Computers in Biology and Medicine 122 (2020) 103794

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Automated integration of facial and intra-oral images of anterior teeth

Mengxun Li [a], Xiangyang Xu [b], Kumaradevan Punithakumar [c], Lawrence H. Le [c], Neelambar Kaipatur [d], Bin Shi [a,*]

[a] Department of Implantology, School and Hospital of Stomatology, Wuhan University, Wuhan, China
[b] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China
[c] Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton AB, Canada
[d] School of Dentistry, University of Alberta, Edmonton, AB, Canada

ARTICLE INFO

Keywords:
Aesthetic analysis
Digital smile design
Intra-oral image
Image integration
Tooth segmentation

ABSTRACT

Background and Objective: Digital smile design is the technique that dentists use to analyze, design, and visualize therapeutic results on a computing workstation prior to actual treatment. Despite it being a crucial step in digital smile design, the process of labeling and integrating the information in facial and intra-oral images is laborious. Therefore, this study aims to develop an automated photo integrating system to facilitate this process.

Methods: The teeth in intra-oral images were distinguished by their curvature and finely segmented using an active contour model. The facial keypoints were detected by a sophisticated facial landmark detector algorithm; these keypoints were then overlaid on the corresponding intra-oral image by extracting the contour of the teeth in the facial and intra-oral photographs. With this system, the tooth width-to-height ratios, smile line, and facial midline were automatically marked in the intra-oral image. The accuracy of the proposed segmentation algorithm was evaluated by applying it to 50 images with 274 maxillary anterior teeth.

Results: The proposed algorithm recognized 96.0% (263/274) of teeth in our selected image set. The results were then compared to those obtained by applying manual segmentation to the remaining 263 recognized teeth. With a 95% confidence interval, a Jaccard index of 0.928 ± 0.081, average distance of 0.128 ± 0.109 mm, and Hausdorff distance between the results and ground truth of 0.461 ± 0.495 mm were achieved.

Conclusions: The results of this study show that the proposed automated system can eliminate the need for dentists to employ a laborious image integration process. It also has the potential for broad applicability in the field of dentistry.

1. Introduction

The current boom in emerging technology, including that for digitalization in dentistry, has been very beneficial for dentists and patients. In the field of restorative dentistry, digital smile design (DSD) has become popular in recent years. The term DSD refers to the process in which dentists analyze, design, and visualize the predicted therapeutic result on a computing workstation prior to the actual treatment; thus, it can serve as an excellent practice and management tool. This technique not only introduces predictability and quantization to the designing process, but also facilitates communication between the clinician, technician, and patient [1].

The workflow of DSD involves intra-oral and facial image/video acquisition, aesthetic analysis, and design [2]. The aesthetic analysis and design process typically entails the use of generalized software with image processing functions such as Photoshop (Adobe Inc. Mountain View, California, USA), Keynote (Apple Inc., Cupertino, California, USA), and Powerpoint (Microsoft Corporation, Redmond, Washington, USA) [2–4]. Some existing software platforms have been specifically designed for smile analysis or equipped with DSD functionality, such as Smile Mesh and TRIOS (3Shape A/S, Copenhagen, Denmark); a review of these software platforms is available in [5]. Despite the functionalities available to clinicians in existing software packages, the process of overlapping the photographs, drawing reference lines and generating measurements is labor-intensive.

The aesthetic outcome of the maxillary anterior region is of paramount importance to the clinician and patient. This region encompasses the teeth between the right and left maxillary canines; therefore, because it is the most visible region, it is referred to as *the aesthetic zone*.

* Correspondence to: Department of Implantology, School and Hospital of Stomatology, Wuhan University, 237 Luoyu Road, Wuhan 430079, China
E-mail addresses: mengxunli@whu.edu.cn (M. Li), shibin_dentist@whu.edu.cn (B. Shi).

https://doi.org/10.1016/j.compbiomed.2020.103794
Received 5 January 2020; Received in revised form 27 April 2020; Accepted 27 April 2020
Available online 23 May 2020

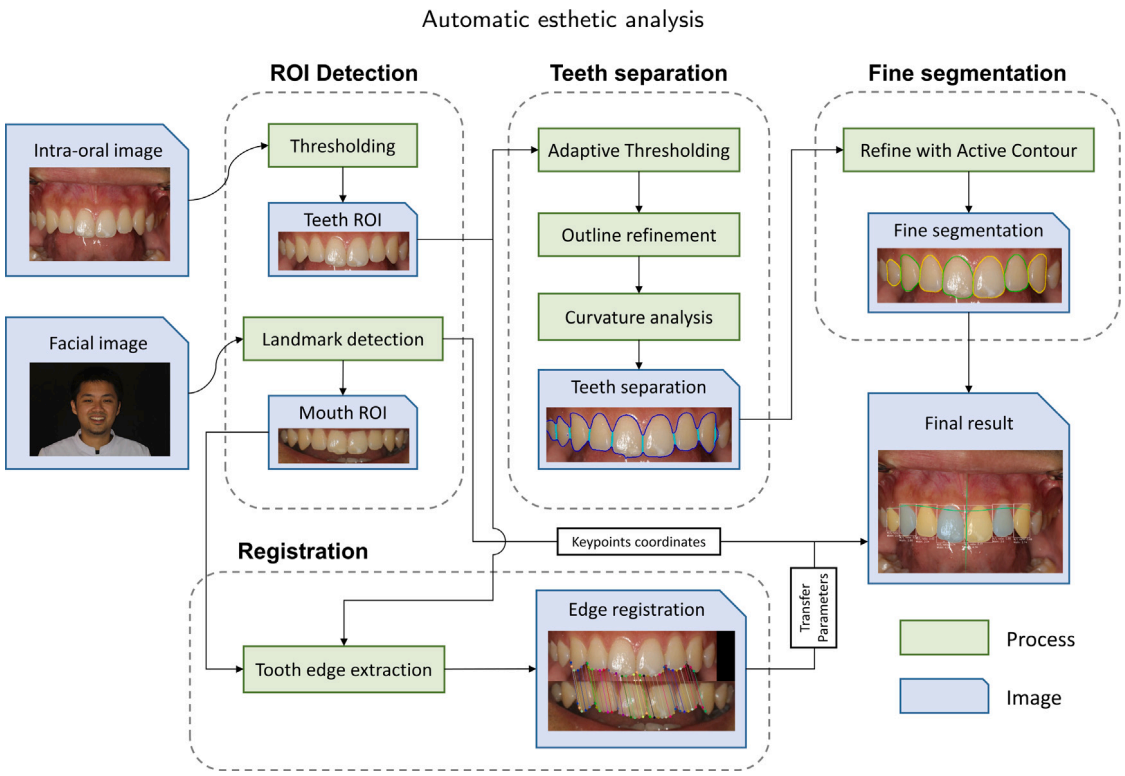


Fig. 1. The workflow of proposed method.

To achieve a better outcome, it is essential that the dentist perform a comprehensive examination and aesthetic analysis of the aesthetic zone during the initial visit. Various standards have been developed for aesthetic evaluation. Among them, tooth size and shape, the facial midline, and the smile line are some of the most important indicators of aesthetic appeal. Stefan et al. [6] analyzed smile image ranks proposed by various people; they concluded that a 75%–80% width-to-height ratio for the central incisor and a 56%–68% tooth-to-tooth proportion were generally regarded as the most attractive. Honn and Goz [7] reviewed factors that influenced people's perception of facial beauty; they found that facial symmetry and averageness play an important role in attractiveness. Altay et al. [8] illustrated that patients with a lower smile line were less satisfied with single-tooth implant restorations in the maxillary anterior region. However, what could be considered to be the standard for facial beauty has always been a debatable topic. In addition to these indicators, numerous aesthetic criteria have been proposed; a detailed review of them can be found in [9]. In general, the process of outlining the teeth in an intra-oral image and integrating the result into a facial image are required for smile analysis. Using the relevant information shown in an intra-oral image, a dentist can proceed with a very precise and detailed treatment plan to restore the aesthetics of the relevant teeth. For instance, when a patient does not have an ideal gingival contour and has a high smile line, a surgical guide can be made by utilizing the acquired information as a reference to perform precise gingival recontouring.

We have modified two tasks to improve smile analysis: (1) the transfer of facial information to intraoral photographs, and (2) the segmentation of teeth in intraoral photographs. Regarding the first task, sophisticated facial landmark detectors [10] can aid in detecting facial key-points. This facial information can then be transferred with the help of an algorithm that aligns the two images. Regarding the second task, there are relatively few reports on performing tooth segmentation using intra-oral photographs. To date, the majority of teeth segmentation studies have focused on X-ray and CBCT images. Regarding X-ray images, Jain and Chen et al. [11] distinguished individual teeth by using integral projection and subsequently segmented the teeth using intensity information. Nomir et al. [12] improved the segmentation result by applying iterative thresholding and adaptive thresholding, followed by integral projection. Shah et al. [13] achieved segmentation by applying active contour without edges, whereas Lin et al. [14] performed segmentation with singularity analysis to acquire textural information about the teeth and alveolar bone. Regarding CBCT images, Mortaheb et al. [15] defined a new feature space and used a mean shift algorithm for CT segmentation. Alternatively, Hosntalab et al. [16] distinguished individual teeth in reconstructed panoramic images and used level-set methods to find tooth boundaries. Ji et al. [17] and Gao et al. [18] segmented anterior teeth in CBCT images by applying a coupled level set method, starting from the appropriate section slice.

The process of tooth segmentation using intra-oral photographs is different from that of X-ray or CBCT images and requires development of a new approach as existing methods proposed for medical images are not practically applicable. Specifically, non-uniform illumination and specular reflection are typically present in the image. Additionally, the teeth may have caries, fillings, incomplete mineralization which can manifest as color abnormalities. Furthermore, orthodontic bracket may be bonded to the teeth, and the teeth may have defects which result in shape abnormalities. These unfavorable conditions increase the difficulty of using intra-oral photographs to perform tooth segmentation.

This study proposes an esthetic region teeth segmentation algorithm based on curvature analysis and active contour. The proposed method also entails the use of a 68-point dlib facial landmark detector [19] to detect facial landmarks and incorporate the smile line and facial midline into intra-oral image by registering the teeth edges of the two types of images. The proposed approach consists of an assistive esthetic region analyzer to automatically detect and integrate essential facial and intra-oral image information to help dentists analyze the shape and position of maxillary anterior teeth.

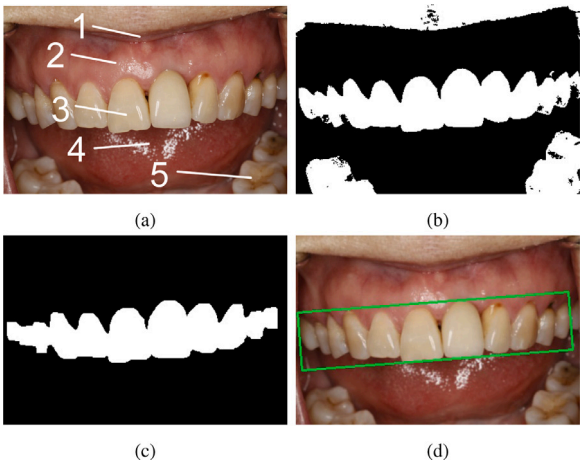


Fig. 2. Identifying the ROI: (a) original intra-oral image: 1. lip and skin, 2. gingiva and mucosa, 3. maxillary teeth, 4. tongue, and 5. mandibular teeth; (b) image after applying the threshold; (c) portion with the largest area; (d) ROI-marked image.

2. Methods

The workflow of the proposed approach is as follows: (1) identification of the region of interest (ROI) in the intra-oral and facial images, (2) tooth delineation, and (3) alignment of the two types of images (refer to Fig. 1). The delineation step can be further separated into tooth separation and shape refinement phases. The details of each workflow step are described below.

2.1. Finding the ROI in the intra-oral image

Clinicians that perform DSD are primarily interested in the maxillary anterior teeth because this area has the largest impact on the aesthetic outcome of the restoration. However, the original intra-oral photographs may contain features other than the maxillary anterior teeth because of the height-to-width ratio of the camera; thus, the photographs typically include redundant information such as lips, tongue, mucosa and mandibular teeth (refer to Fig. 2(a) for an example). The first step of the proposed method is to identify the ROI, which includes the maxillary anterior teeth. The purpose of this step is to make the model more robust and to reduce the computational cost.

Considering that the color of teeth is quite different from that of soft tissue, this step can be completed by implementing intensity-based image processing techniques. The original RGB photograph was resized to a 512-pixel height for uniformity and converted to HSV color space; then, Otsu thresholding [20] was applied to the hue channel of the HSV image (Fig. 2(b)). Assuming that the maxillary anterior teeth constitute the largest portion of the image, the largest connected area in the binary image was identified and the smallest enclosing rectangle was applied as our ROI (refer to Fig. 2(d) for an example).

2.2. Tooth separation

This section describes how individual teeth were distinguished. Often, the gradient of the adjacency of a tooth is not an ideal feature for tooth separation. Additionally, the specular highlight, small fissures, and dental plaque on the surface of teeth can also introduce strong gradients in the image. Therefore, segmentation solely based on the image gradient could be inadequate for this application. Our proposed method is based on the curvature of teeth outline, it also provides a coarse segmentation of individual teeth.

2.2.1. Determining the teeth outline

In the second step, the ROI image produced in the previous step was first resized to a 128-pixel height to obtain a better gradient and accelerate computation. In this step, simply applying global thresholding is insufficient because of non-uniform illumination and tooth color variation (refer to Fig. 3(c)).

First, adaptive thresholding was applied to the hue channel of the image. Let (i, j) be the location of a pixel in the image domain. The threshold $T(i, j)$ was determined according to the average pixel intensity values over a surrounding window of size $n \times n$ centered at (i, j) the adaptive thresholding process can be expressed as

$$T(i,j) = \sum_{(p,q)\in[-\frac{n-1}{2},\frac{n-1}{2}]} \frac{I(i+p, j+q)}{n^2}$$

$$P(i,j) = \begin{cases} 0, \text{ if } I(i,j) \geq T(i,j) \\ 1, \text{ if } I(i,j) < T(i,j) \end{cases}$$

where $I(i, j)$ is the image intensity at (i, j) and P is the resulting binary image. An empirical value of $n = 41$ was used in the process. After binarization, each connected area was filled, and areas with less than 2000 pixels were eliminated (refer to Fig. 3(e) for an example).

An active contour model was used to refine the outline where the mask produced in the previous step was used to construct the initial contour. We implemented a combined version of two classical models, *i.e.*, Caselles et al. [21] and Chan and Vese's work [22] geodesic active contour and active contour without edges, respectively; here, we present the curve evolution equation for that in the form of level set. Suppose ϕ is the level set function which is the signed distance function (SDF) for the contour, the evolution equation can be expressed as

$$\begin{aligned} \frac{\partial \phi}{\partial t} = \delta(\phi)\{ & \underbrace{\alpha_1 [\nu + \gamma div(\frac{\nabla \phi}{|\nabla \phi|})] g |\nabla \phi|}_{\text{geodesic active contour}} \\ & + \underbrace{\alpha_2 [\mu div(\frac{\nabla \phi}{|\nabla \phi|}) - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2]}_{\text{active contour without edges}} \} \end{aligned} \tag{1}$$

where δ is the Dirac delta function, α_1 and α_2 are two factors that regulate the influence of gradient and color on curve evolution, ν is mean curvature motion term, and u_0, c_1 and c_2 represent the color along the contour, and the average color inside and outside of the zero level set, respectively. In Eq. (1), $\alpha_1, \alpha_2, \nu, \gamma, \mu, \lambda_1$, and λ_2 are empirically determined parameters, and g is the gradient map, which is given by

$$g = \frac{1}{1 + |\nabla (G_\sigma * I)|^2} \tag{2}$$

where I is the hue channel of the image, and G_σ is the Gaussian kernel with standard deviation σ.

2.2.2. Separation

The third step was established in consideration of the following facts: (1) individually, teeth have a convex shape, and (2) the concave areas in the teeth outline are typically adjacency. Let $\boldsymbol{p}_i$ be a point along the contour obtained by using the active contour method. Then, the approximate curvature κ_i at point $\boldsymbol{p}_i$ can be computed using:

$$\begin{aligned} \boldsymbol{t}_i &= \frac{\boldsymbol{p}_{i+d} - \boldsymbol{p}_{i-d}}{|\boldsymbol{p}_{i+d} - \boldsymbol{p}_{i-d}|} \\ \boldsymbol{t}'_i &= \boldsymbol{t}_{i+1} - \boldsymbol{t}_i \\ \kappa_i &= \boldsymbol{t}'_i \times \boldsymbol{t}_i \end{aligned} \tag{3}$$

where d is a parameter that determines the precision of our approximation. Although a higher d value eliminates more noise, it may result in key-points being disregarded. We empirically determined that $d = 10$ to be appropriate for our problem.

After calculating the curvature (refer to Fig. 4(a)), a threshold was applied such that only points with a large concave curvature were

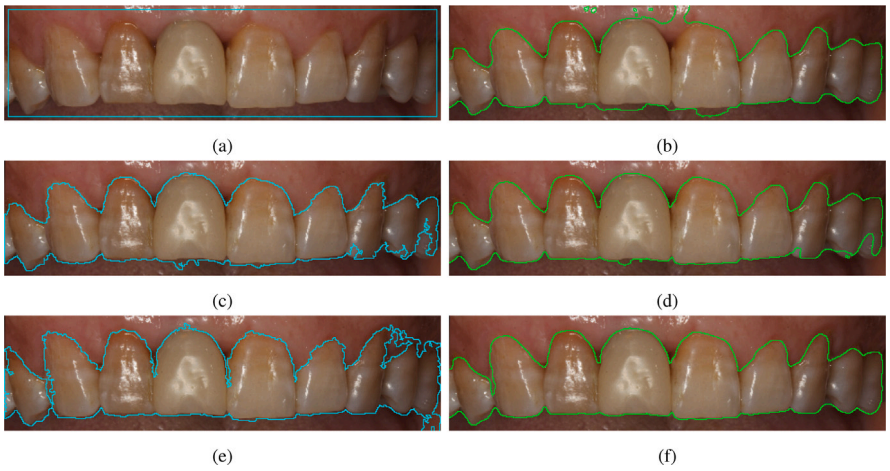


Fig. 3. Outline segmentation: (a) Contour initialization using ROI; (b) Active contour result of ROI based initialization; (c) Contour initialization following Otsu thresholding with elimination of small area; (d) Active contour result of Otsu thresholding-based initialization; (e) Contour initialization following adaptive thresholding with hole-filling and small-area elimination; (f) Active contour result of adaptive thresholding-based initialization.

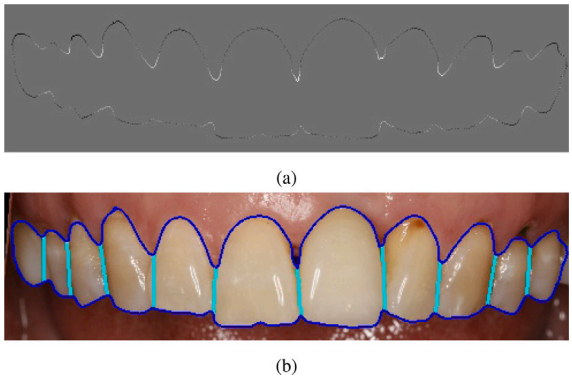


Fig. 4. Tooth separation: (a) The curvature of outline; (b) Separation result.

retained as the keypoints. The final process in this step is to connect key points to distinguish individual teeth. A moving window of size 128 × 40 pixels and a stride of 20 pixels were used to make the connections. for each window, the following criteria were applied to connect two points:

- The distance between the two points along the contour $d_{contour} > 50\,px$
- The angle between the connecting line and x axis, $\theta > 70°$
- For each window, only the line with the shortest length is retained
- When two lines (one from the previous window) are too close to each other (*i.e.*, the distance from their center is smaller than the pre-set threshold), the lines are scored based on their angle θ, length l, and average color c along the line using their weighted average:

  $$score = w_1\theta + w_2 l + w_3 c$$

  The line with the larger angle, shorter length and lower intensity is preserved.

2.3. Shape refinement

Although the previously described step roughly segments teeth, the results for adjacent areas were found to be inadequate (refer to Fig. 5(a)); therefore, the contour was further refined.

The original ROI image was resized to a height of 256 pixels for finer segmentation. a morphological erosion process was applied to the mask of the coarse segmentation to obtain an initial contour. Let ϕ be the SDF of the contour. Then, the energy functional for the active contour can be defined as follows:

$$E(\phi) = \lambda_0 E_{penal} + \lambda_1 E_{grad} + \lambda_2 E_{shape} + \lambda_3 E_{color} + \lambda_4 E_{couple} \quad (4)$$

where E_{penal} is the penalty term that stabilizes the SDF by constraining $|\nabla\phi|$ such that it is close to 1, as described by Li et al. [23]

$$E_{penal}(\phi) = \int_\Omega \frac{1}{2}(|\nabla\phi| - 1)^2 dxdy$$

E_{grad} is the edge gradient term that slows down the contour at strong gradient edges

$$E_{grad}(\phi) = \int_\Omega g\delta(\phi)|\nabla\phi| dxdy$$

where δ is the delta function and g is the same as (2) with I being the intensity value obtained from the grayscale image. It should be noted that using the gradient values from the grayscale image yielded better results for the adjacent-area refinement process than those from the hue channel image.

E_{shape} is the shape energy that minimizes the difference between the contour and the shape prior.

$$E_{shape}(\phi) = \int_\Omega \delta(\phi)(\phi_{sp}^2(x,y) + \zeta|\phi_{pp}(x,y)|)|\nabla\phi| dxdy$$

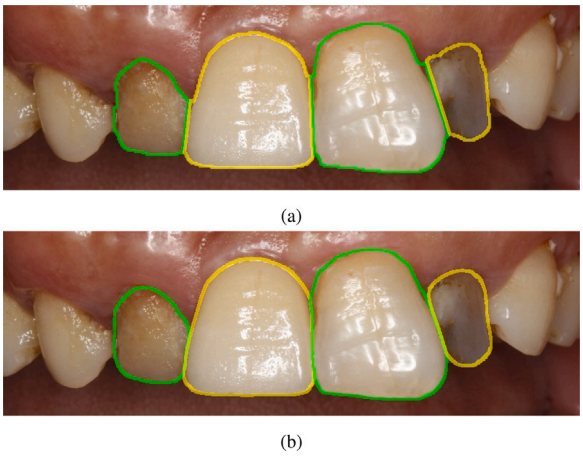
where ϕ_{sp} is the SDF of the coarse segmentation, ϕ_{pp} is the SDF of landmarks obtained form the coarse segmentation step. The term ϕ_{pp} was used to better fit the contour to the landmarks of the concave curvature. This term was derived based on the observation that these landmarks are typically accompanied by sharp edges. The absolute distance was used in ϕ_{pp} instead of the squared distance; this is because it has a strong localized impact, but less impact than the squared distance when it is applied to the contour at distances relatively far from key points; consequently, the process of evolution for these points is still determined by the full contour.

E_{color} represents the color energy that prevents the contour from including bright areas, because tooth gaps tend to be darker than other tooth surfaces in an image.

$$E_{color}(\phi) = \int_\Omega \delta(\phi) I' |\nabla\phi| dxdy$$

I' was obtained by applying adaptive histogram equalization, *i.e.*, CLAHE [24], to grayscale images to eliminate the effects of non-uniform illumination.

The functional defined here is similar to those introduced by [17] and [18] where the segmentation of tooth roots was performed slice by



(a)

(b)

Fig. 5. Shape refinement: (a) Result of coarse segmentation, which show that the contours are not ideal before refinement; (b) Post-refinement result, showing smoother and more accurate contours.

slice in CT images. The proposed method differs from the methods described by [17] and [18] because it introduces ϕ_{pp} into the shape prior term and an edge color term E_{color}. In previous approaches, a region energy term that takes into account the different intensity distributions of the foreground and background pixels was used. Alternatively, in our case, the different color distribution of tooth and gingiva does not help much in the segmentation process; because the boundary between them is typically associated with a strong gradient, It is generally difficult to distinguish neighboring teeth and the adjacency has the same intensity distribution as the teeth body.

The last term in (4) is the same as that in [18]. Tooth segmentation can be achieved by using two individual SDFs and every other tooth are represented by same function. Conversely, adjacent teeth are represented by different functions to ensure that the two different level sets do not merge. The coupled energy E_{copule} is defined as presented below to ensure that there is no overlap between two zero level sets.

$$E_{couple} = \int_{\Omega} H(-\phi)H(-\phi')dxdy$$

where H is the Heaviside function, and ϕ' is the secondary SDF.

By solving the Euler–Lagrange equation (4) and adding a mean curvature motion term, $cg\delta(\phi)$, we obtain the following evolution equation:

$$\frac{\partial\phi}{\partial t} = \lambda_0 div[(1 - \frac{1}{|\nabla\phi|})\nabla\phi] + cg\delta(\phi) + \delta(\phi)div\left[[\lambda_1 g + \lambda_2(\phi_{sp}^2 + \zeta|\phi_{pp}|) + \lambda_3 I']\frac{\nabla\phi}{|\nabla\phi|}\right] + \lambda_4 H(-\phi')$$

2.4. Image registration

The final step in our automated image integration process is to transfer facial information to the intraoral image. To achieve this, facial landmarks have to be identified and the facial image should be registered with the intra-oral image.

We used the 68-point dlib facial landmark detector proposed by Kazemi et al. [19] to detect facial landmarks. With these landmarks identified, it is easy to locate the mouth region (Fig. 6); we refer to this region as the mouth ROI, and it was registered with the teeth ROI that was acquired in the previous step.

The registration was completed by matching the tooth edges in two different images. In the strict sense, the tooth edges in two different images show different projections of the dental arch. However, in reality, they are both captured at the position directly in front of the patient. To simplify this problem, we assumed that there is only similarity transformation between two ROI images and that the registration can be regarded as rigid. Therefore, we used the method proposed in [25] to solve this problem.

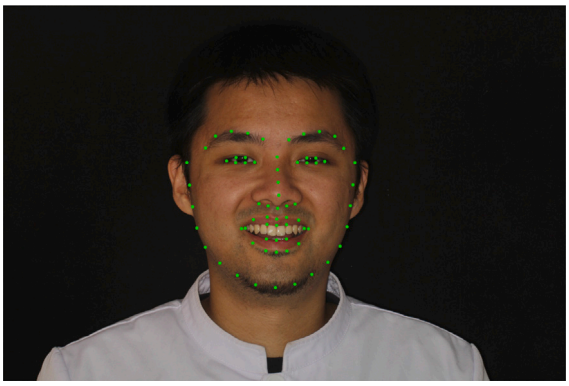


Fig. 6. Facial landmarks.

The two images were resized to have a 128-pixel height; then, the tooth edges were extracted by using techniques similar to those previously described, and cubic B-splines were used to represent the curve. The curvature along two edges was calculated by using (3). According to [25], the curvature calculated for a curve sampled at equally spaced intervals of the absolute integral of the curvature (Fig. 7) can act as a similarity invariant signature, which is independent of rotation, translation, and scaling. This signature was obtained for both curves, and the cross-correlation between them was calculated to perform the registration.

3. Experiments and results

3.1. Data acquisition

The images used to evaluate the performance of our algorithm were randomly chosen from routine images captured by resident dentists in the Department of Implantology within the School and Hospital of Stomatology at Wuhan University. All images were captured directly in front of patients, *i.e.*, perpendicular to the arch, at a distance that ensure at least four teeth between two lateral incisors could be observed in the photograph. The frontal facial photograph used to perform registration was captured positioning the camera at the nose level and approximately 2 meters from the volunteer. The volunteer was asked to give a "natural smile" to expose the teeth for the photograph.

3.2. Outcome presentation

The image results, which were derived based on tooth and face information, were automatically organized for presentation upon completion of the tooth segmentation process and the alignment of the face and intra-oral images. The system evaluated the width-to-height ratio of the teeth and incorporated the estimated smile line and facial midline into the corresponding intra-oral image (refer to Fig. 8 for an example). With the proposed method, the smile line was estimated by applying cubic interpolation to five upper-lip inferior margin points, and the facial midline was estimated by creating a linear model for the nasion, subnasale, vermillion and menton points.

3.3. Accuracy of proposed algorithm

A total of 50 intra-oral photographs focused on the maxillary anterior teeth were chosen to evaluate the proposed segmentation algorithm; all photographs were pre-inspected to confirm that there was no contact between the upper and lower jaws and that there were no

(a) (b)

Fig. 7. ROIs and teeth edges: (a) Teeth ROI; (b) Mouth ROI. The teeth edges and points sampled at equal intervals under similarity invariant parametrization are marked in the images.

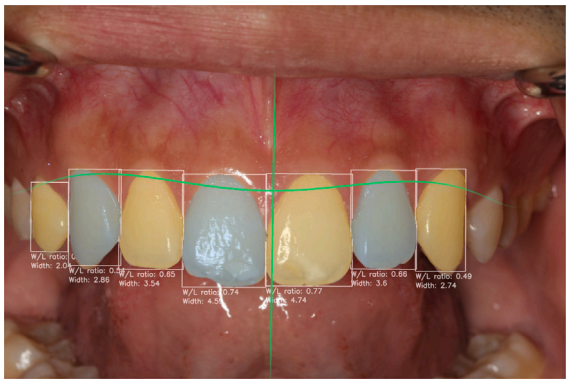


Fig. 8. Final result: the smile line and facial midline are depicted in green, and the width-to-height ratio and tooth width (100 px) are marked in the image.

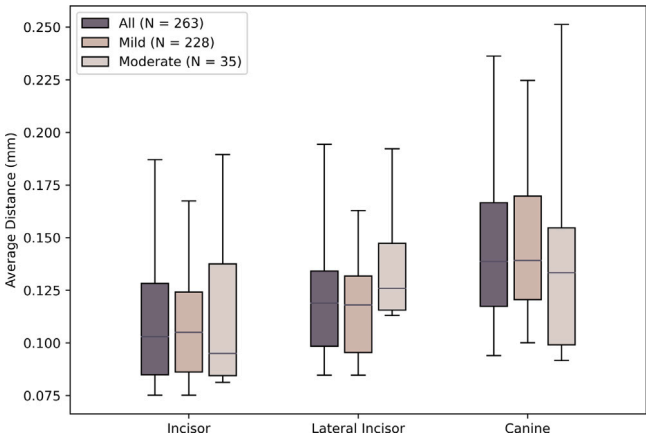


Fig. 9. AVD box-and-whisker plots: the boxes corresponding to data from the first to the third quartiles, with outliers 5 to 95% of the data.

missing teeth. The photographs were then had the teeth segmented manually to obtain the ground truth. Considering tooth crowding may affect the segmentation results, the maxillary anterior teeth in all photographs were visually classified by a licensed orthodontist into mild, moderate or severe degree of misalignment. Among the 50 images, 44, 6, and 0 images were determined to have mild, moderate, and severe misalignment, respectively.

A total of 274 teeth were included in the analysis. Teeth were regarded as unrecognized when a single tooth was incorrectly identified as two teeth or when two adjacent teeth were not distinguished. The proposed approach resulted in the correct detection of 263 of 274 teeth and achieved a tooth recognition accuracy rate of 96.0% when applied to our selected image set. The approach failed to segment 11 teeth in seven images; all of these 11 teeth, were in the category of mild crowing dentition.

Tooth-segmentation accuracy was evaluated by using the results for all 263 correctly recognized teeth to determine the Jaccard index, Hausdorff distance (HD), and Average distance (HD) [26] under 95% confidence interval (CI).

Let U_t and U_p be the segmented regions, as according to the ground truth and proposed method, respectively. The Jaccard Index is defined as follows:

$$Jaccard = \frac{U_t \cap U_p}{U_t \cup U_p}$$

Let G and S be the point sets of the edges of the ground truth and our segmentation result respectively, the HD is defined as:

$$HD(G,S) = \max(h(G,S), h(S,G))$$

where

$$h(G,S) = \max_{g \in G} \min_{s \in S} \|g - s\|$$

The AVD is defined as:

$$AVD(G,S) = \max(d(G,S), d(S,G))$$

where

$$d(G,S) = \frac{1}{N} \sum_{g \in G} \min_{s \in S} \|g - s\|$$

The minimum distance of two points comprising the edges was calculated by using the fast marching method. Because the height of teeth ROI image has already been unified at this point, the edge distance can be estimated from the pixel distance (in *mm*) by assuming the crown of the central incisor has an approximate height of 10 mm.

The results comparing the automated and manual segmentation results for the 263 teeth in terms of tooth position and degree of misalignment are presented in Table 1. The proposed method yielded an average Jaccard index of 0.928 ± 0.081 for the 263 teeth in the 50 images. Fig. 9 presents the box-and-whisker plots that show the segmentation accuracy in terms of AVD results. The automated segmentation method was more accurate for mesial teeth; however, in terms of the degree of misalignment, the results did not significantly differ from those of manual segmentation when there was no severe crowding. The automated segmentation method yielded a mean AVD of 0.128 ± 0.109 mm and mean HD of 0.461 ± 0.495 mm (Table 1) for all tooth types (CI = 95%); therefore, the performance can be regarded as clinically acceptable.

4. Discussion

In this paper, we proposed a system that automatically segments and analyzes tooth features and incorporates relevant facial image information into the corresponding intra-oral image to facilitate smile analysis. The proposed system allows dentists to show patient their current state and explain the treatment plan during their first visit. It also serves to build a foundation for the development of a fully automated smile analysis method and smile design process, thereby helping clinicians to be more time-efficient. In our study, we focused on using tooth width-to-height ratio, facial mid-line and smile line

Table 1
Quantitative evaluation of the proposed segmentation approach with respect to tooth position and degree of misalignment using the Jaccard index, HD and AVD.

	Tooth position			Degree of misalignment		Total
	Incisor	Lateral incisor	Canine	Mild	Moderate	
Count	95	96	72	228	35	263
Jaccard Index	0.950 ± 0.051	0.928 ± 0.080	0.901 ± 0.083	0.931 ± 0.066	0.911 ± 0.140	0.928 ± 0.081
HD (mm)	0.414 ± 0.477	0.455 ± 0.489	0.531 ± 0.493	0.444 ± 0.385	0.573 ± 0.903	0.461 ± 0.495
AVD (mm)	0.116 ± 0.111	0.125 ± 0.102	0.149 ± 0.102	0.125 ± 0.081	0.152 ± 0.208	0.128 ± 0.109

to demonstrate the use of this assistive system. With this framework, more functions, such as tooth color, tooth-to-tooth ratio, and extent of gingiva exposure [9] can be easily implemented to analyze various aesthetic attributes. Beyond DSD aesthetic analysis, the process of tooth segmentation and facial information transfer have the potential to be used for various applications that could be improved by the identification of individual teeth and transfer of facial information, such as digital mock-up, vision-based articulator, and disease recognition.

As compared to X-ray- or CBCT-imaged teeth, it may appear that the teeth in intra-oral images are easier to segment because color is shown. However, intra-oral imaging have its own difficulties, as the images are more likely to be inconsistent because of changes in shooting conditions. Numerous factors can prevent the algorithm from yielding a sufficiently accurate segmentation result, among which non-uniform illumination, specular reflection, and abnormal tooth color/shape are the most significant. Wu et al. [27] used a boosted edge learning approach to detect the edges of teeth photographs and obtained a very good result. However, the training process for the tooth edge detector requires a large amount of data, and they did not report the accuracy of the algorithm. In this study, we developed an adaptive thresholding based active contour method to overcome the problem of non-uniform illumination (Fig. 3(f)). Individual teeth were distinguished by identifying critical points along the outline that have a concave curvature. Additionally, the proposed method does not require a training process, and it is not affected by a weak gradient between individual teeth and the specular highlight. We evaluated the impact of different initialization techniques on the outline segmentation (Fig. 3). Based on our visual experiment, the outline segmentation result that was obtained by applying the adaptive thresholding-based contour initialization was consistently better than that obtained by using global thresholding or ROI-based contour initialization (Fig. 3(b), 3(d)). Although, in many cases these techniques may yield similar results, our method was better equipped to handle non-uniform illumination than the global thresholding-based technique; furthermore, as compared to the ROI-based contour technique, because the product of initialization is closer to the desired result, less time is required for convergence to the final solution.

In the maxillary anterior teeth region, 263 of 274 teeth were correctly detected. Among the correctly detected teeth, the proposed segmentation approach yielded an average Jaccard index and average error of 0.928 ± 0.081 and 0.128 ± 0.109 mm, respectively. The actual distances in the results were derived from the pixel distance in consideration of the fact that the clinical crown height of an central incisor for a typical Asian person has been reported to be approximately 10 mm [28]. Because all the intraoral photographs were taken directly in front of the patient perpendicular to the tooth arch, the estimation was minimally affected by the projection angle and can be regarded as acceptable. The performance of the algorithm was found to be worse for distal teeth (Table 1 and Fig. 9), where the illumination conditions tend to be worse than those for mesial teeth. Additionally, the shape of the canine and the overlapping of teeth may create sharp edges that negatively impact the performance of the segmentation algorithm. In our selected dataset the algorithm did not result in significantly different levels of performance for images with a mild or moderate degree of misalignment. However, it should be noted that this finding does not mean that the proposed approach will work well on severely misaligned teeth, as we did not include cases of severe crowding in this image set and the number of moderate crowding cases was small.

It is also worth mentioning that all the photographs used for evaluation were captured in our department under standard clinical conditions and with relaxed exclusion criteria. Thus, the test images were very diverse and a number of patients had periodontitis or defected tooth that affect the shape of their teeth and gingiva. Regardless, the overall error or the maxillary anterior region was below 0.5 mm, which is sufficient for clinical use. Note that we only confirmed the accuracy for these teeth because, as previously mentioned, this is typically the only region that requires aesthetic analysis in restorative dentistry. Despite this, theoretically, the same segmentation algorithm can be applied to posterior teeth, as long as the outline of those teeth are a wave form.

In recent years, many sophisticated facial detectors have been developed to recognize facial landmarks [10]. Owing to the rapid development of these techniques, the nontrivial problem of identifying facial key points has become quite simple. However, existing techniques have yet to be able to correctly align facial images with intra-oral images. In general, there are two main approaches to performing image registration, *i.e.*, feature based and region based approaches [29]. In our case, a region based approach seemed to be less feasible, because the two types of images were captured under different conditions. For example, if a patient has a low smile line, the amount of tooth exposure in the facial image would be much smaller; this could result in unstable registration. Regarding feature-based alignment, global feature detectors such as SIFT [30] or ORB [31] are not ideal because of the relatively few distinctive features available in the teeth and gingiva. Distinctive features could theoretically be detected from highlight patches or in places where color changes drastically, such as the gingival papilla. However, highlight patches is different in two photographs and the gingival papilla tends to manifest as a repetitive pattern which may result in image mismatch. The registration of the two types of images was achieved by using a similarity invariant signature [25] to register tooth edges. Similar approaches that use curve registration have previously been applied in the field of dentistry; for example, Destrez et al. [32] registered singular points along the tooth neck line between intraoral images and scanned plaster models, Wang and Suenaga et al. [33,34] applied an epipolar constraint search to register tooth edges and align images of teeth that was captured by stereo cameras. In our case, the idea of curve registration can be feasibly applied. For simplification, the transformation model was assumed to be a similarity transformation model; this assumption is in agreement with the existing DSD procedure that entails manual overlap of the dental arch in two types of photographs.

There are some limitations of our work: (1) All photographs have to be captured under the condition of an open bite, as no occlusion between the maxillary and mandibular teeth in the image is required for successful ROI recognition. This condition is typically achieved when a patient is laughing, and we plan to adapt the proposed approach to a gentle smile in the future; (2) Although the proposed segmentation algorithm is not negatively influenced by mild unfavorable conditions such as color abnormality (Fig. 5(b)) or shape defect (Fig. 4(b)), it does not perform well under certain circumstances. For instance, the proposed model assumes that there are no missing teeth or large gaps between teeth and all neighboring teeth has to be in direct contact in

the image to allow them to be distinguished by the curvature of their outline; (3) The segmentation algorithm relies on several parameters, which makes it difficult to determine the set of parameter values that will optimize performance; (4) The sample size used to evaluate the performance of the algorithm was small, and the samples did not include severely misaligned teeth.

5. Conclusions

In this study, we developed a system that automatically integrates information in facial and intra-oral images. The teeth in the intra-oral image were segmented according to the curvature of the teeth outline and subsequently refined by applying an active contour model. The facial image was registered with the corresponding intra-oral image to ensure that the relevant facial information was automatically incorporated into the intra-oral image. This preliminary work toward fully automated smile analysis is expected to be widely applicable in the field of dentistry. Future work will focus on enhancing the robustness of this system, testing and optimizing for severely misaligned teeth, and adding a design phase component to the system.

CRediT authorship contribution statement

Mengxun Li: Software, Methodology, Visualization, Writing - original draft. **Xiangyang Xu:** Methodology. **Kumaradevan Punithakumar:** Writing - review & editing. **Lawrence H. Le:** Writing - review & editing. **Neelambar Kaipatur:** Validation, Writing - review & editing. **Bin Shi:** Supervision, Project administration, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Gabriele Cervino, Luca Fiorillo, Alina Arzukanyan, Gianrico Spagnuolo, Marco Cicciù, Dental restorative digital workflow: Digital smile design from aesthetic to function, Dent. J. 7 (2) (2019) 30, http://dx.doi.org/10.3390/dj7020030.

[2] W.S. Lin, A. Zandinejad, M.J. Metz, B.T. Harris, D. Morton, Predictable restorative work flow for computer-aided design/computer-aided manufacture–fabricated ceramic veneers utilizing a virtual smile design principle, Oper. Dent. 40 (4) (2015) 357–363, http://dx.doi.org/10.2341/13-295-S.

[3] Edward A. McLaren, David A. Garber, Johan Figueira, The photoshop smile design technique (part 1): Digital dental photography, Compend. Cont. Educ. Dent. 34 (10) (2013) 772–774.

[4] Christian Coachman, Eric Van Dooren, Galip Gürel, Cobi J. Landsberg, Marcelo A. Calamita, Mitzan Bichacho, Smile design: From digital treatment planning to clinical reality, Interdiscip. Treat. Plan. 2 (2012) 119–174.

[5] Doya Omar, Carolina Duarte, The application of parameters for comprehensive smile esthetics by digital smile design programs: A review of literature, Saudi Dent. J. 30 (1) (2018) 7–12, http://dx.doi.org/10.1016/j.sdentj.2017.09.001.

[6] Stefan Wolfart, Helen Thormann, Sandra Freitag, Matthias Kern, Assessment of dental appearance following changes in incisor proportions, Eur. J. Oral Sci. 113 (2) (2005) 159–165, http://dx.doi.org/10.1111/j.1600-0722.2005.00206.x.

[7] Mirjam Hönn, Gernot Göz, The ideal of facial beauty: A review, J. Orofac. Orthop. 68 (1) (2007) 6–16, http://dx.doi.org/10.1007/s00056-007-0604-6.

[8] Mehmet Ali Altay, Alper Sindel, Hüseyin Alican Tezerişener, Nelli Yıldırımyan, Mehmet Mustafa Özarslan, Esthetic evaluation of implant-supported single crowns: A comparison of objective and patient-reported outcomes, Int. J. Implant Dent. (2019) 1–8, http://dx.doi.org/10.1186/s40729-018-0153-3.

[9] Cornelia Frese, Hans Joerg Staehle, Diana Wolff, The assessment of dentofacial esthetics in restorative dentistry: A review of the literature, J. Am. Dent. Assoc. 143 (5) (2012) 461–466, http://dx.doi.org/10.14219/jada.archive.2012.0205.

[10] Tadas Baltrusaitis, Peter Robinson, Louis-Philippe Morency, OpenFace: An open source facial behavior analysis toolkit, in: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, ISBN: 978-1-5090-0641-0, 2016, pp. 1–10, http://dx.doi.org/10.1109/WACV.2016.7477553.

[11] Anil K. Jain, Hong Chen, Matching of dental X-ray images for human identification, Pattern Recognit. 37 (7) (2004) 1519–1532, http://dx.doi.org/10.1016/j.patcog.2003.12.016.

[12] Omaima Nomir, Mohamed Abdel-Mottaleb, A system for human identification from X-ray dental radiographs, Pattern Recognit. 38 (8) (2005) 1295–1305, http://dx.doi.org/10.1016/j.patcog.2004.12.010.

[13] Samir Shah, Ayman Abaza, Arun Ross, Hany Ammar, Automatic tooth segmentation using active contour without edges, in: 2006 Biometrics Symposium: Special Session on Research At the Biometric Consortium Conference, IEEE, ISBN: 978-1-4244-0486-5, 2006, pp. 1–6, http://dx.doi.org/10.1109/BCC.2006.4341636.

[14] P.L. Lin, P.Y. Huang, P.W. Huang, H.C. Hsu, C.C. Chen, Teeth segmentation of dental periapical radiographs based on local singularity analysis, Comput. Methods Programs Biomed. 113 (2) (2014) 433–445, http://dx.doi.org/10.1016/j.cmpb.2013.10.015.

[15] Parinaz Mortaheb, Mehdi Rezaeian, Hamid Soltanian-Zadeh, Automatic dental CT image segmentation using mean shift algorithm, in: 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), IEEE, ISBN: 978-1-4673-6184-2, 2013, pp. 121–126, http://dx.doi.org/10.1109/IranianMVIP.2013.6779962.

[16] Mohammad Hosntalab, Reza Aghaeizadeh Zoroofi, Ali Abbaspour Tehrani-Fard, Gholamreza Shirani, Segmentation of teeth in CT volumetric dataset by panoramic projection and variational level set, Int. J. CARS. 3 (3–4) (2008) 257–265, http://dx.doi.org/10.1007/s11548-008-0230-9.

[17] Dong Xu Ji, Sim Heng Ong, Kelvin Weng Chiong Foong, A level-set based approach for anterior teeth segmentation in cone beam computed tomography images, Comput. Biol. Med. 50 (C) (2014) 116–128, http://dx.doi.org/10.1016/j.compbiomed.2014.04.006.

[18] Hui Gao, Oksam Chae, Individual tooth segmentation from CT images using level set method with shape and intensity prior, Pattern Recognit. 43 (7) (2010) 2406–2417, http://dx.doi.org/10.1016/j.patcog.2010.01.010.

[19] Vahid Kazemi, Josephine Sullivan, One millisecond face alignment with an ensemble of regression trees, Comput. Vision Pattern Recogn. (2014) 1867–1874, http://dx.doi.org/10.13140/2.1.1212.2243.

[20] Nobuyuki Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66, http://dx.doi.org/10.1109/TSMC.1979.4310076.

[21] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, in: ICCV-95, IEEE Comput. Soc. Press, ISBN: 0-8186-7042-8, 1997, pp. 694–699, http://dx.doi.org/10.1109/ICCV.1995.466871.

[22] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277, http://dx.doi.org/10.1109/83.902291.

[23] Chunming Li, Chenyang Xu, Changfeng Gui, M.D. Fox, Level set evolution without re-initialization: A new variational formulation, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE, ISBN: 0-7695-2372-2, 2005, pp. 430–436, http://dx.doi.org/10.1109/CVPR.2005.213.

[24] Stephen M. Pizer, E. Philip Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, Karel Zuiderveld, Adaptive histogram equalization and its variations, Comput. Vis. Graph. Image Process. 39 (3) (1987) 355–368, http://dx.doi.org/10.1016/S0734-189X(87)80186-X.

[25] Ming Cui, John Femiani, Jiuxiang Hu, Peter Wonka, Anshuman Razdan, Curve matching for open 2D curves, Pattern Recognit. Lett. 30 (1) (2009) 1–10, http://dx.doi.org/10.1016/j.patrec.2008.08.013.

[26] Abdel Aziz Taha, Allan Hanbury, Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool, BMC Med. Imag. 15 (1) (2015) 29, http://dx.doi.org/10.1186/s12880-015-0068-x.

[27] Chenglei Wu, Derek Bradley, Pablo Garrido, Michael Zollhöfer, Christian Theobalt, Markus Gross, Thabo Beeler, Model-based teeth reconstruction, ACM Trans. Graph. 35 (6) (2016) 1–13, http://dx.doi.org/10.1145/2980179.2980233.

[28] Jae-Won Song, Richard Leesungbok, Su-Jung Park, Se Hun Chang, Su-Jin Ahn, Suk-Won Lee, Analysis of crown size and morphology, and gingival shape in the maxillary anterior dentition in Korean young adults, J. Adv. Prosthodont. 9 (4) (2017) 315–320.

[29] Barbara Zitová, Jan Flusser, Image registration methods: A survey, Image Vis. Comput. 21 (11) (2003) 977–1000, http://dx.doi.org/10.1016/S0262-8856(03)00137-9.

[30] David G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110, http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94.

[31] Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski, ORB: An efficient alternative to SIFT or SURF, in: 2011 IEEE International Conference on Computer Vision (ICCV), IEEE, ISBN: 978-1-4577-1101-5, 2011, pp. 2564–2571, http://dx.doi.org/10.1109/ICCV.2011.6126544, http://ieeexplore.ieee.org/document/6126544/.

[32] Raphael Destrez, Sylvie Treuillet, Yves Lucas, Benjamin Albouy-Kissi, Semi-automatic registration of 3D orthodontics models from photographs, in: Sebastien Ourselin, David R. Haynor (Eds.), SPIE Medical Imaging, SPIE, 2013, p. 86691E, http://dx.doi.org/10.1117/12.2006900.

[33] Junchen Wang, Hideyuki Suenaga, Kazuto Hoshi, Liangjing Yang, Etsuko Kobayashi, Ichiro Sakuma, Hongen Liao, Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery, IEEE Trans. Biomed. Eng. 61 (4) (2014) 1295–1304, http://dx.doi.org/10.1109/TBME.2014.2301191.

[34] Hideyuki Suenaga, Huy Hoang Tran, Hongen Liao, Ken Masamune, Takeyoshi Dohi, Kazuto Hoshi, Tsuyoshi Takato, Vision-based markerless registration using stereo vision and an augmented reality surgical navigation system: A pilot study, BMC Med. Imag. (2015) 1–11, http://dx.doi.org/10.1186/s12880-015-0089-5.

Mr. Mengxun Li got his Bachelors of Dentistry degree from Wuhan University in 2018 and currently a dentistry doctoral candidate in Wuhan University, specializing in implantology. He has been working at the University of Alberta as a research intern in 2016. His research interest includes imaging processing systems and artificial intelligence in dental area and evidence-based research in dental implantology.

Dr. Xiangyang Xu received his B.S. degree, M.S. degree and Ph.D. degree in computer science from Huazhong University of Science and Technology (HUST), in 1988, 1991 and 2010 respectively. From 1991 to 2019, he worked at the School of Computer Science and Technology, HUST. He is currently an associate professor in HUST. His research interests lie in the area of image processing and analysis, pattern recognition and machine learning.

Dr. Kumaradevan Punithakumar received his B.Sc.Eng. (with First class Hons.) degree in Electronic and Telecommunication Engineering from the University of Moratuwa, and his M.A.Sc and Ph.D. degrees in Electrical and Computer Engineering from McMaster University. From 2001 to 2002 he was an Instructor in the Department of Electronic and Telecommunication Engineering at the University of Moratuwa. From 2008 to 2012, he was an Imaging Research Scientist at GE Healthcare, Canada. He is currently an Associate Professor (Research) at the University of Alberta, Department of Radiology and Diagnostic Imaging, and an Operational and Computational Director at the Servier Virtual Cardiac Centre, Mazankowski Alberta Heart Institute. Dr. Punithakumar was the recipient of the Industrial Research and Development Fellowship by the National Sciences and Engineering Research Council of Canada in 2008, and the GE Innovation award in 2009. Areas of interest include medical image analysis and visualization, machine learning, information fusion, object tracking and nonlinear filtering.

Dr. Lawrence H. Le received his Ph.D. in earth physics from the University of Alberta, Edmonton, Canada, in 1991 under the mentorship of late Professor Ernest R. Kanasewich. He held a NSERC postdoctoral position in Schlumberger-Doll Research Lab, Ridgefield, Connecticut under Professor Robert (Bob) Burridge. He started his medical physics residency in the Department of Radiology and Diagnostic Imaging (DRDI) in 1994, completed a MBA degree in finance and technology commercialization at the University of Alberta in 1999, and joined the Department of Radiology and Diagnostic Imaging at the University of Alberta and Capital Health in 2000. He leads the radiation safety and quality assurance X- ray program within Alberta Health Services and also the graduate program in DRDI. He directs the Ultrasonic Bone Tissue Characterization and Imaging group and guides his students to use vigorous geophysical and signal processing principles to study long bones and recover their properties. His research interests are in ultrasound imaging, signal and imaging processing, simulation, inversion, and artificial intelligence. Lawrence is a member of AAPM (The American Association of Physicists in Medicine) and COMP (The Canadian Organization of Medical Physicists).

Dr. Neelambar Kaipatur is a clinician scientist and has completed his Dental degree (2006) from McGill University and a Ph.D. in Orthodontics (2015) from University of Alberta. He has extensive experience in bench side and translational research and as an assistant clinical professor at University of Alberta since 2015, focuses on research projects that have direct impact on clinical care. He has more than 10 peer-reviewed manuscripts published in high impact journals. He is involved in the development of innovative products and techniques that would be ready for testing in clinical setting. As an end-knowledge user, he provides immediate feedback to the research team on ways to improve and deliver a quality product or technique that is both user-friendly and provide better quality of care.

Dr. Bin Shi graduated from Hubei Medical College in 1984 and received Ph.D. degree at the School & Hospital of Stomatology, Wuhan University in 2005. Now he is professor, supervisor of Oral Implantology, director of Oral Implantology Department, School & Hospital of Stomatology, Wuhan University. He is also Vice Chairman of Society of Implantology, Chinese Stomatological Association, Chairman of Society of Implantology, Hubei Stomatological Association. His research interests include evidence based dental implantology, the development and application of new dental materials, pathogenesis and treatment of peri-implantitis and digitalization in prosthodontic and dental implant area.