jh190077-NAJ

Development of artificial intelligence systems to predict facial morphology after orthodontic treatment

Chihiro Tanikawa, DDS, PhD

Dept. of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Osaka University

Background

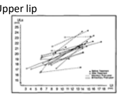
Conventional 2D software

Fig. 1. Background

Based on the premises that

(1) **the degree of movement of hard tissue is proportional to that of soft tissue, and**

(2) **a proportional constant number can be selected by orthodontists, due to which the results among different operators show significant variations**



Coefficient of determination R^2
(A measure of the global fit of the model)

Upper lip and U1 = 0.28
Lower lip and L1 = 0.40

Low predictability

Yogosawa F. Angle Orthod. 1990.

Objective

To develop artificial intelligence (AI) systems that predict the 3D facial topography after orthognathic surgery and fixed edgewise orthodontic treatment.

Methods

- **Samples:** A total of 137 patients who underwent orthognathic surgery (n =72, mean age = 23.5 years old) and orthodontic treatment with premolar extractions (n = 65, mean age = 15.6 years old) were enrolled. Three-dimensional facial images and lateral cephalograms were obtained before and after the treatment.
- **Wire mesh fitting:** For each patient at each time point, a wire mesh fitting on the face was conducted. This method generated 6,017 semi-landmarks on the wire mesh (Figs. 2 and 3).
- **AI systems :** Based on the deep learning method, we developed two AI systems (S and E) to predict changes in the coordinate values of the semi-landmarks due to orthognathic surgery and orthodontic treatment with premolar extractions, respectively. Cephalometric changes during treatment and coordinate values of semi-landmarks on the faces before treatment were employed as predictor variables. The predicted post-treatment facial morphology was calculated as the sum of the coordinate value on the face before treatment and the predicted change for each semi-landmark (Fig. 4).
- **Evaluation:** The system performance was evaluated using eleven-fold cross-validation for each AI system. The system error was determined as the difference between actual post- and predicted post-treatment coordinates of the semi-landmarks in Z-axis (Table 1). Furthermore, the total success rate was examined, wherein success cases were those with an average error of less than 1 mm or 2 mm.

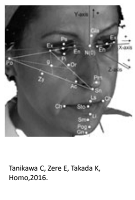


Fig. 2. Coordinate system

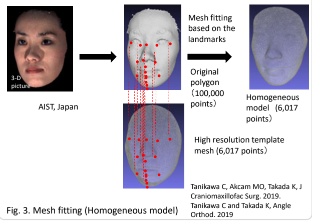


Fig. 3. Mesh fitting (Homogeneous model)

The data was divided into 11 parts, one of which was employed as test data and the remaining 10 as learning data. The test and learning data were exchanged 11 times, and the prediction accuracy of all cases was evaluated.

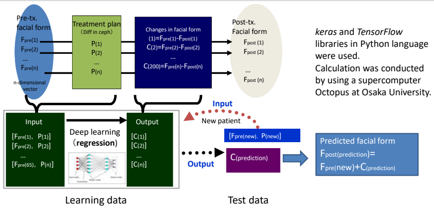


Fig. 4. Mathematical model

Table 1. Calculation of the averaged error d (AveEachPt, AveEachPc, TotalAve)
d(I,j) indicates the error of facial point *j* in the patient *i*.

	No. of facial point						Average
Patient	d(1,1)	d(1,2)	⋯	d(1,j)	⋯	d(1,6017)	AveEachPc(i=1)
	d(2,1)			d(2, j)		d(2,6017)	AveEachPc(i=2)
	⋮			⋮		⋮	⋮
	d(i,1)	d(i,2)	⋯	d(i,j)	⋯	d(i,6017)	AveEachPc(i)
	⋮			⋮		⋯	⋮
	d(n,1)	d(n,2)	⋯	d(n,j)	⋯	d(n,6017)	AveEachPc(i=6017)
Average	AveEachPt(j=1) StdEachPt(i=1⋯n,j=1)	AvePtEach(j=2) StdPtEach(j=2)	⋯ ⋯	AveEachPt(j) StdEachPt(j)	⋯ ⋯	AveEachPt(j=n) StdEachPt(j=n)	TotalAve

Results

Soft- and hard- tissue changes after treatment were shown in Figs 5-6. The system error was 0.89 ± 0.30 mm (S) and 0.69 ± 0.18 mm (E; Table 2). Maximum errors were observed in the nasal alar (S), the chin (S) , the upper lip (S), and the lower lip (S and E; Figs. 7-8). The total success rate of <1mm was 82% (74% [S] and 92% [E]), and the total success rate of <2mm was 100%. A system that recognize the cephalometric landmarks was developed separately [1], and a GUI was developed for the clinical application (Fig. 7).

[1] *Lee C, Tanikawa C, Lim JY, Yamashiro T. "Deep Learningbased Cephalometric Landmark Identification using Landmarkdependent Multi-scale Patches",* *http://arxiv.org/abs/arXiv:1906.02961, 2019.*

System (S) : Orthognathic surgery

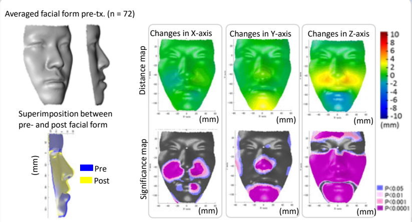


Fig. 5A. Averaged actual facial changes (n = 72)

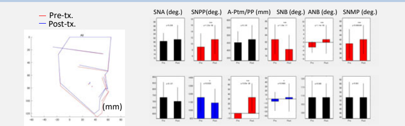


Fig. 6A. Averaged cephalometric changes (n = 72)   * P<0.05; ** P<0.01; t-test

Table 2A. Averaged error, standard deviation (S.D.), minimum value, and maximum value (mm). Please see Table 1 for the definition of AveEachPt and AveEachPc.

(unit, mm)	Average	S.D.	Min	Max	
AveEachPc	0.89	0.36	0.42	4.02	Variation among the points
AveEachPt		0.30	0.33	1.77	Variation among the patients

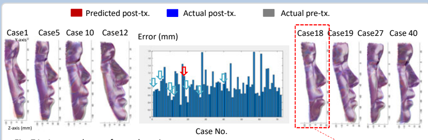


Fig. 7A. Averaged error for each patient

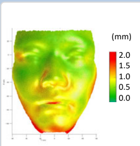


Fig. 8A. Averaged error (Absolute differences in Z-axis between predicted and actual facial form at post treatment)

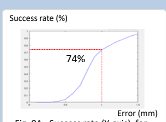


Fig. 9A. Success rate (Y-axis) for each threshold between 0.1mm and 1.5 mm (X-axis; n = 72.)

System (E) : Extraction of four premolars

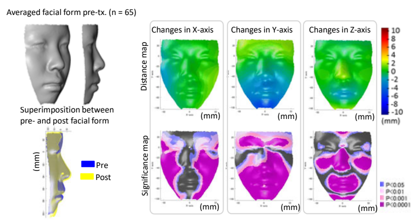


Fig. 5B. Averaged actual facial changes (n = 65)

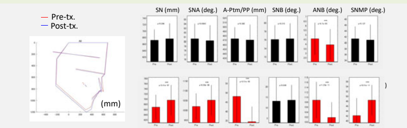


Fig. 6B. Averaged cephalometric changes (n = 65)   ** P < 0.01 ; t-test

Table 2B. Averaged error, standard deviation (S.D.), minimum value, and maximum value (mm). Please see Table 1 for the definition of AveEachPt and AveEachPc.

(unit, mm)	Average	S.D.	Min	Max	
AveEachPc	0.69	0.22	0.24	1.77	Variation among the points
AveEachPt		0.18	0.30	1.02	Variation among the patients

Fig. 7B. Averaged error for each patient

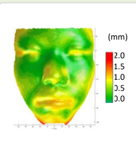


Fig. 8B. Averaged error (Absolute differences in Z-axis between predicted and actual facial form at post treatment)

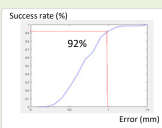
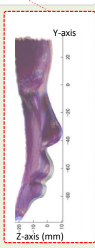


Fig. 9B. Success rate (Y-axis) for each threshold between 0.1mm and 1.5 mm (X-axis; n = 65.)



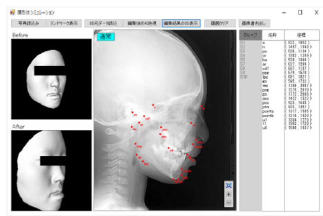


Fig. 7. GUI for the clinical application. The system was combined with a deep Learning model that can recognize the cephalograms.

Conclusions

The AI systems were confirmed to be accurate and reliable in predicting the 3D facial topography following orthognathic surgery and fixed edgewise orthodontic treatment.