Creating a visualization system for changes in facial shape and color

Reimei koike^{1)*}, Ikumi nomura¹⁾, Keiko Ogawa-Ochiai²⁾, Norimichi tsumura³⁾, and Naoki rikihisa⁴⁾

1) Department of imaging science, Graduate school of science and engineering, Chiba University, Chiba, Japan

2) Kampo Clinical Center, Department of General Medicine, Hiroshima University Hospital, Hiroshima, Japan

3) Department of imaging science, Graduate school of Engineering, Chiba University, Chiba, Japan

- 4) Oyumino Central Hospital, Chiba, Japan
- *E-mail: reimei0217@chiba-u.j

Abstract: During the course of treatment in plastic surgery and dermatology, especially aesthetic surgery, it is essential to regularly observe and record any appearance changes of the body. Physicians typically use an interview, photographs of the patient's face before and after treatment, and anatomical data including size, position, and color. However, it is difficult to photograph the affected area under the same conditions every time because the range of the photograph varies depending on the angle and distance. In addition, measurements with a measuring tape are similarly difficult to reproduce. In this study, we used an infrared depth sensor and RGB sensor built into a smartphone to measure the shape of the face and the color of the skin in three dimensions. The face shape and color were measured before and after modulation, and the changes in volume and skin pigmentation of the skin color were calculated and visualized. It was then possible to analyze the shape and skin tone independent of the viewing angle and direction of illumination light. This research shows the possibility of monitoring changes in facial shape and skin tone with a depth camera built into a smartphone. This technology is expected to contribute to the development of telemedicine, where patients can measure their faces at home and receive treatment consultations remotely.

Keywords: visualization, 3D measurement, skin pigment separation, smart phone

1 INTRODUCTION

Regular observation and recording of the changes in body appearance are essential for plastic surgery and dermatology treatments, especially aesthetic surgery. In these departments, one of the main targets of treatment is the skin.

The physician primarily observes the condition of the affected area directly and uses measurements and photographs to record the status. Physical measurement and records should be made by the same physician, and diagnosis naturally depends on their experience because the measurement accuracy depends on the physician's experience. In diagnostic imaging, it is difficult to thoroughly photograph the same position of the affected part because of differences in angles, distances, and lighting conditions. As a result, both methods are imperfect in accuracy and reproducibility. Plastic surgery, dermatology, and aesthetic surgery departments often perform treatments on the face. Even a slight change in the face can significantly change its impression; as a result, accurate measurements are needed. In current clinical practice, 3D measurement systems based on a stereo image have been proposed. Obtaining a 3D model eliminates the need to physically measure the face. Additionally, mapping facial texture maps to the 3D model solves some of the problems with angles and distances, which makes it easier for both the doctor and patient to understand the current condition of face and plan for future treatment.

Technological advances over the last several decades have replaced classic direct anthropometry (using rulers and calipers) and two-dimensional (2D) photography with noninvasive three-dimensional (3D) surface imaging methods [3]. One of the most common techniques for 3D facial surface imaging is digital stereophotogrammetry, such as the Vectra H1 [2]. This technique involves capturing images of the facial surface from multiple cameras with overlapping fields of view and then using software, based on the Multi-View Stereo (MV) [4] method, to merge these images into a single 3D model, with facial geometry represented as a dense point cloud and realistic facial skin texture [3]. Therefore, 3D shape data and a texture map representing detailed skin color information are achieved and can be used to quantitatively and quantitatively evaluate the changes in volume and skin color of multiple 3D shapes.

Infantile hemangiomas occur on the face, arms, and back in approximately 5% of infants, making them the most common benign tumor occurring during infancy. Most infantile hemangiomas are small, innocuous, self-resolving, and require no treatment. However, because of their size or location, some infantile hemangiomas are potentially problematic. These include infantile hemangiomas that may cause permanent scarring and disfigurement, hepatic or airway infantile hemangiomas, and infantile hemangiomas with the potential for functional impairment, ulceration, and associated underlying abnormalities. Early intervention and/or referral (ideally by 1 month of age) is recommended for infants who have potentially problematic infantile hemangiomas. Surgery and/or laser treatment are most useful for the treatment of residual skin changes after involution and, less commonly, may be considered earlier to treat some infantile hemangiomas [1]. If infantile hemangiomas occur on the face, prompt treatment is advisable because they can greatly affect the appearance of the face.

Masui et al. measured lower leg edema using depth cameras, such as Kinect v1 (Microsoft, U.S.) and Structure Sensor, and visualized the change in shape of the edema [5]. In this research, the face was measured using an infrared depth camera mounted on a smartphone, and the change of the shape of the face was analyzed as in the existing system using the method of Masui et al. In addition, a texture map was created and its changes in skin color were analyzed. To confirm whether the system can be used practically, the shape and skin color of an actual human face was modulated in an experiment, and the changes were measured and visualized using the proposed system.

2 Method

2.1 Acquisition of 3D face models

In this study, a 3D shape of the face was measured using an iPhone XR as a depth camera instead of a Kinect because of access. The iPhone XR is a smartphone with a structured light 30,000 dot depth sensor of infrared light projected on the face, and it measures the shape by sensing the distortion [6]. The 3D facial model was reconstructed using Bellus3D software [7]. Bellus3D allows the iPhone XR to act as a 3D scanner in acquiring a point cloud, triangle meshes, and texture map. Measurement and reconstruction of the 3D shape were completed in approximately 60 seconds. The 3D shape was obtained in .obj format, and the 4096×4096 texture map was obtained in .jpg format. Figure 1(a) shows an example of the acquired 3D shape of the face. The corresponding texture map is shown in Figure 1(b).



(a) 3D model of the face
(b) Texture map
Figure 1. Acquisition of a 3D face model and texture map.

2.2 Visualization of the 3D shape

In this section, we calculate the distance between the corresponding meshes after 3D registration; to visualize the volume change of the two 3D face models, the models were colored according to distance. In this paper, the before or after treatments were termed the "earlier model" or "later model" to distinguish between the two 3D shape models. 2.2.1 3D registration

First, two 3D shape models were applied to the RANSAC and ICP algorithm to match each position. Because the ICP algorithm, which is a general 3D registration method, strongly depends on the initial position between point clouds [8], the RANSAC method [9] was applied to remove the effect of outliers before the ICP algorithm. In this paper, the right cheek was filled with a cotton ball to modulate the facial shape to resemble an effect of surgery, as discussed in detail later. Figure 2 shows the 3D registration results of the two point clouds. The cyan colored points indicate the earlier model, and the magenta-colored points indicate the later model. With the above modulation, it is expected that two point clouds are matched at any of the parts, except for the right cheek. According to Figure 2, the RANSAC and ICP algorithm yielded an accurate 3D registration.



Figure 2. The 3D registration result.

2.2.2 Calculation of distance between two models

The shape change from the earlier model to the later model was computed. In this research, as in the method by Masui et al., the collision detection from the earlier model is applied to the later model and the distance between the two models is calculated. The Möller–Trumbore ray intersection algorithm [10], as shown in Figure 3, was used for the collision detection. In three-dimensional space, a point \mathbf{R} on a ray whose length is $t \ge 0$ in the direction of unit vector d from the ray starting point \mathbf{o} is represented by Equation 1.

$$\boldsymbol{P}(t) = \boldsymbol{o} + t\boldsymbol{d} \tag{1}$$

When a ray hits a point T on a mesh consisting of points v_0 , v_1 , and v_2 , the colliding points are expressed by the center of the mass coordinate system as Equation 2.

$$T(u, v) = (1 - u - v)v_0 + uv_1 + vv_2$$
(2)

 $u \ge 0, v \ge 0$, and $u+v \le 1$. When a ray and a mesh collide at point **P** on the mesh, the point **R** on the ray in Equation 1 and point **T** on the mesh in Equation 2 have the same coordinates, and thus are expressed by Equation 3.

$$\boldsymbol{R}(t) = \boldsymbol{T}(u, v) \tag{3}$$

As a result of Equation 3, the ray and mesh intersection problem becomes a simple ternary linear equation problem solving three variables (t, u, v). The above equations can be solved using Cramer's formula. If three variables (t, u, v)exist, the ray and mesh are intersected. Thus, we obtained the distance t between the corresponding meshes for the earlier and later models.



Figure 3. Möller–Trumbore intersection algorithm.

2.2.3 Visualization of distances

Based on the distance t between the corresponding meshes, the change in shape was visualized by coloring the meshes according to the distance. The color scale shown in Figure 4 was used for coloring. Red indicates an increase in volume, and blue indicates a decrease in volume. The range of the scale, 5 mm, was determined empirically from the results.



Figure 4. Color scale.

2.3 Visualization of skin texture

Next, we visualize the change in skin color in three dimensions. The texture map obtained in Section 2.1 has different coordinates (u, v) for each 3D shape model. Specifically, the facial landmarks in the texture map have a different coordinate for each measurement. Therefore, if we simply overlay the texture maps, the size and position of the faces are different, even on the same person. In this study, the texture map was converted by an affine transformation using the facial landmarks. This transformation makes it possible to compare and calculate pixel values for skin color. Next, the separation method for skin pigment [12] is applied to the texture map to obtain a map of skin pigment on the face that eliminates the unevenness of lighting. The change is visualized by calculating the change in concentration of skin pigment between corresponding pixels. This step is explained in the next section.

2.3.1 Conversion of the texture map

In this method, the texture map is transformed by performing the affine transformation for each triangle divided by the facial landmarks so that the coordinates of the source matches the target, taking into account the shape of the face. The facial landmarks were considered first for the transformation. In the acquired texture map, it is difficult to obtain the landmarks on the contour because the contour and neck are drawn continuously, as shown in Figure 1(b). Therefore, following a face morphing system, such as FUTON [11], we set a total of 19 facial landmarks for the eyes, mouth, ears, etc., and define a triangular patch with these landmarks as the vertices so the texture map can be transformed by the affine transformation of triangular patches. Figure 5 shows the results of the transformation of



(a) Source

(b) Target



(c) Source after conversion to the target form

Figure 5. Conversion of the texture map with 19 facial landmarks.

Figure 1(b). Figure 5(a) is the same texture map as Figure 1(b). The 19 points in the figure manually represent set facial landmarks, and the dotted lines represent the triangular patches created by each point. The texture map is transformed using an affine transformation matrix in which the triangular patches in Figure 5(a) match the coordinates of Figure 5(b), which is the target. The affine transformation matrix includes scaling, translation, and rotation of the image. The transformed texture map is shown in Figure 5(c). The texture in Figure 5(c) is the same as the source (Figure 5(a)); however, the feature point coordinates are identical to those of the target (Figure 5(b)). The texture map can be transformed to match landmarks with arbitrary coordinates. 2.3.2 Estimation of the skin pigment map

By applying the pigment separation method based on the independent component analysis proposed by Tsumura et al. [12] to the texture map image, the skin pigment maps of melanin and hemoglobin components, which constitute the skin color, are extracted from the texture map image. The pigment separation method assumes that melanin pigments in the epidermis and hemoglobin pigments in the dermis are spatially distributed independently to form the skin tone. A map of the skin pigment component is estimated by applying independent component analysis to the pixel values of the



(a) Melanin map(b) Hemoglobin mapFigure 6. Separation of skin pigment using Tsumura's method [12].





(a) Melanin difference (b) Hemoglobin difference map
Figure 7. Visualization of differences in skin pigment between Figure 5(a) and Figure 5(b).

skin image. Changes in the concentration of pigment components are visualized by calculating the differences in pigment concentration between the corresponding texture maps and coloring them.

In this study, the skin replicating the hemangioma was measured and analyzed, as will be described later. The acquired texture map is the same, as shown in Figure 5(b). The results of the skin pigment separation [12] in Figure 5(b) are shown in Figure 6. Figure 6 is colored using melanin and hemoglobin pigment vectors for separation. Based on the skin pigment separation method, shades were mapped in the direction of the shade vector, and the concentration distributions of melanin and hemoglobin pigment components were less affected by the unevenness of illumination. The hemangioma reproduced on the right cheek of the original image shown in Figure 2 was nearly separated as a hemoglobin pigment component by Figure 6(b) and had almost no effect on the melanin pigment component shown in Figure 6(a), although its outline was slightly visible.

2.3.3 Visualization of differences in skin pigment

The obtained concentrations of the skin pigments were subtracted in each pixel to calculate its change, and then colored with the vector of the skin color used in the separation to visualize the amount of its change. Figure 7 shows the visualization result of subtracting Figure 5(a) from Figure 5(b) after the separation of skin pigment components, and the change in skin pigment concentration. Whereas the melanin pigment component was unchanged in Figure 7(a) with the entire face region appearing white, the hemoglobin pigment component shown in Figure 7(b) shows the presence of a pigmented spot on the right cheek. The changes in skin color were visualized by the above procedure. In the actual system, this differential skin pigment map was texture mapped to a 3D shape model and displayed.

3 Experiment

In this study, we aimed to propose a system to measure the 3D shape of a face and compute and visualize the minute changes in shape and skin color in two different states. Therefore, in this chapter, faces with different shapes and skin tones were measured and analyzed to confirm the behavior of the proposed system on real subjects. Section 3.1 describes the measurement procedure of this experiment and Section 3.2 describes the measurement data and analysis results.

3.1 Experimental setup

In this study, we used the iPhone XR (Apple, U.S) as a method to measure facial shape and skin texture. The faces were not fixed, and they were measured in a realenvironment under a white LED light source (Viltrox, China). Four subjects were measured: Subject 1: female, in their 20s, Subject 2: male, in their 20s, Subject 3: male, in their 50s, and Subject 4: male, in their 50s. The subjects were all Japanese. The following manipulations were performed with the help of a physician to reproduce the modulated face shape and skin color. For the shape, two or three medical cotton balls were stuffed into the right cheek to create a puffy state of the cheek. Figure 8(a) shows how the cotton balls were packed in the mouth. For color, artificial hemoglobin was applied to reproduce hemangiomas or acne-like redness. Figure 8(b) shows the application of artificial hemoglobin. Each state and measurement procedure is explained below. State (1) is the initial state. In State (2), the right cheek is filled with a cotton ball and the shape modulation state is measured. In State (3), artificial hemoglobin is applied to the puffed skin of the right cheek, and the modulated shape and color is measured. Finally, in State (4), the cotton ball of the right cheek is removed and the color modulation state is measured. The time taken by the iPhone for each state is



(a) Cheek shape change using cotton balls





Figure 8. Changes in shape and color.

approximately 1 minute. Changing the subject's state takes approximately 5 minutes. The total time of the experiment for one participant is approximately 20 minutes.

3.2 Experimental result

The four states of the face were measured for four subjects using Bellus3D on the iPhone XR. The visualization results of the changes in the face between the four states are shown in 3D geometry. The results are described in Sections 3.2.1 and 3.2.2.

3.2.1 Results of changes in shape of the 3D model

Figure 9 visualizes the shape change in each condition for each subject. State (1) is used as the target for 3D registration, and the results are shown in Figure 9. The facial shape was modulated when changed from State (1) to (2) and from State (1) to (3). In States 1 and 2, which contain cotton, the right cheek of all subjects is colored red. Only the right cheek, where the shape has changed, is colored red to indicate the change, while the rest of the cheek is colored green. This is a correct visualization. Specifically, the 3D registration and intersection algorithms were performed correctly. A similar trend was observed in all four subjects. 3.2.2 Results of changes in the texture map of the 3D model

Figure 10 visualizes the color change in each condition for each subject. The melanin and hemoglobin components were obtained by applying the skin pigment separation method to the texture map images, and the amount



Figure 9. Results of changes in shape: (a) Subject 1, female in their 20s, (b) Subject 2, male in their 20s (c) Subject 3, male in their 50s, and (d) Subject 4, male in their 50s.

of change was visualized. In comparison with the results shown in Figure 7 in Section 2.3.2, less change in melanin was observed this time; only results for hemoglobin were observed. When comparing before and after application of artificial hemoglobin (from States 1 to 3, and 1 to 4), Figure 10(a, b, c, d-2) and Figure 10(a, b, c, b-3) confirm a difference in hemoglobin dye concentration at the applied site. Alternatively, positional misalignment occurred between texture maps when artificial hemoglobin was applied with different shapes, such as the changes in States 3 to 4 shown in Figure 10(a, b, c, d-4). These results showed a similar trend in all four subjects.

4 Discussion

In this study, we proposed a system to visualize shape and color changes in the treatment of skin diseases on the face. A 3D shape model and texture map of the face were obtained using a structured optical depth sensor on the iPhone XR. The distance of the shape modulation was calculated by performing the 3D registration and intersection algorithm for the 3D shapes of multiple faces. The shape modulation was successfully visualized by coloring the mesh according to the distance. In addition, it was possible to compare the texture maps of each pixel by transforming the texture maps so that the (u, v) coordinates were the same.

First, we discuss the visualization results of the change in the shape. Figure 9 shows that the entire face was colored green, which means there was no change in shape, and only the right cheek area, where the shape was modulated, was colored red as a change less than 5 mm. The red-colored area corresponds to the area where the cotton ball was inserted. The shape change was visualized as predicted.

Next, we discuss the visualization results of the change in skin color. When comparing before and after color modulation, we found that the modulation was visualized in the right cheek, mainly for the hemoglobin pigment component, as shown in Figure 6 and Figure 10. However, as the change from the State (4) to (3) shown in Figure 10(a, b, c, d-4), when the shape is modulated between states with pigmented spots, it is confirmed that the pigmented spots do not completely overlap when the texture map is transformed, resulting in a slight deviation during subtraction. The proposed transformation of the texture map using the landmarks is only a 2D image processing of the texture map, which represents the 3D view in two dimensions. In other words, the transformation is not strictly based on the facial 3D shape. Therefore, a transformation based on the shape modulation of the acquired 3D shape model is necessary to transform with more strict shape considerations. The separation was not perfect for the melanin pigment component in Figure 6.

This result relates to the skin structure of the face. The skin has a complex layered structure comprising translucent media and can be roughly divided into three layers: the epidermis, dermis, and subcutaneous tissue. The epidermis contains melanin pigments that protect the skin from ultraviolet rays, and the dermis contains a large amount of hemoglobin in capillaries. The skin pigment separation adopted in this paper was an estimation assuming the skin is a two-layered model of the dermis and epidermis [12]. In this study, however, we focused on conditions that are exposed on the skin surface, such as hemangiomas. The results in Fig. 6 were poor because we dealt with conditions that were beyond the two-layer model of the skin.

This transformation made it possible to compute change s in skin pigment concentration and visualize skin color mo dulation. The results were similar when performed on the fo ur subjects. The changes in face shape and color were easily

analyzed using the depth camera of a smartphone. The sha pe and color information of the face has been obtained in pl astic and aesthetic surgery using expensive measuring instr uments or through manual measurement with low



(a-1) From State (1) to (2)



(**b-1**) From State (1) to (2)



(c-1) From State (1) to (2)



(**d-1**) From State (1) to (2)



(a-2) From State (1) to (3)



(**b-2**) From State (1) to (3)



(**c-2**) From State (1) to (3)



(**d-2**) From State (1) to (3)



(**a-3**) From State (1) to (4)



(**b-3**) From State (1) to (4)



(**c-3**) From State (1) to (4)



(**d-3**) From State (1) to (4)



(a-4) From State (4) to (3)



(**b-4**) From State (4) to (3)



(c-4) From State (4) to (3)



(**d-4**) From State (4) to (3)

Figure 10. Results of changes in hemoglobin: (a) female in their 20s, (b) male in their 20s, (c) male in their 50s, (d) male in their 50s.

reproducibility. However, it may become easier to measure the face in three dimensions using a measurement system like the one described above. In our method, the shape and color of the face are accurately obtained using a smartphone. This method is expected to be used in various medical fields and in the establishment of new telemedicine applications, such as home follow-ups. Sharing the analysis results between the physician and patient with devices that many people already have clarify the lesions to be corrected and their solutions, and will lead to strengthened mutual trust in the treatment on going.

However, there are many problems in terms of accuracy for practical use as a medical device. We have not been able to verify the accuracy of the acquired shapes. In addition, the participants maintained neutral facial expressions in the experiment, whereas it is predicted that the measurement accuracy will decrease if there is a change in facial expression during measurement. The variability in the facial expression of the subject during the recording may be reduced by the subject wearing his /her dental mouthpiece.

It is thus necessary to verify the accuracy of the 3D shape measured by the iPhone XR by comparing it with the data obtained using a 3D scanner with guaranteed accuracy.

Additionally, we only captured Japanese people with no appreciable differences in skin color, and we need to capture a greater number of participants with various facial colors. The monitoring system can be used to distinguish whether the color change is due to melanin or hemoglobin.

Color correction was not considered because the measurements were made in the same environment under a white light source. It is predicted that the measurement accuracy will decrease if the lighting environment varies. It

is necessary to measure the face with a color chart under various light sources, evaluate the color reproducibility of each light source, and perform color correction.

For the transformation of texture maps, the use of meshto-mesh correspondences obtained by the intersection algorithm was considered for achieving more accurate 3D geometry. The purpose of this study was to visualize the skin color change in the texture map. In the future, we will consider using different skin models for the regions of interest to cover specific skin structures such as hemangiomas. We aim to prove that it is possible to use this method in clinical practice.

It is difficult to evaluate deep anatomy such as bones and muscles only by optical photography. In addition to this optical evaluation, X-rays and ultrasonography are also required to evaluate deep tissue. However, these duplicate tests do not match our proposed simple, cheap, and reproducible assessment. Of course, when bone evaluation is required, X-rays are taken, and osteotomy will be planned if it is necessary. The necessary tests and evaluations will be left to the discretion of an experienced physician. Our proposed method is one of those tests.

5 Conclusion

In this study, the face shape and skin color were threedimensionally measured using the infrared depth sensor and RGB sensor built into a smartphone. We measured before and after modulations of the shape and color of the face, and calculated the change in volume. The change in pigment of the skin color was calculated and visualized. This method makes it possible to analyze the shape and skin color independent of the viewing angle and direction of the illumination light. The depth camera built into the smartphone can be used to monitor changes in facial shape and skin color. To contribute to the development of telemedicine, in which the patient measures their face at home and receives medical treatment consultations remotely, we need to confirm the accuracy and reproducibility of our system using a 3D scanner for 3D shape and skin color.

ACKNOWLEDGMENTS

We thank Ashleigh Cooper, Phd, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

REFERENCES

[1] Krowchuk, Daniel P., et al. (2019). Clinical practice guideline for the management of infantile hemangiomas. Pediatrics 143.1: e20183475.

[2] Canfield Scientific, Inc. (accessed 27th March 2020), VECTRA H1 3D Imaging System, https://www.canfieldsci.com/imaging-systems/vectra-h1-3d-imaging-system/.

[3] Camison, Liliana, et al. (2018). Validation of the Vectra H1 portable three dimensional photogrammetry system for

facial imaging. International journal of oral and maxillofacial surgery 47.3: 403-410.

[4] Seitz, Steven M., et al. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06). Vol. 1. IEEE.

[5] Kenta Masui, Kaoru Kiyomitsu, Keiko Ogawa-Ochiai, et al. (2020). Technology for visualizing the local change in shape of edema using a depth camera. Artificial Life and Robotics. 24. 10.1007/s10015-019-00541-1.

[6] Apple Inc., (accessed 27th March 2020). iPhone, https://www.apple.com/iphone/.

[7] Bellus3D, Inc., (accessed 27th March 2020). "Bellus 3D:High-quality 3D face scanning", https://www.bellus3d.com/,.

[8] Besl, Paul J., and Neil D. McKay. (1992). Method for registration of 3-D shapes. Sensor fusion IV: control paradigms and data structures. Vol. 1611. International Society for Optics and Photonics.

[9] Fischler, M. A., and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6), 381-395.

[10] Moller, Tomas, and Ben Trumbore. (1997). Fast, minimum storage ray- " triangle intersection., Journal of graphics tools, 2.1 (1997): 21-28.

[11] Mukaida, Shigeru, et al. (2002). Facial image synthesis system: FUTON evaluation as tools for cognitive research on face processing. Trans. IEICE 85.10: 1126-1137.

[12] Norimichi Tsumura, et al. (2003). Image-based skin color and texture analysis / synthesis by extracting hemoglobin and melanin information in the skin. ACM Transactions on Graphics (TOG) 22.3: 770-779.