https://doi.org/10.2352/EI.2022.34.17.3DIA-212

Pose Estimation of Teeth in Pathological Dental Models

Maxime Chapuis[1, 2], Mathieu Lafourcade[1], William Puech[1], Gérard Guillerm[2] and Noura Faraj[1]

[1] LIRMM, Université de Montpellier, CNRS Montpellier, France
[2] Groupe Orqual, Pau, France

Abstract

In this paper, we present a method to estimate the pose of teeth in pathological dental models. Given a 3D pathological model, we aim at computing the orientation and position of each tooth with respect to a healthy dentition. This deviation from a healthy dentition is informative of the orthodontic disorders of the patient. The proposed method is based on the registration of a reference dental model on a patient-specific 3D segmented mesh. Then, dental features, such as the patient's arch forms, are derived. These features, combined with registration information, allow our system to propose a plausible target dental arrangement from which the pose of each tooth is derived.

The key contributions of this work are (a) the use of a registered reference model to automatically estimate a plausible target dental arrangement for a given patient, and (b) the derivation of teeth poses from registration information and the target arrangement.

Introduction

Computer-aided orthodontic treatment planning is used to improve treatment outcome as well as the practitioner's workflow. The goal is to plan the treatment outcome and its different steps on a virtual dental model of the patient. This work has direct applications in this domain, and aims at providing new useful information to the orthodontist. Given a pathological 3D model, we aim at computing the orientation and position of each tooth with respect to a healthy dentition. The *pose* of a tooth corresponds to the *deviation* in *orientation* and *position* from a *healthy* tooth arrangement. We believe that a proper estimation of pathological teeth poses could be valuable to:

- Automatically identify and quantify orthodontic disorders, and indicate the teeth and sectors presenting the most pathological situation.
- Cluster patient models presenting similar pathologies.
- Automatically create treatment animations to help patients visualize their treatment.
- Speed-up the process of designing the outcome of a treatment by using an automatic tooth arrangement as a starting point.

3D pose estimation is the task of predicting the transformation of an object with respect to a reference pose. In the case of teeth, the *reference pose* (a healthy dentition) is patient dependent. This means that a *reference arrangement* has to be estimated for each patient. This problem is closely related to the task of *automatic tooth arrangement* where the goal is to predict a proper tooth arrangement given a pathological dental model (*i.e* a plausible result of the orthodontic treatment). Most of the existing methods for *automatic tooth arrangement* are found in proprietary software, and their details are not easily available. Recently, Wei *et al.* and Li *et al.* proposed two deep-learning based methods for such a task [9, 7]. While offering good-quality tooth alignments on most studied common cases, these methods require large datasets of pre and post treatment tooth arrangements for their training.

In this paper, we propose a pipeline based on the registration of a reference dentition model onto the patient's pathological model. A target tooth arrangement is then estimated by deforming the reference model in a way that its arch forms best matches the patient's arch forms. Note that, the arch form is the shape of the dental arch. Combined with registration information, this estimated target arrangement is used as a reference to compute the pose of each tooth.

The contributions of the proposed approach are (a) the use of a registered reference model to automatically estimate a plausible target dental arrangement for a given patient, and (b) the derivation of teeth poses from registration information and the target arrangement.

Method

The proposed pipeline, illustrated in Fig. 1, has two inputs: a reference dental model, and a segmented and labeled 3D pathological patient model (an example is shown in Fig. 2). Patient models are usually 3D meshes reconstructed from intra-oral scans. To automatically obtain the segmentation of a model, one can use recent deep learning based methods [8, 4].

The pipeline is composed of three main steps, which are the registration, the target estimation and the pose estimation:

- **Registration:** Our reference dental model (shown on Fig. 3a) is registered onto the pathological patient model. The goal of the registration process is to put our reference model in the same pathological configuration as the patient's.
- **Target estimation**: The arch forms are computed on the reference model and its registered version. The target tooth arrangement is then estimated by matching the reference arch forms to the patient arch forms, thus adjusting the reference tooth arrangement to the patient's morphology.
- **Pose estimation**: The pose of each tooth is computed with respect to the estimated target arrangement.

In the next section, we propose to detail each step of the pipeline illustrated in Fig. 1.

Registration

The registration step of the pipeline aims at setting the reference teeth in the patient's pathological arrangement. It decom-

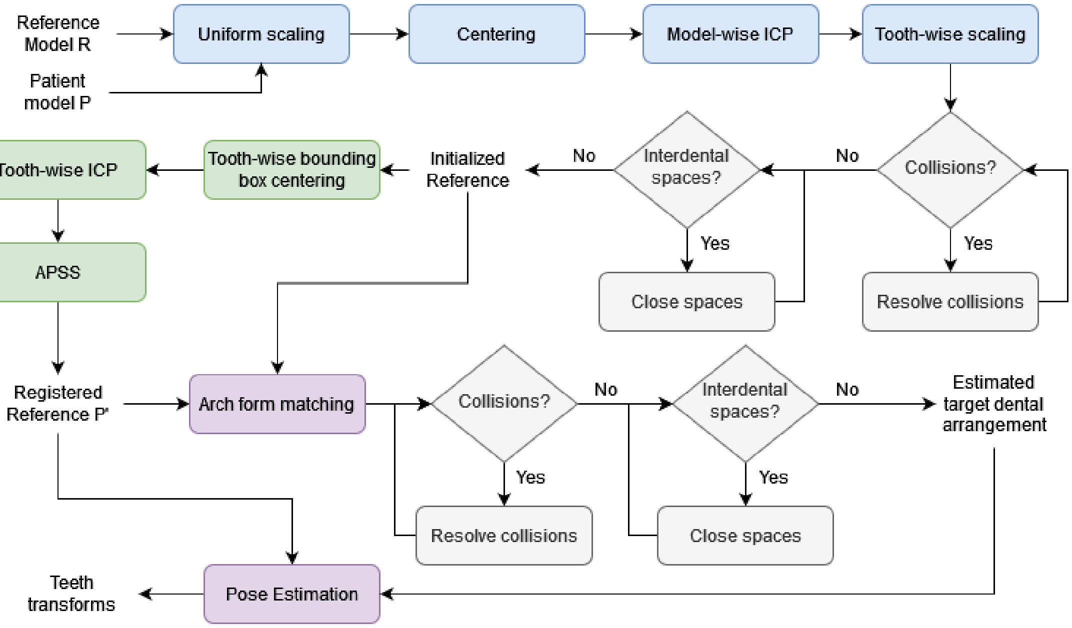


Figure 1: The detailed pipeline used for the pose estimation.

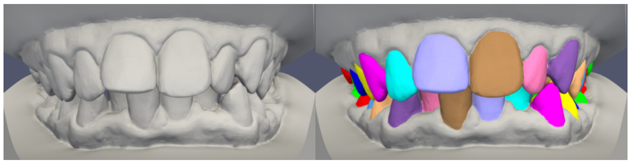

(a) (b)

Figure 2: Front view of a 3D patient model : a) Raw model, b) The segmented and labeled version of the model.

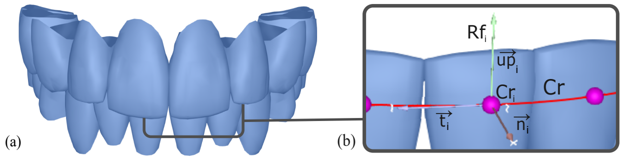


Figure 3: a) Full view of the reference model, b) Each tooth has an associated control point Cr_i, and reference frame Rf_i. Together, the control points Cr_i of an arch make the reference curve Cr.

poses into a *rigid* and a *non-rigid* part. The first one *places* the teeth, whereas the second one *deforms* them so that they closely match the shape and features of the patient's.

Reference Model

Our reference model R, shown in Fig. 3a, is a synthetic model representing generic teeth in a standard configuration (healthy configuration). It is considered as the default target tooth arrangement throughout the pipeline.

To keep track of the teeth poses, control points are placed at the center of each vestibular face. The vestibular side of a tooth faces the lips and cheeks. Together, the control points of an arch are used as the control points of a parametric curve $Cr(t)$ (one per arch). This *reference curve* is a Catmull–Rom spline [2]. The choice to use Catmull–Rom spline is motivated by its interpolating properties (the curve goes exactly through the control points), and its relatively low computational cost. This curve is essential as it gives the ability to define a reference position and orientation at any parameter t. For each control point Cr_i of Cr, a reference frame Rf_i is computed using the tangent at t and the up vector (Fig. 3b). It also has the benefit of being an intuitive primitive for orthodontists, as it resembles an orthodontic wire with braces. Throughout the pipeline, the control points are subject to same the transformations as their corresponding tooth.

Rigid Registration

The main algorithm used to carry out the rigid registration is the Iterative Closest Point (ICP) [1, 3]. This algorithm iteratively constructs a rigid transformation which minimizes the difference between two point-clouds by finding point correspondences in the two datasets. As the ICP algorithm is sensitive to the initial alignment, the registration process starts by an initialization step. Its purpose is to place the reference model R in a good enough starting position to perform an ICP on each tooth. This initialization is done by first scaling and centering R on the patient model P. These operations are done "model-wise", meaning that every tooth is scaled by the same ratio and translated by the same vector. The scaling ratio is estimated by averaging the ratios between the incisors of P and R. The incisors are well suited for this purpose as they are almost always entirely visible in patient models. The two models are then roughly aligned by a model-wise ICP.

The "tooth-wise" operations are then performed. Each tooth of R is scaled to match its corresponding tooth in P. Note that the teeth are not scaled in the up direction because the lower parts of the patient's teeth are hidden under the gum. This scaling operation may introduce mesh collisions in R. In this case, the collisions have to be resolved before moving on to the next step in order to keep a coherent arrangement. One possible strategy is to move each ill-positioned tooth along the reference curve Cr, going

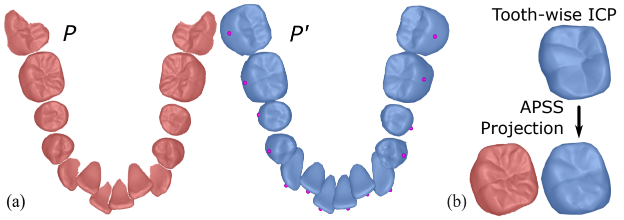


Figure 4: a) Views of the lower arches of the patient model P (red) and the fully registered reference model P' (blue), b) Comparison of the right lower first molar before APSS (top) and after (bottom).

from the incisors to the molars. The pair of central incisors have to be treated first then, the left and right side can be processed independently.

To end the initialization process, the teeth bounding boxes of P and R are aligned. Mandibular bounding boxes[1] are aligned on their upper center, and maxillary bounding boxes[2] on their lower center. The choice to align the top of the crowns is motivated by the fact that patient teeth are not always entirely visible. Therefore, centering them on their bounding boxes centers would result in a poor initialization for the ICP.

Finally, each tooth in R is registered to the corresponding tooth in P by a tooth-wise ICP. From a practical point of view, only the vertices contained in the bounding boxes of the patient's teeth are considered during the ICP.

The result of the rigid registration is a modified version of R called P', with the same tooth arrangement as the patient.

Non-rigid Registration

The teeth of the reference model R are generic teeth and thus, do not match the patients teeth geometry. These differences are mitigated by projecting the reference crowns vertices onto the patient's crowns using the projection method APSS [5]. To avoid unwanted deformations, care should be taken to only project overlapping vertices. The result of this final registration step is shown in Fig. 4a and Fig. 4b.

Target Arrangement Estimation

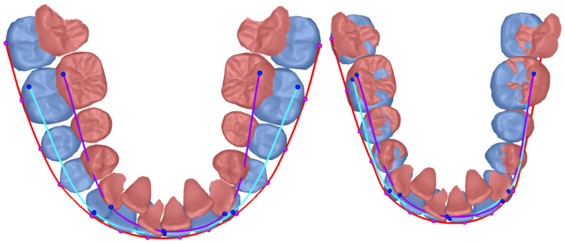
At this point, the reference model R is still in its default arrangement and therefore, does not fit the patient's morphology (as illustrated in Fig. 5a).

Arch forms computation

Following the approach proposed by Kumar *et al.* [6], cusps and incisal edges are computed on R. Theses features are then transferred on P' using the previously computed registration transformations. Since we transfer the feature points, we ensure that they are the same on both models. The respective arch forms of R and P' can then be derived as the curves passing through the outer cusps of the first molars, the canines cusps and the midpoint of the incisors of each arch. In Fig. 5a and Fig. 5b, the arch forms are represented by the blue and the purple curves.

Arch forms matching

To estimate a proper target arrangement, the reference curve Cr of each arch is transformed so that the arch forms of R match the arch forms of P'. The arch forms are represented by a Catmull–Rom spline made of five control points (blue spheres one on Fig. 5a and Fig. 5b).



(a) (b)

Figure 5: Superimposed views of R and P before and after the target tooth arrangement estimation: a) R in the default arrangement, b) R in the target arrangement.

First, the lower reference arch form is centered on its corresponding patient arch form. Then the control points associated to the reference canine cusps and molar cups are translated on the horizontal plane to their matching control points on the patient's arch form. Finally, it is translated back to its original position. A matching upper arch form is created by expanding the lower target arch form by 2.5 mm in the direction of its normals.

Given these patient specific arch forms, the goal is to transfer the changes to the reference curves Cr. On each arch, the anterior teeth[3] and the buccal cusps of posterior teeth[4] should be aligned along the arch form. This is achieved by projecting the edges of the anterior teeth and the buccal cusps of the posterior teeth on the arch form, thus changing the curve Cr. The ideal orientations are then recomputed (as in the Section ”Reference Model”), and the spaces between adjacent pairs of teeth are closed. We treat this operation in the same way as we resolve the collisions in the ”Rigid Registration” subsection , *i.e.* by moving each ill-positioned tooth along the reference curve Cr until there is no space left between two adjacent teeth. The updated curves Cr constitute the *target arrangement* as illustrated in Fig. 5b.

Pose Estimation

The pose of each tooth with respect to the estimated target arrangement is computed as the difference between the pathological arrangement and the target arrangement. For each control point Cr_i, let S_i be its rigid transformation in the initial arrangement, and T_i its rigid transformation in the target arrangement, then the pose P_i of the associated tooth is:

$$P_i = S_i T_i^{-1}. \tag{1}$$

The 4x4 matrix P_i corresponds to the deviation from the target arrangement, and is decomposed in a translation and a rotation. The rotation can then be expressed in the local frame of the associated tooth. The local frame of a tooth is the same as the frame of the corresponding control point in the target arrangement. Note that the origin of the frame is positioned at the center of the tooth.

---

[1] *Mandibular* refers to the lower arch.

[2] *Maxillary* refers to the upper arch.

[3] The fours incisors and the canines.

[4] The premolars and molars.

Experimental Results

As the teeth poses are derived from the registered pathological arrangement and the estimated target arrangement, we have to determine whether:

- The target arrangement is plausible (*i.e.* close to the outcome of an actual treatment),
- The registration produces an satisfactory pathological arrangement (as close as possible to the patient model).

To this end, we designed two experiments, one real patient cases, and one on synthetic cases.

Experiment on real data

The pipeline is run on four different orthodontic cases, two for which we have the post-treatment scans. The patient models come from real intra-oral scans made with a 3Shape TRIOS scanner[5]. The reference model is the synthetic model described in the previous section.

For the first two cases, *a.* and *b.*, the registered reference and the estimated target arrangement are compared to the initial patient arrangement.

Case a. (Fig. 6a), exhibits moderate disorders. Its main issues are the position of right upper premolar and the inclination of the incisors. These disorders are well preserved in the registered reference. The estimated arrangement successfully corrects the position of the upper right second premolar and the forward tip of the incisors.

Case b. (Fig. 6b), suffers a deep overbite (4.8 mm) and severe crowding of the anteriors. The overbite measures the overlapping of the lower incisors by the upper ones. Again, the registered reference is faithful to the patient model. Note that the partial left molar was successfully registered. In the estimated target arrangement, the overbite is reduced and the crowding is resolved. However, the real treatment plan for this patient suggests that the extraction of four premolars is needed. This case illustrates the fact that the expertise of the user is still needed to achieve a proper arrangement on severe or rare cases.

For the two other cases, *c.* and *d.*, the estimated target arrangement is compared the real post-treatment arrangement.

Case c. (Fig. 7a), presents a fair amount of overbite, and several teeth are misaligned. When compared to the post-treatment scan, we observe that the estimated arrangement is very close to the ground truth.

Case d. (Fig. 7b), shows severe disorders in the incisors. The estimated target arrangement seems plausible and fixes these issues, however some differences with the post-treatment scan can be observed. First, the centering of the two arches is sub-optimal and we can notice a slight rotation of the upper arch. Second, the arch forms are narrower than the ground-truth, which leads the to the second molars being out of alignment.

In the four presented cases, the teeth of the estimated target arrangements are well positioned along the arch, with the exception of some second molars. The teeth are leveled on the same horizontal plane, with no spaces between adjacent teeth. However, the occlusion could be improved. Indeed, the arches have matching shapes, but the arrangements lack optimal contacts between opposing teeth. This can be explained by the existence of some contradictions between the occlusion constraints, and leveling and the spacing constraints

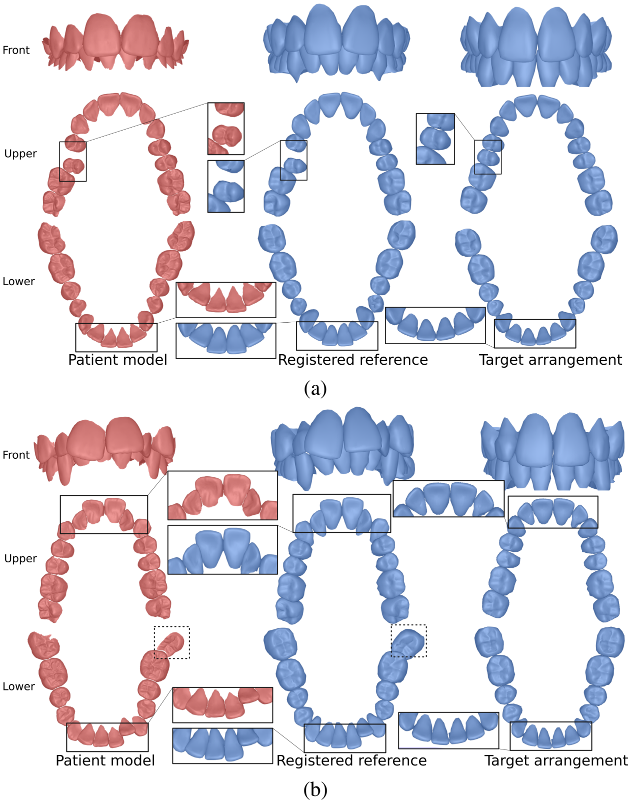


Figure 6: The proposed pipeline applied to two patient models. For each model we show the initial arrangement of the patient model, the registered reference, and the estimated target arrangement: (a) patient case A, (b) patient case B.

This first experiment shows that (a) the registration is accurate enough to produce faithful pathological arrangements which preserve the disorders of the patient, and (b) that the estimated target arrangement are plausible.

Experiment on synthetic data

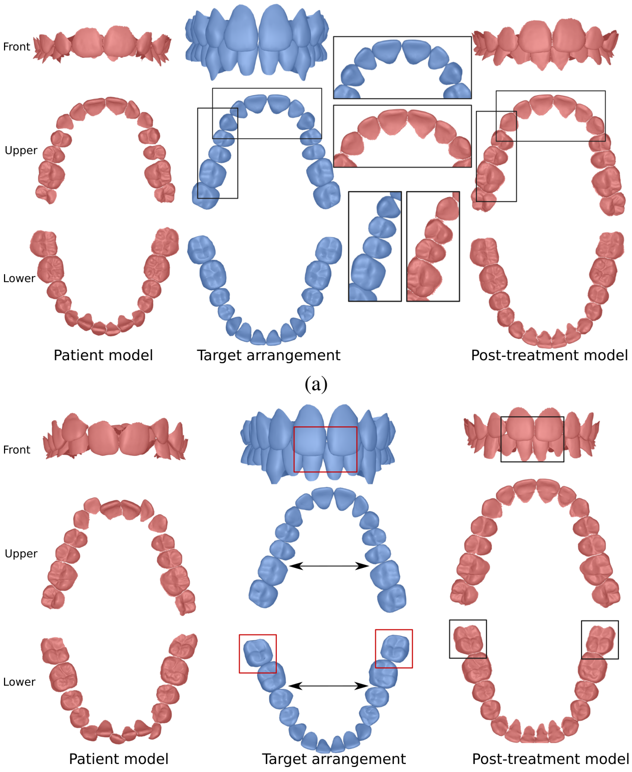
In this second experiment, the pipeline is run on a set of five synthetic patient cases. The goal is to measure the registration error and the pose estimation error. For each synthetic patient, we produce a ground truth target arrangement, and a pathological one. The ground truth is obtained by first changing the shape of the reference teeth, and then altering the shapes of the reference arch forms (using narrowing, widening or shortening operations). To obtain the pathological arrangement, teeth rotations and translations are introduced in the ground truth arrangement.

We first measure the registration error for each model by comparing the registered reference to the corresponding synthetic pathological arrangement. In Table 1, we report, for each type of tooth, the mean registration error in rotation around the three local axes (rx, ry, rz) and the mean registration error in position.

Table 1 shows that the positioning of the teeth is accurate enough for our application, with a mean positioning error way beyond 1 mm (the average length of a tooth is between 7 to 10 mm). The orientation is satisfactory, with a slight deviation from

---

[5] https://www.3shape.com/fr/scanners/trios



(a)

(b)

Figure 7: The proposed pipeline applied to two patient models for which the post-treatment scans were available. For each model we show the initial arrangement of the patient model, and compare the estimated target arrangement to the post-treatment model: (a) patient case C, (b) patient case D.

the ground truth (less than 2° on each axis). Note that, due to their shapes and the fact that they only have one sharp feature, the canines are more difficult to register (with a mean error of 3° on the x axis and 2° on the z axis).

We also measure the error in pose estimation by comparing the estimated pose of each tooth with its real pose (computed with respect to the ground truth arrangement). As for the registration error, we report in Table 2 the mean error in rotation and position for each type of tooth.

As reported in Table 2, most orientation errors occur on the y axis (in our system, it corresponds to the vertical axis), and more particularly on molars (with a mean error of 7°). A mean position error of 2.6 mm is also measured. This confirms what we observed on the patient cases of the first experiment. Even if a portion of the errors are due to the registration, the positions errors and the rotations errors on the y axis suggest that the estimated target arch forms are slightly *wider* or *narrower* than the real arch forms.

Conclusion

In this paper, we proposed an efficient method to estimate teeth poses in pathological dental models with respect to a healthy dentition. Our method is based on the use of a registered *reference dental model* to infer a patient specific healthy arrangement from which teeth poses are derived. The two experiments show that our method is capable of producing *plausible target arrangements* and *informative teeth poses*. They also highlight some limitations. First, the registration is sensitive to quality of the segmentation and partial data. Second, the arch forms "match", but more constraints have to be enforced to obtain fine-grained occlusion between opposing teeth. Third, the need for tooth extraction is not predicted, which can lead, in some cases, to unattainable target arrangements.

In future work, we propose to add a dedicated *occlusion module* to maximize the contacts between opposing teeth in the target arrangement, and tweak the target arch forms to better match the post-treatment results.

In spite of these limitations, the target arrangements and teeth poses could be suitable for the automatic identification of dental disorders, the automatic creation of treatment animations, or for patient clustering. They cannot however be used *as-is* for medical appliances design. For this use case, the arrangement and the registration data have to be used as input of an interactive system to allow an expert to adjust the orientations and positions. This would still be an interesting use case, as it could speed up the treatment outcome design by using our target arrangement as a "first-guess".

	rx (in °)	ry (in °)	rz (in °)	p (in mm)
Incisor	0.97	1.53	0.87	0.35
Canine	**3.24**	1.85	**2.04**	0.34
Premolar	1.01	1.86	1.11	0.23
Molar	1.43	1.81	1.25	0.29
Mean	1.44	1.75	1.21	0.29
Std Dev	1.54	1.67	1.12	0.15

Table 1: Registration error by type of tooth. First, the orientation error on the three local axes (in °), then the position error (in mm).

	rx (in °)	ry (in °)	rz (in °)	p (in mm)
Incisor	1.06	3.95	0.91	2.41
Canine	3.35	4.14	1.87	2.29
Premolar	1.04	3.32	1.17	2.52
Molar	1.67	**7.08**	1.27	**3.05**
Mean	1.56	4.69	1.22	2.61
Std Dev	1.62	3.65	1.10	1.37

Table 2: Errors in pose estimation by type of tooth. First, the orientation error on the three local axes (in °), then the position error (in mm).

References

[1] Besl, P. J., and McKay, N. D. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (1992), 239–256.

[2] Catmull, E., and Rom, R. A class of local interpolating splines. In *Computer Aided Geometric Design*, R. E. BARNHILL and R. F. RIESENFELD, Eds. Academic Press, 1974, pp. 317–326.

[3] Chen, Y., and Medioni, G. Object modelling by registration of multiple range images. *Image and Vision Computing 10*, 3 (1992), 145 – 155. Range Image Understanding.

[4] Cui, Z., Li, C., Chen, N., Wei, G., Chen, R., Zhou, Y., and Wang, W. TSegNet: An efficient and accurate tooth segmentation network on 3D dental model. *Medical Image Analysis 69* (2021),

101949.

[5] Guennebaud, G., and Gross, M. Algebraic point set surfaces. *ACM Trans. Graph. 26*, 3 (July 2007), 23–es.

[6] Kumar, Y., Janardan, R., and Larson, B. Automatic feature identification in dental meshes. *Computer-Aided Design and Applications 9* (08 2013), 747–769.

[7] Li, X., Bi, L., Kim, J., Li, T., Li, P., Tian, Y., Sheng, B., and Feng, D. Malocclusion treatment planning via pointnet based spatial transformation network. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* (Cham, 2020), A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, and L. Joskowicz, Eds., Springer International Publishing, pp. 105–114.

[8] Lian, C., Wang, L., Wu, T.-H., Wang, F., Yap, P.-T., Ko, C.-C., and Shen, D. Deep multi-scale mesh feature learning for automated labeling of raw dental surfaces from 3d intraoral scanners. *IEEE Transactions on Medical Imaging 39*, 7 (2020), 2440–2450.

[9] Wei, G., Cui, Z., Liu, Y., Chen, N., Chen, R., Li, G., and Wang, W. TANet: Towards Fully Automatic Tooth Arrangement. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12360 LNCS. 2020, pp. 481–497.