

# Automatic Detection and Classification of Teeth in CT Data

Nguyen The Duy, Hans Lamecker, Dagmar Kainmueller, and Stefan Zachow  
kainmueller@zib.de

Zuse-Institute Berlin, Germany

**Abstract.** We propose a fully automatic method for tooth detection and classification in CT or cone-beam CT image data. First we compute an accurate segmentation of the maxilla bone. Based on this segmentation, our method computes a complete and optimal separation of the row of teeth into 16 subregions and classifies the resulting regions as existing or missing teeth. This serves as a prerequisite for further individual tooth segmentation. We show the robustness of our approach by providing extensive validation on 43 clinical head CT scans.

## 1 Introduction

Cone beam computed tomography (CBCT) is becoming a preferred imaging technique for three-dimensional diagnosis and therapy planning in dentistry as well as maxillofacial surgery. In dental implantology, for instance, surgical drill guides are individually manufactured based on CBCT data and rapid prototyping techniques. The accuracy of such drill guides highly depends on the quality of the 3D reconstructions of jaw structures. Such reconstructions often end up in tedious image segmentation tasks, in case dental fillings and brackets heavily degrade image quality due to the resulting shadowing artifacts within respective image slices. In most cases, only manual segmentation leads to useful results.

Research on jaw segmentation in CT data has mostly concentrated on the lower jaw [1–4], since the maxilla is generally more difficult to segment automatically. The maxilla exhibits thin bony structures (palate, sinus maxillaris, orbital walls), which are difficult to detect with intensity thresholds alone [5, 6]. To the best of our knowledge, only Kainmueller et al. [7] segment the maxillary bone as part of the midface. Tooth regions, however, are omitted in all previous studies due to the aforementioned artefacts. The problem of detecting teeth in medical images has been well studied for 2D radiographs, but sparsely researched for 3D images, since effects of metal artefacts are less severe in 2D than in 3D. Mahoor et al. [8], Nassar et al. [9] and Lin et al. [10] use a three-step approach consisting of tooth isolation in the row of teeth, independent classification of each isolated tooth and correction of the classification results. All three methods use *integral projection* to separate teeth. Then, each isolated tooth region is classified based on area features [9] or shape features [8, 10]. The latter need a segmentation of each tooth in the isolated region. After independent classification of each region,

the result is corrected by considering tooth order. In this step, string alignment techniques are used. For 3D CT images, Gao et al. [11] and Hosntalab et al. [12] propose algorithms for segmentation, however, they do not perform a classification of teeth. The main limitation of these methods is that they require a segmentation of the tooth region as input. To the best of our knowledge, there exist no approaches which automatically detect teeth in 3D CT data.

This paper contributes an algorithm to reliably segment the maxillary bone and detect the individual dentition state in an automatic way. Therefore, we first perform an accurate and robust segmentation of the bony structure based on statistical shape model (SSM) adaptation, following the approach of [7]. Then, we use this segmentation to detect the 16 tooth regions by fitting 15 separation planes. Subsequently, we classify each separated region as “tooth” or “gap” via histogram analysis, yielding the individual dentition state. Fig. 1 depicts this algorithmic pipeline for an exemplary CT dataset.

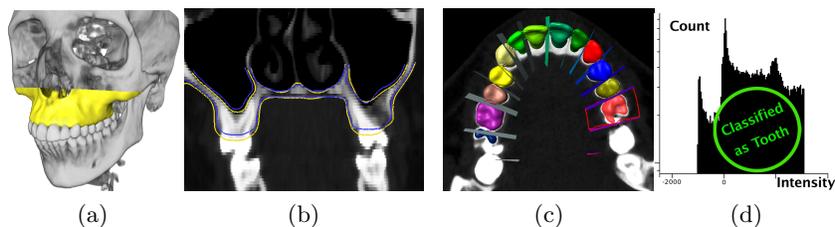


Fig. 1: Algorithmic pipeline in exemplary CT. (a) Maxilla segmentation, shown as yellow surface in volume rendering and (b) as yellow contour in coronal slice (vs. gold standard in blue). (c) Tooth region detection via separation planes. Gold standard teeth shown as colored surfaces. The red contour shows an individual tooth region that yields the intensity histogram (d) used for classification.

## 2 Maxilla segmentation

For maxilla segmentation we follow the three-step approach proposed in [7], consisting of (1) initialisation of shape and pose of a statistical shape model (SSM) via the Generalized Hough Transform, (2) adaptation of the SSM governed by a heuristic bone intensity model, (3) refinement of SSM-based segmentation via locally regularized shape deformation. To cope with the specific properties of the maxilla, as well as the need to accurately reconstruct the tooth region for subsequent tooth separation and classification, we extend this framework as described in the following.

**Automatic bone threshold selection.** The approach in [7] adapts an SSM to given image data by minimizing a cost-function that measures how well the model fits the data. A central parameter in this method is an image-specific

threshold  $t_{Bone}$  characterizing typical bone intensities. For the segmentation of the lower jaw (mandible), a fixed threshold can be applied since the bony structure is rather compact and the cortical region is adequately sampled within the image data. However, for the upper jaw (maxilla), a fixed threshold is insufficient, because thin structures, such as the palate and the walls of the maxillary sinuses are typically undersampled in consecutive image slices. Here, partial volume effects lead to reduced intensities of thin bone structures. Thus, we propose to automatically determine  $t_{Bone}$  by analyzing the intensity histogram  $h(x)$  in the vicinity of the initialised SSM. It exhibits a characteristic peak corresponding to air voxels at about  $-1000 HU$  and another one corresponding to soft tissue around  $0 HU$ . We model the shape of each peak by a Gaussian function  $g_i(x) = a_i \exp\left(-\frac{(x-\mu_i)^2}{\sigma_i^2}\right)$ ,  $i \in \{1, 2\}$  with mean  $\mu_i$ , standard deviation  $\sigma_i$ , and scale  $a_i$ , and the remaining intensity occurrences with a constant  $C$ . The final function fitted to the histogram is  $f(x) = g_1(x) + g_2(x) + C$ . The problem of fitting  $f(x)$  to  $h(x)$  by determining the optimal parameter set  $\Theta^* = (a_1, a_2, \mu_1, \mu_2, \sigma_1, \sigma_2, C) = \arg \min_{\Theta} (h(x) - f(x))^2$  can be solved via Levenberg-Marquardt optimization. Using  $\Theta^*$ , we set the bone threshold as  $t_{Bone} = \mu_2 + 2\sigma_2$ , i.e., roughly above 85% of the intensities captured by the soft tissue peak.

**Segmentation strategy for the tooth region.** The key idea of the algorithm in [7] is to analyze intensity profiles  $P$  sampled in direction of the surface normal on each vertex of the SSM to drive the segmentation process. To segment the bone region, we propose to decide for each intensity profile on the tooth region whether it contains a bone-soft tissue interface or not. After binary classification, we apply an individual segmentation strategy for each class. Using histogram analysis as described above, we heuristically determine lower intensity thresholds of soft tissue  $t_{Tissue} = \frac{\mu_1 + \mu_2}{2}$  (inbetween air- and soft-tissue mean) and teeth  $t_{Teeth} = \arg \min_x \frac{d}{dx} h(x) \wedge x > 1000 HU$ , i.e. the steepest descent above 1000HU. The latter can be interpreted as an upper threshold to bone, and thus serves as a lower threshold to teeth. We use these thresholds to classify each sample point on an intensity profile  $P$ . Afterwards, a rule-based decision is made whether to analyze the profile with the bone intensity model (BIM) as employed in [7]: If there is (1) no region classified as tooth, or (2) one or two regions classified as tooth as well as a transition from bone-classified to air-classified sample points, the BIM is employed. In all other cases, a conservative strategy is employed that slightly prefers the mid-sample point on the profile (i.e. the current position of the respective vertex on the deformable surface) over all others.

### 3 Tooth detection

After successful segmentation of the maxillary bone, we determine the volumetric image region containing the teeth as all voxels within a reasonable distance to the *tooth patch*, which is a predefined region on our maxilla SSM's surface.

We propose the following scheme to detect/classify teeth within that region: In a first step, the region is divided into 16 plausible subregions, each containing either single teeth or representing missing teeth. In a second step, the individual dentition state is classified.

**Separation.** For decomposing the entire tooth region into subregions, we employ 15 separation planes. We propose to formulate the task of finding a suitable position and orientation for each of these planes as a graph optimisation problem. This leads to an optimal solution with respect to an objective function as described in the following. First, we limit the number of possible positions and orientations for each plane to a finite set. Moreover, we limit the distances and angles between adjacent planes to plausible intervals which have been learnt a priori from the sizes of the respective teeth in training data. Hence, a feasible sequence of tooth separation planes (ID  $i$  next to ID  $i + 1$ ) defines a sequence of consecutive tooth cells corresponding to a sequential tooth numbering scheme.

For each potential separation plane  $P_i$ , we define a cost that encodes how likely it separates two teeth given the image data. This cost is derived from the average image intensity  $m_i$  and standard deviation  $s_i$  within the plane as well as the average directional derivative of the image perpendicular to the plane,  $g_i$ . Additionally, a penalty cost  $a_i$  is assigned to highly tilted planes as a regularizing term in case many consecutive teeth are missing. The cost  $c$  of  $P_i$  is then defined as a weighted sum

$$c(P_i) = \alpha \cdot m_i + \beta \cdot s_i + \gamma \cdot \frac{1}{g_i^2 + \epsilon} + \delta \cdot a_i \quad (1)$$

where a small  $\epsilon$  serves for avoiding divisions by zero. Note that the weights  $(\alpha, \beta, \gamma, \delta)$  are fixed for all planes.

The costs are encoded within a graph: Each potential separation plane (i.e. combination of position and orientation) is represented by a graph node; Feasible neighboring separation planes are encoded via a graph edge between the respective nodes; Each graph node is weighted by the cost of the respective separation plane. The optimal path through this graph, computed with Dijkstra’s algorithm, yields our optimal set of separation planes:

$$\left\{ \hat{P}_i, i = 1 \dots 15 \right\} = \arg \min_{\{P_i, i=1 \dots 15\}} \sum_{i=1}^{15} c(P_i) \text{ subj. to } (P_i, P_{i+1}) \in E \forall i < 15 \quad (2)$$

where  $E$  denotes the set of graph edges.

**Classification.** Having partitioned the entire tooth region into distinct cells, we can independently decide whether a cell contains a tooth or not. If we detect a tooth, we can classify it directly from the ID of the cell containing it. Since teeth are imaged with a high intensity value in CT images, we use the intensity histogram of the region as a feature for the decision whether a tooth is present or not. To this end, a support vector machine (SVM) is learnt from training data and employed as a binary classifier.

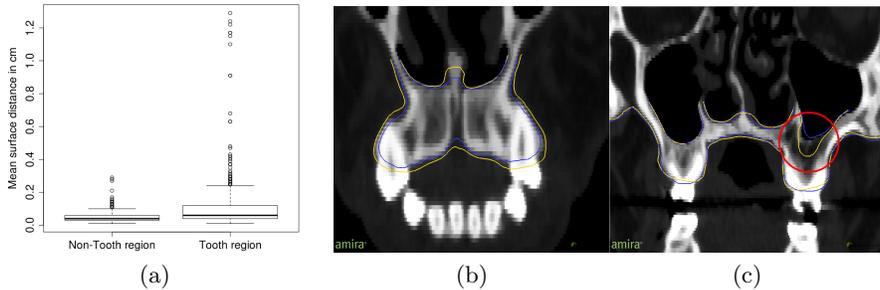


Fig. 2: Maxilla segmentation results. (a) Mean surface distances assessed in 43 datasets. (b,c) Exemplary automatic segmentation (yellow) vs. ground truth (blue). (c) Low intensity at sinus causes inaccuracy.

## 4 Results and Discussion

The data basis for our experiments consists of 43 clinical CT datasets, fully depicting the maxilla and the upper teeth. The individual maxillas differ in number of teeth and bone density. Moreover, the data basis contains “pathological” cases that occur in clinical practise, e.g. with dental prosthesis or metal artefacts due to fillings of implants.

**Maxilla segmentation.** We assess the accuracy of maxilla segmentation in a leave-one-out evaluation. For each dataset, we use the ground truth shapes of the other 42 datasets to generate the SSM applied for segmentation. As quality measure for segmentation, we use mean surface distance to the ground truth. In order to evaluate the performance of the tooth region segmentation strategy, we separately measured mean surface distances for tooth and non-tooth region. Fig. 2 plots the resulting error measures and shows exemplary segmentation results. The median mean surface distance is below 1 *mm* for both tooth- and non-tooth regions. Slightly higher errors for the tooth region can be attributed to a lack of image features within teeth, as well as a stronger influence of metal artefacts which occur mainly in the tooth region.

**Tooth separation.** In order to test our tooth separation approach, we perform a ten-fold cross validation. We learn the tooth sizes in the training set, and apply our separation algorithm to the test set using these size parameters.

As for the weights  $(\alpha, \beta, \gamma, \delta)$  of the per-plane costs that build the separation cost function (cf. Eq. 1), we set  $\alpha = 1$  and determine the best set of remaining weights (“best” w.r.t. volume overlap averaged over all teeth and datasets) via exhaustive search. We had to exclude this search from the cross validation framework for performance reasons, and hence performed it just once, w.r.t. the whole data set, instead of for each training set, in violation of a strict separation between training- and test data.

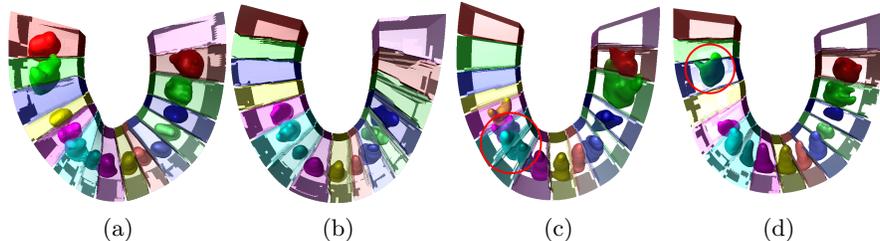


Fig. 3: Separation results. Tooth region and corresponding gold standard tooth are shown in same color. (a,b) Good results despite gaps. (c) Two regions contain three teeth - correspondence error is propagated. (d) One tooth in wrong region.

We measure the separation quality of each tooth region by volume overlap of the respective gold standard tooth: Let  $T_i \subset \mathbf{R}^3$  be the  $i$ -th gold standard tooth and  $R_i \subset \mathbf{R}^3$  the respective tooth region as determined by our separation algorithm. The volume overlap for this region is then computed as  $(T_i \cap R_i)/T_i$ . Table 1(a) lists the results for all datasets containing teeth, which are 27 out of the total 43. In 22 of these 27 datasets, *all* tooth regions determined by fully automatic separation contain at least 70% of the corresponding gold standard tooth (Table 1(a), 2nd row).

The tooth separation algorithm achieves high volume overlap when the row of teeth is complete or only solitary gaps exist inbetween. The method can even deal with larger gaps when they do not split off solitary teeth (Fig. 3(a,b)). If, however, one solitary tooth is split from the remaining row of teeth by a gap significantly larger than a single tooth, the separation algorithm runs the risk of placing it in a non-corresponding region (Fig. 3(d)). This might be partly due to a solitary tooth's naturally enlarged freedom to move. Furthermore, in case subsequent teeth are relatively slim and densely packed, they might be split into too few regions (Fig. 3(c)).

Overlap threshold	Number of datasets with teeth below threshold						Ground truth		
	0	1	2	3	4	5	Auto, ideal sep.	Auto, auto sep.	
< 50%	25	1	0	1	0	0	Tooth	Tooth	61%
< 70%	22	4	0	1	0	0	No tooth	No tooth	38%
< 90%	8	9	4	5	0	1	Tooth	Tooth	58%
							No tooth	No tooth	37%

Table 1: (a) Tooth separation results on 27 datasets. (b) Tooth classification results based on ideal (top) and automatic (bottom) separation. 61% of ground truth regions contain teeth; 39% contain gaps.

**Tooth classification.** To evaluate classification performance, we first run a ten-fold cross-validation on “ideal” regions which perfectly separate teeth. These “ground truth” regions are created from ground truth tooth segmentations. In a second experiment, we run another cross validation on the “real-world” regions computed by our tooth separation method. Table 1(b) lists the results. For ideal separation, 99% of auto-classifications are correct, while we have 0% false negatives and 1% false positives. For automatic separation, 95% of auto-classifications are correct, while we have 3% false negatives and 2% false positives.

In case classification fails, in the “real-world” scenario, it can be attributed mostly to previous errors in the automatic tooth separation step, where either correspondence between tooth and region does not apply (cf. Fig. 3(c,d)), or a tooth is split into two parts, one located in the region which is not supposed to contain a tooth. There are also individual cases where the patient has metal implants in the jaw which have intensity values similar to teeth and are therefore classified as tooth (cf. Fig. 4).

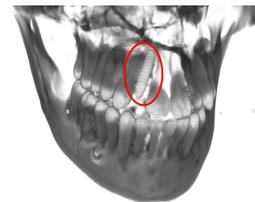


Fig. 4: Metal implant has same intensity as teeth

## 5 Conclusion

We propose a method which is able to segment the maxillary bone and detect and classify upper jaw teeth in CT images. Contrary to previous work, our segmentation approach specifically deals with the tooth regions and thus allows for subsequent tooth detection and classification. Our experiments provide evidence of the robustness of our approach on 43 clinical data sets. Even datasets with strong metal artefacts are processed successfully (cf. Fig. 5).

However, in challenging regions like the thin walls and sinus floor where polyps often occur, maxilla segmentation may still be improved, e.g. via locally adaptive threshold estimation. Furthermore, we observe that our classification scheme for bone/tooth profiles in some cases leaves us with too few profiles to drive the segmentation correctly. This may be overcome by a denser sampling of the tooth region on the SSM.

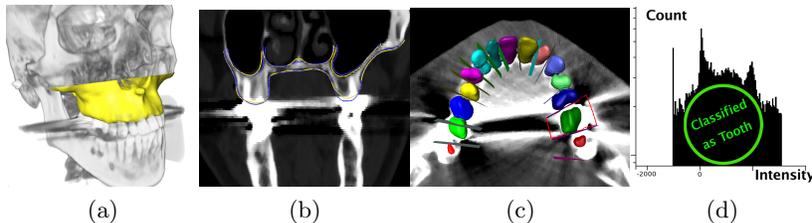


Fig. 5: Exemplary CT with dental fillings: Whole pipeline of segmentation (a,b), separation (c) and classification (d) successful despite strong metal artefacts

As for tooth classification, we expect that the use of volumetric information instead of intensity histograms as feature vectors would reduce false positive classifications. To this end, future work may integrate shape features, e.g. SSMs of individual teeth, into both tooth separation and classification. Most importantly, our tooth detection pipeline, together with tooth-specific SSMs, may serve as input for accurate automatic segmentation of individual teeth.

## References

1. Lilja, M., Vuorio, V., Antila, K., Setälä, H., Järnstedt, J., Pollari, M.: Automatic Segmentation of the Mandible From Limited-Angle Dental X-Ray Tomography Reconstructions. *ISBI: From Nano to Macro (2007)* 964–967
2. Rueda, S., Gil, J.A., Pichery, R., Alcañiz, M.: Automatic segmentation of jaw tissues in CT using active appearance models and semi-automatic landmarking. *MICCAI 9(Pt 1)* (January 2006) 167–74
3. Lamecker, H., Zachow, S., Wittmers, A., Weber, B., Hege, H., Elsholtz, B., Stiller, M.: Automatic segmentation of mandibles in low-dose CT-data. *Int. J. of Comp. Ass. Rad. Surg.* **1** (2006) 393
4. Kainmueller, D., Lamecker, H., Seim, H., Zinser, M., Zachow, S.: Automatic Extraction of Mandibular Nerve and Bone from Cone-Beam CT Data. In et al., G.Z.Y., ed.: *MICCAI (1)*. Volume 5762 of LNCS., Springer (2009) 76–83
5. Barandiaran, I., Macía, I., Berckmann, E., Wald, D., Dupillier, M., Paloc, C., Graña, M.: An automatic segmentation and reconstruction of mandibular structures from CT-data. In: *Proc. Int. Conf. on Intelligent Data Eng. and Automated Learning*, Springer (2009) 649–655
6. Tognola, G., Parazzini, M., Pedretti, G., Ravazzani, P., Grandori, F., Pesatori, A., Norgia, M., Svelto, C.: *IST 2006 - International Workshop on Imaging Systems and Techniques Minori, Italy 29 April 2006*. (April) (2006) 3–6
7. Kainmueller, D., Lamecker, H., Seim, H., Zachow, S.: Multi-object Segmentation of Head Bones. *MIDAS Journal, Contribution to MICCAI Workshop Head and Neck Auto-Segmentation Challenge*. (2009) 1–11
8. Mahoor, M., Abdelmottaleb, M.: Classification and numbering of teeth in dental bitewing images. *Pattern Recognition* **38**(4) (April 2005) 577–586
9. Nassar, D., Abaza, a., Ammar, H.: Automatic Construction of Dental Charts for Postmortem Identification. *IEEE Transactions on Information Forensics and Security* **3**(2) (June 2008) 234–246
10. Lin, P., Lai, Y., Huang, P.: An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information. *Pattern Recognition* **43**(4) (April 2010) 1380–1392
11. Gao, H., Chae, O.: Automatic Tooth Region Separation for Dental CT Images. *2008 Int. Conf. on Conv. and Hybrid Inf. Tech.* (2008) 897–901
12. Hoshtalab, M., Aghaeizadeh Zoroofi, R., Abbaspour Tehrani-Fard, A., Shirani, G.: Segmentation of teeth in CT volumetric dataset by panoramic projection and variational level set. *Int. J. of Comp. Ass. Rad. Surg.* **3**(3-4) (June 2008) 257–265