Robotics Meets Cosmetic Dermatology: Development of a Novel Vision-Guided System for Skin Photo-Rejuvenation

Muhammad Muddassir, *Student Member, IEEE*, Domingo Gómez Domínguez, Luyin Hu, Shujian Chen, and David Navarro-Alarcon, *Senior Member, IEEE*

***Abstract*—In this article, we present a novel robotic system for skin photo-rejuvenation, which uniformly delivers laser thermal stimulation to a subject's facial skin tissue. The robotized treatment is performed by a customized manipulator instrumented with a cosmetic laser generator, a depth sensor, and a thermal camera. To plan the optimal trajectory of the manipulator, our system computes the surface model of the face and segments it into seven regions, then automatically fills these regions with uniform laser irradiation. Finally, we report the detailed results of experiments on multiple human subjects to validate the performance of the system. To the best of the author's knowledge, this is the first time that facial skin rejuvenation has been automated by robot manipulators.**

***Index Terms*—Cosmetic dermatology robots, face models, skin photo-rejuvenation, vision-guided manipulation.**

I. Introduction

THERE are two main ageing processes that affect a person's skin condition [1]. Ageing due to the biological clock, which is currently irreversible, and photo-ageing, which results from exposure to ultraviolet radiation coming from the sun; the latter is widely considered treatable to some extent [2], [3]. With the aim of "reversing" skin damage, in the past decades, people have turned to the so-called *beauty clinics* for receiving various types of noninvasive dermatological treatments. These treatments are typically performed by cosmetic instruments based on laser light [4], intense pulsed light [5], radio-frequency [6], etc. Worldwide, the beauty industry has seen an exponential increase in the demand for aesthetic skin rejuvenation treatments.

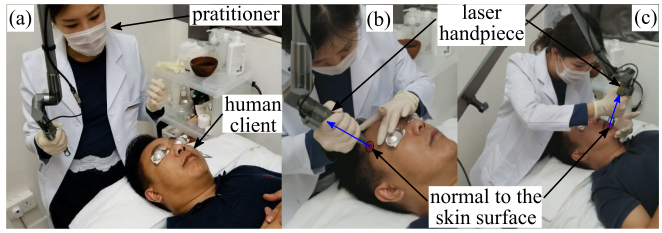


Fig. 1. Conventional skin photo-rejuvenation treatment: (a) examination of the facial skin condition, (b) manipulation of the handpiece over the forehead, and (c) manipulation of the handpiece over the left jaw.

A typical skin photo-rejuvenation treatment conducted in these beauty clinics is shown in Fig. 1, where a dermatologist (or "skin technician" [7]) visually examines the client's skin condition to determine the type of treatment to be performed and the appropriate laser light parameters [8]. Then, a dermatologist manipulates the laser handpiece in repetitive motions over different areas of the face to stimulate the skin tissue. Note that the targeted skin tissue must be stimulated with the *exact amount* of energy and time to produce the expected results without causing damage [9], which is a challenging task for the human operators (dermatologists, practitioners, or skin technicians). In addition, the complete rejuvenation treatment lasts around 25 min—It is a tedious and tiring task for a professional who may perform it several times in a single day, which results in the existing high turnover rate of experienced professionals in the industry [10]. These issues clearly show the need to develop robotic systems that can automate the manipulation of instruments.

Robotics in cosmetic dermatology is currently an underexplored field compared to medical robots [11]–[13]. Thus, it has the potential to present further interesting challenges and opportunities. Only few commercial robotic systems are available for cosmetic and aesthetic dermatological treatments [14]. One of the commercialized systems is ARTAS Follicular Unit Extraction (FUE) robot that can remove healthy follicles from a donor and autonomously transplant them onto the recipient's

Manuscript received October 27, 2020; revised January 14, 2021 and March 5, 2021; accepted April 12, 2021. Date of publication April 23, 2021; date of current version April 18, 2022. Recommended by Technical Editor H. (TE) Wang and Senior Editor H. (SE) Qiao. This work was supported in part by the Research Grants Council under Grant 14203917, in part by PROCORE France/Hong Kong Joint Research Scheme sponsored by the RGC and the Consulate General of France in Hong Kong under Grant F-PolyU503/18, in part by the Jiangsu Industrial Technology Research Institute Collaborative Research Program Scheme, in part by RODS Tech Ltd., and in part by PolyU under Grants 252047/18E and ZZHJ. *(Corresponding author: David Navarro-Alarcon.)*

The authors are with the Department of Mechanical Engineering, The Hong Kong Polytechnic University (PolyU), Kowloon, Hong KongD. Navarro-Alarcon is also with the Research Institute for Smart Ageing (RISA) of PolyU. (e-mail: 18074096r@connect.polyu.hk; correodgd@gmail.com; 19044457r@connect.polyu.hk; shujian.chen@connect.polyu.hk; dna@ieee.org).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TMECH.2021.3075207.

Digital Object Identifier 10.1109/TMECH.2021.3075207

scalp [15]. Another example is reported in [16], where a hair removal dermatology system was developed; This system can automatically conduct laser hair removal by using a manipulator while activating a laser instrument over regions defined by a supervising dermatologist [17]. However, none of these robotic systems was specifically designed for facial skin rejuvenation treatments.

To provide a feasible solution to the open problem, in this article, we present an innovative system capable of autonomously performing skin photo-rejuvenation treatment on a human face. The developed system is composed of a 6-DOF robot manipulator with a custom-built end-effector that carries a laser cosmetic instrument. This system is equipped with an RGB-D camera to reconstruct the facial geometric model and three distance sensors to monitor the surroundings of the end-effector. To the best of the author's knowledge, this is the first time that a robotic approach for facial skin rejuvenation has been reported in the literature. The original contributions are as follows:

1) Development of a specialized mechanical prototype for cosmetic treatments.
2) Design of a new sensor-based method for controlling the thermal stimulation of skin tissues.
3) Experimental validation of the developed robotic system.
4) Perform the photo-rejuvenation treatment with proposed robotics system on multiple human subjects.

The rest of this article is organized as follows: Section II presents the proposed prototype; Section III describes the sensing system; Section IV introduces the control algorithms; Section V reports the experiments; Section VI concludes this article.

II. Robotising Skin Photo-Rejuvenation

A. Common Practice

The operational principle of skin photo-rejuvenation treatment is the thermal stimulation of collagen in the skin; laser energy absorbed by the chromophore in the skin tissue causes a gradual temperature increase in the irradiated region [2]. Fig. 1(a) depicts the conventional setup to manually perform skin photo-rejuvenation at beauty clinics. The treatment starts with a close examination of the client's skin by a dermatologist (or a practitioner), then he/she empirically sets the parameters of the laser generator based on the skin tone and condition. These parameters include laser diameter, laser energy, and fluence (laser energy per centimetre square J/cm^2) [18]. Fig. 1(b) and (c) shows how the laser handpiece is manipulated in a normal direction to the skin surface. To operate on a particular face region, the dermatologist typically follows an S-shaped path and uses a foot pedal to turn the laser ON and OFF. Throughout this manuscript, we refer to the instance of delivering the laser energy to the skin surface as a "laser shot."

During the skin rejuvenation treatment, the dermatologist should monitor the traces of the laser shots to avoid overlappings and gaps between each laser shot (laser irradiation). If the laser shots are not uniformly distributed then the overlapping can cause serious side effects (i.e., erythema, hyperpigmentation, and crusts, whereas gaps can lead to the suboptimal treatment of the skin. In this study, we use a Q-switched Nd: YAG (1064 nm) laser, with a pulse duration of less than 10 ns, and an adjustable repetition rate of 1–10 Hz. The wavelength of 1064 nm lies within the infrared spectrum (invisible to human eyes). The laser generator used in this study also fires a low energy flash of visible light with each laser shot for the convenience of the operator. Even with this visual aids, an experienced operator can easily lose track of the degree of thermal stimulation delivered to the skin.

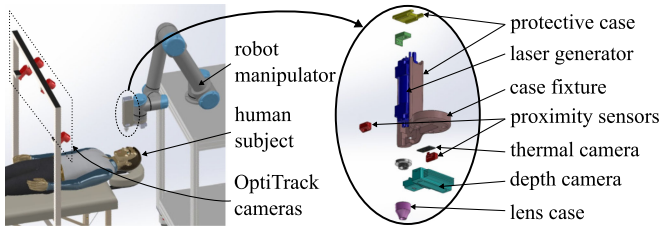


Fig. 2. (on left) Proposed setup. Robotised facial skin photo-rejuvenation system. (on right) Exploded view of the customised end-effector.

B. Proposed Setup

Fig. 2 conceptually illustrates the proposed robotic system and its components. The robot manipulator is placed on top of a wheeled platform that provides mobility to the system, which is an essential feature for an indoor environment like beauty clinics. A custom-made rejuvenation end-effector is also developed for the robotic system. The purpose of this end-effector is to carry all the sensors and laser handpiece during the treatment. This end-effector was designed considering the following requirements:

1) The moving end-effector should have proximity-like behavior to avoid any possible collision with the human face during the treatment.
2) The structure should be rigid enough to hold the laser generator filament, vision, and distance sensors during manipulation.
3) The laser should be able to stimulate any point over the surface of the face.
4) The structure should be compact.

To fulfil the above requirements, a 3-D printed rectangle shell-like structure is designed to house the laser generator and the sensors, as shown in Fig. 2. An extension is designed beside the shell, which has screw interfaces to firmly attach the instrumented customised end-effector to the robot manipulator. The small extended structures around the shell are the housings of TOF10120 distance sensors to continuously monitor the closeness of the end-effector to the obstacle in the environment, as in Fig. 2. The proposed prototype utilizes an Orbbec Astra Mini S depth camera for the facial model reconstruction task. For the sake of simplicity, we denote the custom-made end-effector for the photo-rejuvenation treatment by an end-effector of a robot manipulator throughout this manuscript.

To manipulate the laser generator over the facial skin cosmetic treatment, a UR5 robot manipulator from Universal Robots is used. The control box of the robot is placed inside the mobile platform; a Linux PC communicates with the robot's servo

controller via a TCP/IP socket. All the proposed algorithms run in the Linux PC with ROS [19]. The complete system can be divided into two main subsystems: a sensing system and a control system. Both systems are interconnected with each other and are responsible for dedicated tasks. To control the motion of the robot manipulator, we consider kinematic control instead of learning-based control like [20]–[23], as manipulation of the laser instrument is a delicate task and the error of a few millimetres in motion can cause undesirable treatment results (or even damage to skin tissues). Thus, the complete system is tightly calibrated to minimize the laser positioning errors during operation. In this article, the coordinate frame of the robot base is the inertial frame for the developed system.

III. Sensing System

A. Facial Geometric Model Computation

The facial surface model reconstruction starts with a search of a human face in an RGB image of the scene. A face detector by [24] provides a bounding box around a human face. Then, the system enforces the facial landmark detector by [25] to search the landmarks only inside the bounding box to locate the position of the key facial landmarks (i.e., eyes, eyebrows, nose, lips, and face boundary) from the captured RGB image. To compute the pose of the detected face, the system requires the position and the orientation of the face in the camera coordinate frame. The position of the detected face is assumed to be equivalent to ${}^{C}\vec{r}_F$ of a 3×1 vector. Let $\vec{d}_l$ and $\vec{d}_r$ denote the position vectors of 3×1 that define the position of the left and right eye of the detected face in a camera coordinate. Then, the position vector of the detected face will be ${}^{C}\vec{r}_F = (\vec{d}_l + \vec{d}_r)/2$. The rotation matrix of a coordinate frame consists of three orthonormal column vectors of 3×1, i.e. $R = [\vec{\alpha}\ \vec{\beta}\ \vec{\gamma}] \in SO(3)$. The rotation matrix ${}^{C}R_F$ that defines the orientation of the face in the camera coordinate frame can be computed by obtaining three orthonormal vectors as follows:

$$ {}^{C}R_F = \begin{bmatrix} \vec{\alpha}_F & \vec{\beta}_F & \vec{\gamma}_F \end{bmatrix} \tag{1} $$

where

$$ \vec{\alpha}_F = \frac{\vec{d}_r - {}^{C}\vec{r}_F}{\|\vec{d}_r - {}^{C}\vec{r}_F\|}, \vec{\beta}_F = \frac{[0\ 0\ 1]^T \times \vec{\alpha}_F}{\|[0\ 0\ 1]^T \times \vec{\alpha}_F\|}, $$

and

$$ \vec{\gamma}_F = \frac{\vec{\alpha}_F \times \vec{\beta}_F}{\|\vec{\alpha}_F \times \vec{\beta}_F\|} $$

The reason to equate $\vec{\beta}_F$ with the cross product of $\vec{\alpha}_F$ and $[0\ 0\ 1]^T$ is to constrained $\vec{\beta}_F$. Otherwise, $\vec{\beta}_F$ can be defined by infinite many normal vectors on the plane perpendicular to the $\vec{\alpha}_F$. Whereas $[0\ 0\ 1]^T$ depicts the last column of the rotation matrix or $\vec{\gamma}$ of the observing coordinates frame. Now, the pose of the face in the camera coordinate is

$$ {}^{C}T_F = \begin{bmatrix} {}^{C}R_F & {}^{C}\vec{r}_F \\ 0 & 1 \end{bmatrix} \tag{2} $$

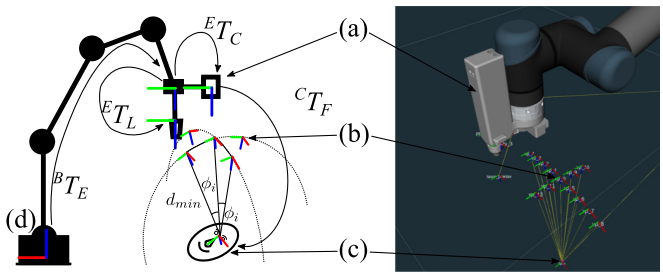


Fig. 3. Estimating the viewpoints. (a) depth camera, (b) estimated viewpoints ${}^{B}T_{v_i}$, (c) pose of the human face ${}^{B}T_F$ and (d) coordinate frame of the robot base (inertial frame of reference).

where the ${}^{C}R_F$ is a 3×3 rotation matrix and ${}^{C}\vec{r}_F$ is a 3×1 translation vector. After estimating the pose of the face in the camera coordinate, it is transformed to the robot's base coordinate using, ${}^{B}T_F = {}^{B}T_E {}^{E}T_C {}^{C}T_F$. Here, ${}^{B}T_F$ defines the pose of the face in the robot's base coordinate, ${}^{B}T_E$ is the homogeneous transformation from the robot's base to the end-effector. ${}^{C}T_F$ defines the pose of the face in the camera coordinates.

The commercial depth cameras have a limited field of view, thus the depth data acquired from one view may contain inconsistencies (holes). To reconstruct the consistent (smooth) facial model, the sensing system estimates N_v number of viewpoints around the detected face to acquire depth data from different viewpoints [N_v is empirically chosen as 13. 1 front view. 4, 4 left and right-hand side. 2, 2 up and down)]. Capturing images from different viewpoints avoids the possible occlusion and provides a detailed surface of a human face. These viewpoints are estimated with the increments and decrements of a predefined angle ϕ_i around the pose of the detected face ${}^{B}T_F$, and the angle ϕ_i will be same for both latitudinal (around x-axis of the ${}^{B}T_F$) and longitudinal viewpoints (around y-axis of the ${}^{B}T_F$), as shown in Fig. 3. It is assumed that the longitudinal viewpoints lies on the circles at $y = 0$ in the zx-plane, which can be estimated using the parametric equation of a circle, as ${}^{F}R_{v_i} = R_y(\phi_i)$ and ${}^{F}\vec{r}_{v_i} = [-d_{\min}\sin(\phi_i), 0, -d_{\min}\cos(\phi_i)]$, resulting a homogeneous transformation matrix ${}^{F}T_{v_i}$ of 4×4. In the inertial frame (the robot base), the transformation matrix will be ${}^{B}T_{v_i} = {}^{B}T_F {}^{F}T_{v_i}$, where $d_{\min}$ in the Fig. 3 defines the minimum range of the depth camera. Similarly, the latitudinal viewpoints are estimated around x-axis on the yz-plane, by ${}^{F}R_{v_i} = R_x(\phi_i)$ and ${}^{F}\vec{r}_{v_i} = [0, d_{\min}\sin(\phi_i), -d_{\min}\cos(\phi_i)]^T$ and in the inertial frame it will be ${}^{B}T_{v_i} = {}^{B}T_F {}^{F}T_{v_i}$. The estimated viewpoints around the face are shown in Fig. 3.

To acquire the visual data from the estimated viewpoints, the control system commands the end-effector to visit each estimated viewpoint ${}^{B}T_{v_i}$. After reaching a viewpoint, the sensing system captures the visual data and extracts only the region that contains a detected face from the captured RGB and depth images. Then, these regions from both images are converted into point cloud form. Each point cloud has an offset transformation from their origin to the point clouds captured from the other viewpoints. To align all the point clouds, the relative pose between each viewpoint should be known. As the control system

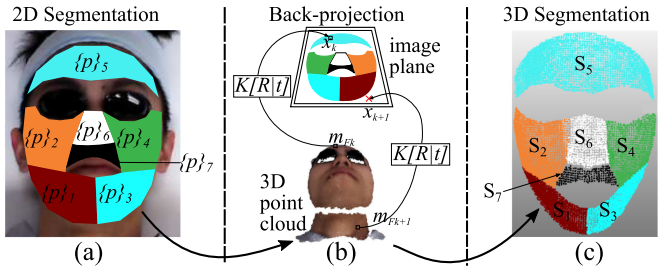


Fig. 4. Segmentation of a facial model using the 2-D facial landmarks. (a) the polygons $\{p\}_j$ defines each region, (b) back-project each 3-D point to 2-D image plane and (c) segments M_F point cloud into S_j segments, where $j = 1, 2, \ldots 7$.

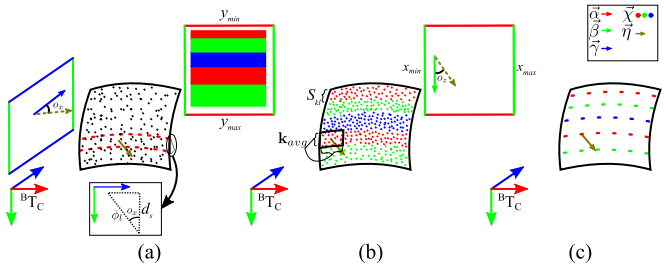


Fig. 5. Path Planning Algorithm. (a) partitioning of a point cloud along y-axis. (b) sweeping of a patch $\mathbf{k}_{\text{avg}}$ from x_{min} to x_{max}. (c) averaging of the points inside the patch $\mathbf{k}_{\text{avg}}$. ϕ_l, laser diameter. d_s, laser shot separation distance. d_p, interpath distance.

receives continuous feedback of the end-effector position from the robot manipulator, the relative pose between each viewpoint can be computed as ${}^{v_i}T_{v_{i+1}} = ({}^{B}T_{v_i})^{-1}\,{}^{B}T_{v_{i+1}}$. Additionally, the Point-to-Plane ICP [26] followed by CICP [27] is applied to the point clouds captured at each viewpoint. To remove the noisy data points, the voxel grid downsampling algorithm, with a voxel leaf length of 2 mm, is applied on the final facial model. This downsampling algorithm keeps the density consistent and removes the holes from the final facial model. Each point in a point cloud is a structure of three vectors, denoted by $\{\vec{x}, \vec{n}, \vec{c}\}$, where $\vec{x} = [x\ y\ z]^T$ is the position vector, $\vec{n} = [n_x\ n_y\ n_z]^T$ represents the normal vector and $\vec{c} = [r\ g\ b]^T$ defines the color.

B. Automatic Face Segmentation

The adipose layer beneath the facial skin is uneven throughout the face, e.g., the cheek region has a thicker fat layer than the forehead region. Therefore, dividing each region into segments can aid to set the appropriate laser light energy level depending on the region to be stimulated. Furthermore, the classification of different parts of a face enables to compute the trajectory separately for the laser instrument; note that if such geometric constraints are not considered within the trajectory generation, the manipulated instrument can potentially collide with protruding areas of the face (e.g., the nose). To cope with these issues, an algorithm to automatically segment the human facial model into seven regions is also proposed, as in Fig. 4. The classification of each region of a face is called "segmented region" in this manuscript, which is comprized of the left cheek, left jaw, right cheek, right jaw, forehead, nose, and upper lips.

The computed facial model is in the form of a point cloud of unorganized 3-D data (i.e., the neighboring points in space are not necessarily adjacent in computer memory indexing, and vice versa). To plan the path for the robot manipulator, the control system should be aware of the correspondence between each point in the model and each segment of the face. As a solution to this problem, we use 2-D facial landmarks to cluster the unstructured data into seven segments (point clouds), see Fig. 4. Let S_j (for $j = 1, \ldots, 7$) denotes the jth region of the facial model and s_i^j is an arbitrary ith point in S_j.

Initially, the autosegmentation algorithm detects the facial landmarks using [24], [25], from the RGB image captured from the first viewpoint ${}^{B}T_{v_0}$, which provides the most reliable view to detect the facial landmarks (since no facial feature will be occluded from the front view). Once these key landmarks are detected, the algorithm draws polygons $\{p\}_j$ on the image plane to define the seven regions on the image plane, as shown in Fig. 4(a). Then, a sorting routine runs seven parallel processes to back-project each 3-D point m_{F_k} of the facial model M_F on the image plane, by using the camera's perspective projection relation:

$$\begin{bmatrix} \vec{\xi} \\ 1 \end{bmatrix} = K[R|t] \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix} \tag{3}$$

where $\vec{\xi}$ is a 2×1 column vector that denotes the projection of a 3-D point $\vec{x}$ on the image plane, K is the 3×3 matrix of intrinsic parameters, and R, t is the extrinsic parameters (the 3×3 rotation matrix and 3×1 translation vector) of the camera, which are calculated by standard camera calibration algorithms. In Fig. 4(b), the algorithm tests whether a given 3-D point lies in the polygon $\{p\}_j$ by using the ray-casting method reported in [28]. Fig. 4(c) depicts the segmented 3-D facial model (in the form of a point cloud).

IV. Control System

A. Path Planning

To plan the paths for each segment, a sampling-based path planning algorithm is devised. This path planning algorithm takes the point cloud of a surface as an input and outputs a set of vertical or horizontal paths (in the form of a point cloud). These paths are objected to filling the input surface with laser shots (according to the diameter of the laser ϕ_l). The path can be planned vertically or horizontally depending on the longest side of the operating segment. For example, the forehead is a wider segment where horizontal paths can decrease the number of lane changes for the robot manipulator. Similarly, the nose region has a longer length than width, in this case, vertical paths are more appropriate.

Fig. 5 illustrates the path planning algorithm in three steps. Initially, the algorithm bins all the points inside the range of $[y_{\text{min}}, y_{\text{min}} + \phi_l)$ into a container, known as strip S_{jk}, where k is the strip index and j the segment index. In Fig. 5(a), y_{max} and y_{min} are the maximum and minimum values along the y-axis among all the points in a segment. When the surface is facing toward

the camera and the z-axis of the camera and the normal inward to the surface are parallel, the range $[y_{\min}, y_{\min} + \phi_l)$ provides equal width strips to the laser diameter. However, if the surface is in line with the y-axis of ${}^{C}T_{C}$, the binning of points to a strip in the range $[y_{\min}, y_{\min} + \phi_l)$ will yield the strip's width greater than the laser diameter ϕ_l, which ultimately introduce gaps among the planned paths, if the surface is curved or not facing the camera. As only two path planning (horizontal or vertical) approaches are adopted, rotation of the strip around x- or y-axis also needs to be compensated. In Fig. 5(a), o_x represents the rotation of a strip around the x-axis, which is computed by a dot product of the projection of averaged normal vector in S_{jk} and z-axis of the camera frame ${}^{B}T_{C}$ on its yz-plane as follows:

$$o_x = \cos^{-1}\left(\frac{(\vec{\eta}_s - (\vec{\eta}_s.\vec{\gamma}_C)\vec{\gamma}_C).\ \vec{\eta}_{s,yz}}{\|\vec{\eta}_s - ((\vec{\eta}_s.\vec{\gamma}_C)\vec{\gamma}_C)\|\ \|\vec{\eta}_{s,yz}\|\|}\right) \tag{4}$$

where $\vec{\eta}_{s,yz}$ denotes the projected vector of averaged normal vector $\vec{\eta}_s$ of a strip on the yz-plane of ${}^{B}T_{C}$ and $\vec{\eta}_{s,yz} = \vec{\gamma}_C \times \vec{\eta}_s$. Whereas, the $\vec{\gamma}_c$ points toward the z-axis of ${}^{B}T_{C}$. Now the incremental step $d_s = \phi_l / \cos(o_x)$ will be adjusted according to o_x and the strip S_{kl} will be resorted with the new range $[y_{\min}, y_{\min} + d_s)$. All the remaining points in a segment will be binned accordingly. The maximum range of the previous strip will become the minimum range of the next strip as well as the maximum range of next strip will be incremented with d_s.

The width of each strip S_{jk} is equal to the diameter of laser shots ϕ_l, so a line passing through the center of each strip can be considered as an optimal path for the robot end-effector (no intralane overlapping of laser shots). This path can be computed using a polynomial fit, but this method can not guarantee that the fitted polynomial will pass from the center of a strip of points. Furthermore, the points in each strip are sparsely distributed, which can also bias the polynomial fit. Then, the fitted polynomial could generate a path deviated from the center of the strip. That is why sweeping-averaging patch method is devised to ensure the planned paths always pass through the center of each strip.

Once the strip S_{jk} is partitioned from a segment of the point cloud, a patch $\mathbf{k}_{\text{avg}}$ (like a rectangular stencil) of width ϕ_l and height d_s is placed on one side of the strip, shown in Fig. 5(b). The patch $\mathbf{k}_{\text{avg}}$ is a buffer of points, which populates the points after every incremental/decremental step and computes the average of all enclosed points inside it. The size of the step is equal to the diameter of the laser ϕ_l. Mathematically, $\mathbf{k}_{\text{avg}}$ can also be defined as a set of points, $\mathbf{k}_{\text{avg}} = \{s_1, \ldots, s_i\}$. where $s_i = \{\vec{x}_i, \vec{n}_i, \vec{c}_i\}$ and $\vec{x}_i = [x_i\ \ y_i\ \ z_i]^T$. $s_i \in \mathbf{k}_{\text{avg}}$ iff $(x_{\mathbf{k}_{\text{avg}}} - d_s/2) < x_i < (x_{\mathbf{k}_{\text{avg}}} + d_s/2)$, where $x_{\mathbf{k}_{\text{avg}}}$ is the position of the patch $\mathbf{k}_{\text{avg}}$ in the strip. In Fig. 5(b), a patch $\mathbf{k}_{\text{avg}}$, in second strip from the top, is sweeping from left to right with d_s incremental step. The position of the patch $\mathbf{k}_{\text{avg}}$ in a strip is defined as $x_{\mathbf{k}_{\text{avg}}} = x_{\min} + ad_s, \quad a = 0, 1, 2 \ldots, L$. where L is the length of the strip and is defined as $L = \text{int}((x_{\min} - x_{\max})/d_s)$. $x_{\min}$ and $x_{\max}$ are the minimum and maximum values along x-axis of a strip S_{jk}.

Now, the P_j are the paths for a segment S_j and a path point p_i^j is a structure of vector defined as, $p_i^j = \{\vec{\chi}_i^j, \vec{\eta}_i^j\}$, where $\vec{\chi}$ and $\vec{\eta}$ are position and normal vectors of 3×1. These two entities are obtained by averaging the position and the normal vectors of the points inside a patch $\mathbf{k}_{\text{avg}}$, as in Fig 5(c)

$$\vec{\chi} = \frac{1}{N_k}\Sigma_i^{N_k}\vec{x}_i, \qquad \vec{\eta} = \frac{1}{N_k}\Sigma_i^{N_k}\vec{n}_i \Big/ \left\|\frac{1}{N_k}\Sigma_i^{N_k}\vec{n}_i\right\| \tag{5}$$

where N_k is the number of points inside the patch. The patch $\mathbf{k}_{\text{avg}}$ moves from the $x_{\min}$ to $x_{\max}$ in the first strip and then returns from $x_{\max}$ to $x_{\min}$. The back and forth motion of the patch in the adjacent strips enforces the algorithm to generate an S-shaped path for a segment S_j. Fig. 6 illustrates the output of the proposed scanning and path planning algorithms.

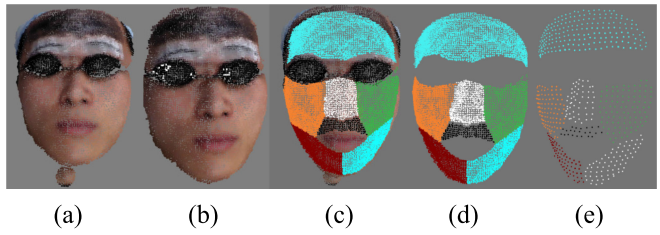


Fig. 6. Facial reconstruction to path planning. (a) raw facial model, obtained from the scanning (see Section III-A). (b) postprocessed facial model, unnecessary point has been discarded. (c) and (d) 3-D segmentation (see Section III-B). (e) planned path for each region (see Section IV-A).

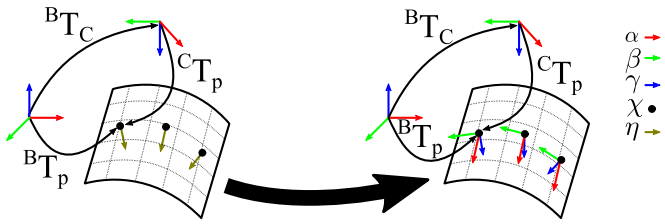


Fig. 7. Calculation of three column vectors $\vec{\alpha}_p$, $\vec{\beta}_p$, and $\vec{\gamma}_p$ of a rotation matrix ${}^{B}R_p$ from a surface normal vectors $\vec{\eta}$.

B. Robot Control Framework

Each point in a planned path is a structure of two vectors $\vec{\chi}$ and $\vec{\eta}$. In this dermatological treatment, the laser light source should project the laser light perpendicularly to the skin surface at the instance of delivering laser light energy. This can maximize the energy transferred to the skin in each laser shot. $\vec{\chi}$ and $\vec{\eta}$ are converted into a position command for the robot manipulator, where $\vec{\chi}$ directly defines the desired position of the robot end-effector. To define the orientation of the robot end-effector, first, the rotation matrix needs to be derived from the normal vector. As a rotation matrix is composed of three column vectors and each vector is orthonormal to others, i.e., ${}^{B}R_p = [\vec{\alpha}_p, \vec{\beta}_p, \vec{\gamma}_p]^T$. Let us consider that $\vec{\gamma}_p$ and $\vec{\eta}$ are parallel normal vectors. Then, the cross product of the second column of the rotation matrix of robot base $\vec{\beta}_B = [0\ 1\ 0]^T$ and the third column $\vec{\gamma}_p$ of desired rotation matrix ${}^{B}R_p$ at a path point gives $\vec{\alpha}_p = \vec{\beta}_B \times \vec{\gamma}_p$. Note $\vec{\beta}_p$ can be computed as $\vec{\beta}_p = \vec{\gamma}_p \times \vec{\alpha}_p$. Fig. 7 provides a better

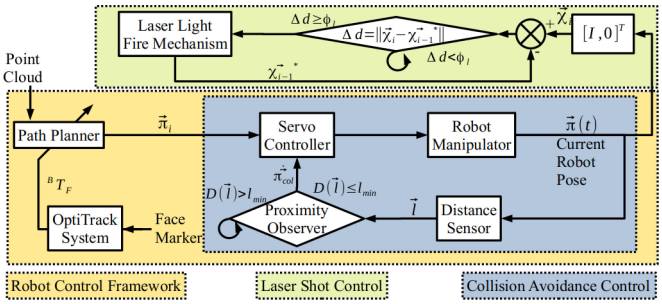


Fig. 8. Three proposed control systems to guarantee a safe and accurate robotic photo-rejuvenation.

illustration of the normal vectors calculation of rotation matrix from a surface normal vector.

The robot manipulator used in the proposed prototype can only recognize the rotation in the axis-angle representation. To convert a rotation matrix in axis-angle Suppose an arbitrary vector $\vec{u}$ is parallel to the rotation axis of a rigid body rotation, which is defined by a rotation matrix ${}^{B}R_{p}$ of 3×3

$$ {}^{B}R_{p} = \left[\vec{\alpha}_p \; \vec{\beta}_p \; \vec{\gamma}_p\right] = \begin{bmatrix} \alpha_{px} & \beta_{px} & \gamma_{px} \\ \alpha_{py} & \beta_{py} & \gamma_{py} \\ \alpha_{pz} & \beta_{pz} & \gamma_{pz} \end{bmatrix}. \tag{6} $$

Then, the vector $\vec{u}$ is calculated as [29]

$$ \vec{u} = \begin{bmatrix} \beta_{pz} - \gamma_{py} & \gamma_{px} - \alpha_{pz} & \alpha_{py} - \beta_{px} \end{bmatrix}^{T}. \tag{7} $$

The magnitude of $\vec{u}$ is equal to $||\vec{u}|| = 2\,sin\,\theta$, where θ is the angle to which a rigid body rotates around $\vec{u}$ [29], where θ is computed as follows:

$$ |\theta| = \cos^{-1}\left(\frac{\mathrm{Tr}({}^{B}R_{p}) - 1}{2}\right). \tag{8} $$

Now the representation of the rotation of a rigid body in axis angle is $\nu = \theta\vec{u}$. To define the pose of the robot end-effector for each path point, the pose vector will be $\vec{\pi}_i = [\vec{\chi}_i^T \quad \vec{\nu}_i^T]^T$, where $\vec{\pi}_i$ is a pose vector of 6×1 for each path point p_i. Fig. 8 shows all control systems and their interconnection with other control systems as well.

The maximum velocity depends on the diameter of the laser ϕ_l and the pulse rate of the laser machine τ_l.

$$ \left\|\dot{\vec{\pi}}_{\max}\right\| = \phi_l \tau_l. \tag{9} $$

This consideration ensures a continuous motion of the manipulator (instead of a position-stepping mode) while delivering laser energy. The generated paths are subjected to update in the case of noticeable motion of the human head. Once the noticeable motion detected (3 mm translational or 4° rotational motion), the robot motion controller halts the robot manipulator and acquire an updated pose of the face ${}^{B}T_{F*}$. Then, the increment in pose will be ${}^{F*}T_{F} = {}^{B}T_{*F}^{-1}\,{}^{B}T_{F}$. The components $\vec{\chi}$ and $\vec{\eta}$ of the planned path p_i^j will be updated accordingly, $[\vec{\chi}^{*'}, 0]' = {}^{F*}T_{F}[\vec{\chi}', 0]'$ and $[\vec{\eta}^{*'}, 0]' = {}^{F*}T_{F}[\vec{\eta}', 0]'$. These updated path points are then converted to the pose vector using the framework described in this section. This technique of updating the paths instead of replanning reduces the computation load and ensures real-time performance during treatment.

C. Laser Shot Control

To deliver the laser energy uniformly over a treating skin surface, the control of the laser firing instance is paramount. To define a laser firing instance, an impulse function $\delta_i(\Delta d - \phi_l)$ is implied , where Δd is the argument of the impulse function and is defined as the distance covered by the robot end-effector from the position of the last laser shot $\vec{\chi}_{i-1}^{*}$ to the current position $\vec{\chi}_i$, so $\Delta d = \|\vec{\chi}_i - \vec{\chi}_{i-1}^{*}\|$, where ϕ_l defines the diameter of a laser shot. This impulse function generates 1 as an output only when its argument becomes zeros. Then, the instance of a laser shot can be defined as follows:

$$ \delta_i(\Delta d - \phi_l) = \begin{cases} 1 & \Delta d \geq \phi_l \\ 0 & \text{else} \end{cases}. \tag{10} $$

The laser shot instance only occurs when the robot is following the pose vectors $\pi_{p_i}^{B}$. Otherwise, the output from $\delta_i(.)$ will not be considered. Accordingly, the pose vector of the end-effector when the laser shot occurs can be defined as $\vec{\psi}_i = \delta_i\vec{\pi}_i$. Where $\vec{\psi}_i$ denotes the pose vector of 6×1 of end-effector when the laser shot instance occurs. This laser fire control relies on Δd, which only depends on the current position of the robot end-effector and the position of the last laser shot. After every laser shot instance, Δd is reset to zero. This simple control of the laser shot performs elegantly to distribute the laser shots uniformly.

D. Collision Avoidance Control

The robot end-effector must maintain a safe distance to the human face while performing rejuvenation treatment. To equip the robotic system with the proximity-like behavior, three distance sensors are housed in the end-effector besides the soft cap, as shown in Fig. 2. This ensures that the system is aware of end-effector's surroundings and can avoid any physical contact between the end-effector and the face, during the treatment. These point distance sensors can only provide the distance from the sensor's receiver to the target surface. To reliably include the distance measurement provided by these sensors in the system, we utilize their accumulated weighted averages as follows:

$$ \vec{l} = \frac{\Sigma_m \omega_m \vec{l}_m}{\Sigma_m \omega_m} \quad , \quad \omega_m = \frac{\vec{l}_m . \vec{\eta}_*}{\left\|\vec{l}_m\right\| \|\vec{\eta}_*\|} \tag{11} $$

where $\vec{l}$ is the weighted averaged vector from the sensors to the surface, ω_m the weight based on the dot product of the two vectors, $\vec{l}_m$ the vector parallel to the measuring direction of the distance sensor, $\vec{\eta}_*$ the normal vector pointing inward to the surface of the face and located at the intersection of $\vec{l}_m$, and the surface of the facial model (m represents the index number of each distance sensor).

To implement a repulsive behavior in order to avoid the collision between the face and the end-effector, a controller in the backend is continuously observing the $\vec{l}$ and intervenes, only

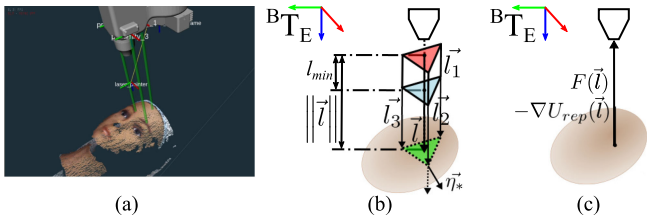


Fig. 9. Collision avoidance control. (a) The simulated realisation of the method. (b) Distance vectors. (c) Force acted on the robot end-effector.

when $\|\vec{l}\| < l_{\min}$, where $l_{\min}$ is the least close allowed distance for the robot end-effector. This behavior is the same as two repelling bodies in an artificial potential field. In this context, if we can consider the robot end-effector as a moving agent and the facial surface as a boundary and there is a repulsive force between them, which is defined by the gradient of the artificial potential function as in [30], i.e.

$$U_{\text{rep}}(\vec{l}) = \begin{cases} \frac{1}{2}\left(\frac{1}{l_{\min}} - \frac{1}{D(\vec{l})}\right)^2 & D(\vec{l}) \le l_{\min} \\ 0 & D(\vec{l}) > l_{\min} \end{cases} \tag{12}$$

and the gradient will be

$$\nabla U_{\text{rep}}(\vec{l}) = \begin{cases} \left(\frac{1}{l_{\min}} - \frac{1}{D(\vec{l})}\right)\frac{1}{D(\vec{l})^2}\Delta D(\vec{l}) & D(\vec{l}) \le l_{\min} \\ 0 & D(\vec{l}) > l_{\min} \end{cases} \tag{13}$$

Here, $U_{\text{rep}}(l)$ is the potential field between the robot end-effector and the facial model, whereas $D(\vec{l})$ represents the distance function of l, as shown in Fig. 9. The artificial force acted on the robot end-effector will be $F(l) = -\nabla U_{\text{rep}}(\vec{l})$ and, the generated velocities to push the end-effector away from the facial surface is $\dot{\vec{\pi}}_{\text{col}} = \kappa F(\vec{l})$, where κ is an arbitrary scaling factor. In this manuscript, the stability analysis of the proposed controllers is considered out of scope, since the motion of the robot, the manipulator has been controlled via a kinematic controller, which has been extensively studied in the literature, and its stability analysis is available in the literature [31]. The collision avoidance control in the Section IV-D works on the principle of virtual potential field and its stability analysis was reported in [32].

V. Results

A. Cross-Calibration

Cross-calibration is a method to validate a transformation link or a coordinates frame from multiple transformation chains (or kinematic chains). The proposed system is relying on multiple sensors and actuators to perform the complete rejuvenation task. Each sensor is measuring in its coordinate frame, e.g., the depth camera, distance sensors, and tracking system. Likewise, the actuators, i.e., the robot manipulator and the laser generator, require the control command in their frame of reference. Thus, without the accurate knowledge of the relationship between each coordinate frame, the complete proposed system is not able to perform the treatment properly.

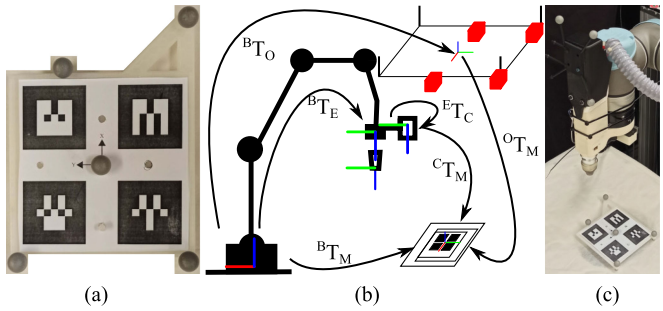


Fig. 10. Setup to acquire and validate the transformations between different coordinate frames. (a) AR/OptiTrack Marker. (b) Conceptual illustration of important transformation frames. (c) Physical illustration.

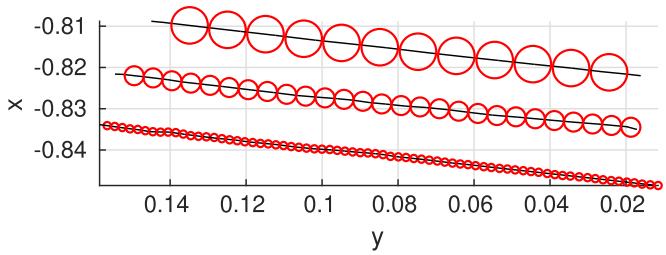


Fig. 11. Laser shot separation 0.01, 0.005, and 0.002 m (top to bottom).

Fig. 10 illustrates the physical setup to conduct cross-calibration. A marker, as shown in Fig. 10 (a), is composed of the AR markers and tracking markers and has a common transformation frame for the depth camera and tracking system. When the tracking system and depth camera are observing this marker at a same time, a transformation from the robot base to the tracking system's origin can be obtained by ${}^{B}T_{O} = {}^{B}T_{E}{}^{E}T_{C}{}^{C}T_{M}({}^{O}T_{M})^{-1}$. Whereas, the superscripts and subscripts are denoting as B the robot base, O the OptiTrack motion tracking system, E the end-effector, C the depth camera, and M the AR/OptiTrack marker. The superscripts and subscripts represent the parent and child transformation frame, respectively. Once ${}^{B}T_{O}$ is known, any transformation frame can be computed or validated.

B. Laser Shot Separation Distance Test

The laser shot control is evaluated by the separation distance test. In this test, a paper printed with black ink is placed on cardboard, and the robot end-effector follows a linear path while firing the laser shot of 3 mm diameter. For each run, different laser separation distance d_s was assigned as $d_s = 0.01,\ 0.005,\ 0.002$ m. In Fig. 11, these laser shot instances are plotted with red circles, and the path followed by the robot end-effector with a black line. Table I shows that in each test, the laser firing control has repeatedly triggered the laser shot instance precisely at the desired distance. That is why the mean value of each test is near to the desired distance and the standard derivation is infinitesimally small. The highest standard derivation is recorded for SDT3, where the predefined laser diameter ϕ_l

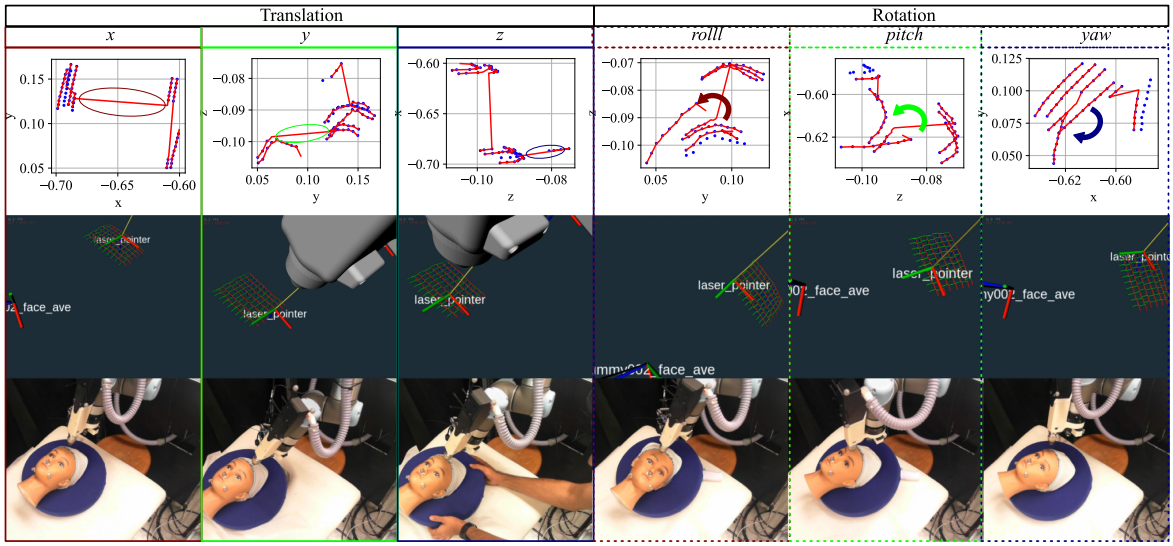


Fig. 12. Realtime adjustment of the planned paths and robot manipulator motion while performing rejuvenation treatment. In the first row, the red line and blue dots are the end-effector motion and planned paths. The ellipses and rounded arrows represent the translational and rotational motions of the treated face, which was detected then adjusted by the system. The second row contains the realtime visualization of the planned path and end-effector motion. The experimental setup is shown in the third row.

TABLE I
EVALUATION OF SEPARATION DISTANCE CONSISTENCY

Tests	N_t	d(m)	$\varnothing_l$(mm)	μ(mm)	σ(nm)	Φ(%)
SDT1	11	0.111	10	10.09	1.44	77.832
SDT2	26	0.131	5	5.05	1.45	77.940
SDT3	71	0.146	2	2.05	2.01	76.388

SDT is the separation distance test, whereas N_t denotes the number of laser shots fired in each test, ϕ_l the radius of the laser, d the distance covered by the end-effector, μ the mean distance of the laser shots, and σ standard deviation in the laser shots and Φ the area covered by the laser shots.

is 2 mm. This can be improved by decreasing the end-effector's speed, which provides the system with a wider window of time to acquire the end-effector position from the robot manipulator. The last column of Table I illustrates the area covered by the laser shots.

C. Real-Time Update of Planned Path

Fig. 12 demonstrates the real-time response of the system in the case of sudden movement of the human head during the treatment. To track the pose of the face continuously, Three markers were attached at the upper lips, left, and right jaw. The system ignores the small linear and angular motion within 3 mm and 4° respectively. In the first row of Fig. 12, the blue dots depict the desired paths and the red shows the following path by the end-effector. The second row illustrates the path updates and follows in parallel real-time simulation. Whereas, the motion of the mannequin's head can be seen in the pictures in the third row. The cherry red, green, and navy blue ellipses in each of the first three graphs in the second row represent the linear motion in x, y, and z-axis, respectively. Similarly, the cherry red, green, and navy blue arrows display the rotation in x, y, and z-axis, respectively (roll, pitch, and yaw).

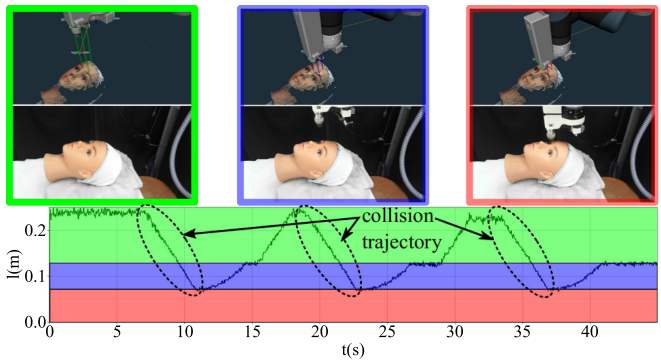


Fig. 13. Collision avoidance control test. Green color represents the free zone, blue the operation zone, and the red danger zone. The first and second rows illustrate the simulated and real-world setting for the test. In the third row, the distance between distance sensor and facial surface $\|\vec{l}\|$ versus time t has been plotted.

D. Collision Avoidance Test

To assess the collision avoidance control proposed in Section IV-D, the robot end-effector is enforced to move near to the mannequin's face (to mimic the possible collision scenario). Fig. 13 illustrates the three levels of closeness according to the desired behavior. In Fig. 13, the green color represents the "free zone," the blue color "operation zone" (the distance required to perform the treatment), and the red color "danger zone" (dangerously close distance between the end-effector and the face). Fig. 13 illustrates the implementation of the collision avoidance control. In this test, we have provided the robot end-effector with a collision trajectory three times. The distance from the robot end-effector and the face l versus time t is plotted in the third row of Fig. 13. It is evident from Fig. 13 that the controller is

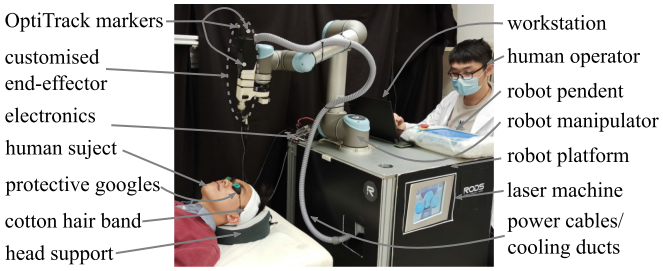


Fig. 14. Detailed illustration of the proposed robotic prototype and its components.

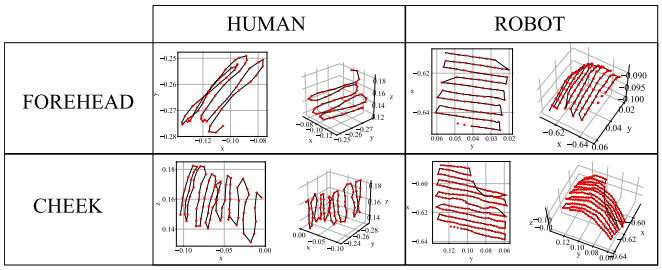


Fig. 15. Motion and laser shot instance plot for both human and robot operator (units are in meters). The black lines represent the motions and red dots the laser shots instances by both operators.

preventing the robot end-effector from entering the red zone and pushing the end-effector back to the operation zone.

E. Test on the Human Subjects

Fig. 14 depicts the proposed robotic prototype. The proposed automated skin photo-rejuvenation treatment has been performed on the human volunteers after receiving the ethical approval from the institute.[1] According to our research protocols, the human subjects wore a set of protected glasses to protect their eyes from the reflected laser light. Their eyebrows were also covered by medical paper tape. Subjects' heads were covered with a strip of thick cotton cloth to avoid laser exposure to their hairs. In this pilot study, the comparison between the robot and human operator were based on the average time, average number of shots, and average area coverage on a drawn patch of charcoal. Initially, we drew square patches of 47 × 47 mm with charcoal on the human subject of fixed size. This selection keeps the comparison unbiased between a human and a robot operator. Then, the human and robot operators perform the treatment inside the drawn patches on the left and right side of the face of the human subject.

The human operator was asked to fill the charcoal patch with laser shots while avoiding overlapping and (the same target set for the robotics prototype), as shown in Fig. 19. For human subjects, the Q-Switch Nd: YAG laser of 1064 nm wavelength with 4 mm laser diameter were used (pulse width). The laser energy was set to 600mJ and the laser fluence was 10 J/cm^2.

[1]Ethics Approval Reference Number: HSEARS20200604002, Human Subjects Ethics Sub-committee, Departmental Research Committee, The Hong Kong Polytechnics University, Hong Kong.

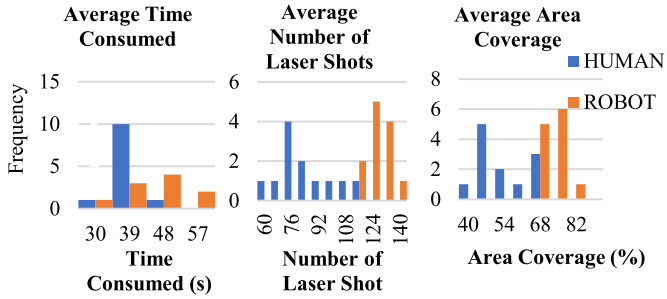


Fig. 16. Approximate distribution (Histogram) of the average time, number of laser shots and area coverage of the data presented in Table II.

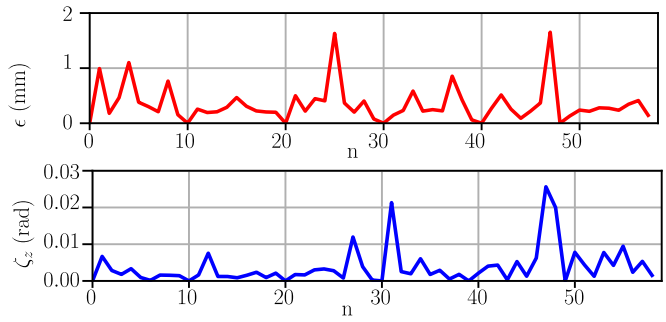


Fig. 17. ϵ denotes the error norm during the motion of the robot manipulator. ζ_z is the angle between the surface normal and the direction of the laser.

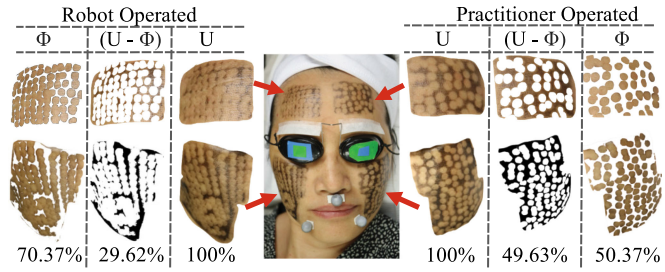


Fig. 18. Comparison of the average area covered Φ by robot and human operators in the operable area U.

The defined laser parameters were similar for both operators. In Fig. 15, the path followed by each operator is represented by a black line and the center of each red circle is the instance of a laser shot. The data for plotting the path and laser shot instances were recorded in the motion capture system in real-time while performing the treatment. The 3-D plots in Fig. 15 shows the comparable difference among both operators in path consistency and laser shot uniformity

Table II shows data of eleven human trails, which includes time, number of laser shots, and operable area coverage by the human and the robot operator. The mean and standard deviation of the area covered Φ_h by the human operator is 50.37% and 9.56%, respectively. Whereas, the robotic system has achieved 70% of the mean in area coverage Φ_r with the standard deviation of 4.13%, which is about 20% more area coverage and half the

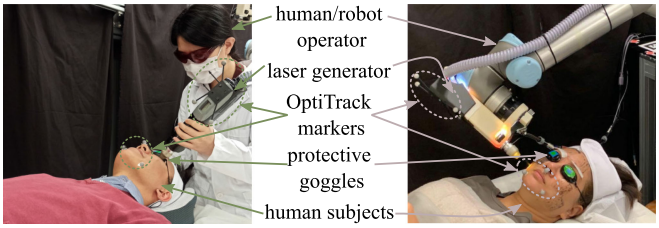


Fig. 19. Human Trails. Skin photo-rejuvenation treatment is conducting by human and robot operators.

TABLE II
EVALUATION OF SEPARATION DISTANCE CONSISTENCY

		HUMAN			ROBOT		
ID	N_p	t_h (s)	N_h	Φ_h (%)	t_r (s)	N_r	Φ_r (%)
1	3	117	323	61.24	102	341	64.66
2	3	133	350	66.36	124	357	67.69
3	3	113	343	65.04	127	346	65.60
3	2	69	162	46.07	89	245	69.68
4	2	63	153	43.51	100	262	74.52
5	3	64	197	37.35	96	381	72.24
6	2	64	197	56.03	96	241	68.54
7	4	156	290	41.24	131	523	74.37
8	3	96	259	49.11	65	355	67.31
9	2	77	172	48.92	160	260	73.95
10	2	63	160	45.50	114	239	67.98
11	2	69	155	44.08	140	274	77.93

[1] N_p denotes the number of charcoal patches applied on the facial skin, t_h and t_r are time taken by a human and robot operator, respectively. Similarly, N_h and N_r are the number of laser shots delivered by a human and robot operator. Φ_h and Φ_r the area covered by the laser shots by a human and robot operator, respectively.

standard deviation compared to the human operator. Here, one fact is worth to mention that *an inscribed circle in a square* can cover only 78.54% of the total area of the circle. Now, if we compare the value Φ_r, it provides a better comparison of covered versus operable area (without overlapping). Fig. 18 depicts the covered versus operable region, where the operable region is denoted by U and covered with Φ. For most of the case reported in Table II, the time consumed by the human operator (t_h) is less of the robotic system (t_r). The average time consumed to treat a patch by the robot operator is 43.35 s, which is higher than the 35 s of the human operator's average time. This trend is also similar for the number of irradiation (N_h and N_r). The average number of irradiation on a patch delivered by the robotic operator is 123.35, where the average for the human operator is 89.06. The robot operator is irradiating 33% more area than the human operator. Fig. 16 show that the robot operator is more consistent in the average number of laser shot and area coverage than the human operator. The reason, human operator completed the treatments in less time because the pulse rate of the laser machine was kept constant at 4 Hz. As the robot operator was delivering more number of laser shots per patch than human, so the robot operator needed more time to complete each treatment.

Fig. 17 illustrates the translational error norm, $\epsilon = \|\vec{\chi} - \vec{\chi}_d\|$, where the $\vec{\chi}$ and $\vec{\chi}_d$ are the position and desired position of the laser projection on the surface. ζ_z is the angle between the surface normal and the orientation of laser tool, $\zeta_z = \cos^{-1}(\frac{\vec{\eta}.\vec{\eta}_d}{\|\vec{\eta}\|\|\vec{\eta}_d\|})$. Where $\vec{\eta}$ is a normal vector and parallel to the laser light beam. $\vec{\eta}_d$ is a surface normal which is equal to

Fig. 20. Human subjects after treatment. The robotic system has performed the skin photo-rejuvenation on the left-hand side of each human subject's face and an experienced human operator has performed on the right-hand side.

the last column of ${}^{B}R_p$. Fig. 19 shows the physical setup to conduct human trails. Both operators were equipped with the same laser generator from the same company. The motion of the laser handpiece during the treatment was recorded using the OptiTrack motion tracking system for both operators. The path followed by the human operator and the robot can be seen in Fig. 15. Fig. 20 provides a visual comparison of the area coverage without overlapping. The video demonstration of the proposed algorithm can be found here.[2]

VI. CONCLUSION

In this article, we demonstrate an automated facial skin rejuvenation robotic prototype. To complete the manipulation task, an industrial-grade robotic manipulator along with a custom end-effector is used. The custom end-effector is equipped with a depth camera, three proximity sensors, and a laser generator. A method is presented to estimate the pose of the human face and the viewpoint around the face. An accurate facial model is constructed from the point cloud data captured at different estimated viewpoints. The 3-D facial model is segmented into seven regions using 2-D facial landmarks. Then, the optimal paths of the robot manipulator are extracted from each segmented area. The optimal path ensures an efficient and uniform distribution of the laser shots. To control the spacing distance of the laser shots while following the optimal path, a control law is devised to fire the laser shot after a predefined laser separation distance. As safety is the priority of service tasks in human environments, proximity-like behavior is implemented to avoid any possible collision between the robot end-effector and the human face.

Initially, the facial model is reconstructed in the camera coordinates, and the motion of the face is tracked by the Optictrack motion system. The transformation between multiple sensor coordinates and the robot coordinate is formulated. To avoid the possible error that could be introduced by a false measurement from single sensor feedback, a cross-calibration technique is utilized. This error reduction is achieved by observing the same AR/OptiTrack marker ${}^{B}T_M$ from the two transformation chains, then minimizing the difference between them. The laser

[2] [Online]. Available: https://github.com/romi-lab/cosmetic_robotics/raw/main/videos/video.mp4

separation distance is also evaluated by firing the laser shot on a plain piece of paper. Then, the separation of laser shots is measured by a vernier calibre. The laser shots were plotted on a graph to check the possible overlapping. The validation of each proposed control system has been covered thoroughly.

The uniform laser distribution demonstrates the potential improvement in the outcomes of the robotized treatment, which can not be achieved by a human operator. However, the proposed methods have some limitations. The laser fire controller only uses the local information of two positions, $\vec{\chi}_i$ and $\vec{\chi}^*_{i-1}$, so the overlapping of the laser shots may appear at the edges of the path. At the current stage, the laser shot energy level can only be preset by a human operator. Future works also include integrating skin tissue temperature as a feedback of the laser firing controller. This will enable the robotics system to decide whether to fire or not to fire on an instance. Furthermore, it is also observed that only temperature measurements of the surface of the skin may not be sufficient to develop a laser fire controller. The temperature of the skin surface can be measured by a thermal camera, which depends on the emittance of the observed object. Thus, the raw temperature measurement may perturb from a real value [33]. The room temperature also influences the temperature of the skin surface. Another reason is a low thermal diffusivity of the skin outer layer "stratum corneum," which makes the outer layer of the skin a good insulator. Thus, the temperature changes occurring inside the skin tissue due to laser-skin interaction cannot be predicted by the temperature values on the skin surface. Furthermore, an in-depth clinical study will be conducted to compare the efficiency of the skin photo-rejuvenation treatment by the proposed robotic prototype and a human operator, where the accuracy, consistency, and postdermatological aesthetic metrics will be evaluated.

REFERENCES

[1] D. Holck and J. Ng, "Facial skin rejuvenation," *Curr. Opin. Ophthalmol.*, vol. 14, no. 5, pp. 246–252, 2003.

[2] J. E. Oblong, "Chapter 15 - Wrinkles: Cosmetics, Drugs, and Energy-Based Systems," in *Cosmetics Applications of Laser & Light-Based Systems, Ser. Personal Care & Cosmetic Technology*, G. S. Ahluwalia, ed. Boston, MA, USA: William Andrew Publishing, 2009, pp. 301–316.

[3] V. A. Narurkar, "Chapter 13 - Global Total Nonsurgical Rejuvenation: Lasers and Light-Based Systems in Combination With Dermal Fillers and Botulinum Toxins," in *Cosmetics Applications of Laser & Light-Based Systems, Ser. Personal Care & Cosmetic Technology*, G. S. Ahluwalia, ed. Boston, MA, USA: William Andrew Publishing, 2009, pp. 281–289.

[4] D. J. Goldberg and J. Whitworth, "Laser skin resurfacing with the q-switched Nd:YAG laser," *Dermatologic Surg.*, vol. 23, no. 10, pp. 903–907, 1997.

[5] P. Babilas, S. Schreml, R.-M. Szeimies, and M. Landthaler, "Intense pulsed light (IPL): A review," *Lasers Surg. Med.*, vol. 42, no. 2, pp. 93–104, 2010.

[6] M. Lolis and D. Goldberg, "Radiofrequency in cosmetic dermatology: A review," *Dermatologic Surg.*, vol. 38, no. 11, pp. 1765–1776, 2012.

[7] J. S. Resneck Jr. and A. B. Kimball, "Who else is providing care in dermatology practices? Trends in the use of nonphysician clinicians," *J. Amer. Acad. Dermatol.*, vol. 58, no. 2, pp. 211–216, 2008.

[8] M. El-Domyati *et al.*, "Radiofrequency facial rejuvenation: Evidence-based effect," *J Amer. Acad Dermatol.*, vol. 64, no. 3, pp. 524–535, 2011.

[9] D. Goldberg and C. Metzler, "Skin resurfacing utilizing a low-fluence Nd:YAG laser," *J. Cosmetic Laser Ther.*, vol. 1, no. 1, pp. 23–27, 1999.

[10] J. Resneck and A. B. Kimball, "The dermatology workforce shortage," *J. Amer. Acad. Dermatol.*, vol. 50, no. 1, pp. 50–54, 2004.

[11] M. Hwang and D. Kwon, "Strong continuum manipulator for flexible endoscopic surgery," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 5, pp. 2193–2203, Oct. 2019.

[12] C. Li, X. Gu, X. Xiao, C. M. Lim, and H. Ren, "Flexible robot with variable stiffness in transoral surgery," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 1, pp. 1–10, Feb. 2020.

[13] A. J. Muñoz-Vázquez, J. D. Sánchez-Torres, E. Jiménez-Rodríguez, and A. G. Loukianov, "Predefined-time robust stabilization of robotic manipulators," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 3, pp. 1033–1040, Jun. 2019.

[14] M. T. Draelos and Z. D. Draelos, "Robotic cosmetic dermatology," *J. Cosmet. Dermatol.*, vol. 10, no. 2, pp. 87–88, 2011.

[15] R. M. Bernstein, "Integrating robotic fue into a hair transplant practice," *Hair Transplant. Forum Int.*, vol. 22, no. 6, 2012, pp. 228–229.

[16] H.-w. Lim *et al.*, "A study on the development of a robot-assisted automatic laser hair removal system," *Photomed. Laser Surg.*, vol. 32, no. 11, pp. 633–641, 2014.

[17] W. Koh *et al.*, "A pilot study on the evaluation of physicians' laser delivery performance using a laser beam detection kit," *Photomed. Laser Surg.*, vol. 35, no. 6, pp. 317–323, 2017.

[18] J. P. Farkas, J. E. Hoopman, and J. M. Kenkel, "Five parameters you must understand to master control of your laser/light-based devices," *Aesthetic Surg. J.*, vol. 33, no. 7, pp. 1059–1064, 2013.

[19] M. Quigley *et al.*, "Ros: An open-source robot operating system," in *ICRA Workshop Open Source Software*, vol. 3, no. 3.2. Kobe, Japan, 2009, Paper no. 5.

[20] H. Su, W. Qi, Y. Hu, H. R. Karimi, G. Ferrigno and E. De Momi, "An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators," *IEEE Trans. Ind. Informat.*, p. 1, 2020.

[21] O. Zahra and D. Navarro-Alarcon, "A self-organizing network with varying density structure for characterizing sensorimotor transformations in robotic systems," in *Proc. Annu. Conf. Towards Auton. Robot. Syst.*, 2019, pp. 167–178.

[22] D. Navarro-Alarcon, O. Zahra, C. Trejo, E. Olguín-Díaz, and V. Parra-Vega, "Computing pressure-deformation maps for braided continuum robots," *Front. Robot. AI*, vol. 6, 2019, Art. no. 4.

[23] H. Su, A. Mariani, S. E. Ovur, A. Menciassi, G. Ferrigno and E. De Momi, "Toward teaching by demonstration for robot-assisted minimally invasive surgery," *IEEE Trans. Automat. Sci. Eng.*, vol. 18, no. 2, pp. 484–494, Apr. 2021.

[24] D. King, "Dlib-ml: A Machine Learning Toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Jul. 2009.

[25] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1867–1874.

[26] K.-L. Low, "Linear least-squares optimization for point-to-plane icp surface registration," *Chapel Hill, Univ. North Carolina*, vol. 4, no. 10, pp. 1–3, 2004.

[27] J. Park, Q.-Y. Zhou, and V. Koltun, "Colored point cloud registration revisited," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 143–152.

[28] J. o'Rourke, *Computational Geometry in C*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[29] A. Baker, *Matrix Groups: An Introduction to Lie Group Theory*. 1st ed. London: Springer Science & Business Media, 2012.

[30] H. M. Choset *et al.*, *Principles of Robot Motion: Theory, Algorithms, and Implementation*. Cambridge, MA, USA: MIT Press, 2005.

[31] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics, Ser. Advanced Textbooks in Control and Signal Processing*. London, U.K.: Springer London, 2009.

[32] D. Navarro-Alarcon, J. Qi, J. Zhu, and A. Cherubini, "A Lyapunov-stable adaptive method to approximate sensorimotor models for sensor-based control," *Front. Neurorobotics*, vol. 14, 2020, Art. no. 59.

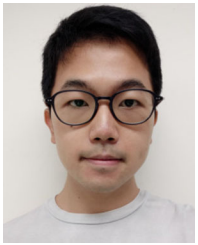
[33] "User's Manual FLIR ETS3xx Series," 2018. [Online]. Avilable: https://www.flir.com/globalassets/imported-assets/document/flir-ets320-user-manual.pdf

Muhammad Muddassir (Student Member, IEEE) received the master's degree in control science and engineering from Beijing Institute of Technology (BIT), Beijing, China, in 2017. Since 2018, he is working toward the PhD degree in mechanical engineering at The Hong Kong Polytechnic University, KLN, Hong Kong.

He joined the Hong Kong Polytechnic University in 2018 first as a Research Associate. His current research interests include the development and control of sensor-guided robotic systems.

Domingo Gómez Domínguez received the B.Eng. degree in energy engineering from the Higher Technical School of Engineering of Huelva, Spain, in 2017. He received the M.Sc. degree in electrical engineering from Automation and Robot Technology, Technical University of Denmark (DTU), Lyngby, Denmark, in 2019.

After one year of automating systems as a Technical Student with the European Organisation for Nuclear Research (CERN). He also worked as a Research Assistance with Automation and Robot Technology, DTU. Later, he moved to Hong Kong to work as a Research Assistant at PolyU. His research interest is to make complex automated systems of easy usage.



Luyin Hu received the B.Eng. degree from the Hong Kong Polytechnic University, Hung Hom, Hong Kong, in 2019, and is currently working toward the Part-Time M.Phil. degree with the Hong Kong Polytechnic University, both in mechanical engineering.

He is currently a Research Assistant with the Mechanical Engineering Department, Hong Kong Polytechnic University. His research interests include multimodal perception, model-based servoing, and control system design.

Shujian Chen received the B.Eng. degree in mechanical engineering from the Hong Kong Polytechnic University, Hung Hom, Hong Kong, in 2019. He is currently working toward the M.S. degree in mechanical engineering(concentrated in robotics) with Columbia University, New York, NY, USA.

David Navarro-Alarcon (Senior Member) received the Ph.D. degree in mechanical and automation engineering from The Chinese University of Hong Kong (CUHK), NT, Hong Kong, in 2014.

Since 2017, he has been with the Hong Kong Polytechnic University (PolyU), KLN, Hong Kong, where he currently is an Assistant Professor with the Department of Mechanical Engineering, and the Principal Investigator of the Robotics and Machine Intelligence Laboratory. He is also an Investigator at the Research Institute for Smart Ageing (RISA) of PolyU. Before joining PolyU, he was a Research Assistant Professor with the CUHK T Stone Robotics Institute, Hong Kong, from 2015 to 2017. His current research interests include perceptual robotics and control theory.