Automated bone segmentation from dental CBCT images using patch-based sparse representation and convex optimization

Li Wang
Department of Radiology and BRIC, University of North Carolina at Chapel Hill, North Carolina 27599

Ken Chung Chen
Department of Oral and Maxillofacial Surgery, Houston Methodist Hospital Research Institute, Houston, Texas 77030 and Department of Stomatology, National Cheng Kung University Medical College and Hospital, Tainan, Taiwan 70403

Yaozong Gao, Feng Shi, Shu Liao, and Gang Li
Department of Radiology and BRIC, University of North Carolina at Chapel Hill, North Carolina 27599

Steve G. F. Shen and Jin Yan
Department of Oral and Craniomaxillofacial Surgery and Science, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University College of Medicine, Shanghai, China 200011

Philip K. M. Lee and Ben Chow
Hong Kong Dental Implant and Maxillofacial Centre, Hong Kong, China 999077

Nancy X. Liu
Department of Oral and Maxillofacial Surgery, Houston Methodist Hospital Research Institute, Houston, Texas 77030 and Department of Oral and Maxillofacial Surgery, Peking University School and Hospital of Stomatology, Beijing, China 100050

James J. Xia
Department of Oral and Maxillofacial Surgery, Houston Methodist Hospital Research Institute, Houston, Texas 77030; Department of Surgery (Oral and Maxillofacial Surgery), Weill Medical College, Cornell University, New York, New York 10065; and Department of Oral and Craniomaxillofacial Surgery and Science, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University College of Medicine, Shanghai, China 200011

Dinggang Shen[a)]
Department of Radiology and BRIC, University of North Carolina at Chapel Hill, North Carolina 27599 and Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea 136701

Purpose: Cone-beam computed tomography (CBCT) is an increasingly utilized imaging modality for the diagnosis and treatment planning of the patients with craniomaxillofacial (CMF) deformities. Accurate segmentation of CBCT image is an essential step to generate three-dimensional (3D) models for the diagnosis and treatment planning of the patients with CMF deformities. However, due to the poor image quality, including very low signal-to-noise ratio and the widespread image artifacts such as noise, beam hardening, and inhomogeneity, it is challenging to segment the CBCT images. In this paper, the authors present a new automatic segmentation method to address these problems.

Methods: To segment CBCT images, the authors propose a new method for fully automated CBCT segmentation by using patch-based sparse representation to (1) segment bony structures from the soft tissues and (2) further separate the mandible from the maxilla. Specifically, a region-specific registration strategy is first proposed to warp all the atlases to the current testing subject and then a sparse-based label propagation strategy is employed to estimate a patient-specific atlas from all aligned atlases. Finally, the patient-specific atlas is integrated into a *maximum a posteriori* probability-based convex segmentation framework for accurate segmentation.

Results: The proposed method has been evaluated on a dataset with 15 CBCT images. The effectiveness of the proposed region-specific registration strategy and patient-specific atlas has been validated by comparing with the traditional registration strategy and population-based atlas. The experimental results show that the proposed method achieves the best segmentation accuracy by comparison with other state-of-the-art segmentation methods.

Conclusions: The authors have proposed a new CBCT segmentation method by using patch-based sparse representation and convex optimization, which can achieve considerably accurate segmentation results in CBCT segmentation based on 15 patients. © *2014 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4868455]

1. INTRODUCTION

Segmentation of the cone-beam computed topographic (CBCT) image is an essential step for generating three-dimensional (3D) models in the diagnosis and treatment planning of the patients with craniomaxillofacial (CMF) deformities. It requires segmentation of bony structures from soft tissues, as well as separation of mandible from maxilla. CBCT scanners are becoming popularly used in the clinic, even in the private practice settings, due to its lower cost and lower dose compared to the conventional spiral multislice CT (MSCT) scanners.[1] However, the quality of CBCT images is significantly lower than that of spiral MSCT [Fig. 1(a)]. Also, there are many imaging artifacts [Fig. 1(b)], including noise, beam hardening, inhomogeneity, and truncation, that are inherent in CBCT units due to the nature of imaging acquisition and reconstruction process.[2] For example, CBCT machines for the purpose of dose reduction are often operated at milliamperes that are approximately one order of magnitude below those of medical CT machines. Thus, the signal-to-noise ratio of CBCT is much lower than that of CT. These artifacts affect image quality, and eventually the accuracy of subsequent segmentation.[3,4] Furthermore, based on the clinical requirement for correctly quantifying the deformity, CBCT scans are often acquired when the maxillary (upper) and mandibular (lower) teeth are in maximal intercuspation, i.e., the upper and lower teeth closely bite together as shown in Fig. 2. This leads to the display of both upper and lower teeth on the same cross-sectional slice as in Fig. 2(c), thus making it extremely difficult to separate them in the same slice.[5] To date, in order to use CBCT clinically for the diagnosis and treatment planning, the segmentation has to be done completely manually by the experienced operators.

However, manual segmentation is tedious, time-consuming, and error-prone.[6,7] Although automated segmentation methods have been previously developed for CBCT segmentation, they were mainly based on simple thresholding and morphological operations,[8] thus sensitive to the presence of the artifacts. Recently, statistical shape model (SSM) has been utilized for robust segmentation of the mandible.[9,10] However, these approaches are only applicable

to the objects with relatively regular and simple shapes (e.g., mandible). They are not applicable to the objects with complex shapes (e.g., maxilla). On the other hand, interactive segmentation approaches[5,11] were also proposed to integrate automatic segmentation with manual guidance. For example, Le *et al.* proposed an interactive geometric segmentation technique to separate upper and lower teeth in CT images.[5] However, these interactive methods[5,11] also have difficulty in producing plausible segmentations due to the presence of artifacts such as strong metal artifacts of dental implant.[5] In Ref. 12, Duy *et al.* proposed a fully automatic method for tooth detection and classification in CT image data. However, they only focus on the upper teeth extraction in the CT image and their separation algorithm does not work as well on CBCT as it does on CT due to the low contrast in CBCT.[13] To the best of our knowledge, our work is the first study aiming to fully automatically segment (and separate) the mandible from the maxilla on CBCT images.

Recently, there is a rapidly growing interest in using patch-based sparse representation.[14–16] This approach makes an assumption that image patches can be represented by sparse linear combination of image patches in an overcomplete dictionary.[17–20] This strategy has been applied to a good deal of image processing problems, such as image denoising,[17,18] image in-painting,[21] image recognition,[19,22] image super-resolution,[20] and deformable segmentations,[23] achieving promising results.

In this paper, we propose a new method for fully automated CBCT segmentation by using patch-based sparse representation to (**1**) segment bony structures from the soft tissues and (**2**) further separate the mandible from the maxilla. Specifically, we first propose a region-specific registration strategy to warp all the atlases to the current testing subject and then employ a sparse-based label propagation strategy to estimate a patient-specific atlas from all aligned atlases. Then, the estimated patient-specific atlas is integrated into a convex segmentation framework based on *maximum a posteriori probability* (MAP) for accurate segmentation. In this paper, the atlases consist of both spiral MSCT and CBCT images. The reasons for also using spiral MSCT images as a part of the atlases include (1) although image formation process is different between spiral CT and CBCT, they share the same patterns of anatomical structures (e.g., the bones generally have high intensities while the soft tissues have low intensities), which are captured in a patch fashion in our method to estimate the probability; (2) the spiral MSCT images have better contrast and less noise than the CBCTs, thus needing less processing time to construct atlases by manual segmentations (i.e., taking around 30 min for each MSCT image), compared to the CBCT images (i.e., requiring around 12 h for each CBCT image even by an experienced operator).

In summary, the novelty of our work includes the following:

- To clinically use CBCT for the dental diagnosis and treatment planning, we propose a **fully automated segmentation method** for the dental CBCT images, which

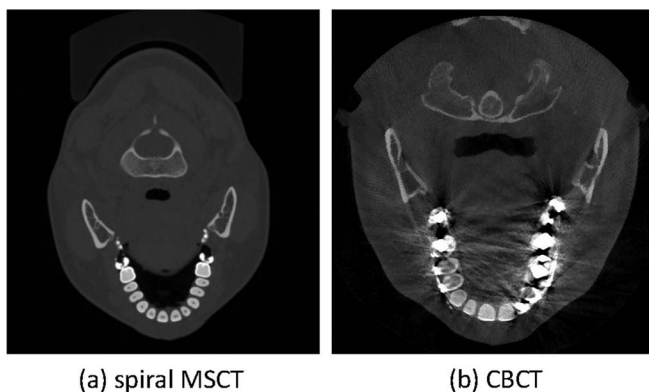(a) spiral MSCT   (b) CBCT

Fig. 1. Comparison between (a) spiral MSCT and (b) CBCT images. Comparing to the MSCT, CBCT scans have severe image artifacts, including noise, beam hardening, inhomogeneity, and truncation.
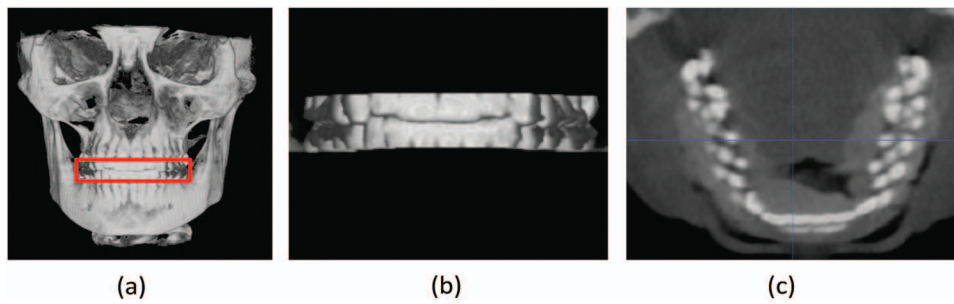
FIG. 2. (a) Maximal intercuspation. (b) Zoomed view of the teeth. (c) Both upper and lower teeth are displayed on the same cross-sectional slice.

currently has to be done completely manually by the experienced operators (taking around 12 h as mentioned above). This automated method shows its potential in solving the important clinical problem based on our diverse data from 15 patients with different ages and conditions.

- We also propose using the MSCT samples to estimate a patient-specific atlas for each CBCT image by employing a recently proposed patch-based sparse representation technique, since the basic patterns in MSCT and CBCT were found to be very similar.
- We further integrate our estimated patient-specific atlas into a convex segmentation framework, based on maximum *a posteriori* probability, for more accurate CBCT segmentation. This type of atlas-guided level sets segmentation method has not been previously developed for the challenging CBCT image segmentation.
- Most importantly, the performance of our proposed method is much better than any of other state-of-the-art methods, including auto segmentation, and multi-atlas based labeling with majority voting (MV) or conventional patch-based method (CPM).

Note that partial results in this paper were reported in our recent conference paper.[24] The remainder of this paper is organized as follows. The proposed method is introduced in Sec. 2. The experimental results are then presented in Sec. 3, followed by the discussion and conclusion in Sec. 4.

2. METHOD

In this study, we aim to segment a CBCT scan into three structures/regions: mandible, maxilla (the skull without the mandible), and background. Let Ω be the image domain and C be a closed subset in Ω, which divides the image domain into three disjoint partitions $\Omega_{i=1}^3$, such that $\Omega = \bigcup_{i=1}^3 \Omega_i$. For every voxel $\chi \in \Omega$, we first estimate its patient-specific probability belonging to each class $p_i(\chi) \triangleq p(\chi \in \Omega_i)$, and then use this patient-specific probability map (also called as atlas) as a prior to segment the CBCT image based on *maximum a posteriori probability*. Finally, a convex segmentation framework is proposed for precise segmentation. The flowchart of the proposed method is shown in Fig. 3.

2.A. Subjects

The CBCT scans of 15 patients (6 males/9 females) with nonsyndromic dentofacial deformity and treated with a double-jaw orthognathic surgery were included in this study. Their average age at the time of surgery was 26 ± 10 years (range: 10–49 years). These CBCT scans were acquired with a matrix of 400×400, a resolution of 0.4 mm isotropic voxel, and the time of exposure of 40 s. The details of imaging protocol are provided in Table I. On the other hand, additional 30 subjects with normal facial appearance scanned at maximal intercuspation were randomly selected from our HIPAA deidentified MSCT database. Their ages were 22 ± 2.6 years (range: 18–27 years). The MSCT images were acquired with a matrix of 512×512, a resolution of $0.488 \times 0.488 \times 1.25$ mm,[3] and the time of exposure of less than 5 s. All the MSCT and CBCT images were HIPAA deidentified prior to the study. The study was approved by The Methodist Hospital Institutional Review Board. These 30 CTs and 15 CBCTs were manually segmented to serve as the ground-truth segmentations by the two experienced CMF surgeons using Mimics 10.01 software (Materialise NV, Leuven, Belgium).

2.B. Region-specific registration with guidance of landmarks

Traditionally one can align the template image with the to-be-segmented image based on image intensity. However, it imposes two limitations in the application of MSCT/CBCT registration: (1) Head MSCT is usually acquired with different field of view (FOV) from CBCT [cf. Figs. 4(a) and 4(b)]. If we directly register them using image intensity, the registration error could be very large. (2) Precise separation of the closely bitted mandibular and maxilla teeth requires accurate registration of two images on the teeth region. Since teeth region only takes a small part of the entire image, by using global image registration, the registration accuracy of teeth region is usually limited. To overcome these limitations,
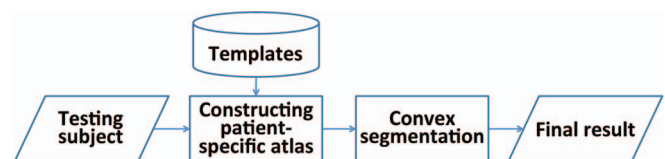

FIG. 3. The flowchart of the proposed method.

TABLE I. Summary of dental imaging protocols for MSCT and CBCT used in this study.

Image	kV	mA (s)	FOV, mm X, Y	FOV, mm Z	Voxels, mm X, Y	Voxels, mm Z	Dimensions X, Y	Dimensions Z
MSCT: GE LightSpeed RT	120	120	250	300	0.488	1.25	512	240
CBCT: i-CAT	120	0.25	160	160	0.4	0.4	400	400

we propose a region-specific, landmark-guided registration, with the landmark guidance coming from the two levels as described next. **In the whole head level**, we automatically detect 15 anatomical landmarks [Fig. 4(a)] using a robust landmark detection algorithm[25] to affine align MSCT atlases [e.g., Fig. 4(b)] with the CBCT testing image [Fig. 4(c)]. In this way, the affine alignment is not subject to different FOVs, since it is purely estimated from the detected landmarks. After landmark-guided affine alignment, MSCT/CBCT would have the same FOV and thus we can use deformable registration algorithm to further improve the registration accuracy. In fact, there are many registration methods[26–30] that we can employ. In this paper, B-spline registration algorithm based on the mutual information[31] is adopted for deformable registration. This algorithm was implemented in the free-access Elastix toolbox.[32] Figures 4(d) and 4(e) show the warped results without and with head-level landmarks, respectively. It can be seen that the result without guidance of head-level landmarks miss

matching with the testing CBCT. **In the teeth level**, we also rely on the automatically detected anatomical landmarks to improve the registration accuracy. Specifically, to align teeth regions of MSCT/CBCT, we use only eight teeth landmarks [Fig. 4(a)] for estimating the affine registration. Ideally, after head-level landmark-guided affine registration, we could adopt deformable registration to further improve the registration accuracy. However, as shown in Fig. 2, due to the beam hardening and complex bone structure, it is difficult for the current state-of-the-art deformable registration algorithms to work well. Figure 4(e) shows a typical result after deformable registration on the teeth region, which is worse than the registration result using only teeth-landmark-guided registration as shown in Fig. 4(f). In Sec. 2.C, we will use two separate registrations to estimate the patient-specific atlas at two different regions, i.e., teeth region and other remaining regions. Specifically, to estimate the patient-specific atlas at teeth region, we will use only those eight teeth landmarks to derive the registration. To estimate the patient-specific atlas for other regions, we will use all 15 landmarks to derive an initial affine registration and then adopt Elastix[32] for further refinement.

2.C. Estimating a patient-specific atlas

Atlas-based segmentation has demonstrated its robustness and effectiveness in many medical image segmentation problems.[33] Specifically, a population-based atlas, in the context of this work, is defined as the pairing of a template image and its corresponding tissue probability maps.[34] The
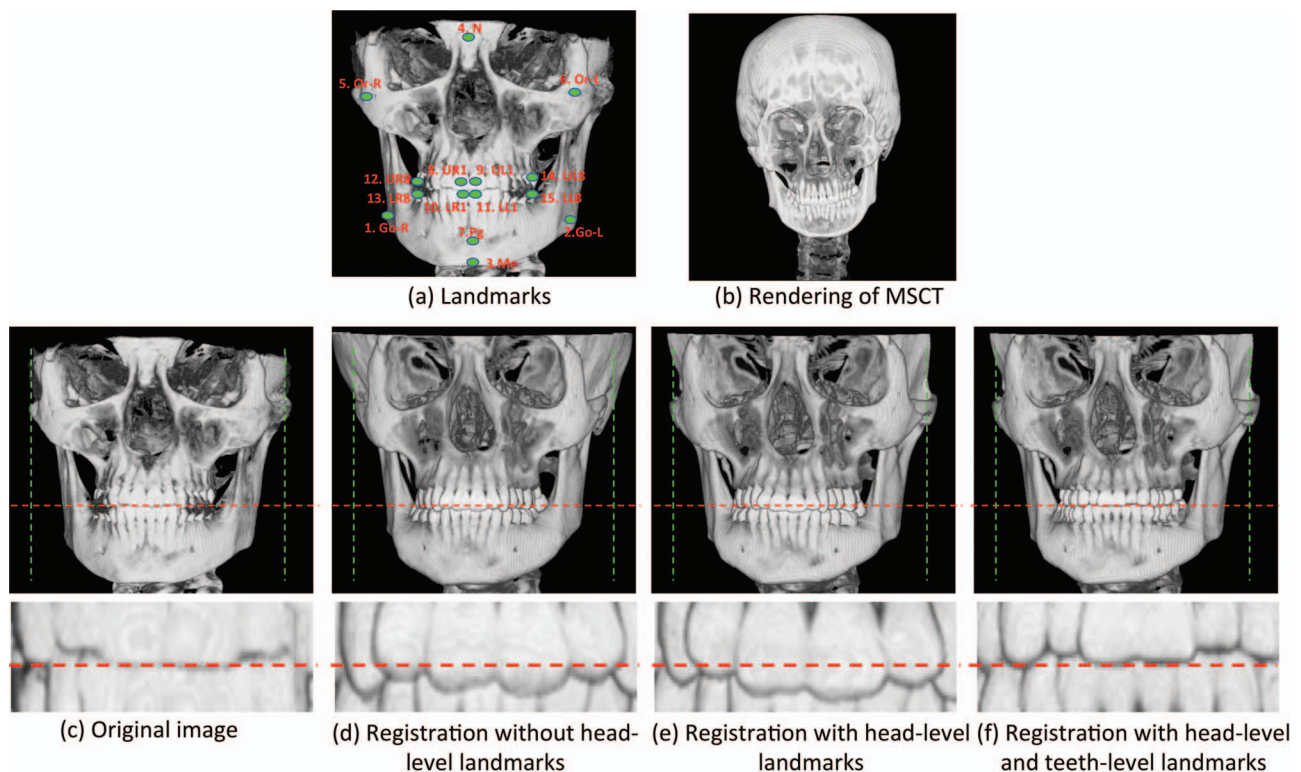

FIG. 4. (a) Fifteen anatomical landmarks: #1–#15 are used to estimate the affine matrix to warp the atlases onto the testing subject; #8–#15 are used to estimate the affine matrix to warp the teeth joint regions between the template images and the testing image. (c) is the same as (a). (d) and (e) show the registration results without and with head-level landmarks, respectively. (f) shows the registration result with both head-level and teeth level landmarks.
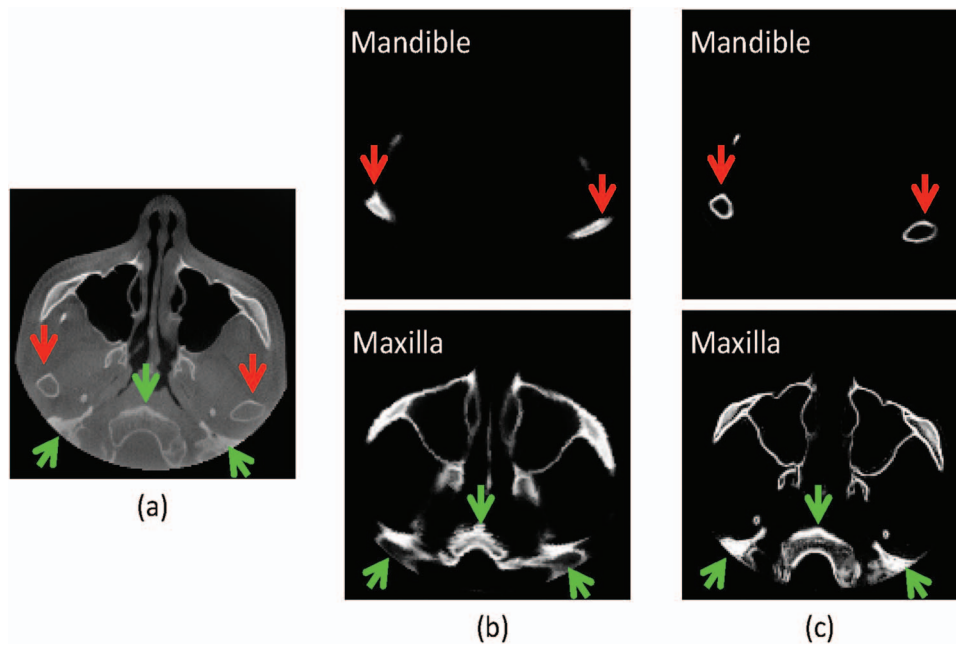
FIG. 5. (a) Original image. Comparison between population-based atlas (b) and patient-specific atlas (c). The proposed patient-specific atlas produces a more accurate estimation than the population-based atlas, especially in the regions indicated by the arrows.

template image often refers to an intensity image, equally or weighted averaged from a number of aligned intensity images of the training subjects. Given a population-based atlas, image registration could be used to warp this atlas to the query image, to yield a transformation that allows the atlas's tissue probability maps to be transformed and treated as tissue probabilities for the query subject. However, a population-based atlas often fails to provide useful guidance, especially in the regions with high intersubject anatomical variability, thus leading to unsatisfactory segmentation results [Fig. 5(b)]. One way to overcome this problem is to integrate the patient-specific information[35–37] in the atlas construction. To this end, we propose to construct a patient-specific atlas by combining both population and patient information as detailed below.

Specifically, we propose to estimate the patient-specific atlas by using a patch-based representation technique.[38,39] The rationale is that an image patch generally provides richer information, e.g., anatomical pattern, than a single voxel. First, N intensity images of atlases $I_j(j = 1, \ldots, N)$ and their corresponding segmentations S_j are aligned onto the space of the testing image I according to the registration method described in Sec. 2.B. Then, for each voxel χ in the testing image I, its corresponding intensity patch with size of $w \times w \times w$ can be represented as a column vector $Q(\chi)$. Similarly, at the same location, we can obtain patches $Q_j(\chi)$ from the jth aligned atlas. An initial codebook B can be constructed using all these atlas patches, i.e., $B(\chi) = [Q_1(\chi), Q_2(\chi), \ldots, Q_N(\chi)]$. To alleviate the effect of possible registration errors, the initial codebook B is extended to include more patches from the neighboring search window (i.e., a $w' \times w' \times w'$ cubic) of all N aligned atlases, thus generating an overcomplete codebook with superior representation capacity.[40] In the codebook $B(\chi)$, each patch is represented by a column vector and normalized with unit $\ell 2$ norm.[40,41] To represent the patch

$Q(\chi)$ by the codebook $B(\chi)$, its coding vector c could be estimated by many coding schemes, such as vector quantization, locality-constrained linear coding,[42] and sparse coding.[43,44] In this paper, we utilize the sparse coding scheme[43,44] to estimate the coding vector c by minimizing a non-negative Elastic-Net problem,[45]

$$\min_{c \geq 0} \frac{1}{2} \| Q(\chi) - B(\chi)c\|_2^2 + \lambda_1 \|c\|_1 + \frac{\lambda_2}{2} \|c\|_2^2, \qquad (1)$$

where the first term is the least square fitting term, the second term is the $\ell 1$ regularization term used to enforce the sparsity constraint on the reconstruction vector c, and the last term is the $\ell 2$ smoothness term used to enforce the similarity of coding coefficients for similar patches. Equation (1) is a convex combination of $\ell 1$ lasso[46] and $\ell 2$ ridge penalty, which encourages a grouping effect while keeping a similar sparsity of representation.[45] Each element of the coding vector c, i.e., $c_j(y)$, reflects the similarity between the target patch $Q(\chi)$ and the patch $Q_j(y)$ in the codebook. Based on the assumption that the similar patches should share similar labels, we use the sparse coding c to estimate the prior probability of the voxel χ belonging to the ith structure/region (e.g., mandible, maxilla, soft-tissue, or background), i.e.,

$$p_i(\chi) = \frac{1}{Z} \sum_j \sum_{y \in W(\chi)} c_j(y)\delta_i(S_j(y)), \qquad (2)$$

where Z is a normalization constant to ensure $\sum_i p_i(\chi) = 1$, and $\delta_i(S_j(y)) = 1$ if the label $S_j(y) = i$; otherwise, $\delta_i(S_j(y)) = 0$. By visiting each voxel in the testing image I, we can finally build a patient-specific atlas p_i. The constructed patient-specific atlas for the image shown in Fig. 5(a) is provided in Fig. 5(c), which captures more accurate patient-specific anatomical structures than the traditional population-based atlas [Fig. 5(b)], which was constructed using Joshi *et al.*'s
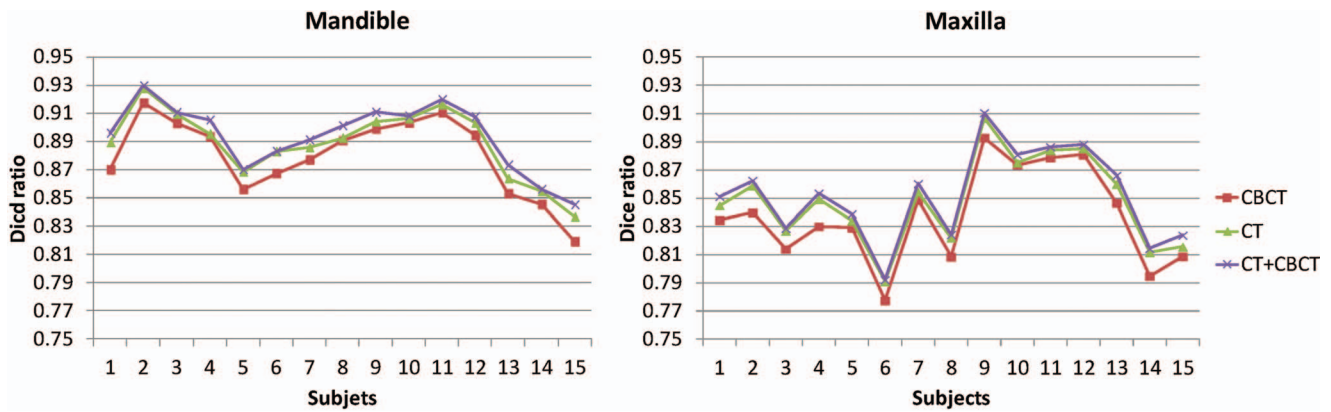
FIG. 6. Evaluation of Dice ratios on mandible and maxilla with respect to different sets of atlases: 15 CBCT subjects (excluding the testing one), 30 MSCT subjects, and 15 CBCT + 30 MSCT subjects (excluding the testing one).

groupwise registration method.[47] Finally, the patient-specific atlas can be used as a prior to accurately guide the subsequent level-sets based segmentation.

2.D. Convex segmentation based on MAP

The patient-specific atlas could provide a rough segmentation if simply thresholding the priori probability for each voxel. However, it may result in spatially inconsistent segmentation since the patient-specific atlas is estimated in a voxelwise manner. Moreover, in the construction of the patient-specific atlas, it is difficult to enforce the geometrical constraints, e.g., no overlapping between mandible and maxilla. To address these issues, we propose the following strategy to integrate the patient-specific atlas into a level-set framework based on the *maximum a posteriori probability* rule.

Specifically, to accurately label each voxel χ in the image domain Ω of the testing image, we jointly consider its neighboring voxels $y \in \mathcal{O}(\chi)$, where $\mathcal{O}(\chi)$ is the neighborhood of voxel χ. In fact, we can use a Gaussian kernel K_ρ with scale ρ to control the size of the neighborhood $\mathcal{O}(\chi)$.[48] The regions $\{\Omega_i\}$ produce a partition of the neighborhood $\mathcal{O}(\chi)$, i.e., $\{\Omega_i \cap \mathcal{O}(\chi)\}_{i=1}^3$. We first consider the segmentation of $\mathcal{O}(\chi)$ based on maximum *a posteriori* probability. According to the
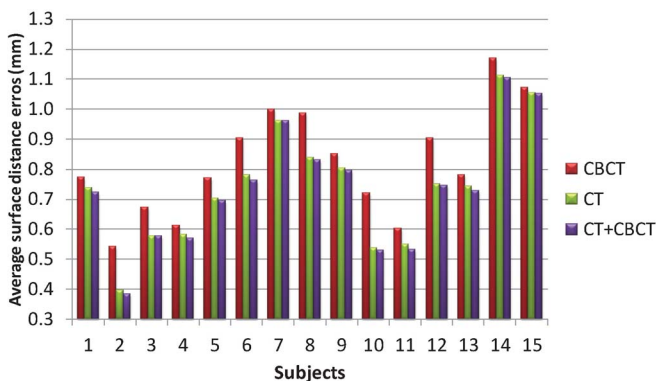

FIG. 7. Evaluation of average surface distance errors on mandible with respect to different sets of atlases: 15 CBCT subjects, 30 MSCT subjects, and 15 CBCT + 30 MSCT subjects.

Bayes rule:

$$p(y \in \Omega_i \cap \mathcal{O}(\chi)|I(y))$$
$$= \frac{p(I(y)|y \in \Omega_i \cap \mathcal{O}(\chi))p(y \in \Omega_i)p(y \in \mathcal{O}(\chi))}{p(I(y))}, \quad (3)$$

where $p(I(y)|y \in \Omega_i \cap \mathcal{O}(\chi))$, denoted by $p_{i,\chi}(I(y))$, is the structure probability density in region $\Omega_i \cap \mathcal{O}(\chi)$. Note that $p(y \in \Omega_i)$, i.e., $p_i(y)$, is the *a priori* probability of y belonging to the region Ω_i, which has been estimated in Sec. 2.B. Note that the prior probability was not utilized in the previous work[49] (all partitions were assumed to be equally possible) and only population-based probability was utilized.[50] $p(y \in \mathcal{O}(\chi)) = \mathbf{1}_{\mathcal{O}(\chi)}(y)$ is the indicator function, and $p(I(y))$ is independent of the choice of the region and can therefore be neglected. Accordingly, Eq. (3) can be simplified as

$$p(y \in \Omega_i \cap \mathcal{O}(\chi)|I(y)) \propto p_{i,\chi}(I(y))p_i(y). \quad (4)$$

Assuming that the voxels within each region are independent, the MAP will be achieved only if the product of $p_{i,\chi}(I(y))p_i(y)$ across the regions $\mathcal{O}(\chi)$ is maximized: $\prod_{i=1}^3 \prod_{\Omega_i \cap \mathcal{O}(\chi)} p_{i,\chi}(I(y))p_i(y)$. Taking a logarithm transformation and then integrating all the voxels χ, the maximization can be converted to the minimization of the following energy:

$$\mathcal{E}^{\mathcal{D}}(C) = -\sum_{i=1}^3 \iint_{\Omega\,\Omega_i} K_\rho(x - y) \log p_{i,\chi}(I(y))p_i(y)dydx.$$
$$(5)$$

In this way, multiple level set functions could be used to represent the regions $\{\Omega_i\}$ as in the work of Li *et al.*[48] However, the minimization problem with respect to the level set function is usually a nonconvex problem and there is a risk of being trapped in the local minima. Alternatively, based on the work of Goldstein and Bresson,[51] multiple variables could be used by taking values between 0 and 1 to derive a convex formulation. Since, in our project, there are only three different regions of interest: mandible, maxilla, and background, we need just two segmentation variables $u_1 \in [0\ 1]$ and $u_2 \in [0\ 1]$ to represent the partitions $\{\Omega_i\}$:$M_1 = u_1, M_2 = u_2$, $M_3 = (1 - u_1)(1 - u_2)$. Therefore, we convert Eq. (5) as
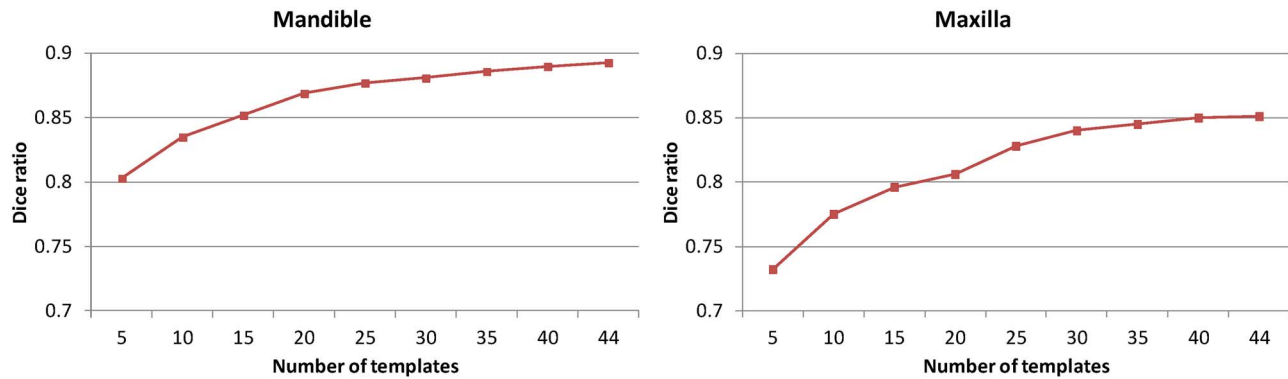
FIG. 8. Changes of Dice ratio of segmentation with respect to the number of atlases used. Experiment is performed by leave-one-out strategy on a set of 45 atlases.

follows:

$$\mathcal{E}^{\mathcal{D}}(u_1, u_2) = -\sum_{i=1}^{3} \iint_{\Omega \ \Omega} K_\rho(x - y)$$

$$\times \log p_{i,x}(I(y))p_i(y)M_i(y)dydx. \qquad (6)$$

There are many options to estimate $p_{i,x}(I(y))$. In this paper, we utilize a Gaussian distribution model with the local mean $\mu_i(x)$ and the variance $\sigma_i^2(x)$ (Ref. 52) to estimate it:

$$p_{i,x}(I(y))$$
$$= \exp(-(\mu_i(x) - I(y))^2 / 2\sigma_i^2(x)) / (\sqrt{2\pi}\sigma_i(x)). \qquad (7)$$

Based on the assumption that there should be no overlap between mandible and maxilla, we propose the following penalty constraint term:

$$\mathcal{E}^{\mathcal{P}}(u_1, u_2) = \int u_1(x)u_2(x)dx. \qquad (8)$$

In addition, the length regularization term[53] is defined as the weighted total variation of functions u_1 and u_2,

$$\mathcal{E}^{\mathcal{R}}(u_1, u_2) = \int g(I(x))(|\nabla u_1(x)| + |\nabla u_2(x)|)dx, \qquad (9)$$

where g is a nonedge indicator function that vanishes at object boundaries.[53]

Finally, we define the entire energy functional below, which consists of the data fitting term $\mathcal{E}^{\mathcal{D}}$, the overlap penalty term $\mathcal{E}^{\mathcal{P}}$, and the length regularization term $\mathcal{E}^{\mathcal{R}}$:

$$\min_{u_1, u_2 \in [0 \ 1]} \{\mathcal{E}(u_1, u_2) = \mathcal{E}^{\mathcal{D}} + \alpha\mathcal{E}^{\mathcal{P}} + \beta\mathcal{E}^{\mathcal{R}}\}, \qquad (10)$$

where α and β are the positive coefficients. Based on Ref. 51, the energy functional in Eq. (10) can be easily minimized in a fast way with respect to u_1 and u_2.

3. EXPERIMENTAL RESULTS

3.A. Evaluation metrics

In the following, we mainly employ Dice ratio to evaluate the segmentation accuracy, which is defined as
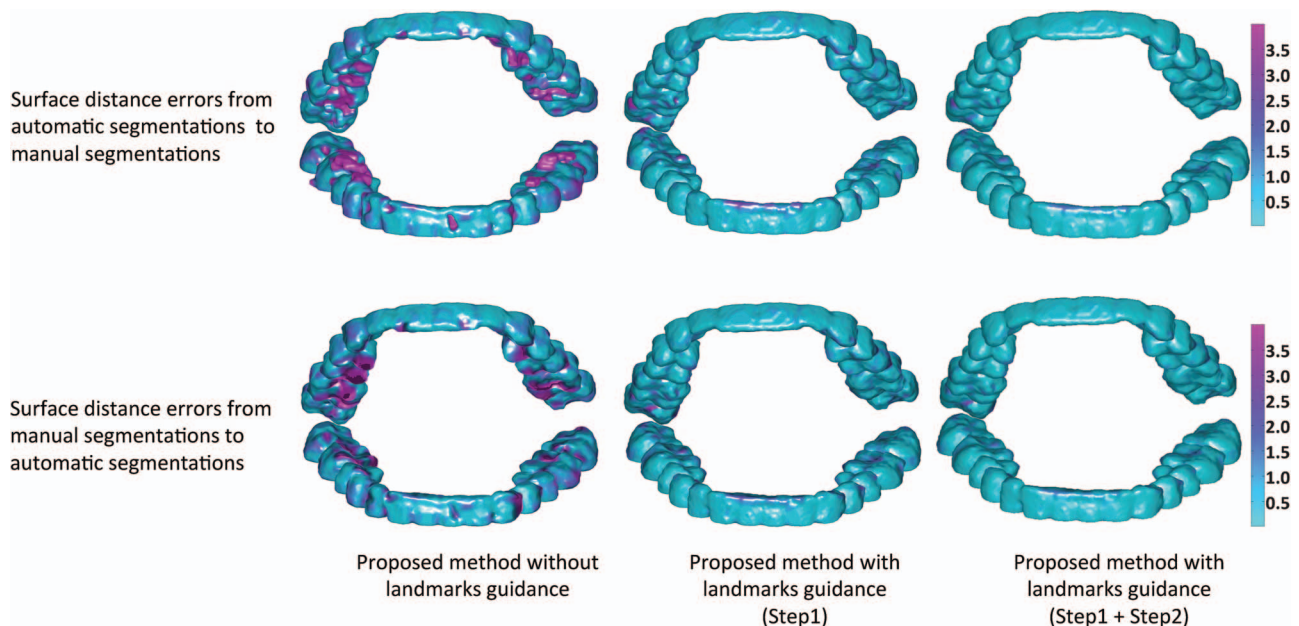
$$DR = 2|A \cap B|/(|A| + |B|),$$

FIG. 9. Comparison of results by the proposed method without or with landmark guidance.
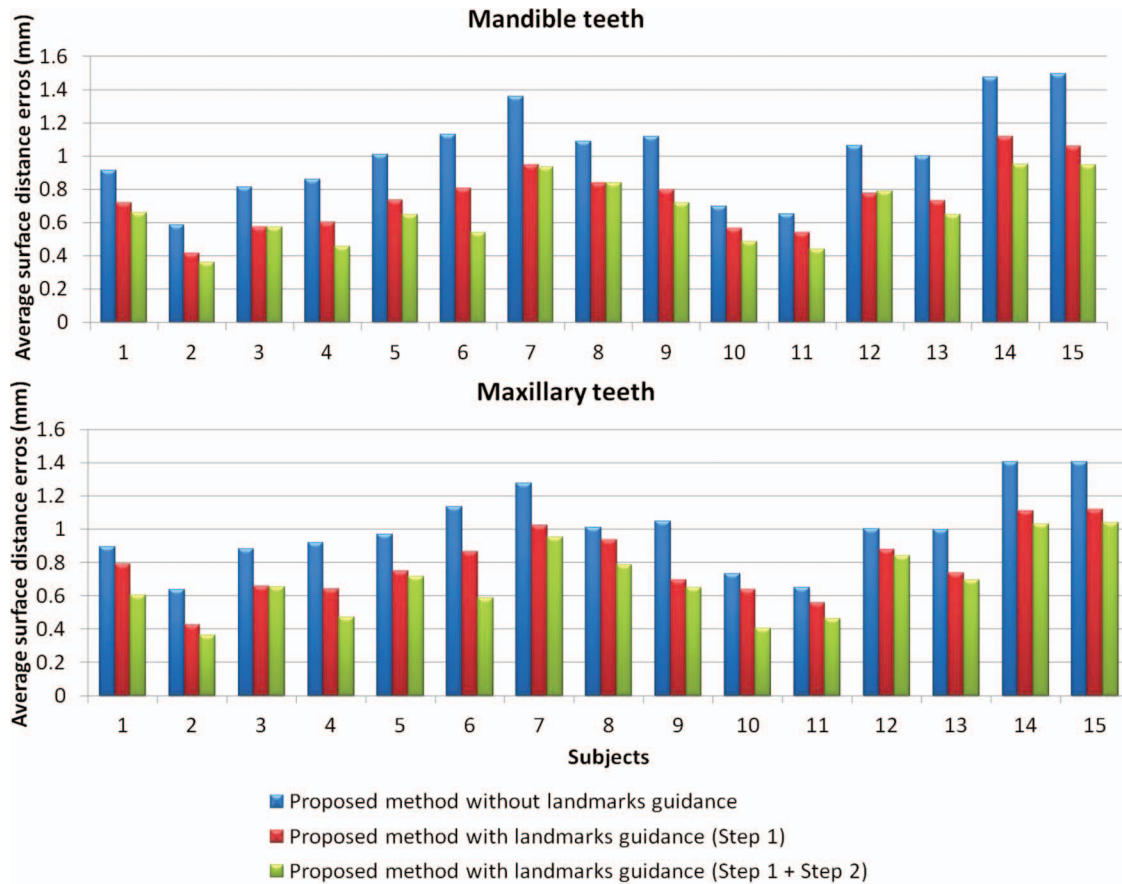
FIG. 10. Average surface distances between the surfaces obtained by three different settings of our method and the manual segmented surfaces, on 15 subjects in the teeth joint regions.

where A and B are two segmentation results of the same image. We also evaluate the accuracy by measuring the average surface distance error, which is defined as

$$D(A, B) = \frac{1}{2}\left(\frac{1}{n_A} \sum_{a \in surf(A)} dist(a, B) \right.$$

$$\left. + \frac{1}{n_B} \sum_{b \in surf(B)} dist(b, A) \right),$$

where $surf(A)$ is the surface of segmentation A, n_A is the total number of surface points in $surf(A)$, and $dist(a, B)$ is the nearest Euclidean distance from a surface point a to the surface B.

3.B. Parameter optimization

The parameters in this paper were determined based on 30 MSCT subjects via cross-validation. For example, we tested the weight for $\ell1$-term $\lambda_1 = \{0.01, 0.1, 0.2, 0.5\}$, the weight for $\ell2$-term $\lambda_2 = \{0, 0.01, 0.05, 0.1\}$, the patch size $w = \{5, 7, 9, 11, 13\}$, and the size of neighborhood $w' = \{3, 5, 7, 9\}$. We then compare the sum Dice ratios of mandible and maxialla with respect to the different combinations of these parameters, and found that the best accuracy is achieved with the parameter $\lambda_1 = 0.1$, $\lambda_2 = 0.05$,

$w = 9$, $w' = 5$. Cross-evaluation was also performed to determine other parameters, such as $\rho = 3$ for the Gaussian kernel K_ρ, and the weights $\alpha = 10$ for the overlap penalty term $\mathcal{E}^P$ and $\beta = 10$ for the length regularization term $\mathcal{E}^R$. Empirically, we found the performance of our algorithm is insensitive to the small perturbation of these parameters. However, the above range for each parameter was empirically chosen in our experiments, which could lead to local minimum results as we will discuss later.

3.C. Atlas selection

In this section, we evaluate the performance of segmentation accuracy using different sets of atlases, i.e., using (1) 15 CBCT subjects (excluding the testing one), (2) 30 MSCT subjects, and (3) 15 CBCT + 30 MSCT subjects (excluding the testing one) as atlases. The Dice ratios on mandible and maxilla, and the average surface distance errors on mandible, with respect to different sets of atlases, are shown in Figs. 6 and 7, respectively. It can be seen that the use of CBCT atlases alone result in slightly lower Dice ratios and higher surface distance errors than those obtained using spiral MSCT atlases. One reason may be due to the limited number of CBCT atlases. It is not surprising that the combination of MSCT and CBCT achieves the highest Dice ratios and also the smallest surface
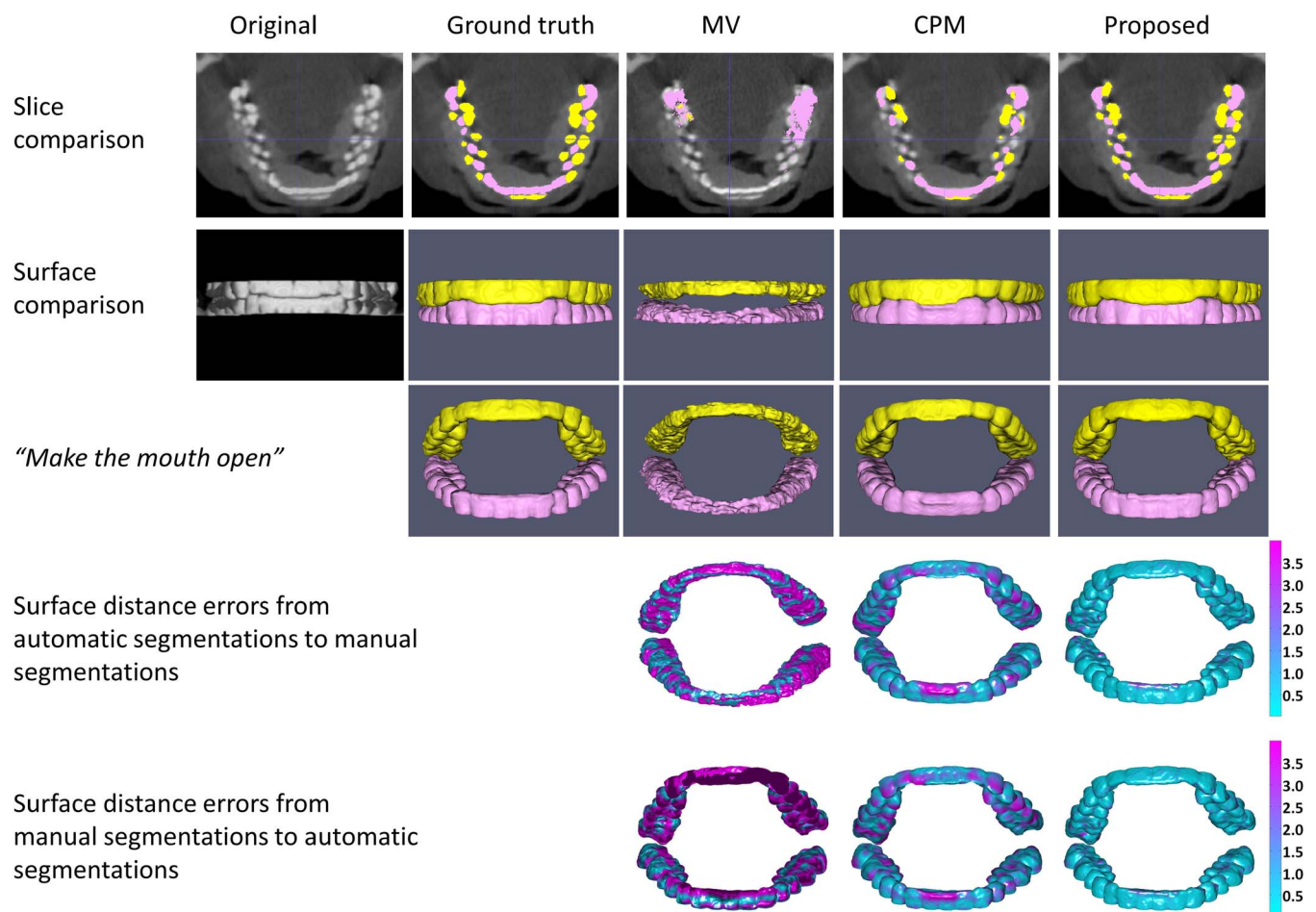
FIG. 11. Segmentation results of three different methods for the case with maximal intercuspation. The first and second rows show the comparison on slice and surfaces, respectively. In the third row, we rotate the upper teeth and lower teeth to virtually make the mouth open for better visualization. The last two rows show the surface distance errors by three different methods.

distance errors, due to inclusion of more atlases. Therefore, in this paper, we finally choose CBCT+ MSCT as the atlases.

3.D. Number of atlases vs accuracy

In the following, we explore the relationship between the number of atlases and the segmentation accuracy. Figure 8 shows the Dice ratios as a function of the number of atlases. As we can see, increasing the number of atlases generally improves the segmentation accuracy, as the average Dice ratio increased from 0.803 ($N = 5$) to 0.894 ($N = 44$) for mandible, and 0.732 ($N = 5$) to 0.855 ($N = 44$) for maxilla. However, the inclusion of more atlases would also bring in larger computational cost. Meanwhile, we found that, in our tests, when the number of atlases reaches 40, the performance of CBCT segmentation converges.

3.E. Performance for the case of maximal intercuspation

As shown in Fig. 2(b), due to the closed-bite position, the maxillary and mandibular teeth are touched to each other, which makes the automatic segmentation very difficult to

achieve, as mentioned in the Introduction. To demonstrate the importance of employing the landmark-guided region-specific registration for separation of the maxillary (upper) and mandibular (lower) teeth, we first show the comparison results of the proposed method (Step 1: estimation of the patient-specific atlas) without and with landmark guidance for the teeth region in Fig. 9. The surface distance errors between the manual segmentations and automatic segmentations are calculated for better comparison. We also rotate the upper and lower teeth in different directions for a better visualization of the opened mouth. It can be clearly seen that the proposed method with the landmark-guided registration produces much more accurate results than that without landmark guidance. To show the advantage of the convex formulation in segmentation, we further show the result of the proposed method with both Step 1 (estimation of the patient-specific atlas) and Step 2 (convex segmentation) in the third column of Fig. 9. The average surface distance errors for the mandibular teeth and the maxillary teeth on 15 subjects are plotted in Fig. 10, which clearly demonstrates the advantage of the landmark-guided registration in estimation of the patient-specific atlas and the subsequent convex segmentation.
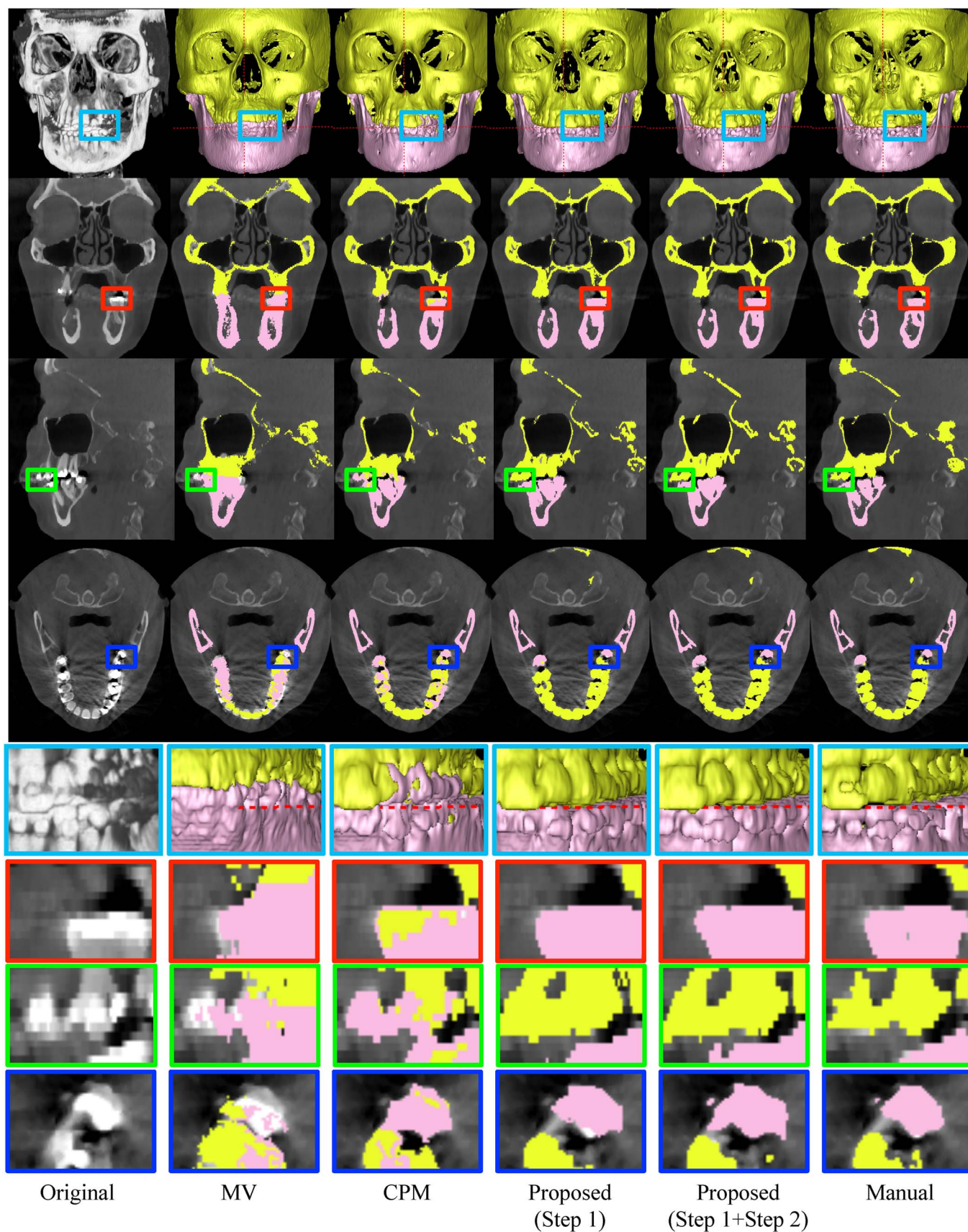
FIG. 12. Comparison of segmentation results by four different methods on a typical CBCT image. Note that the fourth row shows the segmentation of four different methods on the image shown in Fig. 1(b).
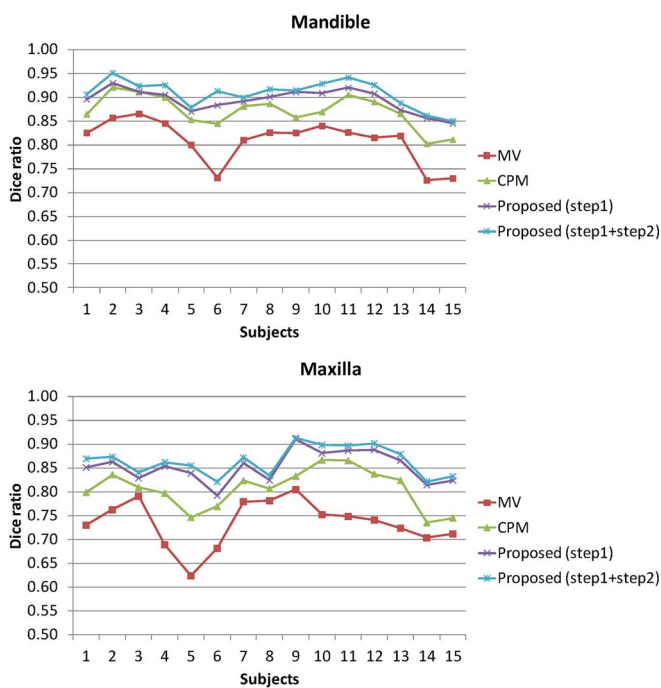
FIG. 13. Dice ratios of segmented mandible and maxilla by four different methods on 15 CBCT subjects.

3.F. Comparisons with the state-of-the-art methods

We first demonstrate the performance of different methods for the cases with maximal intercuspation [shown in Figs. 2(b) and 2(c)] in Fig. 11. We compare with other multi-atlases-based automatic segmentation methods,[38,39,54,55] using ma-

jority voting scheme, and conventional patch-based method.[39] Note that, for a fair comparison, we perform a similar cross-validation as in Sec. 3.B to derive the optimal parameters for the CPM, thus obtaining finally the patch size of $9 \times 9 \times 9$ and the neighborhood size of $5 \times 5 \times 5$. The segmentation results on a selected slice and two surfaces by different methods are shown in the first and second rows, respectively. It can be seen that the proposed method successfully delineates the maxillary and mandibular teeth. We further calculate the surface distance errors between the manual segmentations and automatic segmentations in the fourth and fifth rows, which demonstrates that the proposed method achieves the best accuracy. Note that, for fair comparison, the results from the proposed method were directly derived from the patient-specific atlas (Step 1) by simple thresholding, similar to the CPM.

We then demonstrate the performance of different methods on a typical subject as shown in Fig. 12. In the first row, from left to right show the volume rendering of the original intensity image, and the surfaces obtained by MV, CPM,[39] the proposed method by directly using the maximum class probability from Step 1 (prior estimation) as segmentation result, the proposed method with both Step 1 (prior estimation) and Step 2 (convex segmentation), and the manual segmentation. For better visualization, the corresponding results on slices and zoomed views are also provided in the bottom rows. Due to the errors in image registration process, the surface by MV is far from accurate, due to incorrect labeling of some upper teeth as lower teeth. Due to closed-bite position and the large intensity variations, the patch-based fusion method[39] cannot accurately separate the mandible from maxilla, thus
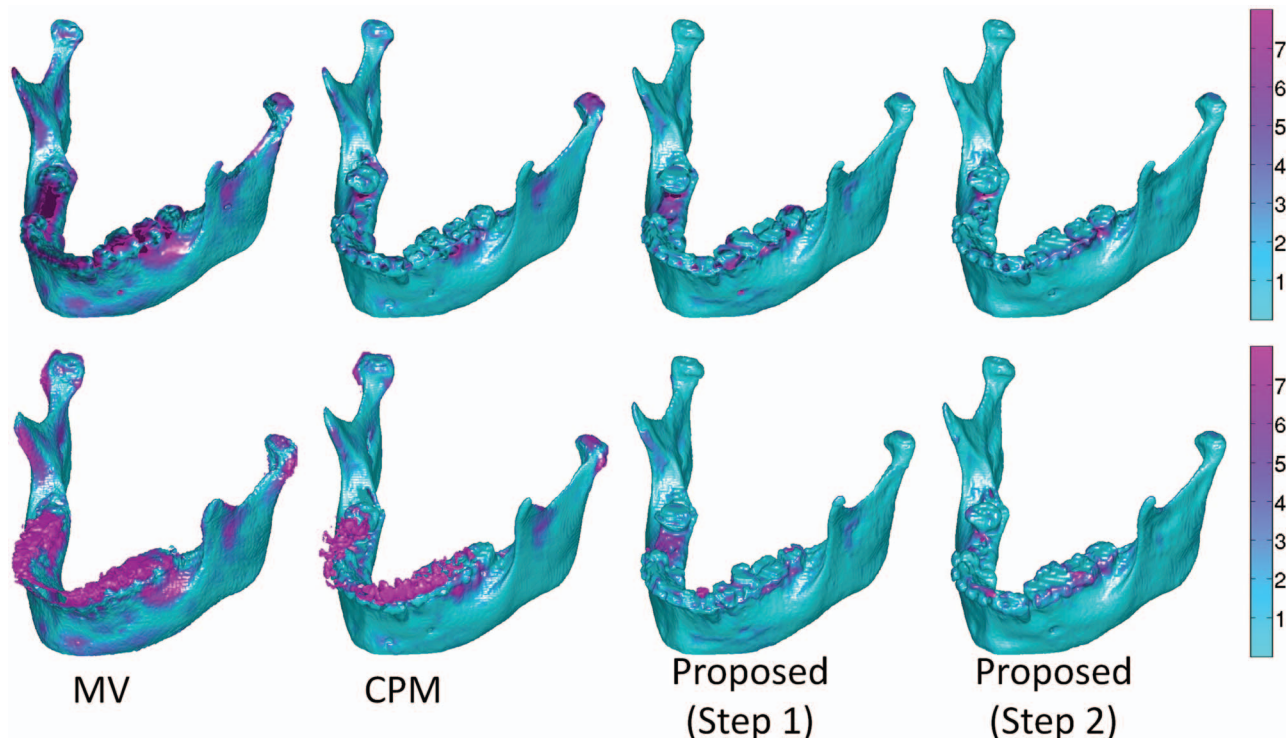

FIG. 14. The upper row shows the surface distances in mm from the surfaces obtained by four different methods to the manual segmented surfaces. The lower row shows the surface distances from the manual segmented surfaces to the surfaces segmented by four different methods.
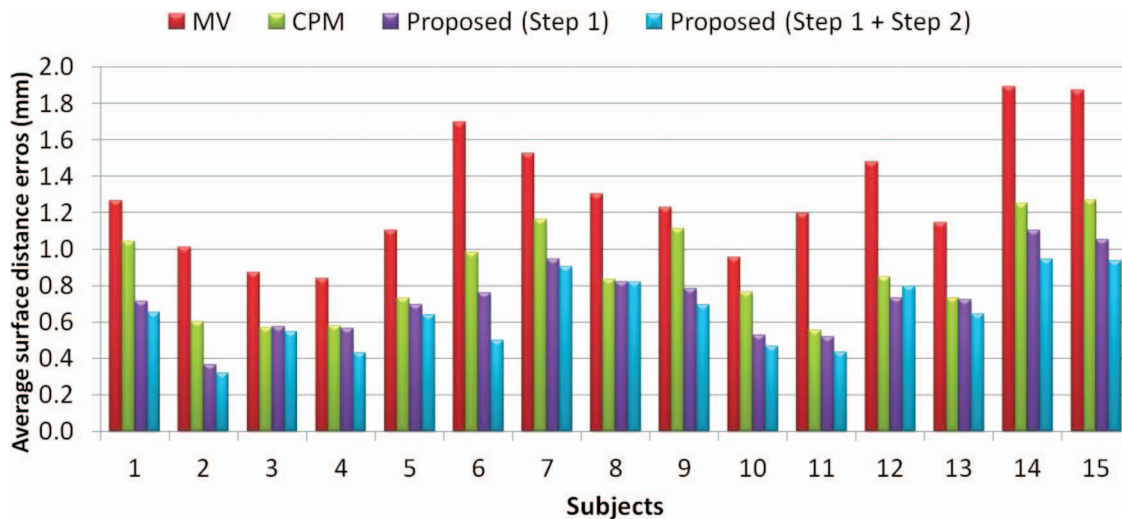
FIG. 15.   Average surface distances from the surfaces obtained by four different methods to the manual segmented surfaces on 15 subjects.

mislabeling the upper teeth and lower teeth. Overall, the proposed method produces much more reasonable results.

We finally use Dice ratio to quantitatively evaluate the overlap ratio between manual segmentations and automatic segmentation on 15 subjects, as shown in Fig. 13 and Table I. As can be observed, by integrating both patient-specific atlas and convex segmentation, the proposed method achieves the highest Dice ratios. The upper row of Fig. 14 shows the surface distances from the surfaces obtained by these methods to the manual segmented surfaces, and the lower row shows the surface distances from the manual segmented surfaces to the automatically segmented surfaces. As can be seen from Fig. 14, the proposed method achieves superior results over all other methods. The average surface distance errors on 15 subjects are plotted in Fig. 15. Additionally, the Hausdorff distance

$$D_H(A, B) = max\{dist(A, B), dist(B, A)\}$$

was also used to measure the maximal surface-distance errors of each of 15 subjects. The average Hausdorff distance on all 15 subjects are shown in Table II, which again demonstrates the advantage of our proposed method.

4. DISCUSSIONS

4.A. Normal subjects vs patients

The success of applying spiral MSCT atlases of the normal subjects to the CBCT images of the patients with CMF de-

formity can be mainly attributed to the following two factors: (1) The deformation between the subject with CMF deformity and the normal subject is first alleviated by the image registration; (2) After registration, the probabilities for the CBCT subject with CMF deformity are then robustly estimated by the proposed patch-based sparse technique in Sec. 2.C.

4.B. Image quality between MSCT and CBCT

Although the quality is different between MSCT and CBCT, the influence of the image quality is minimized due to the following reasons. First, our method works on image patches, where similar local patterns can be captured although the whole images may have large contrast differences. Second, in the following sparse representation, all the patches are normalized to have the unit $\ell2$-norm to alleviate the intensity scale problem.[40,41] Third, the testing patch from CBCT can be well represented by the overcomplete patch dictionary constructed from the MSCT, as they share the same patterns of anatomical structures (e.g., the bones generally have high intensities while the soft tissues have low intensities).

4.C. Patch-based sparse representation vs conventional patch matching method

The conventional patch-based matching methods[38,39,56-58] are less dependent on the accuracy of registration and this technique has been successfully validated on brain labeling[38] and hippocampus segmentation[39] with promising results. However, these methods[38,39,56-58] mainly use a simple

TABLE II. Average Dice ratios and surface distance errors (in mm) on 15 subjects (The best performance is indicated in boldface).

		MV	CPM (Ref. 39)	Proposed (Step 1)	Proposed (Step 1+2)
Dice ratio	Mandible	0.82 ± 0.04	0.88 ± 0.02	0.90 ± 0.03	**0.92 ± 0.02**
	Maxilla	0.74 ± 0.05	0.82 ± 0.03	0.86 ± 0.02	**0.87 ± 0.02**
Average distance error	Mandible	1.30 ± 0.33	0.87 ± 0.25	0.72 ± 0.20	**0.65 ± 0.19**
Hausdorff distance error	Mandible	3.71 ± 1.68	2.45 ± 1.43	1.25 ± 0.62	**0.96 ± 0.53**

intensity difference based similarity measure (Sum of the Squared Difference, SSD). Consequently, they are sensitive to the variance of tissue contrast and luminance, which is often viewed in CBCTs. By contrast, in the proposed method, we assume that image patches can be represented by sparse linear combination of image patches in an overcomplete dictionary. It is a representation problem instead of a directly patch-matching problem as in Refs. 38 and 39. In the proposed method, all the patches are normalized to have the unit ℓ2-norm to alleviate the variance of tissue contrast and luminance.[40,41] The testing patch is well represented by the overcomplete patch dictionary with the sparse constraint. The derived sparse coefficients are utilized to (1) measure the patch similarity, instead of directly using the intensity difference similarity as used SSD in Refs. 38 and 39 and (2) estimate a patient-specific atlas, instead of a population-based atlas.

4.D. Computational cost

In our implementation, we use the LARS algorithm,[59] which was available in the SPAMS toolbox (http://spams-devel.gforge.inria.fr), to solve the Elastic-Net problem. The average computational time is around 5 h for segmentation of a $400 \times 400 \times 400$ image with a spatial resolution of 0.4 mm isotropic voxel on a linux server with 8 CPUs and 16G memory. In our future work, we will further optimize the code to reduce the computational time.

4.E. Summary

We proposed a new method for automatic segmentation of CBCT images. We first estimated *a priori* probability from spiral MSCT atlases by using a sparse label fusion technique. Then, the priori probability was integrated into a convex segmentation framework based on MAP. Comparing to the state-of-the-art label fusion methods, our method achieved more accurate segmentation results. In our future work, we will validate the proposed method on more subjects and will also improve its robustness by increasing the variability of the atlases, such as including more training subjects with different CMF deformities.

ACKNOWLEDGMENTS

a)Author to whom correspondence should be addressed. Electronic mail: dgshen@med.unc.edu; Telephone: 919-966-3535; Fax: 919-843-2641.

[1]M. Loubele, R. Bogaerts, E. Van Dijck, R. Pauwels, S. Vanheusden, P. Suetens, G. Marchal, G. Sanderink, and R. Jacobs, "Comparison between effective radiation dose of CBCT and MSCT scanners for dentomaxillofacial applications," Eur. J. Radiol. **71**, 461–468 (2009).

[2]R. Schulze, U. Heil, D. Groβ, D. Bruellmann, E. Dranischnikow, U. Schwanecke, and E. Schoemer, "Artefacts in CBCT: A review," Dentomaxillofacial Radiol. **40**, 265–273 (2011).

[3]M. Loubele, F. Maes, F. Schutyser, G. Marchal, R. Jacobs, and P. Suetens, "Assessment of bone segmentation quality of cone-beam CT versus multislice spiral CT: A pilot study," Oral Surg., Oral Med., Oral Pathol., Oral Radiol., Endodontol. **102**, 225–234 (2006).

[4]W. Li, S. Liao, Q. Feng, W. Chen, and D. Shen, "Learning image context for segmentation of the prostate in CT-guided radiotherapy" Phys. Med. Biol. **57**, 1283–1308 (2012).

[5]B. H. Le, Z. Deng, J. Xia, Y.-B. Chang, and X. Zhou, "An interactive geometric technique for upper and lower teeth segmentation," in *MICCAI 2009*, Vol. 5762, edited by G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor (Springer, Berlin, 2009), pp. 968–975.

[6]D. Shen and H. H. S. Ip, "A Hopfield neural network for adaptive image segmentation: An active surface paradigm," Pattern Recognit. Lett. **18**, 37–48 (1997).

[7]Y. Zhan and D. Shen, "Automated segmentation of 3D US prostate images using statistical texture-based matching method," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, Vol. 2878, edited by R. Ellis and T. Peters (Springer, Berlin, 2003), pp. 688–696.

[8]B. A. Hassan, "Applications of cone beam computed tomography in orthodontics and endodontics," Thesis, Reading University, VU University Amsterdam, 2010.

[9]D. Kainmueller, H. Lamecker, H. Seim, M. Zinser, and S. Zachow, "Automatic extraction of mandibular nerve and bone from cone-beam CT data," in *MICCAI 2009*, Vol. 5762, edited by G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor (Springer, Berlin, 2009), pp. 76–83.

[10]S. T. Gollmer and T. M. Buzug, "Fully automatic shape constrained mandible segmentation from cone-beam CT data," in *9th IEEE International Symposium on Biomedical Imaging ISBI* (IEEE, Barcelona, 2012), pp. 1272–1275.

[11]S. Suebnukarn, P. Haddawy, M. Dailey, and D. Cao, "Interactive segmentation and three-dimension reconstruction for cone-beam computed-tomography images," NECTEC Tech. J. **8**, 154–161 (2008).

[12]N. T. Duy, H. Lamecker, D. Kainmueller, and S. Zachow, "Automatic detection and classification of teeth in CT data," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, Vol. 7510, edited by N. Ayache, H. Delingette, P. Golland, and K. Mori (Springer, Berlin, 2012), pp. 609–616.

[13]N. T. Duy, "Automatic segmentation for dental operation planning," Diploma thesis, Zuse Institut Berlin, 2013.

[14]L. Wang, F. Shi, G. Li, Y. Gao, W. Lin, J. H. Gilmore, and D. Shen, "Segmentation of neonatal brain MR images using patch-driven level sets," NeuroImage **84**, 141–158 (2014).

[15]L. Wang, F. Shi, Y. Gao, G. Li, J. H. Gilmore, W. Lin, and D. Shen, "Integration of sparse multi-modality representation and anatomical constraint for isointense infant brain MR image segmentation," NeuroImage **89**, 152–164 (2014).

[16]Y. Gao, S. Liao, and D. Shen, "Prostate segmentation by sparse representation based classification," Med. Phys. **39**, 6372–6387 (2012).

[17]M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process. **15**, 3736–3745 (2006).

[18]J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," IEEE Trans. Image Process. **17**, 53–69 (2008).

[19]J. Winn, A. Criminisi, and T. Minka, "Object categorization by learned universal visual dictionary," in *Tenth IEEE International Conference on Computer Vision*, Vol. 2 (IEEE Computer Society, Beijing, 2005), vol. 1802, pp. 1800–1807.

[20]J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Trans. Image Process. **19**, 2861–2873 (2010).

[21]M. J. Fadili, J. L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," Comput. J. **52**, 64–79 (2009).

[22]J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Discriminative learned dictionaries for local image analysis," in *Discriminative Learned Dictionaries for Local Image Analysis (CVPR)* (IEEE Computer Society, Anchorage, AK, 2008), pp. 1–8.

[23]S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou, "Deformable segmentation via sparse shape representation," in *MICCAI 2011*, Vol. 6892, edited by G. Fichtinger, A. Martel, T. Peters (Springer, Berlin, 2011), pp. 451–458.

[24]L. Wang, K. Chen, F. Shi, S. Liao, G. Li, Y. Gao, S. Shen, J. Yan, P. M. Lee, B. Chow, N. Liu, J. Xia, and D. Shen, "Automated segmentation of CBCT image using spiral CT atlases and convex optimization," in *Medical Image Computing and Computer-Assisted Intervention – MIC-*

CAI 2013, Vol. 8151, edited by K. Mori, I. Sakuma, Y. Sato, C. Barillot, N. Navab (Springer, Berlin, 2013), pp. 251–258.

[25] Y. Gao, Y. Zhan, and D. Shen, "Incremental learning with selective memory (ILSM): Towards fast prostate localization for image guided radiotherapy," in *Medical Image Computing and Computer-Assisted Intervention – MIC-CAI 2013*, Vol. 8150, edited by K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab (Springer, Berlin, 2013), pp. 378–386.

[26] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," IEEE Trans. Med. Imaging **21**, 1421–1439 (2002).

[27] E. I. Zacharaki, D. Shen, S.-k. Lee, and C. Davatzikos, "ORBIT: A multiresolution framework for deformable registration of brain tumor images," IEEE Trans. Med. Imaging **27**, 1003–1017 (2008).

[28] D. Shen, W.-h. Wong, and H. H. S. Ip, "Affine-invariant image retrieval by correspondence matching of shapes," Image Vis. Comput. **17**, 489–499 (1999).

[29] H. Jia, G. Wu, Q. Wang, and D. Shen, "ABSORB: Atlas building by self-organized registration and bundling," NeuroImage **51**, 1057–1070 (2010).

[30] S. Tang, Y. Fan, G. Wu, M. Kim, and D. Shen, "RABBIT: Rapid alignment of brains by building intermediate templates," NeuroImage **47**, 1277–1287 (2009).

[31] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," IEEE Trans. Image Process. **9**, 2083–2099 (2000).

[32] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, "Elastix: A toolbox for intensity-based medical image registration," IEEE Trans. Med. Imaging **29**, 196–205 (2010).

[33] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy," NeuroImage **46**, 726–738 (2009).

[34] F. Shi, P. Yap, G. Wu, H. Jia, J. H. Gilmore, W. Lin, and D. Shen, "Infant brain atlases from neonates to 1- and 2-year-olds," PLoS One **6**, e18746 (2011).

[35] Y. Shi, F. Qi, Z. Xue, L. Chen, K. Ito, H. Matsuo, and D. Shen, "Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics," IEEE Trans. Med. Imaging **27**, 481–494 (2008).

[36] Q. Feng, M. Foskey, S. Tang, W. Chen, and D. Shen, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009 (ISBI'09)* (IEEE, Boston, MA, 2009), pp. 282–285.

[37] Q. Feng, M. Foskey, W. Chen, and D. Shen, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," Med. Phys. **37**, 4121–4132 (2010).

[38] F. Rousseau, P. A. Habas, and C. Studholme, "A supervised patch-based approach for human brain labeling," IEEE Trans. Med. Imaging **30**, 1852–1862 (2011).

[39] P. Coupé, J. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," NeuroImage **54**, 940–954 (2011).

[40] J. Wright, M. Yi, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," Proceedings of the IEEE **98**, 1031–1044 (2010).

[41] H. Cheng, Z. Liu, and L. Yang, "Sparsity induced similarity measure for label propagation," in *IEEE 12th International Conference on Computer Vision (ICCV)* (IEEE, Kyoto, 2009), pp. 317–324.

[42] J. Wang, J. Yang, K. Yu, F. Lv, T. S. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, San Francisco, CA, 2010), pp. 3360–3367.

[43] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation.," IEEE Trans. Pattern Anal. Mach. Intell. **31**, 210–227 (2009).

[44] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Miami, FL, 2009), pp. 1794–1801.

[45] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," J. R. Stat. Soc., Ser. B **67**, 301–320 (2005).

[46] R. J. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Stat. Soc., Ser. B **58**, 267–288 (1996).

[47] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," NeuroImage **23**(Suppl. 1), S151–S160 (2004).

[48] C. M. Li, C. Y. Kao, J. C. Gore, and Z. H. Ding, "Minimization of region-scalable fitting energy for image segmentation," IEEE Trans. Image Process. **17**, 1940–1949 (2008).

[49] L. Wang, L. He, A. Mishra, and C. Li, "Active contours driven by local Gaussian distribution fitting energy," Signal Process. **89**, 2435–2447 (2009).

[50] L. Wang, F. Shi, W. Lin, J. H. Gilmore, and D. Shen, "Automatic segmentation of neonatal images using convex optimization and coupled level sets," NeuroImage **58**, 805–817 (2011).

[51] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split Bregman method: Segmentation and surface reconstruction," Journal of Scientific Computing **45**, 272–293 (2010).

[52] T. Brox and D. Cremers, "On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional," Int. J. Comput. Vis. **84**, 184–193 (2009).

[53] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," Int. J. Comput. Vis. **22**, 61–79 (1997).

[54] H. Wang, J. W. Suh, S. R. Das, J. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," IEEE Trans. Pattern Anal. Mach. Intell. **35**, 611–623 (2013).

[55] M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," IEEE Trans. Med. Imaging **29**, 1714–1729 (2010).

[56] S. F. Eskildsen, P. Coupé, V. Fonov, J. V. Manjón, K. K. Leung, N. Guizard, S. N. Wassef, L. R. Østergaard, and D. L. Collins, "BEaST: Brain extraction based on nonlocal segmentation technique," NeuroImage **59**, 2362–2373 (2012).

[57] W. Bai, W. Shi, D. O'Regan, T. Tong, H. Wang, S. Jamil-Copley, N. Peters, and D. Rueckert, "A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images," IEEE Trans. Med. Imaging **32**, 1302–1315 (2013).

[58] P. Coupé, S. F. Eskildsen, J. V. Manjón, V. Fonov, and D. L. Collins, "Simultaneous segmentation and grading of anatomical structures for patient's classification: Application to Alzheimer's disease," NeuroImage. **59**, 3736–3747 (2012).

[59] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," Ann. Stat. **32**, 407–499 (2004).