33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Automatic Dent-landmark Detection in 3-D CBCT Dental Volumes

Erkang Cheng, Jinwu Chen, Jie Yang, Huiyang Deng, Yi Wu,
Vasileios Megalooikonomou, Bryce Gable, Haibin Ling

Abstract— **Orthodontic craniometric landmarks provide critical information in oral and maxillofacial imaging diagnosis and treatment planning. The Dent-landmark, defined as the odontoid process of the epistropheus, is one of the key landmarks to construct the midsagittal reference plane. In this paper, we propose a learning-based approach to automatically detect the Dent-landmark in the 3D cone-beam computed tomography (CBCT) dental data. Specifically, a detector is learned using the random forest with sampled context features. Furthermore, we use spacial prior to build a constrained search space other than use the full three dimensional space. The proposed method has been evaluated on a dataset containing 73 CBCT dental volumes and yields promising results.**

I. INTRODUCTION

Orthodontic craniometric (cephalometric) anatomical landmarks provide important symmetry measurements for oral and maxillofacial imaging diagnosis and treatment planning [1]. For example, facial asymmetry is a widely used feature in the analysis of craniofacial deformities. Another example is the analysis of superimposition of cranial base structures, which is used to monitor temporal changes associated with orthodontic treatment and growth. Detection or localization of these anatomical landmarks is a key asset for clinic diagnosis and treatment planning in orthodontics.

Two dimensional cephalometric measurements from lateral and/or frontal cephalograms were widely studied in last several decades [2]. As a projection of 3D volumes, 2D images suffer from several disadvantages including non-homogenous enlargement and distortion on lateral structures, inaccurate landmark locations due to overlapping structures, and landmarks that appear on the lateral may not appear on the frontal image or vice versa [3]. Three dimensional computed tomography (CT)-based cephalometric analysis were recently introduced as a new method for 3D diagnosis and as an aid in 3D virtual planning for orthognathic surgery. Recently, cone-beam computed tomography (CBCT) data has been studied to supply more advance opportunities in orthognathic analysis [4]. Technological advances in CBCT offer significant advantages in both quality and quantity of data representation. The most recent studies involving CBCT scans have shown 3D measurements are much more accurate than 2D measurements.

Anatomic landmarks for symmetry measurements are reported as being manually annotated by experts in both 2D images and 3D volumes. These landmarks are used to construct reference planes for symmetry evaluation. In 2D images, landmarks are traced and measured in lateral and/or frontal cephalograms. Axial, sagittal and coronal views help locating landmarks in 3D volumes [5]. Existing methods for localizing these landmarks, however, mainly relies on manual annotation that are tedious and time consuming.

We aim at solutions for automatically detecting anatomical landmarks in 3D CBCT volumes. We expect these solutions to enable us to take full advantage of CBCT scans in performing 3D measurements and analysis. Object detection from visual input has achieved great progress in the past decade and has been successfully applied to tasks such as face and vehicle detection [6] using machine learning approaches. Similar approaches have recently been applied to anatomical structure localization tasks as well [8], [10].

In this paper, we focus on detection of the *Dent-landmark* (or *Dent* in short), which is the most superior point of the odontoid process of the epistropheus. Together with the sella turcica (S), the junction of the nasal and frontal bones in the midline (N), Dent is selected as the midsagittal (x) reference plane [4]. Fig. 1 gives an example of the Dent-landmark in a 3D CBCT volume, where the Dent (in red) is annotated in Axial, Sagittal and Coronal planes in the volume.

We propose a learning-based Dent-landmark detection by employing the random forest [7] as the discriminative learning framework with sampled context features. Random forest has the capability to handle large training samples with high feature dimension. We use sampled context features to capture the rich context information around Dent-landmarks in the three dimensional space. Furthermore, spatial prior is used to boost the efficiency of search procedure.

To validate the proposed method, it is tested using a dataset containing 73 CBCT volumes. Our detector is learned using 50 training volumes with Dent-landmark groundtruth. The rest 23 volumes are used for testing. We compare the proposed method with the method using AdaBoost [9]. The results demonstrate the effectiveness of our approach.

The rest of the paper is organized as follows: In Section II, we present the methodology used in the paper. Experiments and analysis are proposed in Section III. Finally, conclusion

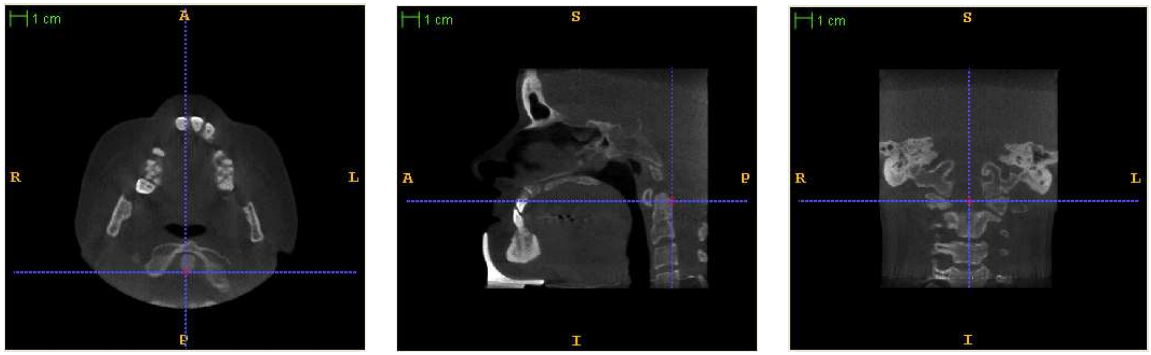
Erkang Cheng is with Center for Data Analytics & Biomedical Informatics, Computer & Information Science Department, Temple University, Philadelphia, PA, 19122 USA (email: erkang.cheng@temple.edu)

Jinwu Chen, Bryce Gable, and Jie Yang are with the Division of Oral and Maxillofacial Radiology, School of Dentistry, Temple University, Philadelphia, PA, 19140 USA (email: jinwuchen@yahoo.com.cn, bryce.gabler@dental.temple.edu, jyang@temple.edu)

Yi Wu is with College of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China (email: yi.wu@temple.edu)

Vasileios Megalooikonomou is with Data Engineering Laboratory, Temple University, Philadelphia, PA, 19122 USA (email: vasilis@temple.edu)

Haibin Ling is with Center for Data Analytics & Biomedical Informatics, Computer & Information Science Department, Temple University, Philadelphia, PA, 19122 USA (corresponding author; email: hbling@temple.edu)



(a) Axial (b) Sagittal (c) Coronal

Fig. 1. Orthogonal views of Dent-landmark in a 3D CBCT volume

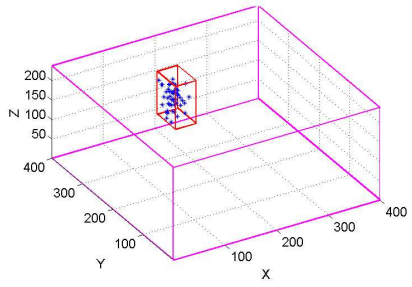


Fig. 2. Constrained search space of Dent-landmark

is given in Section IV.

II. METHODOLOGY

In this section, we describe the methodology for automatical Dent-landmark detection. A discriminative detector is learned using the random forest framework with sampled context features. To utilize the strong spacial prior for Dent-landmark, we constrain the searching in a subspace to improve the efficiency.

A. Spatial prior

We observed that, the spatial distribution of Dent-landmark is clustered in a small region other than the whole three dimensional space. The uninform and exhaustive searching in the original three dimensional space is not necessary and time consuming. By taking count of the strong spatial prior of Dent-landmark, we apply our proposed method in the constrained search space other than the whole three dimensional space. Note that similar ideas have been used in anatomic structure detection in previous studies [10].

Fig. 2 demonstrates a constrained search space of a Dent-landmark. The full three dimensional space in the 3D CBCT data is of size $400 \times 400 \times 327$. The blue dots are groundtruth of Dent-landmarks. The constrained search space is represented by the red cube.

Based on this observation, we define the constrained search space by $S_{constrained} = \{(X,Y,Z)|X_0 \leq X \leq X_0 + X_{size}, Y_0 \leq Y \leq Y_0 + Y_{size}, Z_0 \leq Z \leq Z_0 + Z_{size}\}$, where X_{size}, Y_{size} and Z_{size} define the size of the constrained search space. Here, the constrained search space is simply the bounding box of groundtruth of the training volumes. In the experiment, the constrained search space is chosen to cover 90% training volumes in comparison with the clustered center.

We apply the spatial prior of Dent-landmark in both training and testing stages. A detector is trained using the small constrained region other than the whole three dimensional space. Searching the Dent-landmark candidates is restricted in the same constrained space.

B. Training samples

With 3D CBCT dental data, to train a detector, we need to obtain positive and negative samples from these training volumes based on their distance to the groundtruth. Both positive and negative samples are extracted using the spatial prior of Dent-landmark.

Let (X_t, Y_t, Z_t) be an annotated location of a Dent-landmark, a positive sample (X_p, Y_p, Z_p) and a negative sample (X_n, Y_n, Z_n) should satisfy the following condition:

$$\begin{aligned} ||(X_p - X_t, Y_p - Y_t, Z_p - Z_t)|| &\leq 2 \quad voxels, \\ ||(X_n - X_t, Y_n - Y_t, Z_n - Z_t)|| &> 4 \quad voxels. \end{aligned} \tag{1}$$

In other words, in the training stage, we draw positive instances around the groundtruth within a certain distance. Negative training samples are randomly extracted in the constrained search space $S_{constrained}$. These negative samples are with distances between groundtruth larger than a threshold (say four voxels).

After sampling positive and negative samples in the constrained search space in the training volumes, a detector is trained using the random forest with sampled context features.

C. Sampled context feature

A Dent-landmark is surrounded of different hard and soft tissues. The context information is significant for determine its location. Such context information would help us in detecting Dent-landmark in CBCT dental data. In our work, we apply sampled context features to capture such context information.

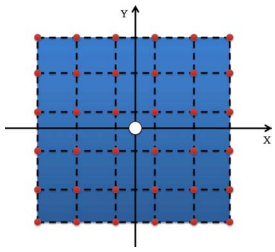


Fig. 3. Example of sampling pattern for sampled context feature (projected in 2D space)

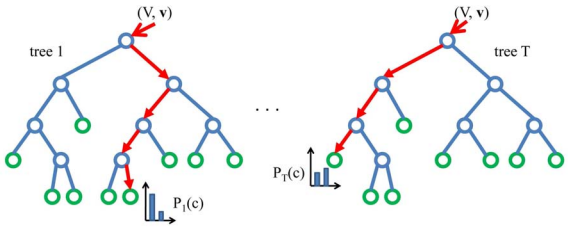


Fig. 4. The learning and inference framework

Basically, for a potential candidate position, we sample a few points around it in the volume under a sampling pattern. The sampling pattern is centered at the voxel being calculated. Then, a few local features are extracted for each sampling point (e.g. voxel intensity and gradient) from the original volume. The descriptive power of sampled context features is achieved by using flexible sampling patterns. We can manipulate the sampling pattern by changing the size of the pattern and the number of sampled points.

Fig. 3 shows the sampling pattern (illustrated in 2D case for better visualization) used in our work. The sampling points are represented by red dots. The sampling pattern is centered at a (candidate) Dent-landmark (shown in white) and the size of sampling pattern is indicated by a blue rectangle.

Suppose that the sampling pattern is of size $s_1 \times s_2 \times s_3$ in the three dimensional space. We extract P points from that pattern. At each sampling point (x,y,z), we include its intensity I and gradient $\|g\|$ in the feature representation. In addition to the sampled context features, the normalized coordinate $(x/X_{size}, y/Y_{size}, z/Z_{size})$ is also added into the feature pool. Here, $(X_{size}, Y_{size}, Z_{size})$ define the size of the being used volume. Overall, we get a feature vector $\mathbf{F}$ with size of N $(N = 2P+3)$ for one training sample.

D. Learning framework

Combined with sampled context features in the constrained search space, we employ the random forest as the discriminative learning framework in our work. Random forest [7] has been successfully applied to medical image analysis in various tasks [8]. As illustrated in Fig. 4, a random forest is an ensemble classifier consisting of T decision trees, with each of them consisting of split (in blue) and leaf nodes (in green). A split node consists of a feature f_θ and a threshold π. A leaf node, in contrast, is associated with an estimated class distribution. Each tree is trained on a different subset randomly sampled from the original training dataset. The trees grown are not pruned. The algorithm of training each tree is given in Algorithm 1.

To classify a voxel $\mathbf{v}$ in volume V, the feature extracted from $\mathbf{v}$ goes through all trees in the forest (e.g., the red paths in Fig. 4). In each tree, it starts at the root and branches left or right according to the comparison of threshold π. At the leaf node reached in tree t, a previously learned distribution $P_t(c|V,\mathbf{v})$ is recorded. The distributions from all trees are averaged to generate the final classification,

$$P(c|V,\mathbf{v}) = \frac{1}{T}\sum_{t=1}^{T} P_t(c|V,\mathbf{v})), \tag{2}$$

where $c = 1$ indicates $\mathbf{v}$ is a Dent-landmark, and $c = 0$ indicates that $\mathbf{v}$ is not a Dent-landmark.

We compare the random forest and Adaboost [9] in the experiments. Adaboost is an efficient ensemble learning method that builds a strong classifier as a weighted linear combination of a set of weak classifiers. Specifically, let $x \in \Re^d$ be a d dimensional input feature vector, the final (strong) classifier $h(x) : \Re^d \to \{-1,1\}$ has the following form:

$$h(x) = \sum_{i=1}^{n} c_i h_i(x), \tag{3}$$

where $h_1(x), \cdots, h_n(x)$ are the n weak classifiers, c_i is the weight of each selected weak classifier. The probabilistic responses (through logistic transformation) utilized in the inference stage is

$$P(c|V,\mathbf{v}) = \frac{e^{h(\mathbf{v})}}{1+e^{h(\mathbf{v})}}. \tag{4}$$

Algorithm 1 Training a decision tree in a random forest

1: Randomly drawn a subset samples with replacement from the original training data for each tree.
2: Randomly propose a set of splitting candidates $\phi = (\theta, \pi)$ (feature parameter θ and thresholds π)
3: Partition the set of sample $Q = \{(V_i, \mathbf{v}_i)\}$ into left and right subsets by each ϕ:

$$Q_l(\phi) = \{(V,\mathbf{v}) | f_\theta(V,\mathbf{v}) < \pi\}$$
$$Q_r(\phi) = Q \backslash Q_l(\phi)$$

4: Find the ϕ that maximize the Gini index:

$$\phi^* = \arg\max_{\phi} gini(\phi)$$
$$gini(\phi) = \sum_{i=1}^{c} \hat{p}_i(1-\hat{p}_i)$$

where $\hat{p}_i$ is the probability of class i estimated from the samples in the subset Q.
5: If the largest Gini index gini(ϕ^*) is sufficient, and the depth in the tree is below a maximum, then iterate recursively for left and right subsets $Q_l(\phi^*)$ and $Q_r(\phi^*)$

TABLE I

STATISTICS OF DETECTION ERRORS COMPARED TO DENT-LANDMARK GROUNDTRUTH (IN MM)

Learning framework	average	min	max
Random Forest	**3.1532**	**0.8402**	**6.2716**
Adaboost	3.4888	1.2768	8.1552

III. EXPERIMENT

In this section we describe the experiments to evaluate the proposed method.

A. Data

The dataset used in the experiment has 73 CBCT volumetric data. The CBCT scan was obtained with 400×400 matrix taken at $0.4mm$ slice thickness. The number of slices in one CBCT volume is around 327. We randomly split the dataset into a training set with 50 volumes and a testing set with the rest 23 volumes.

B. Experiment setup

By considering the constrained search space, we sample training samples from region centered at $P(195,320,135)$ with size of $40 \times 60 \times 110$. The sampling pattern for sampled context features in the experiment is of size $100 \times 100 \times 100$. In particular, $216 = 6 \times 6 \times 6$ points are uniformly drawn from the sampling pattern. Intensity I and gradient $\|g\|$ of each point are used to form the feature vector. Together with normalized coordinate of each point, the feature dimension is $216 \times 2 + 3 = 435$. Using Eq. 1, 125 positive and 150 negative samples are extracted from the constrained search space. In summary, we use $13750 = (125 + 150) \times 50$ samples for training.

C. Experimental results

A random forest with 10 trees is learned with training samples from the constrained search space. Each tree is with depth of 6. For each node in a tree, 10 features are randomly selected as input. For Adaboost as discriminative learning framework, 50 weak classifiers are selected to build a strong classifier.

To locate the Dent-landmark in a testing volume, we search in the constrained search space. Each candidate returns a probability of being a Dent-landmark. Candidates with probability value larger than 80% are used to evaluate the final result. The average of coordinates of these selected candidates is chosen as Dent-landmark location.

We evaluate the results by random forest and Adaboost as learning framework. The detection error (in mm) is taken to analyze the results. Detection errors of all 23 testing CBCT volumes are given in Fig. 5. The statistics of these detection errors is listed in Table I. The average detection error by random forest as learning framework yield to $3.15mm$. The results demonstrate that the proposed method is effective to detect Dent-landmark in CBCT dental volume.

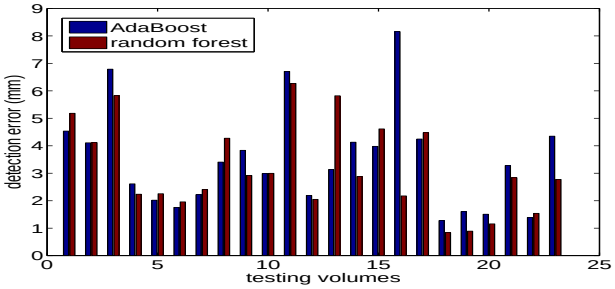


Fig. 5. Detection error of each testing CBCT volume

IV. CONCLUSIONS

In this paper we propose a discriminative method for automatic Dent-landmark detection in 3D CBCT dental volumes. Specifically, we use the random forest combined with sampled context features for this purpose. In addition, spacial prior is integrated to form a constrained search space for increase the efficiency. Promising results are obtained in the dataset with 73 CBCT volumes. In the future, we plan to explore high-level features for landmark detection in 3D CBCT dental data. More landmarks will be studied in our future work as well.

V. ACKNOWLEDGMENTS

This work was supported in part by NSF Research Grant IIS-0916624. The funding agency specifically disclaims responsibility for any analysis, interpretations and conclusions.

REFERENCES

[1] M. Maeda, A. Katsumata, Y. Ariji, A. Muramatsu, K. Yoshida, S. Goto, K. Kurita, and E. Ariji, "3d-ct evaluation of facial asymmetry in patients with maxillofacial deformities," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 102, no. 3, pp. 382–390, 2006.

[2] M. Moldez, K. Sato, J. Sugawara, and H. Mitani, "Linear and angular filipino cephalometric norms according to age and sex," *The Angle orthodontist*, vol. 76, no. 5, p. 800, 2006.

[3] W. Bholsithi, W. Tharanon, K. Chintakanon, R. Komolpis, and C. Sinthanayothin, "3d vs. 2d cephalometric analysis comparisons with repeated measurements from 20 thai males and 20 thai females," *Biomedical Imaging and Intervention Journal*, vol. 5, no. 4, p. e21, 2009.

[4] A. Katsumata, M. Fujishita, M. Maeda, Y. Ariji, E. Ariji, and R. Langlais, "3d-ct evaluation of facial asymmetry," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 99, no. 2, pp. 212–220, 2005.

[5] B. Gribel, M. Gribel, D. Frazã, J. McNamara Jr, and F. Manzi, "Accuracy and reliability of craniometric measurements on lateral cephalometry and 3d measurements on cbct scans," *The Angle orthodontist*, vol. 81, no. 1, pp. 28–37, 2011.

[6] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[7] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[8] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu, "Regression forests for efficient anatomy detection and localization in ct studies," in *Recognition Techniques and Applications in Medical Imaging, MICCAI workshop*, 2010, pp. 106–117.

[9] Y. Freund and R. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," *Computational Learning Theory*, pp. 23–37, 1995.

[10] Y. Zheng, B. Georgescu, H. Ling, S. K. Zhou, M. Suehling, and D. Comaniciu, "Constrained Marginal Space Learning for Efficient 3D Anatomical Structure Detection in Medical Images," *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 194-201, 2009.