We developed the leave-4-out to evaluate the models.

To calculate the optimized parameters in shape, we fixed the texture variance at 40% and the combined variance at 95%. We then took into account the results obtained and fixed the shape variance at 25% to calculate the optimized parameters in texture. The results can be seen in Fig. 6.

The best configuration is for 25% of the shape variance and 50% of the texture variance. For the optimal configuration, results in mean and standard deviation for each structure are shown in Table 1.

For a mean pt.crv distance lower than 5mm for at least 90% of the points for each image or structure, a good initialization can be defined. Accordingly, 90% of cortical bones, 66% of trabecular cores, 70% of mandibular canals and 72% of dental nerves were found correctly for all the 215 images. An example of good segmentation can be seen in Fig. 7.

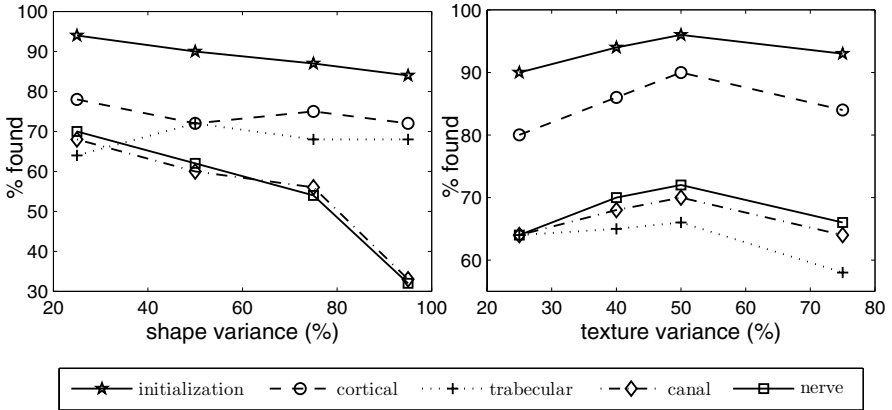


Fig. 6. Influence of shape and texture variance considered

Table 1. Results for the optimal configuration

| pt.crv | mean (mm) | | | | std (mm) | | | |
|--------|-----------|------------|-------|-------|----------|------------|-------|-------|
| | Cortical | Trabecular | Canal | Nerve | Cortical | Trabecular | Canal | Nerve |
| | 1.63 | 2.90 | 4.76 | 3.40 | 1.63 | 2.16 | 0 | 1.11 |

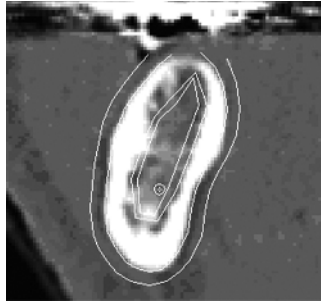


Fig. 7. Successful fitting

We do not expect more accuracy in the trabecular core because this structure has a high shape variability and we annotated the training set to avoid trabecular details during the segmentation.

5 Conclusion

AAMs perform accurate and automatic segmentation of cortical bone, trabecular core, mandibular canal and dental nerve taking into account all the variability of real cases. None of the previous work undertaken had achieved a real, automatic jaw tissues segmentation. In the research presented in this paper, semi-automatic

landmarking was developed to improve the AAM method and decrease annotation time.

References

1. Moreno, A., Bloch, I., Dutreuil, J. and Deval, M.: 3D Reconstruction from Radiologic Projections applied to Dental Implantology and to oral-facial surgery, ENST Report (2004).
2. Verstreken, K.: An Image-Guided Planning System for Endosseous Oral Implants. *IEEE Transactions on Medical Imaging*, vol.**17** (1998), 5.
3. Fütterling, F., Klein, R., Straber, W., Weber, H.: Automated Finite Element Modeling of a Human Mandible with Dental Implants. 6-th International Conference in Central Europe on Computer Graphics and Visualization, 1998.
4. Stein, W., Hassfeld, S., and Muhling, J.: Tracing of thin tubular structures in computer tomographic data. *Computer Aided Surgery* vol.**3** (1998), 83-88
5. DeBruijne, B.M., Van Ginneken, B., Niessen, W.J., Viergever, M.A., Wiro, J.N.: Adapting Active Shape Models for 3D Segmentation of Tubular Structures in Medical Images, *Information Processing in Medical Imaging Vol.2732* (2003). *Lecture Notes in Computer Science*, 136-147
6. Cootes, T.F., Edwards, G.J. and Taylor, C.J.: Active Appearance Models. *Proc. European Conference on Computer Vision*, vol.**2** (1998) 484-498
7. Cootes, T.F. and Taylor, C.J.: Statistical Models of Appearance for Computer Vision, Report (2004)
8. Cootes, T.F. and Taylor, C.J.: Statistical models of appearance for medical image analysis and computer vision, *Proc. SPIE Medical Imaging* (2001)
9. Stegmann, M.B.: Active Appearance Models - Theory, Extensions & Cases, Master Thesis IMM-EKS-2000-25 (2000)
10. Stegmann, M.B., Ersboll, B.K. and Larsen, R.: FAME - a flexible appearance modelling environment. *IEEE Transactions on Medical Imaging* (2003)