COMPUTER-AIDED DESIGN & APPLICATIONS
2018, VOL. 15, NO. 6, 863–871
https://doi.org/10.1080/16864360.2018.1486962

Alignment of dental depth images from an intraoral scanner

Min S. Ko and Sang C. Park

Ajou University of Republic of Korea, Korea

ABSTRACT

This paper proposes an automated alignment procedure of dental depth images captured by intraoral scanners like the Microsoft Kinect. The proposed procedure identifies an acquired order from initial scanning data and attempts to compute the registration sequence for aligning all dental depth images into a completed model. A core part of the proposed procedure is an algorithm that computes the similarity of each dental depth images, considering only the adjacent dental depth images. The algorithm was designed carefully by considering two major technological requirements of the problem: accuracy and efficiency. To satisfy the accuracy requirement, the proposed algorithm uses the concept of feature vector based on fast point feature histogram (FPFH) to compute the similarity index of each dental depth images. In this phase, the computation of nonadjacent dental depth image was excluded by the axis aligned bounding box (AABB) detection, satisfying the required efficiency. The data used in experiment is consisted of 1,021 dental depth images of mandibular teeth.

KEYWORDS

3D scanning; Intraoral scanner; 3D alignment; registration

1. Introduction

Owing to the increased demand for biocompatibility and esthetically pleasing dental restorations, all-ceramic restorations such as glass-boned porcelain and phase-stabilized zirconia have been introduced in restorative dentistry [6], [21]. The development of all-ceramic restoration has allowed wider expanded application in dental CAD/CAM. It is because of the manufacturing of all-ceramic restoration requiring a 3D dental model for machining operations. Today, various digitizing systems have been proposed to obtain 3D dental models. More recently, there are two methods to digitize a 3D model. First one is performing directly in the patient's mouth (intraoral) and second one is performing indirectly after taking an impression and fabricating master cast (extraoral) [7], [14].

A 3D model captured by extraoral optical scanner have been used widely in dental diagnosis, treatment planning and fabrication of restorations. An extraoral optical scanner based on machine vision cameras, called structured light systems (SLSs) is popular for using in dental CAD/CAM due to their fast measuring speed, simple optical arrangement, non-contact arrangement, moderate accuracy, and low cost. But prior to the extraoral digitization, an impression of the clinical situation should be taken [19]. The impression technique as well as the properties of the impression material could affect the fitting precision of fixed restorations as well. The fabrication of dental plaster model is time-consuming and error-prone method that requires the presence of collaboration with a dental laboratory. Moreover, shadowing effect and virtual hole problem limit the use of the extraoral optical digitization method [5].

A 3D dental model obtained from intraoral scanner can be used as an alternative to conventional plaster casts. Intraoral digitization allows the dental-care provider to obtain directly the data from the patient's teeth (Fig. 1). This method creates 3D models in real-time by registration of each depth image [7]. Thus, it is not necessary to take an impression and fabricating plaster model. Further, this method shows real-time intermediate results of 3D Model during measurement. Also, it can be stored easily, and transmitted digitally [12]. An intraoral scanning technology that provides ease of measurement has more attractive than conventional dental impression system. However, the most critical issue in dental scanning system is accuracy of the final dental restoration [16]. Unfortunately, previous approaches in intraoral dentistry for an accurate 3D dental model is not accurate enough. The accuracy of 3D dental model is a major factor for dental restorations, which has an influence on the fit of fixed restorations. Because the final product of the manufacturing procedure is intended to replace a person's missing tooth structure, the required accuracy level

CONTACT Sang C. Park Scpark@ajou.ac.kr; Min S. Ko Killay@ajou.ac.kr

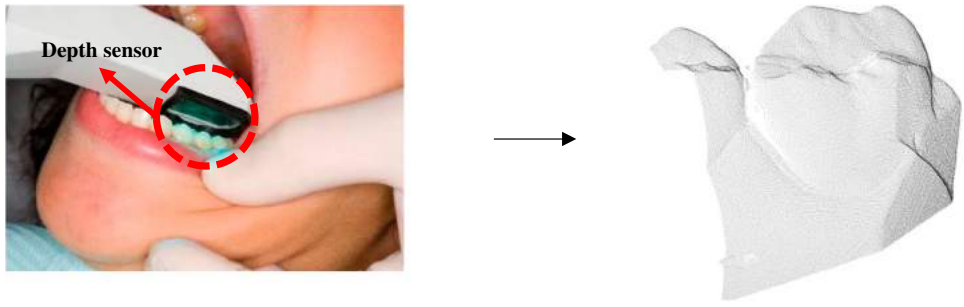

Figure 1. Dental depth image captured by intraoral scanner.

(50μm) should be ensured throughout all steps of manufacturing [17]. Thus, recent studies in intraoral dentistry fields are consist of various digitization techniques for proper accuracy. Among them, the most influential digitization factor affecting accuracy is registration: the process of finding a spatial transformation that aligns multiple dental depth image.

One of the probably most well-known examples for aligning point cloud data is Kinectfusion [9]. Kinectfusion uses the Microsoft Kinect device for real-time 3D reconstruction. Kinectfusion uses a Z-buffer ICP algorithm for fast registration, but it is relatively not accurate [1]. This method focuses on reconstruction speed, cannot provide the accuracy level which is required in dentistry (Tab. 1). Furthermore, it may cause accumulative errors which are fatal to a final dental restoration. Besl and Mckay propose an ICP (*Iterative closest point*) algorithm for registration which requires the specification of an appropriate procedure to find the closest point on a geometric entity to a given point [2]. But single ICP method in intraoral scanning may cause computation time loss. Because each dental depth image is consisted of many points, and dental image shapes are consisted of complex curves. Kim and Chang propose a simultaneous registration of multiple depth images using Global hessian matrix to appropriate each depth images [4]. It works relatively stable, and lead to superior results in previous tests, but it is relatively slow to use in dentistry. T. Weber et al. proposes an automatic registration of unordered point clouds using feature characteristic [22]. It works relatively fast, and lead to superior results in previous tests. But this method, which considers a shape of all dental depth images, may not be suitable for patient with a similar tooth shape in the mouth.

Previous studies not suitable for aligning a dental depth images, because they do not consider the characteristics of intraoral scanning. Before presenting the proposed procedure in detail, it is necessary to regard some characteristics of intraoral scanning. An intraoral scanning method uses a hand-held scanner for real-time 3D reconstruction and capture a dental depth images from a part of patient's tooth. Each dental depth images are captured by swing a handheld scanner to left and right sides of the tooth (Fig. 2). A registration method of intraoral scanning uses sequential processing because it focuses on real-time reconstruction which is required fast computation speed [12]. Each depth images are aligned with other depth image which is their previous sequence. However, this method does not help in terms of accuracy. Because of the complicated shape of the teeth, there is no guarantee that the first and second depth images are good correspondence.

Table 1. Requirements of intraoral scanning specification.

Vertex size of Complete 3D Model(Total)	Up to 90,000,000
Accuracy level of Complete 3D Model	Less than 50μm
Frame per second(fps) during intraoral scanning	Up to 15 fps

The proposed method in this paper is based on the above considerations. It requires that each depth image of the total set shares some overlap with at least one other depth image. The input element of proposed method is a roughly aligned dental depth images from an intraoral scanning (i.e. Kinectfusion method) method and the output element is dental depth images which is precisely aligned.

The objective of this paper is implement an automated registration method for achieving an accurate 3D dental model that is satisfied requirement level of dentistry. Tab. 1 shows in detail [16]. The overall structure of the paper is as follows. Section 2 addresses the proposed approach to automated alignment method as well as the overall framework of the proposed procedure. Section 3 describes the details of the core algorithms of the proposed procedure. Illustrative examples are provided in Section 4. Finally concluding remarks are given in Section 5.

2. Approach to automated alignment

Fig. 4 shows an overview of the proposed procedure. The first step involves roughly aligned depth images captured by handheld intraoral scanners (the order of acquisition

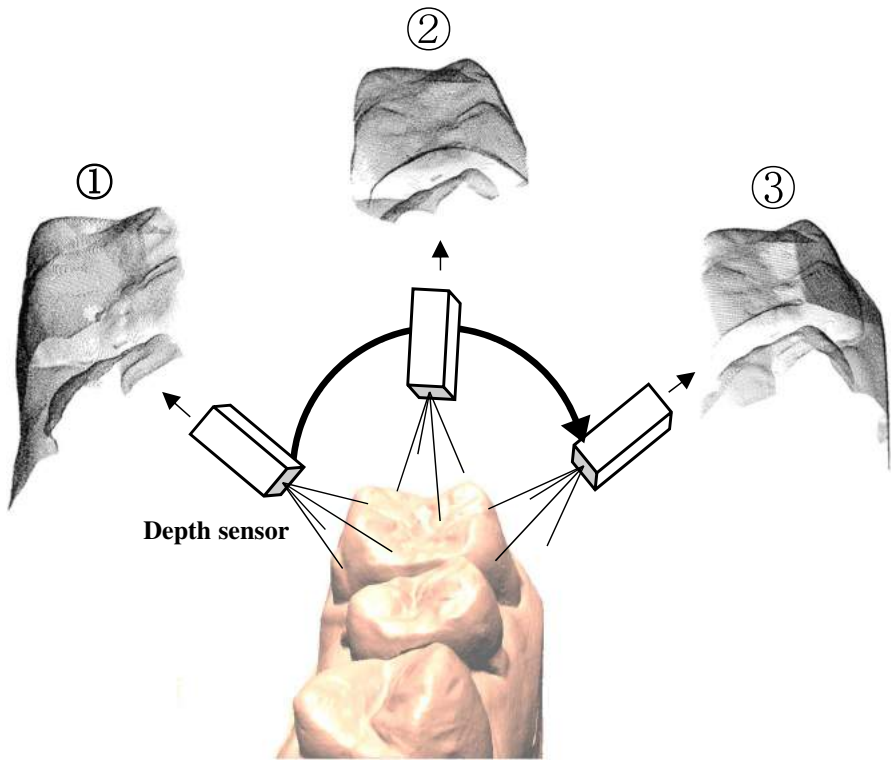

Figure 2. Dental range images acquired from different directions.

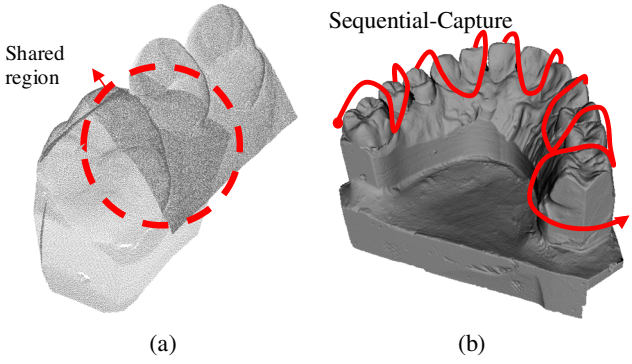

Figure 3. Requirements of proposed method, from left to right; (a) Each depth image has shared some overlapping region, (b) Each depth image is sequentially captured.

is shown in Fig. 3(b)). Once all depth images are obtained, each dental depth images are passed through 3 major steps. During preparation phase (Section 3.1), unnecessary or disturbing points are deleted and fast point feature histogram (FPFH) descriptors are computed [19]. In the first of second step, nonadjacent depth images between selected depth images are rejected in consider list by axis aligned bounding box (AABB) detection. (Section 3.2). Then a Robust heuristic is used to select the dental depth image pairs, which are used within the pairwise initial align (Section 3.2). In Registration step, a rigid transformation of each depth images is computed. (Section 3.3) All computed transformations are used to align the original dental depth images into a single view. The output is aligned dental depth images that contain all points of the original depth images in one global coordinate system.

3. Methodology

This section gives a detailed description of proposed procedure reconstruct a registration order. Each dental depth images are propagated through three major steps.

3.1. Preparation

The purpose of this section is to prepare all depth images for further analysis. First step is sampling. All parts of a Depth image are not equally useful to estimate the pose. And, the relative orientation between two depth images can be calculated from a fraction of all points. When the more points are used, the accuracy could be increased, but the computational load is increased much more [8], [13]. Effective sampling, which minimizes the points while retaining the specific shape, becomes a stepping stone for good registration. These problems have traditionally been addressed in the CAD/CAM domain. In this paper, a uniform voxel grid is used for the Point

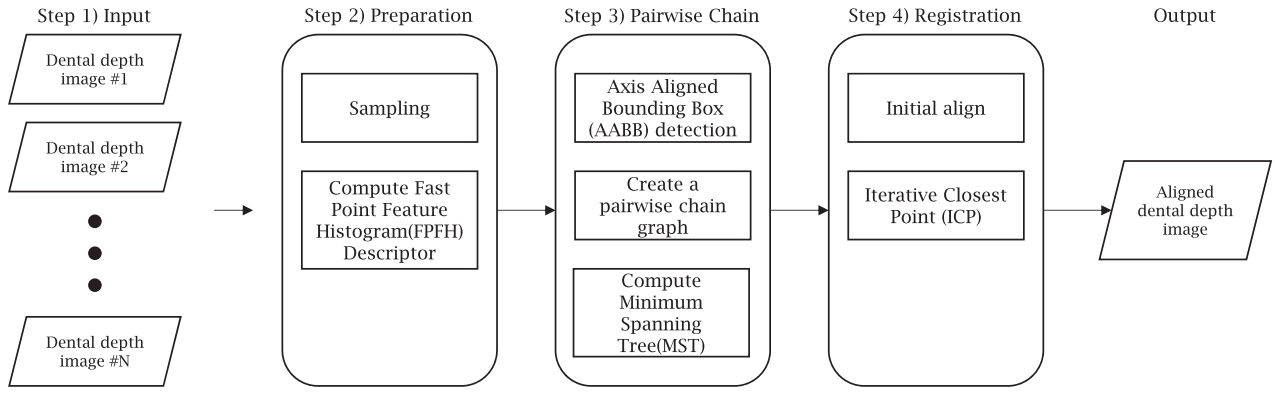

Figure 4. Overview of the proposed procedure.

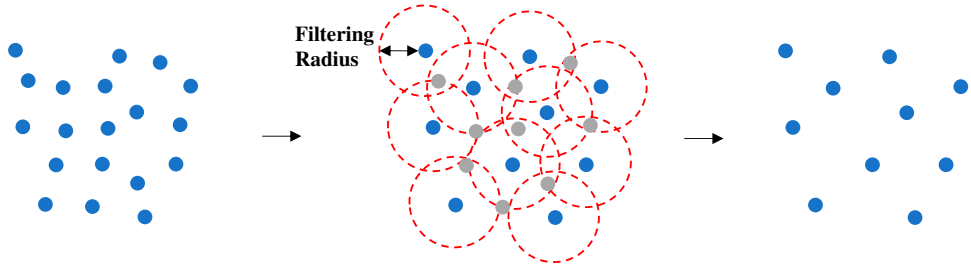

Figure 5. Radius sampling.

sampling. The sampling grid is laid over the point cloud and partitions all points into different cells [15].

Second step is distance sampling. From the experience of the authors, captured data beyond 220 mm become noisier (e.g. tongue), which disturb the further alignment process. The distance sampling filter deletes points beyond a specified distance to the intraoral sensor.

The last step of in this section is compute point descriptor computation. Point descriptor characterize the local environment of a point within the one depth image. A good point descriptor should be robust against noise, fast to compute, and invariant against rotation and translation of the local environment of a point to avoid mismatches [11]. This paper uses Fast Point Feature Histograms (FPFH), because they are fast to compute, relatively stable, and lead to superior results. FPFH method uses the angles and normal of adjacent points to create Darboux frame. It used to calculate FPFH Descriptors that represent the relationship to other points within the user specified radius [19].

Darboux frame is created between a query point p_q and each gathered adjacent point. Fig. 6 shows in detail. The source point p_s of a specific Darboux frame is the point of the point pair with the smaller angle between its normal and the connecting line between the point pair. The target point p_t is the second point of the point pair. If n_s is the normal at point p_s, then the Darboux frame [u, v, w] between the point pair p_s and p_t with the origin

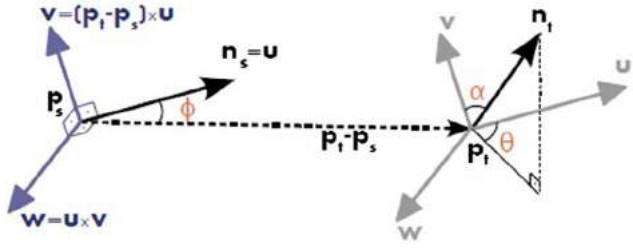

Figure 6. Darboux frame between a point pair in one depth image [18].

in p_s is created as follows:

$$u = n_s \tag{1}$$

$$\mathrm{v} = \mathrm{u} \times \frac{(p_t - p_s)}{\|p_t - p_s\|} \tag{2}$$

$$\mathrm{w} = \mathrm{u} \times \mathrm{v} \tag{3}$$

As the next step three angular valuesα, ϕ and θ are calculated, by using the distance between p_s and p_t, the normal n_t of the point p_t, and the Darboux frame:

$$\alpha = \mathrm{v} \cdot n_t \tag{4}$$

$$\phi = \mathrm{u} \cdot \frac{(p_t - p_s)}{d} \tag{5}$$

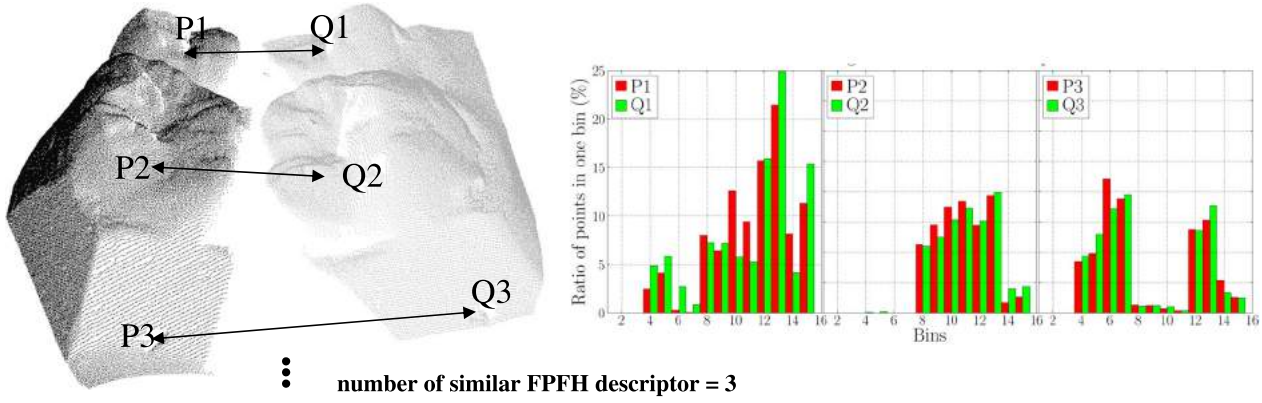

Figure 7. Correspondence matching with similar FPFH descriptors [19].

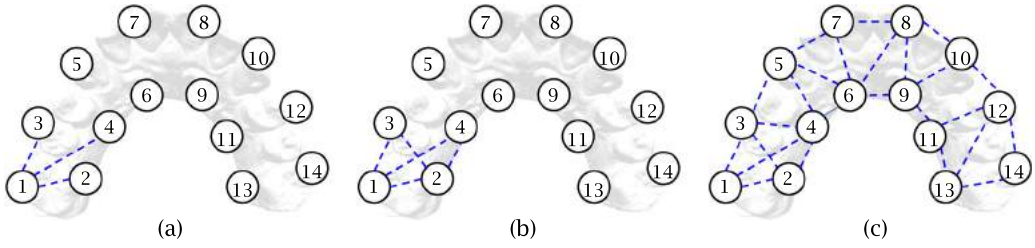

Figure 8. The process of creating pairwise chain based on AABB overlapped neighborhoods. from left to right; (a) Depth image #1 Step; (b) Depth image #2 Step; (c) Depth image #14 step.

$$\theta = \arctan(w \cdot n_t, u \cdot n_t) \tag{6}$$

The three angular values indicate the angles between the point normal vectors. Then, twelve values from the two points and two normal vectors are decreased to three angular values, which describe their relationships. All calculated angular values from the relationship between p_q and its point neighbors are now binned into a Simplified Point Feature Histogram(SPFH) of p_q. Each of the three angular dimensions of a point pair relationshipα, ϕ and θ is divide into s intervals. As a result, there are 3 kinds of histogram bins. The final FPFH of p_q is calculated by adding the weighted SPFH values of its point neighbors p_k (See Eq. 8, point neighbors in user specified boundary). The weight w_k is defined as the Euclidean distance from p_q to the point p_k[22].

$$\text{FPFH}(p_q) = SPFH(p_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{w_k}SPFH(p_k) \tag{7}$$

As mentioned above, FPFH descriptor can be a feature factor that indicates a specific point. Therefore, if the certain FPFH descriptor of different depth images are the same, these point pair can be a good matching factor [22]. In this phase, FPFH descriptor is calculated for all points in all depth images. This will be used as a reference for finding the correspondence in the back-end phase.

3.2. Pairwise chain

The purpose of this section is to create a pairwise chain of all depth images with FPFH descriptors. The proposed method uses consecutively applied transformations between two depth images to align all depth images of the input set. The main idea is to create a new matching sequence order through the number of similar FPFH descriptors between the depth images.

First, AABB detection performed to select a candidate group. Because there are too many depth image samples in this study, it is inefficient to consider all of image. When only the overlapping depth images are selected through the AABB detection, about 10 to 15 depth images are taken into candidate group. The process of creating a pairwise chain based AABB is shown in (Fig. 8).

Second, a similarity index number of FPFH descriptors is calculated between all candidate depth images (Fig. 9). The similarity index indicates the similarity between the depth images. After all similarity index between candidate depth image is computed, all depth images are represented by connected, edge-weighted undirected graph. In this paper, we call it pairwise-chain. In pairwise-chain, the edge is weighted by the similarity index. After all edges are weighted, the Minimum Spanning Tree (MST) of pairwise-chain is computed. The edges between each node (Fig. 9(a)) are weighted according to number of similar FPFH Descriptors. Thus, the

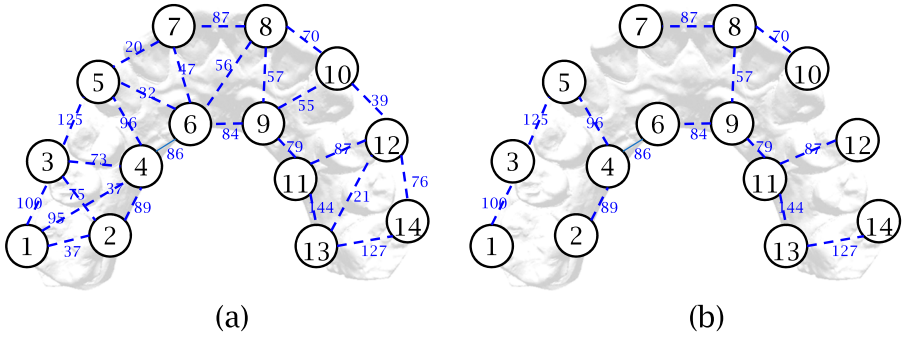

Figure 9. Pairwise chain, from left to right; (a) Weighted by Similar FPFH Descriptor; (b) Completed Pairwise chain, after MST.

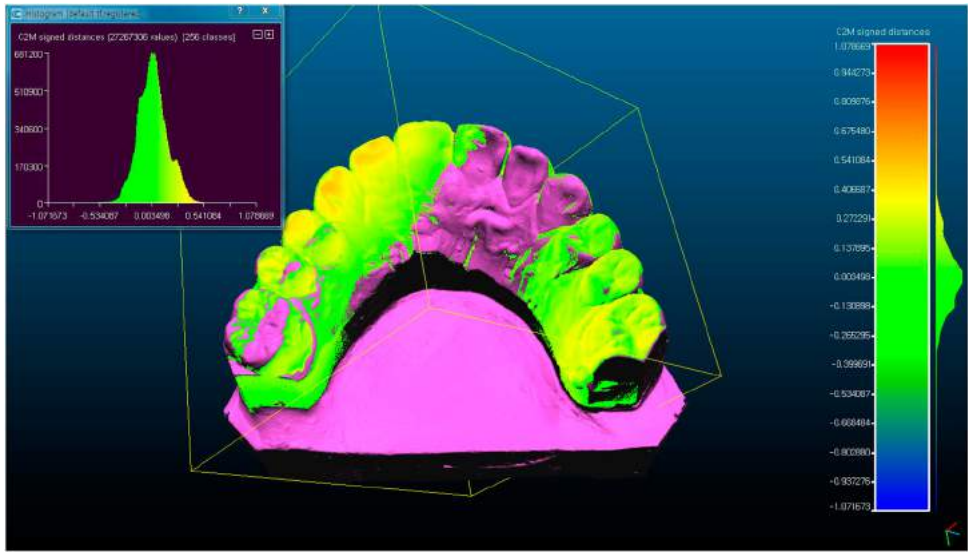

Figure 10. Cloud to mesh(C2M) signed distance and its distribution : intraoral scanning (Kinectfusion).

larger the number, the more similar the two images are. The weight of an edge in the pairwise chain is defined to be inversely proportional to the number of found point correspondences between the connected depth images (for MST calculations). After all edges are weighted, the Minimum Spanning Tree (MST) of pairwise chain is computed (Fig. 9(b)). In this paper, using Kruskal's algorithm for solving MST [10]. The resulting of MST represents the optimal solution of all possible depth image pair registration combinations.

3.3. Registration

In this section, a registration based on the pairwise chain that previously redefined is performed. The transformation matrix is applied to the original source depth image in order to align it with the original target depth image into a common coordinate system. First step is initial alignment. If the two depth images are too far away or the angle is too twisted, it will be hard to proceed with a registration [18]. In order to prevent this in advance, initial alignment performs to translate early pose. In this paper, initial alignment based on the matched FPFH descriptor used in section 3.2 is performed. The computed transformation is applied to the source depth image to roughly alignment to its target depth image. Second step is iterative closest point (ICP) algorithm [2]. Even if initial alignment is performed well, ICP algorithm should be applied because there is a high probability that a fine error will occur. The proposed method follows the modified procedure of the iterative closest point algorithm [3]. The authors propose to carry out the sparsity of the vector of residuals, i.e. maximize the number of correspondence with a distance closer to zero, instead of minimizing the squared distance of all correspondences. The nonconvex relaxation of l_p norm is used for l_p norm (p is between 0 and 1), which is usually used as convex approximation of l_0 norm. This approach yields to superior results in terms of accuracy, if the point cloud does not include too may outliers. More sufficient results for decreasing values of p are reported at the cost of more iterations until convergence [22]. From experience of authors, the entire phase of Registration Section is: initial alignment once, and ICP algorithm for 15 times.

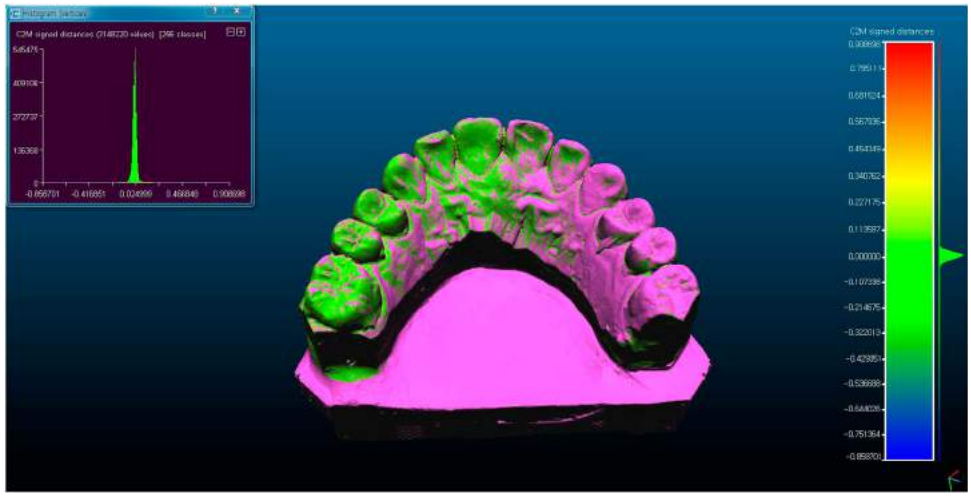

Figure 11. C2M signed distance and its distribution : The Chang's method [4].

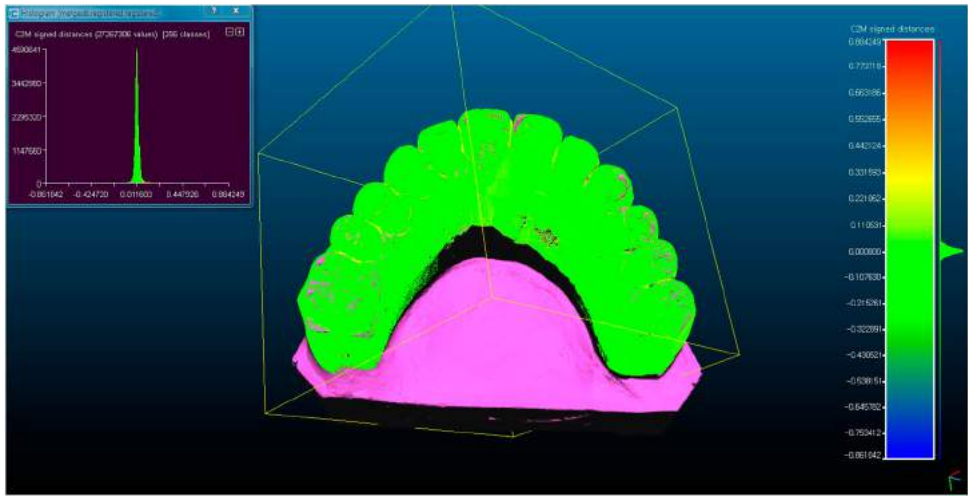

Figure 12. C2M signed distance and its distribution : The Proposed method.

Nevertheless, the search for the better ICP variants and the experiment of repetition value is needed.

4. Results

The data used in the experiment consisted of 1,021 dental depth images of mandibular teeth. Each image consists of 80,000 to 140,000 vertices, and it is an experimental data set that have been scanned from the right third molar to the left third molar. The Kinecfusion[9] method and T.huber[22] method were available on the web page, and Chang's method[4] is implemented by seeing the contents of the paper. The proposed algorithm was implemented in C++ language and test runs were made on a personal computer with a i5-2500 processor with 8 GB memory with geforce gtx 760ti(6gb) and a Windows 7 operating system. The cloud to mesh(C2M) signed distance and its distribution from *CloudCompare* software is used to measure the accuracy of registration.

Fig. 10–12 shows the comparison of the three methods. The model colored by pink is a reference model, and the other shapes are registered depth images. In Tab. 2, Mean dist column means the average error between the reference model and the registered depth image. Also, it means precision of registration and the numbers in parentheses indicate dispersion. Kinectfusion, which is the most similar method to intraoral scanning, is a real-time 3D reconstruction method performed by KINECT, which is a depth sensor of *Microsoft*. Since the alignments is performed in the order of captured, an accumulated error occurs and the overall accuracy is quite low (Fig. 12(a)). Chang's method is to solve the global hessian matrices which accumulates single pose transformations at once, and show the best results except runtime. Since this method performs mutual matching with all the depth images, it provides a precise result that is suitable for dentistry. But it has a disadvantage that runtime is comparatively long (Fig. 12(b)). T.weber's

Table 2. Result of tests.

Method	Mean dist (μm)	Runtime (s)
KinectFusion	91.00 (0.1)	Real-time
Chang's method	3.36 (0.05)	552
Proposed method	5.19 (0.05)	30

method is similar to this study, but it uses global point cloud graph for all point clouds. but this method, which considers a shape of all dental depth images, may not be suitable for patient with a similar tooth shape in the mouth. In addition, even if the registration is achieved by tuning parameters, a distortion occurs in the shape of the teeth by global registration. The detailed distribution of each result can be confirmed by the upper left graph of (Fig. 10–12).

5. Conclusion

This paper proposes an automated alignment procedure of dental depth images captured by intraoral scanners. A core part of the proposed procedure is an algorithm that computes the similarity of each dental depth images, considering only the adjacent dental depth images. The algorithm was designed carefully by considering two major technological requirements of the problem: accuracy and efficiency. To satisfy the accuracy requirement, the proposed algorithm consists of three steps: (1) find a adjacent dental depth image by AABB detection, (2) compute a FPFH descriptor between adjacent dental depth images, (3) make a pairwise-chain, (4) solve a pairwise-chain using MST, and (5) reconstruct a registration sequence. The proposed method provides better result than conventional approach and satisfied a required level of result in dentistry. For the implementation of the proposed algorithm, a detailed description of the algorithm is provided and several illustrations are provided to prove the usefulness of the algorithm. Although the main application of the proposed algorithm is dental depth images from an intraoral scanner, the proposed algorithm can be applied to other small objects.

Acknowledgements

This work was supported by a grant (17CTAP-C129828-01) from Infrastructure and transportation technology promotion research program funded by Ministry of Land, Infrastructure and Transport of Korean government.

ORCID

Min S. Ko http://orcid.org/0000-0001-7088-3150

References

[1] Benjemaa, R.; Schmitt, F.: Fast global registration of 3D sampled surfaces using a multi-z-buffer technique, Image and Vision Computing, 17(2), 1999, 113–123. http://doi.org/10.1016/S0262-8856(98)00115-2

[2] Besl, P-J.; McKay, N-D.: Method for registration of 3-D shapes, Robotics-DL tentative, International Society for Optics and Photonics, 1992, 586-606. http://doi.org/10.1117/12.57955

[3] Bouaziz, S.; Tagliasacchi, A.; Pauly, M.: Sparse iterative closest point, computer graphics forum, 32(5) 2013, 113–123.

[4] Chang, M-H.; Kim, T-W.; Seo, Y-H.; Lee, S-C.; Yang, Z.: Simultaneous Registration of Multiple Range Views of Large 3D Scanned Data with Markers, In Proceedings of the 2008 CAD/CAM KOREA, 77–84.

[5] Chang, M-H.; Oh, J-W.; Park, S-C.: Next viewing directions for the scanning of dental impressions, Computer-aided design, 66, 2015, 24–32. http://doi.org/10.1016/j.cad.2015.04.007

[6] Denry, I.; Kelly, J-R.: State of the art of zirconia for dental applications, Dental materials, 24(3), 2008, 299–307. http://doi.org/10.1016/j.dental.2007.05.007

[7] Güth, J.; Keul, C.; Stimmelmayr, M.; Beuer, F.; Edelhoff, D.: Accuracy of digital models obtained by direct and indirect data capturing, Clinical oral investigations, 17(4), 2013, 1201–1208.

[8] Hänsch, R.; Weber, T.; Hellwich, O.: Comparison of 3D interest point detectors and descriptors for point cloud fusion, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(3), 2014, 57–64.

[9] Izadi, S.; Kim, D.; Hiliges, O; Molyneaux, D.; NewCombe, R.; Kohli, P.; Fitzgibbon, A.: KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera, Proceedings of the 24th annual ACM symposium on User interface software and technology, 2011, 559–568.

[10] Kruskal.; Joseph, B.: On the shortest spanning subtree of a graph and the traveling sales problem, Proceedings of the American Mathematical society, 7(1), 1956, 48–50.

[11] Lowe, D.G.: Distinctive image features from scal-invariant keypoints, International journal of computer vision, 60(2), 2004, 91–110 http://doi.org/10.1023/B:VISI.0000029664.99615.94

[12] Luthardt, R.G.; Loos, R.; Quaas, S.: Accuracy of intraoral data acquisition in comparison to the conventional impression, International journal of computerized dentistry, 8(4), 2005, 283–294.

[13] Moreels, P.; Perona, P.: Evaluation of features detectors and descriptors based on 3d objects, International Journal of Computer Vision, 73(3), 2007, 263–284.

[14] Patel, N.: Integrating three-dimensional digital technologies for comprehensive implant dentistry." The Journal of the American Dental Association, 2010, 20S-24S. http://doi.org/10.14219/jada.archive.2010.0357

[15] Pauly, M.; Gross, M.; Kobbelt, L.: Efficient simplification of point-sampled surfaces, Proceedings of the conference on Visualization'02, IEEE Computer Society, 2002, 163–170.

[16] Quass, S.; Rudolph, H.; Luthardt, R-G.: Direct mechanical data acquisition of dental impressions for the manufacturing of CAD/CAM restorations, Journal of

dentistry, 35(12), 2007, 903-908. http://doi.org/10.1016/j.jdent.2007.08.008

[17] Rudolph, H.; Luthardt, R-G.; walter, M-H.: Computer-aided analysis of the influence of digitizing and surfacing on the accuracy in dental CAD/CAM technology, Computers in biology and medicine, 37(5), 2007, 579–587. http://doi.org/10.1016/j.compbiomed.2006.05.006

[18] Rusinkiewicz, S.; Levoy, M.: Efficient variants of the ICP algorithm, 3-D Digital Imaging and Modeling, 2001, 145–152.

[19] Rusu, R-B.; blodow, N.; Beetz, M.: Fast point feature histograms (FPFH) for 3D registration, Robotics and Automation, 2009, 3212–3217.

[20] Seelbach, P.; Brueckel, C.; Wostmann, B.: Accuracy of digital and conventional impression techniques and workflow, Clinical oral investigations, 17(7), 2013, 1759-1764. http://doi.org/10.1007/s00784-012-0864-4

[21] Tan, P.L.B.; Dunne, J.T.: An esthetic comparison of a metal ceramic crown and cast metal abutment with an all-ceramic crown and zirconia abutment: a clinical report, The journal of prosthetic dentistry, 91(3), 2004, 215-218. http://doi.org/10.1016/j.prosdent.2003.12.024

[22] Weber, T.; Hänsch, R.; Hellwich, O.: Automatic registration of unordered point clouds acquired by Kinect sensors using an overlap heuristic, ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 96–109. http://doi.org/10.1016/j.isprsjprs.2014.12.014