

Large-Scale Vision-Based Tactile Sensing for Robot Links: Design, Modeling, and Evaluation

Lac Van Duong and Van Anh Ho

Abstract—The sense of touch allows individuals to physically interact with and better perceive their environment. Touch is even more crucial for robots, as robots equipped with thorough tactile sensation can more safely interact with their surroundings, including humans. This study describes a recently developed large-scale tactile sensing system for a robotic link, called TacLINK, which can be assembled to form a whole-body tactile sensing robot arm. The proposed system is an elongated structure comprising a rigid transparent bone covered by continuous artificial soft skin. The soft skin of TacLINK not only provides tactile force feedback but can change its form and stiffness by inflation at low pressure. Upon contact with the surrounding environment, TacLINK perceives tactile information through the three-dimensional (3-D) deformation of its skin, resulting from the tracking of an array of markers on its inner wall by a stereo camera located at both ends of the transparent bone. A finite element model was formulated to describe the relationship between applied forces and the displacements of markers, allowing detailed tactile information, including contact geometry and distribution of applied forces, to be derived simultaneously, regardless of the number of contacts. TacLINK is scalable in size, durable in operation, and low in cost, as well as being a high-performance system, that can be widely exploited in the design of robotic arms, prosthetic arms, and humanoid robots, etc. This paper presents the design, modeling, calibration, implementation, and evaluation of the system.

Index Terms—Tactile sensing skin, large scale tactile sensor, vision-based force sensing, stereo camera, non-rigid registration, calibration, finite element method (FEM), soft robotics.

I. INTRODUCTION

NOWADAYS, robots are not all confined within safety fences inside factories performing highly accurate repetitive operations at high speed. Many are being increasingly used in activities of humans' daily lives, such as caretaker, medical, and therapeutic robots. These robots are therefore expected to frequently interact with humans, especially via physical contact. Increased need for safe and intelligent interactions between humans and robots has led to attempts to equip robots with whole-body robotic skin that can sense multiple modalities, especially tactile sensations. Equipping robots with suit-

Manuscript received February 23, 2020; accepted August 30, 2020. Date of publication September 10, 2020; date of current version September 10, 2020. This article was recommended for publication by Associate Editor and Editor M. Yim upon evaluation of the reviewers' comments. This work was supported in part by JSPS KAKENHI Grant Number 18H01406. (*Corresponding authors:* Lac Van Duong and Van Anh Ho.)

Duong and Ho are with Japan Advanced Institute of Science and Technology, Japan 923-1211. (e-mail: lacduong@jaist.ac.jp; van-ho@jaist.ac.jp).

This article has supplementary downloadable multimedia material available at <http://ieeexplore.ieee.org>, provided by the authors. The material consists of a demonstration video showing the operation of the proposed sensing system.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2020

able tactile feedback facilitates their awareness of surroundings through touch sensations, in the same manner that humans feel and interact with theirs. This study was motivated by a desire to develop a whole robot arm with soft and highly deformable tactile skin that conveys detailed information regarding contact geometry and force from dynamic interactions between the robot and its surroundings.

Tactile sensing devices for robotic skin ?, ? must not only be sensitive and flexible ?, ?, ?, but be highly scalable ?, ? and stretchable ?, ?, enabling them to be wearable on a wide and curved surface of a robot's body, from fingertips to larger areas, including hands, arms, and chest. Tactile sensors were designed as a matrix of sensing elements (taxels), whose working principle was based on physical transducing phenomena, from applied force/pressure to changes in, for example, resistance, capacitance, inductance, electromagnetic field strength, and light density ?. Most of these designs, however, focused only on the structure and principle, without considering the bulk of wires and analog-to-digital converters, etc. required for a large number of taxels. Also, high cost and manufacturing complexity have partly constrained the number of commercial tactile devices.

To date, few studies have effectively provided a robot with large areas of tactile sensors. Typical designs involve integrated electronic skin formed by many spatially distributed modular sensing points, allowing data to be processed locally and sent through a serial bus, thereby reducing the number of wires ?. For example, a network of rigid hexagonal printed circuit boards (PCB) provided a variety of sensory modalities including proximity, vibration, temperature, and light touch ? useful in many control strategies and applications ?, ?. However, based on the integration of various sensors, electronic components, and a local microcontroller, this design was low spatial resolution and relatively expensive. Moreover, the embedded electronic sensors inside the robotic skin are considered not durable enough for frequent collisions during contact with the surrounding environment.

Vision-based sensing provides advantages to soft tactile skin, including high spatial resolution and sensitivity, in that a camera is employed to document deformation of artificial skin for conversion to tactile information ?, ?. Such technology can sense deformation of a large area of soft robotic skin without embedded sensors, markedly reducing wiring, electronics, and risk of damage. For example, the GelSight sensor can measure the detailed surface texture of a touched object, where a lookup table of each sensor that maps observed light intensity of the reflective surface to geometry gradients was experimentally build ?. The GelForce sensor comprises two layers of markers

on a transparent elastic body for measuring surface traction fields, resulting from experimental methods to calibrate the relationship between applied force and the movements of markers [2]. The TacTip family can adapt to different 3-D shapes that utilize change in the positions of markers to perform tactile perception, manipulation, and exploration [3]. Rather than determine the actual location and direction of an applied force, TacTip focused on the statistical ability of the derived data of a dotted pattern to provide the approximate edge of a contacted surface for helping robots to localize and follow the contours. Most of the aforementioned sensors were designed for fingertips or grippers, their scalability to large and curved areas has not been reported.

Vision-based tactile sensing systems have difficulty in determining a clear relationship between measured deformation and distribution of applied contact force. Vision-based force sensing has been evaluated, including estimations of the external force that can solve equilibrium equations based on calculation of elastic membrane tensional force [4], minimize energy using visually measured contour data of a deformed object [5], and solve Cauchy's problem in elasticity estimating forces acting on a gripper [6]. Other strategies using machine learning have been investigated [7, 8]. However, the aforementioned studies focusing on estimating the total force applied could not determine the location and distribution of applied forces upon multiple contact points or across a wide surface. Although external forces on soft robots could be estimated using an open-source simulation framework (SOFA) with an optical tracking system [9], this system is not compact, and the technique is complex with several limitations.

An efficient tactile sensing system should minimize the amount of wires and electronics and maximize tactile sensing capability at low cost. To this end, we developed TacLINK, a large-scale tactile sensing system for a robotic link, with soft artificial skin using vision-based sensing technology (Fig. ??). This study elaborates our preliminary work in a conference paper [10]. Notably, the finite element (FE) model of the skin was formulated enabling TacLINK to determine the distribution of applied forces on the skin surface under the condition of multiple simultaneous contacts. This study describes the entire development of TacLINK, from system design to modelling and implementation, as well as experiments evaluating this system. The contributions of this paper include:

- 1) Development of a low-cost (about US\$150) efficient design for robotic links equipped with large-scale tactile force sensing skin. The soft tactile skin is durable, comfortable to touch, highly deformable, and inflatable to change its form and stiffness.
- 2) Robust algorithms to extract information from images of three-dimensional deformation of elongated and curved skin based on the views supplied by a stereo camera.
- 3) A generalized FE approach to compute contact force that is powerful in spatial force reconstruction. Results can be utilized for similar structures.

The open-source TacLINK and a demonstration video can be found here: <https://github.com/lacduong/TacLINK>

The remainder of this paper is organized as follows. Section II provides an overview of system design. Section III presents a

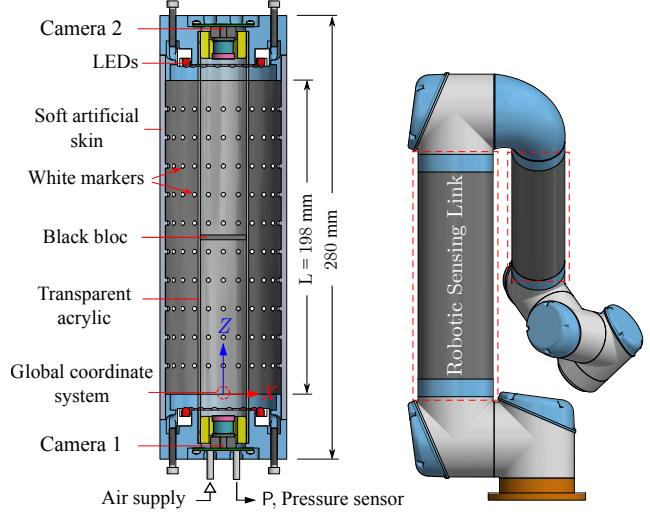


Fig. 1. (Left) Configuration of the robot links (i.e., TacLINK) with large-scale tactile sensation. (Right) Sketch of a collaborative robot (UR5) equipped with TacLINK, enabling safe and intelligent interaction through the sense of touch.

mathematical model of the proposed stereo camera, followed by algorithms for stereo implementation and calibration in Section IV. Section V presents the finite element model for the artificial skin used to analyze and calculate tactile forces in Section VI. Section VII describes experiments performed to evaluate stereo vision and tactile force reconstruction. Sections VIII and IX discuss and conclude this paper, respectively.

II. SYSTEM DESIGN

A. Structure and Principle

Figure ??(a) depicts an overall view and the interior structure of TacLINK. It comprises a transparent acrylic tube that connects its two ends, acting like a bone frame for maintaining its rigidity. Each end has a connecting part that can accommodate a fisheye lens CMOS camera (ELP USBFHD01M-L21: resolution 640×480 pixels, frame rate YUV2 30 fps, field of view $\sim 150^\circ$), and a series of high-intensity LEDs for illumination. The LED source uses a polarizing filter to produce uniform light and minimize reflected light. To block direct light from LEDs to facing camera, a black bloc is set in the middle of the acrylic tube. TacLINK is covered by soft continuous artificial skin (sensing area $\sim 49,763 \text{ mm}^2$) dyed black to completely isolate the inner space with ambient light. Based on detectability of the cameras, a total of 240 white markers of diameter $D_{\text{marker}} = 2.8 \text{ mm}$ sit on the inner wall of artificial skin, which is firmly attached affording total sealing of the inner space between the tube and skin. An inlet supplies air to the inner space, and an outlet is connected to an external pressure sensor. Altering the inner air pressure (0–1.5 kPa) inflates or deflates the skin to change its stiffness. Note that if f is the focal length of a camera in pixels, and d_{min} is the minimum detectable pixel size of a marker on the image, the design length of skin should satisfy $L < f \frac{D_{\text{marker}}}{d_{\text{min}}}$.

The sensing principle of TacLINK involves the two co-axial cameras, one at each end, that form a stereo camera. These cameras take consecutive pictures of markers on the inner wall

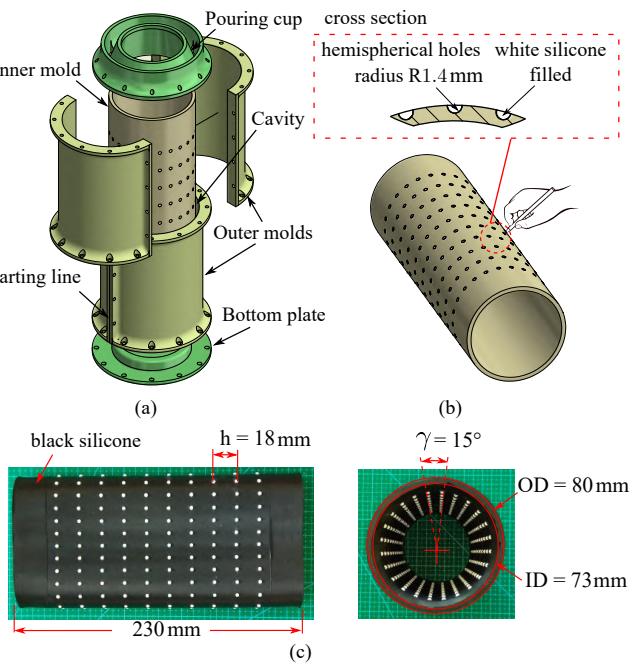


Fig. 2. Casting for skin fabrication. (a) The 3-D printable parts of casting mold. (b) Filling holes of the inner mold with silicone to cast the markers. (c) Inside and top view of the actual soft skin body with distributed markers.

of the skin, enabling the calculation of the 3-D positions of all markers on the global coordinate system. This approach took advantage of the finite element model of the skin to calculate the distribution of applied forces based on a structural stiffness matrix and extracted displacements of the markers. The contact forces can be treated as the acting concentrated force resulting from the nodal forces of the FE model.

B. Artificial Skin

TacLINK perceives tactile information through its skin, the properties of which play an important role in sensing characteristics. In this paper, artificial skin of thickness $t = 3.5$ mm was made of silicone rubber material Ecoflex 00-50 (Smooth-On Inc., USA) with good elasticity and relatively low mixed viscosity. The skin was fabricated by a casting method, as illustrated in Fig. ???. First, parts of the mold were designed by 3-D CAD software (Autodesk Inventor, Autodesk Inc., USA) and printed using a 3-D printer (M200, Zortrax S.A., Poland). The top funnel-shaped mold was customized as the pouring cup. To cast the markers, an array of hemispherical holes, of pitches $h = 18$ mm, $\gamma = 15^\circ$, was printed on the surface of the inner mold and filled with white silicone (Fig. ??(b)). After the white markers were formed, all parts of the mold were assembled. The black silicone was poured into the inner space and allowed to cure at room temperature for about three hours. Finally, the skin body was removed manually from the casting mold. Fig. ??(c) shows a prototype of artificial skin with actual markers evenly distributed.

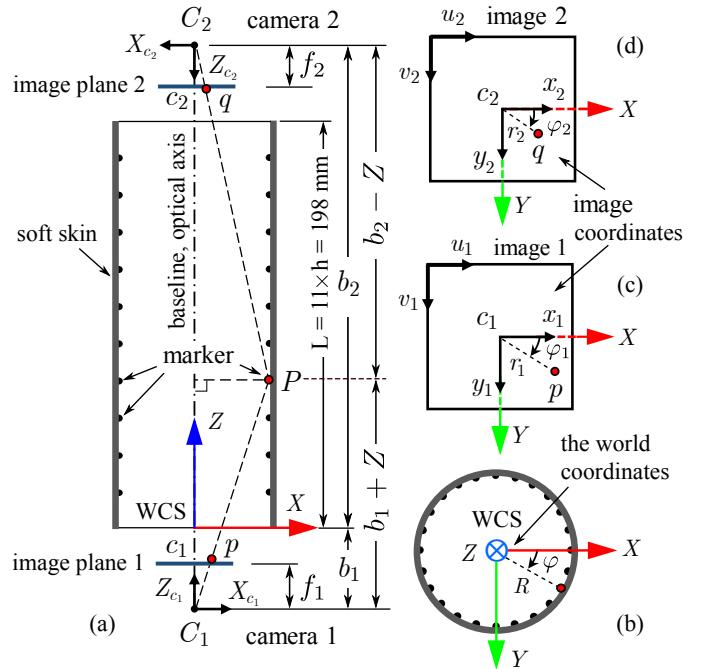


Fig. 3. Configuration of the proposed stereo camera. We imposed that the world coordinate system (WCS) XYZ located on the baseline. Its origin coincided with the center of the first end, and X -axis was parallel to the X_{c_1} - and X_{c_2} -axes of the camera frames $X_{c_1}Y_{c_1}Z_{c_1}$ and $X_{c_2}Y_{c_2}Z_{c_2}$. Note that the image of camera 2 was flipped over its y_2 -axis to unify the direction of the coordinates of the two images. Thus, the image coordinates on the x_1 , x_2 - and y_1 , y_2 -axes were in the same directions as the X and Y -axes of WCS, respectively.

III. VISION-BASED MODEL

This section addresses the configuration of the stereo camera, consisting of two co-axial cameras, used to produce a 3-D geometrical space. The camera was modeled by a usual pinhole, with the pixel sensors assumed to be square in shape (i.e., the focal lengths on the x - and y -axes are similar).

A. Stereo Camera Model

The objective of the vision-based system was to construct the 3-D shape of the soft skin by tracking the position of markers. Fig. ??(a) illustrates the stereo camera consisting of two co-axial cameras with its baseline and optical axes along the centerline of the module. The Z -axis of the world (global) coordinate system (WCS) XYZ coincided with the baseline, and its origin located at the center of the first end (with camera 1). The stereo model describes the 3-D position of a point $P = [X, Y, Z]^\top \in \mathbb{R}^3$ in WCS captured by cameras on two image planes with corresponding 2-D points, $p = [x_1, y_1]^\top \in \mathbb{R}^2$, and $q = [x_2, y_2]^\top \in \mathbb{R}^2$ (Fig. ??(c)–(d)). Based on the geometrical relationships, the projective equation was determined to be:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \frac{f_1}{b_1 + Z} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \frac{f_2}{b_2 - Z} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad (1)$$

with

$$x_1 = u_1 - c_{x_1}, \quad y_1 = v_1 - c_{y_1}, \quad (2)$$

$$x_2 = u_2 - c_{x_2}, \quad y_2 = v_2 - c_{y_2}. \quad (3)$$

where f_1 and f_2 correspond to the focal lengths of the two cameras in pixels; b_1 and b_2 denote the position of WCS origin in

two camera frames, respectively; $c_1(c_{x_1}, c_{y_1})$ and $c_2(c_{x_2}, c_{y_2})$ are the pixel locations of the principal points on the two pixel coordinate systems, $\{u_1, v_1\}$ and $\{u_2, v_2\}$, respectively.

In this application, the stereo camera can be fully modelled by four intrinsic parameters $\{f_1, f_2, c_1, c_2\}$ and two extrinsic parameters $\{b_1, b_2\}$. The method for self-automatic calibration is presented in Section IV-D.

The problem was to determine the 3-D location of a point through information supplied by the two cameras. For convenience, Cartesian coordinates were converted to polar and cylindrical coordinates on the image and world coordinates, respectively, where R and φ represent radial and angular coordinates of WCS (Fig. ??(b)). Mathematically, the other form of (??) can be expressed as

$$r_1 = \frac{f_1}{b_1 + Z} R, \quad r_2 = \frac{f_2}{b_2 - Z} R, \quad (4)$$

where r_1 and r_2 are the radial coordinates, and the angular coordinates φ_1 and φ_2 are two arcs from $-\pi$ to π , computed from data supplied by the two cameras as (see Fig. ??(c)–(d))

$$r_1 = \sqrt{x_1^2 + y_1^2}, \quad \varphi_1 = \arctan 2(y_1, x_1), \quad (5)$$

$$r_2 = \sqrt{x_2^2 + y_2^2}, \quad \varphi_2 = \arctan 2(y_2, x_2). \quad (6)$$

The solutions of (??) can be derived as follows:

$$Z = \frac{f_1 b_2 r_2 - f_2 b_1 r_1}{f_1 r_2 + f_2 r_1}, \quad (7)$$

$$R = (b_1 + b_2) \frac{r_1 r_2}{f_1 r_2 + f_2 r_1}. \quad (8)$$

Figure ??(b)–(d) reveals that the ideal angular coordinates obey the relationship as $\varphi = \varphi_1 = \varphi_2$. In fact, the angular φ_1 and φ_2 may differ slightly, due to misalignment of the axes of the two installed cameras. Because this discrepancy is constant, it can be corrected by compensation. The angular coordinate could be approximated as $\varphi = \frac{1}{2}(\varphi_1 + \varphi_2)$. However, because the angular coordinates discontinue at $-\pi$ and π , φ can be calculated using the function:

$$\varphi = \begin{cases} \frac{1}{2}(\varphi_1 + \varphi_2) & |\varphi_1 - \varphi_2| < \pi \\ \frac{1}{2}(\varphi_1 + \varphi_2) - \pi \text{sign}(\varphi_1 + \varphi_2) & |\varphi_1 - \varphi_2| > \pi \end{cases} \quad (9)$$

After φ is obtained, X and Y can be calculated:

$$X = R \cos(\varphi), \quad (10)$$

$$Y = R \sin(\varphi). \quad (11)$$

It should be noticed that equations (??)–(??) represent the straightforward solution of the triangulation of the stereo camera. This result enhances the ability of the proposed 3-D vision-based system to perform in real-time.

B. Uncertainty of 3-D Measurement with Digital Camera

The objective of this section is to analyze the uncertainty of stereo reconstruction, including the effects of camera parameters. This can help to optimize the best working space for design. In particular, the localization error depends on the detectability in both images associated to the projections $p(r_1, \varphi_1)$ and $q(r_2, \varphi_2)$, and is characterized by the sensitivity

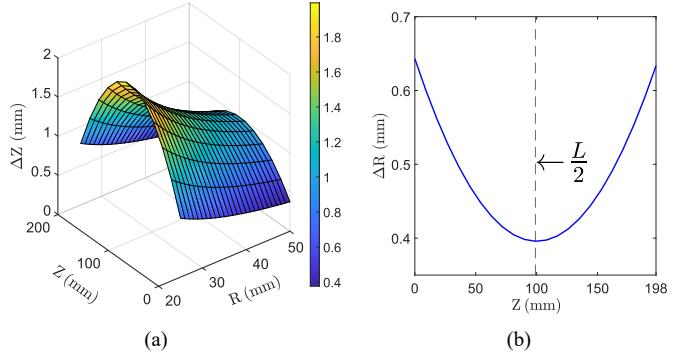


Fig. 4. Stereo variation simulation results illustrating the highest variations of error in the working space. (a) Uncertainty of Z coordinate with deviations $\Delta r_1 = -1/\sqrt{2}$, $\Delta r_2 = 1/\sqrt{2}$. (b) Uncertainty of radial coordinate ΔR with deviations $\Delta r_1 = 1/\sqrt{2}$, $\Delta r_2 = 1/\sqrt{2}$.

of imaging parameters to errors. Let Δr_1 and Δr_2 be the image plane coordinate error variables, independent from each other. Because the standard deviation of pixel error of a digital camera varies from 0 to 1 pixels, the normalized error variables of Δr_1 and Δr_2 are uniformly distributed within the interval of erroneously perturbed $[-1/\sqrt{2}, 1/\sqrt{2}]$ pixels ?. For ease of analysis, the ideal calibrated stereo parameters were assumed to be known beforehand (Table I). From (??) and (??), the derivative of the coordinate Z and the radial coordinate R with respect to r_1 and r_2 , resulted in the uncertainty of ranges Z and R being:

$$\Delta Z = f_1 f_2 \frac{b_1 + b_2}{(f_1 r_2 + f_2 r_1)^2} (r_1 \Delta r_2 - r_2 \Delta r_1), \quad (12)$$

$$\Delta R = \frac{b_1 + b_2}{(f_1 r_2 + f_2 r_1)^2} (f_1 r_2^2 \Delta r_1 + f_2 r_1^2 \Delta r_2). \quad (13)$$

Substituting r_1 and r_2 from (??) into (??) and (??) results in the expressions of uncertainty as

$$\Delta Z = \frac{(b_2 - Z)^2 (b_1 + Z)}{f_2 (b_1 + b_2) R} \Delta r_2 - \frac{(b_1 + Z)^2 (b_2 - Z)}{f_1 (b_1 + b_2) R} \Delta r_1, \quad (14)$$

$$\Delta R = \frac{(b_1 + Z)^2}{f_1 (b_1 + b_2)} \Delta r_1 + \frac{(b_2 - Z)^2}{f_2 (b_1 + b_2)} \Delta r_2, \quad (15)$$

where $Z \in [0, L]$, $R \in [R_{\min}, R_{\max}]$ are in the working ranges.

According to (??) and (??), the localization errors were nonlinear in assessing the deviations of normalized pixel misalignments Δr_1 and Δr_2 . Fig. ?? shows the highest uncertainties in ranges of ΔZ and ΔR in the designed working space. The uncertainty ΔR was found to depend solely on the Z -coordinate, with greater certainty in the middle region ($Z \rightarrow 0.5L$, see Fig. ??(b)). Generally, markers located in the middle tend to be measured in R with higher accuracy than markers near the two ends. In contrast, Fig. ??(a) indicates that the range uncertainty ΔZ increases significantly on the middle and reduces on the positive direction of R . Thus, markers close to the two ends and away from the centerline can be localized with higher accuracy of Z . In addition, increasing f_1 and f_2 can enhance precision (reduce the variations ΔZ and ΔR) by installing camera with smaller pixel size, or adjusting the focal length of the lens, although the latter will reduce the view angle of the lens.

IV. STEREO IMPLEMENTATION

This section describes a robust algorithm to implement the stereo registration that unifies tracking and matching. This is followed by techniques used to measure the 3-D surface of the skin under large deformation despite occlusion or limited sight, providing the precise location of every marker.

A. Configuration

The stereo vision is designed to measure the 3-D positions of all nodes on the 24×11 mesh of skin (see Fig. ??(b)). Let $\ell \in \{1, \dots, n_1\}$ denotes the path and $j \in \{0, \dots, n_2\}$ represents the cell, corresponding to the grid locations on the φ - and Z -axes, here $n_1 = 24$, $n_2 = 11$. The node (ℓ, j) can be identified by labeling with an index ℓ_j defined as (see Fig. ??(a))

$$\ell_j := n_1 \times j + \ell. \quad (16)$$

Hence, the nodes belonging to the mesh are represented by the set $\mathcal{N} = \{\ell_j\}, |\mathcal{N}| = N = 288$, and they were numbered from 1 to N . Herein, \mathcal{N} was divide into three sets of nodes (i.e., $\mathcal{N} = \mathcal{B}_1 \cup \mathcal{M} \cup \mathcal{B}_2$): The fixed nodes were located on the two clamped edges $\mathcal{B}_1 = \{\ell_0\}, |\mathcal{B}_1| = n_1$, and $\mathcal{B}_2 = \{\ell_{n_2}\}, |\mathcal{B}_2| = n_1$ (see Fig. ??(b)); and the free nodes or markers were tracked by cameras $\mathcal{M} = \{\ell_j\}_{j=1, \dots, n_2-1}, |\mathcal{M}| = n = 240$. Moreover, the full 3-D coordinates $\mathbf{X} \in \mathbb{R}^{3N}$ of all N -nodes was defined, with the sub-vector $\mathbf{X}_i = P(i) \in \mathbb{R}^3$ being the 3-D position of node $i \in \mathcal{N}$. Similarly, the sets $I_1 = \{p(i)\}$, and $I_2 = \{q(i)\}$ corresponded to the projections of all nodes on stereo images, here $\forall i \in \mathcal{N}, |I_1| = |I_2| = N$.

B. Non-Rigid Registration of Stereo

Because the data obtained from image processing were the unorganized sets of 2-D points, implementation of stereo reconstruction required matching a marker's projections on the stereo images. This process is known as image registration. The earliest approaches were based on frame to frame updating ?, although this tracking method may fail under conditions of fast movement, or be affected by noise. Another approach consists of using the coherent point drift (CPD) for non-rigid registration ?; however, this iterative optimization method is rather time-consuming with large point set. To overcome these drawbacks, inspired by guidewire tracking for concentric tubes in fluoroscopic images ?, this study proposes an active tracking algorithm for non-rigid point set registration of stereo.

Fig. ??(a) shows distribution of all markers in stereo images, in which each 2-D path contains 10 makers, resulting in 24 paths. In each image, markers appear close to each other when approaching the center, making it difficult to detect markers at the far distance. Thus, markers should be traced from the boundary area of each image to its center, corresponding to a reduction in possibility of detection. By similarity, each path can be modeled by a finite set of nodes $\vartheta_i \in \mathbb{R}^2, i = 0, 1, \dots, n_2$, where ϑ_0 and ϑ_{n_2} are the fixed nodes as the corresponding to the start and end points, as illustrated in Fig. ???. Thus, assumed that under condition of deformation, the path is strongly constrained along the line $\vartheta_0\vartheta_{n_2}$, such that the constructed path connecting nodes $\{\vartheta_i\}$ should be continuous and smooth. This hypothesis is used through-out the registration algorithm.

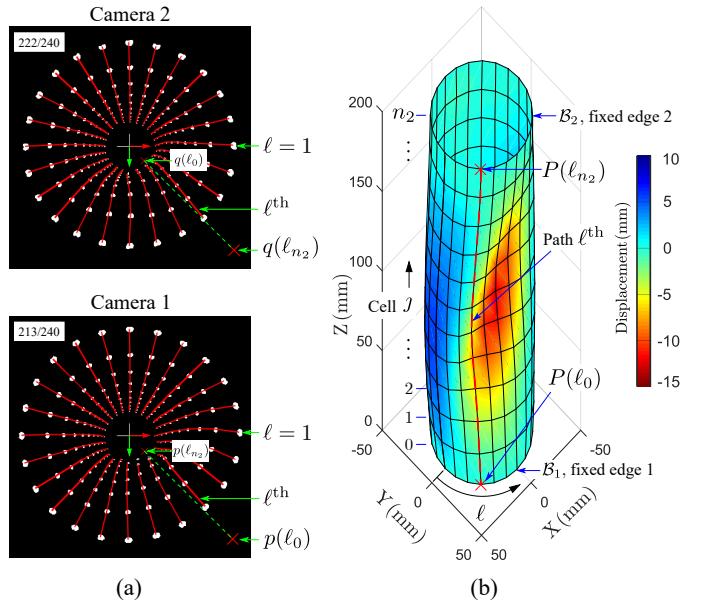


Fig. 5. A scenario for verifying the proposed stereo implementation, in which the TaCLINK was in contact with a long cylindrical surface. (a) Stereo images with the proposed intuitive path tracking algorithm. (b) Reconstruction from stereo images of the 3-D deformation of artificial skin.

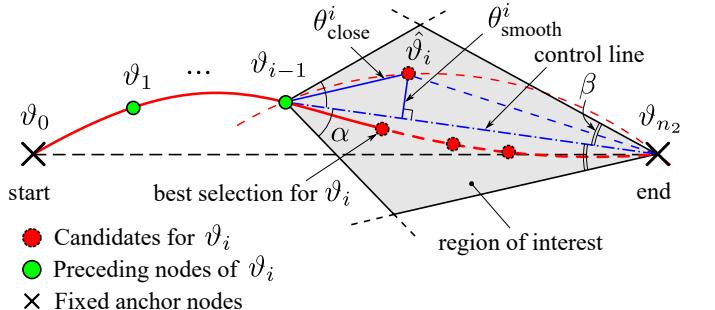


Fig. 6. The intuitive curve tracking algorithm. Using the proposed objective function with two known anchor points, ϑ_0 and ϑ_{n_2} , the tracking process can trace unknown nodes sequentially $\vartheta_1, \vartheta_2, \dots, \vartheta_{n_2-1}$.

The problem of stereo registration can be formulated as seeking the optimal set of nodes for every path. Note that these 3-D paths through disparate stereo views are projected opposite on the radial direction of two image planes (see Fig. ??(a)). That is, when observed from the boundary area to the center of stereo images, the nodes on the first and second images vary as $p(\ell_0) \rightarrow p(\ell_{n_2})$, and $q(\ell_{n_2}) \rightarrow q(\ell_0)$. These curve projections of an arbitrary path ℓ are very constrained by two fixed nodes laid on two boundary edges (i.e., $\ell_0 \in \mathcal{B}_1$ and $\ell_{n_2} \in \mathcal{B}_2$, where $\ell = 1, \dots, n_1$ (see Fig. ??(b)).

Let $\Phi \subset \mathbb{R}^2$ denotes the generalized set of detected markers in a frame of a camera. If ϑ_i is being tracked with former tracking markers as $\vartheta_1, \dots, \vartheta_{i-1}$ (see Fig. ??), then to guarantee the node ϑ_i is the node closest to node ϑ_{i-1} , the closeness term for every candidate $\hat{\vartheta}_i \in \Phi$ can be calculated as

$$\theta_{\text{close}}^i(\hat{\vartheta}_i, \vartheta_{i-1}) = \|\hat{\vartheta}_i - \vartheta_{i-1}\|, \quad (17)$$

where $\|\cdot\|$ denotes the standard Euclidean distance.

In contrast, when considering the smoothness of the path, it may be sufficient to characterize the curvature of the arc

$\vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2}$ at vertex $\hat{\vartheta}_i$ as $\frac{2 \sin \angle \vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2}}{\|\vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2}\|}$. Based on the triangle $\triangle \vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2}$ with a determined closeness term of distance $\vartheta_{i-1}\hat{\vartheta}_i$, then the curvature at $\hat{\vartheta}_i$ would depend only on the distance of $\hat{\vartheta}_i$ and the control line $\vartheta_{i-1}\vartheta_{n_2}$ (i.e., small deformation $\angle \vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2} \approx \pi$, then $\sin \angle \vartheta_{i-1}\hat{\vartheta}_i\vartheta_{n_2} \propto d(\hat{\vartheta}_i, \vartheta_{i-1}\vartheta_{n_2})$). Hence, the smoothness term can be formulated as

$$\begin{aligned} \theta_{\text{smooth}}^i(\hat{\vartheta}_i, \vartheta_{i-1}, \vartheta_{n_2}) &= d(\hat{\vartheta}_i, \vartheta_{i-1}\vartheta_{n_2}) \\ &= \frac{\|\vartheta_{i-1}\hat{\vartheta}_i \times \vartheta_{i-1}\vartheta_{n_2}\|}{\|\vartheta_{i-1}\vartheta_{n_2}\|}. \end{aligned} \quad (18)$$

This led to a proposed objective function to search ϑ_i , while satisfying the criteria of closeness and smoothness. This function can be expressed as

$$\begin{aligned} \theta_{\text{obj}}^i(\hat{\vartheta}_i, \vartheta_{i-1}, \vartheta_{n_2}) &= (1 - \lambda)\theta_{\text{close}}^i(\hat{\vartheta}_i, \vartheta_{i-1}) \\ &\quad + \lambda\theta_{\text{smooth}}^i(\hat{\vartheta}_i, \vartheta_{i-1}, \vartheta_{n_2}), \end{aligned} \quad (19)$$

where $\lambda \in [0, 1]$ is the weight that determines the relative contribution of the two factors.

In addition, obviously, not all nodes belonging to Φ are the ideal candidates for tracking ϑ_i . For example, points $\hat{\vartheta}_i$, which possess the geometric relationship $\angle \hat{\vartheta}_i\vartheta_{i-1}\vartheta_{n_2} > 90^\circ$ or $\angle \hat{\vartheta}_i\vartheta_{n_2}\vartheta_{i-1} > 90^\circ$, are unlikely to be actual marker node ϑ_i . Thus, to ensure the reliability of the results, we created a limited search area (region of interest), constrained by two searching angles with respect to the control line $\vartheta_{i-1}\vartheta_{n_2}$ with upper limits α and β shown in Fig. ???. For this purpose, only points within region of interest were considered as candidate nodes. Therefore, for all points $\hat{\vartheta}_i \in \Phi$, the optimal tracking node ϑ_i is the node that minimizes the objective function:

$$\vartheta_i = \arg \min_{\hat{\vartheta}_i \in \Phi} \theta_{\text{obj}}^i(\hat{\vartheta}_i, \vartheta_{i-1}, \vartheta_{n_2}), \quad (20)$$

$$\text{subject to: } \angle \hat{\vartheta}_i\vartheta_{i-1}\vartheta_{n_2} < \alpha, \quad \angle \hat{\vartheta}_i\vartheta_{n_2}\vartheta_{i-1} < \beta. \quad (21)$$

Note that (??) is also a criterion to determine the breakpoint of a path tracking once no candidate is within a region of interest. If Φ_1 and Φ_2 are the sets of detected markers on stereo images after image processing at each camera frame, then the tracking process would start at the boundary area and extend to the center of the image, generating the organized sets I_1 and I_2 . This algorithm running in $\mathcal{O}(n^2)$ time is shown in Algorithm ??.

C. 3-D Reconstruction

This section presents techniques used to calculate the 3-D position of all markers based on the 2-D projections provided by the proposed registration algorithm. During physical contact, uncertainty in detection of markers may be due to a lack of sight and occlusion. Three scenarios often occur in the detection of a marker: First, the 3-D coordinates of markers captured on both cameras' images can be easily computed using (??)–(??). Second, for markers missed or not well captured by one camera, $Z(\ell_j) \approx j \times h$ can be estimated as the initial value (which is sufficient due to the small axial Z -deflection of a cylinder), and the X - and Y -positions can be calculated by either camera in (??). Third, the coordinates of markers not detected on both images can be estimated by linear interpolation through

Algorithm 1 STEROREGISTRATION(Φ_1, Φ_2)

Require: Given the sets of extracted markers Φ_1 and Φ_2 on two images after image processing.

```

1: START Initialize:  $\alpha, \beta$ ,  $\{p(\ell_0), p(\ell_{n_2})\}_{\ell=1-n_1}, \{q(\ell_{n_2}), q(\ell_0)\}_{\ell=1-n_1}$   $\triangleright$  searching angles
2: for  $\ell \leftarrow 1$  to  $n_1$  do
3:    $I_1(\ell_{0:n_2}) \leftarrow \text{pTRACK}(\Phi_1, p(\ell_0), p(\ell_{n_2}), \alpha, \beta)$ 
4:    $I_2(\ell_{n_2:0}) \leftarrow \text{pTRACK}(\Phi_2, q(\ell_{n_2}), q(\ell_0), \alpha, \beta)$ 
5: end for
6: return Organized sets  $I_1$  and  $I_2$ 

function pTRACK( $\Phi, \vartheta_0, \vartheta_{n_2}, \alpha, \beta$ )  $\triangleright$  path tracking
1: Initialize:  $\vartheta \leftarrow \{\vartheta_0, \emptyset, \dots, \emptyset, \vartheta_{n_2}\}$   $\triangleright |\vartheta| = n_2 + 1$ 
2: for  $i \leftarrow 1$  to  $n_2 - 1$  do
3:    $\vartheta_i \leftarrow \text{using (??) and (??)}$ 
4:   if  $\vartheta_i = \emptyset$  then break  $\triangleright$  breakpoint
5:   Update:  $\Phi = \Phi \setminus \{\vartheta_i\}$ 
6: end for
7: return Optimal set of nodes  $\{\vartheta_i, i = 0, \dots, n_2\}$ 
end function

```

its neighbors in the same path. If j^- and j^+ are the nearest determined markers of the missing marker j on path ℓ , where $j^- < j$ and $j^+ > j$, then the linear interpolation can be formulated as

$$\mathbf{X}(\ell_j) = w\mathbf{X}(\ell_{j-}) + (1 - w)\mathbf{X}(\ell_{j+}), \quad 0 < j < n_2. \quad (22)$$

where $w = (j^+ - j)/(j^+ - j^-)$ is weight function. In practice, based on the geometric constraints of 2-D paths on stereo images and stability, the tracking parameters α and β equal $2\gamma = 30^\circ$ (Fig. ??), and weight $\lambda = 0.5$ in (??) was selected. The proposed stereo implementation worked efficiently, even in case of large deformation (e.g., Fig. ??), with the evaluation of accuracy verified in Section VII-B. However, the limitations of this system associated with elongated shape mean some markers may be occluded or overlap upon external contact e.g., multiple contacts on the same path, resulting in a 3-D reconstruction that may not well represent its actual position.

D. Stereo Self-Calibration

This vision-based sensing system requires predetermination of the parameters of the stereo camera. Although cameras can be calibrated with reference planar patterns (e.g., ?), this method is unsuitable for two facing cameras due to restrictions in pattern observation. This study utilized the initial geometrical position of markers in known positions to determine the stereo parameters with self-calibration ability to enable automatic calibration. In this scenario, $(R_{0i}, \varphi_{0i}, Z_{0i}), \forall i \in \mathcal{M}$ were regarded as a set of markers at their initial state and (r_{1i}, φ_{1i}) and (r_{2i}, φ_{2i}) as its measured projections.

1) *Estimation of Camera Parameters:* Geometrically, due to the symmetric distribution of 3-D markers around the Z -axis, the projections of markers on each image plane should be symmetric relative to the origin of its coordinates. Therefore, the principal points c_1 and c_2 can be simply determined by computing the mean centroids of all markers in pixel coordinates as $c_{x_1} = \frac{1}{n} \sum_i u_{1i}$, $c_{y_1} = \frac{1}{n} \sum_i v_{1i}$, and $c_{x_2} = \frac{1}{n} \sum_i u_{2i}$, $c_{y_2} =$

TABLE I
SELF-CALIBRATION RESULTS OF STEREO CAMERA PARAMETERS ($\mu \pm \sigma$)

Camera i	f_i [pixels]	b_i [mm]	c_{xi} [pixels]	c_{yi} [pixels]	$k_{1i} [\times 10^{-7}]$	$k_{2i} [\times 10^{-11}]$
1	226.91 ± 0.09	26.50 ± 0.04	322.76 ± 0.01	225.23 ± 0.01	27.85 ± 0.09	-14.09 ± 0.04
2	226.54 ± 0.06	226.49 ± 0.02	299.70 ± 0.00	251.89 ± 0.01	21.08 ± 0.10	-11.58 ± 0.05

$\frac{1}{n} \sum_{i \in \mathcal{M}} v_{2i}$, where $\forall i \in \mathcal{M}$, $n = 240$ represents the entire set of markers. For ease of implementation, after retrieving the principal points, a center square of area $S \times S$ pixels² around each principal point was cropped to yield a uniform image size before image processing, here $S = 400$ pixels. The principal points were redefined as $(\frac{S}{2}, \frac{S}{2})$, and the calibration objective was to determine the primary stereo parameters $\mathcal{C} = [f_1, f_2, b_1, b_2]^\top$. From the projective relationship (??), the stereo parameters can be determined by minimizing the cost function:

$$\min_{\mathcal{C}} \sum_{i \in \mathcal{M}} (\|R_{0i}f_1 - r_{1i}b_1 - r_{1i}Z_{0i}\|^2 + \|R_{0i}f_2 - r_{2i}b_2 + r_{2i}Z_{0i}\|^2). \quad (23)$$

Note that (??) requires at least two reference points to derive four unknown parameters \mathcal{C} . This minimization problem, a linear least-squares $\min_{\mathcal{C}} \|\mathcal{A}\mathcal{C} - \mathbf{b}\|^2$, can be solved with MATLAB's built-in mldivide function i.e., $\mathcal{C} = \text{mldivide}(\mathcal{A}, \mathbf{b})$.

2) *Correction of Lens Distortion*: When using wide angle lenses, it is necessary to consider the lens distortion of camera. In this camera configuration, we only assessed the term for radial distortion. If \check{r}_1 and \check{r}_2 are the actual observed coordinates, and r_1 and r_2 are the ideal radial image coordinates according to the pinhole model described in (??) and (??), then the radial division distortion model ? can be written as

$$r_1 = \frac{\check{r}_1}{1 + k_{11}\check{r}_1^2 + k_{21}\check{r}_1^4}, \quad (24)$$

$$r_2 = \frac{\check{r}_2}{1 + k_{12}\check{r}_2^2 + k_{22}\check{r}_2^4}, \quad (25)$$

where $\{k_{11}, k_{21}\}$ and $\{k_{12}, k_{22}\}$ are the coefficients of the radial distortion of the two camera lenses, respectively.

To determine the distorted coefficients, the preliminary parameters \mathcal{C} in (??) were estimated by ignoring distortion. Then, the ideal image coordinates r_1 and r_2 were determined using the equations in (??). Thus, from (??) and (??), the coefficients $\mathcal{L} = [k_{11}, k_{21}, k_{12}, k_{22}]^\top$ can be estimated by minimizing the cost function as

$$\min_{\mathcal{L}} \sum_{i \in \mathcal{M}} (\|r_{1i}\check{r}_{1i}^2k_{11} + r_{1i}\check{r}_{1i}^4k_{21} + (r_{1i} - \check{r}_{1i})\|^2 + \|r_{2i}\check{r}_{2i}^2k_{12} + r_{2i}\check{r}_{2i}^4k_{22} + (r_{2i} - \check{r}_{2i})\|^2). \quad (26)$$

Once \mathcal{L} is obtained, the parameters \mathcal{C} can be recomputed by recalling (??), with this procedure repeated until reaching convergence. The root-mean-square error (RMSE) of coordinates was used to assess the calibration process, as shown in Fig. ???. This figure shows rapids convergence of calibrated coordinates RMSE within fewer than four iterations and a significant reduction in erroneous results. The calibration process was performed a total of 20 times, yielding the camera parameters

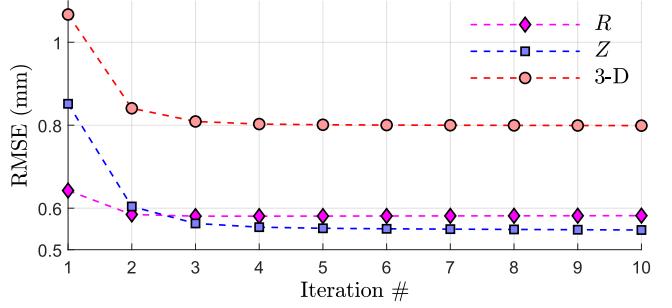


Fig. 7. Calibration results for each iteration: RMSE of coordinates for estimating stereo parameters and lens distortion.

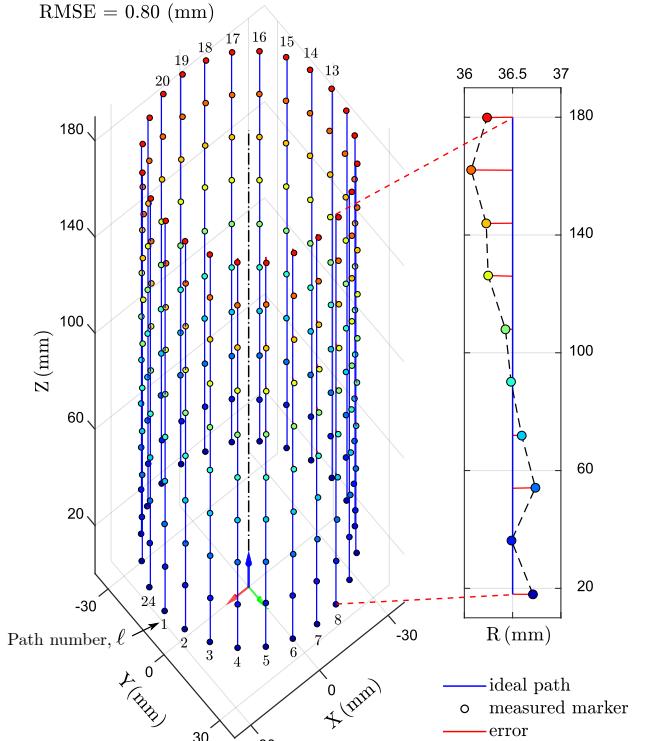


Fig. 8. Illustration of the initial 3-D position of markers (roundly marked) using the calibrated parameters.

and distortion coefficients shown in Table I. These results were used to verify some parameters of the stereo camera configuration. For example, the length of artificial skin after calibration can be calculated as $L = b_2 - b_1 \approx 200.0$ mm (see Fig. ??), while the designed value is 198.0 mm (see Fig. ??), i.e., the absolute was 2.0 mm, or 1.01%. With calibrated parameters, Fig. ?? illustrates the initial geometric 3-D paths of markers and their measured locations. The average RMSE of the calibrated coordinates R , Z and 3-D were 0.58 mm,

0.55 mm and 0.80 mm, respectively.

By referring to the pattern of markers, we could determine the intrinsic and extrinsic parameters of the stereo camera and lens distortion. Note that although light traveling through the acrylic tube would be refracted (incident and refracted rays are shifted), the system's calibration used refracted images, resulting in the parameters of lens distortion covered this effect.

V. FE MODEL OF ARTIFICIAL SKIN

This section introduces the finite element model of the skin to establish the relationship between nodal displacements of the markers and external forces. For ease of modeling, the skin was assumed to be linearly elastic. The skin was modeled by the flat shell elements combining membrane and bending behaviors based on Reissner-Mindlin theory ?. Only the static model was considered, ignoring the gravitational effect on the skin, as we hypothesized that the effect of gravity on the skin was much smaller than the effects of the range of measured forces.

A. Meshing and Geometry of the Element

Owing to its 3-D symmetric shape, the skin can be equally discretized into a mesh of $N_e = 264$ elements (Fig. ??(a)). The rectangular domain Ω of a shell element is estimated as $2a = h$ and $2b = 2R_0 \sin(\gamma/2)$ (Fig. ??(c)). To build the elemental and global vectors and matrices, it was necessary to define mesh topology providing a crucial nodal connectivity of elements. By positioning an element $e \in \{1, \dots, N_e\}$ by the node associated with the left bottom corner (Fig. ??(b)). The element at (ℓ, j) is numbered $e = \ell_j$ (??) and formed by four nodes distributed counter-clockwise, $\hat{1}(\ell, j+1)$, $\hat{2}(\ell, j)$, $\hat{3}(\ell+1, j)$, and $\hat{4}(\ell+1, j+1)$, where $j = 0, \dots, n_2$, $\ell = 1, \dots, n_1 - 1$; if $\ell = n_1$, then nodes 3 and 4 are $\hat{3}(1, j)$ and $\hat{4}(1, j+1)$, respectively. Fig. ??(a) shows the FE mesh of the skin with node and element numbering.

For ease of computation, the displacement field of each element should be described in its local domain. In this paper, besides the global coordinate system, the four-node shell element was formulated in the local coordinates $\{\hat{x}, \hat{y}, \hat{z}\}$, and the natural coordinates $\{\xi, \eta\}$ shown in Fig. ??(c). If the \hat{x} -axis is defined as the opposite of the Z -axis, and the edge 2-3 defines the \hat{y} -direction, then the \hat{z} -axis is obtained by the cross product of \hat{x} - and \hat{y} -axes. Each shell element node has five degrees of freedom i.e., in-plane displacements $u_{\hat{x}}, u_{\hat{y}}$, lateral displacement u_z , and rotations $\theta_{\hat{x}}$ and $\theta_{\hat{y}}$ of \hat{z} -axis in the planes $\hat{x}\hat{z}$ and $\hat{y}\hat{z}$, respectively ?.

B. FE Equations and Simulation

The state equation of an element in static equilibrium can be determined from its local domain (Fig. ??(c)) by applying the principle of virtual work (PVW) as (see Table II and Appendix for a detailed explanation and derivation of variables)

$$\underbrace{\int_{\Omega} \delta \hat{\epsilon}^T \hat{\sigma} d\Omega}_{\text{Internal virtual strain energy}} = \underbrace{\int_{\Omega} \delta \hat{u}^T \hat{s} d\Omega + \delta \hat{d}^T \hat{f}^{\text{ext}}}_{\text{External virtual work}}. \quad (27)$$

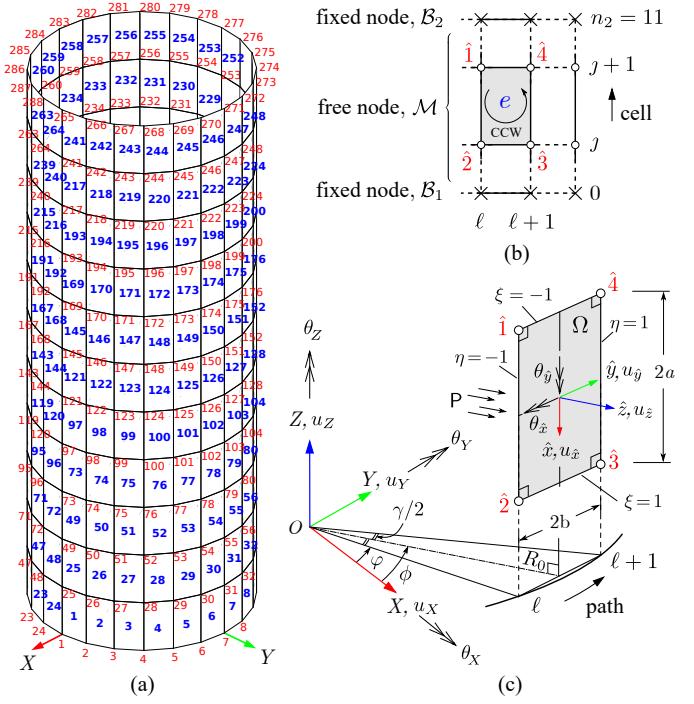


Fig. 9. Three-dimensional FE model for artificial skin. (a) Finite element mesh of $N_e = 264$ elements and $N = 288$ nodes. (b) Mesh topology. (c) Geometry of a rectangular flat shell element.

The element equation in equilibrium can be expressed as

$$\hat{k}\hat{d} = \hat{f}_{\text{pressured}} + \hat{f}^{\text{ext}}. \quad (28)$$

The element equation (??) represents the relationship between nodal displacements \hat{d} , the distributed force $\hat{f}_{\text{pressured}}$ due to inner air pressure P (??), and the external concentrated force \hat{f}^{ext} . These elements locate on different local orientations, thus, to build the global equation, the element equation (??) must be expressed in the global coordinates as

$$k\mathbf{d} = \mathbf{f}_{\text{pressured}} + \mathbf{f}^{\text{ext}}. \quad (29)$$

where details of its derivation can be found in Appendix.

Finally, the system equation of the FE model is established by assembling together all (??) of elements $e = 1, 2, \dots, N_e$ based on the connected nodes in the meshing topology as

$$\mathbf{K}\mathbf{d} = \mathbf{F}_{\text{pressured}} + \mathbf{F}^{\text{ext}}. \quad (30)$$

Besides, as the skin of the TacLINK is clamped at both ends, with the degrees of freedom of all nodes set at zero, then the boundary conditions of the FE equation (??) are expressed as

$$\mathbf{D}_i = \mathbf{0}, \forall i \in \mathcal{B}_1 \cup \mathcal{B}_2. \quad (31)$$

To verify the ability of the FE model to describe the physical characteristics of the skin, the model was built numerically and implemented in MATLAB. Young's modulus E of silicone material was calibrated to be 0.1 MPa (see Section VII-C) and Poisson's ratio ν was set as 0.5. TacLINK was successfully simulated under multiple contacts at different locations with varied pressure. The displacement vector \mathbf{D} in (??) was found to be solvable with the boundary conditions (??), as indicated in ?. Fig. ?? shows results of subjecting the TacLINK to a

TABLE II
NOMENCLATURE FOR SECTION V AND VI

Symbol	Description
$\hat{\mathbf{u}}$	Displacement field vector, $= [u_{\hat{x}} \ u_{\hat{y}} \ u_{\hat{z}} \ \theta_{\hat{x}} \ \theta_{\hat{y}}]^T$.
$\hat{\mathbf{s}}$	Surface load of inner air pressure P , $= [0 \ 0 \ P \ 0 \ 0]^T$.
$\hat{\boldsymbol{\epsilon}}$	$\in \mathbb{R}^8$ Strain vector.
$\hat{\boldsymbol{\sigma}}$	$\in \mathbb{R}^8$ Stress vector.
$\hat{\mathbf{k}}, \mathbf{k}$	$\in \mathbb{R}^{20 \times 20}, \in \mathbb{R}^{24 \times 24}$ Element stiffness matrices.
$\hat{\mathbf{d}}, \mathbf{d}$	$\in \mathbb{R}^{20}, \in \mathbb{R}^{24}$ Nodal displacement vectors of an element,
$\hat{\mathbf{d}}_i$	$= [u_{\hat{x}_i} \ u_{\hat{y}_i} \ u_{\hat{z}_i} \ \theta_{\hat{x}_i} \ \theta_{\hat{y}_i}]^T, i = 1-4$,
\mathbf{d}_i	$= [u_{X_i} \ u_{Y_i} \ u_{Z_i} \ \theta_{X_i} \ \theta_{Y_i} \ \theta_{Z_i}]^T, i = 1-4$.
$\hat{\mathbf{f}}, \mathbf{f}$	$\in \mathbb{R}^{20}, \in \mathbb{R}^{24}$ Nodal force vectors of an element,
$\hat{\mathbf{f}}_i$	$= [f_{\hat{x}_i} \ f_{\hat{y}_i} \ f_{\hat{z}_i} \ m_{\hat{x}_i} \ m_{\hat{y}_i}]^T, i = 1-4$,
\mathbf{f}_i	$= [f_{X_i} \ f_{Y_i} \ f_{Z_i} \ m_{X_i} \ m_{Y_i} \ m_{Z_i}]^T, i = 1-4$.
$\hat{\mathbf{K}}$	$\in \mathbb{R}^{6N \times 6N}$ Global stiffness matrix.
$\hat{\mathbf{D}}$	$\in \mathbb{R}^{6N}$ Global displacement vector,
\mathbf{D}_i	$= [u_{X_i} \ u_{Y_i} \ u_{Z_i} \ \theta_{X_i} \ \theta_{Y_i} \ \theta_{Z_i}]^T, i = 1-N$.
$\hat{\mathbf{F}}$	$\in \mathbb{R}^{6N}$ Global force vector,
\mathbf{F}_i	$= [F_{X_i} \ F_{Y_i} \ F_{Z_i} \ M_{X_i} \ M_{Y_i} \ M_{Z_i}]^T, i = 1-N$.

Overscript $\hat{\cdot}$ indicates variables expressed in the local coordinates \hat{x} , \hat{y} , and \hat{z} . All quantities are presented in SI units.

normal force 0.2 N at node #127(7, 5) (Fig. ??(a)), and with the application of additional inner pressure of 0.5 kPa (Fig. ??(b)). These findings indicate that the FE model can provide a realistic response of the skin, including 3-D deformation, stress-strain, and reaction force at fixed ends. Based on this model, we can evaluate the overall stiffness of the skin to optimize the design and choose suitable skin material for particular applications.

VI. FE MODEL-BASED TACTILE FORCE SENSING

Tactile force \mathbf{F}^{tac} can be defined as the external force \mathbf{F}^{ext} acting on the skin surface, excluding the distributed forces generated by inner pressure $\mathbf{F}_{\text{pressured}}$. Based on (??), we could derive a straightforward relationship as

$$\mathbf{F}^{\text{tac}} = \mathbf{K}\mathbf{D}_{\text{measured}} - \mathbf{F}_{\text{pressured}}. \quad (32)$$

The scope of this paper considers only the measured translations (u_X, u_Y, u_Z) with forces (F_X, F_Y, F_Z), ignoring the degrees of freedom of nodal rotations and moments. Thus, the measured nodal displacement vector $\mathbf{D}_{\text{measured}}$ can be computed from the output vector \mathbf{X} of the stereo camera as

$$\mathbf{D}_{\text{measured},i} := [\Delta\mathbf{X}_i^T \ \underbrace{0 \ 0 \ 0}_{\text{rotations}}]^T, \forall i \in \mathcal{N}. \quad (33)$$

where $\Delta\mathbf{X} = \mathbf{X} - \mathbf{x}$, and \mathbf{x} is the coordinates vector of nodes in the undeformed state of artificial skin.

The computed tactile force vector \mathbf{F}^{tac} is based on sampling a large number of points i.e., $N = 288$ nodes with $3 \times N$ force components. In practice, errors during 3-D reconstruction induce errors in force calculations. Based on cylindrically shaped skin, we simulated evaluation of this effect in the tangential, axial and radial directions. For ease of analysis, we assumed that δ_{err} mm was the absolute error in every direction. Node #127 (see Fig. ??) was chosen and constrained in the tangential (X -axis), radial (Y -axis) and axial (Z -axis) deflections as $\mathbf{D}_{127} = [\delta_{\text{err}}, \delta_{\text{err}}, \delta_{\text{err}}, 0, 0, 0]^T$. Solving with (??) and (??), we calculated the reaction force at node #127 as $\mathbf{F}_{127}^{\text{ext}} = [0.08\delta_{\text{err}}, 0.02\delta_{\text{err}}, 0.30\delta_{\text{err}}, 0, 0, 0]^T$. These findings

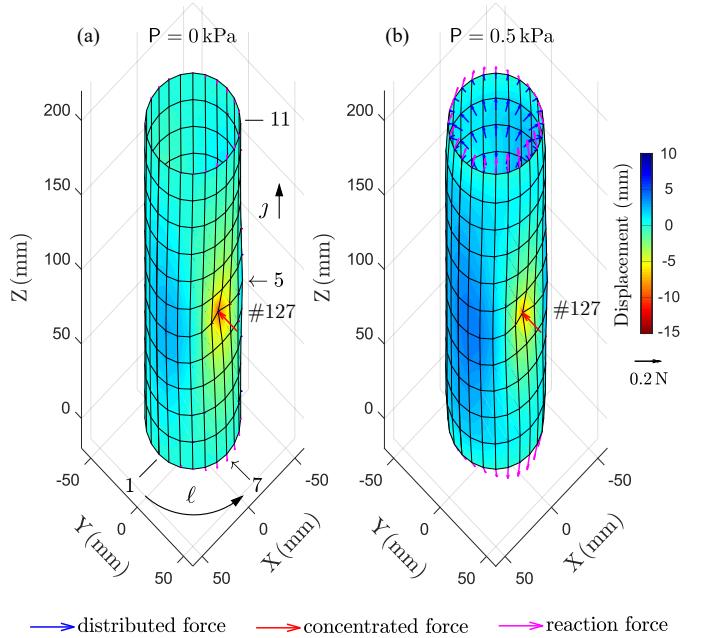


Fig. 10. FE model simulation results for soft artificial skin. (a) Skin deformation by a normal force 0.2 N at node #127(7,5). (b) Response of skin to additional inner air pressure $P = 0.5$ kPa.

indicated that spatial errors result in much higher errors of force in the tangential (~ 4 times) and axial (~ 15 times) directions than in the radial direction. Under conditions of normal contact with the cylindrical skin surface, the radial displacements should be greater than displacements in other directions. Thus, to ensure the reliability of each nodal force, only the radial component was considered as

$$F_R = F_X \cos \varphi + F_Y \sin \varphi. \quad (34)$$

Also, for practical reasons, we only considered the computed force F_R above a certain threshold i.e., $f_{\text{thresh}} = 0.02\delta_{\text{err}}$ N. Finally, the tactile forces were regarded as the concentrated forces acting at free nodes $\mathbf{F}_i^{\text{tac}}, \forall i \in \mathcal{M}$, and as the reaction forces acting at two edges $\mathbf{F}_i^{\text{tac}}, \forall i \in \mathcal{B}_1 \cup \mathcal{B}_2$.

VII. EXPERIMENTAL VALIDATION

A. Experimental Platform

The experimental platform set up to calibrate and evaluate the operation of the TacLINK is shown in Fig. ???. Within the platform, TacLINK is fixed in a rigid vertical position. The inner pressure was adjusted using a pneumatic throttle valve, and measured with a pressure sensor (MPXV7007, NXP Inc., USA) through a 16-bit ADC data acquisition board (USB-231, Measurement Computing Corp., USA), with a measurement resolution of pressure of 0.85 Pa. To provide precise displacement of external contact with the skin, we set up a 6-DOF robotic arm (VP-6242, Denso Robotics, Japan) which can precisely displace with high repeatability of ± 0.02 mm. A probe was attached to the center of the robot's end-effector through a 3-axes force sensor (USL06-H5-50N, Tec Gihan, Japan). The robot moved the probe in plane OYZ and pushed it horizontally against the skin at different nodes on the seventh path ($\ell = 7$).

To accurately push the entire cross-sectional area of the marker, we utilized a cylindrical probe as a standard screw M3 with a larger diameter $D_{\text{probe}} = 5.5$ mm.

The experiment was run on a desktop PC with an i7-7700 processor at 3.60 GHz and 8GB of RAM. The system performed in real-time at about 13 Hz, with time to capture stereo pictures ~ 0.003 s, image-processing ~ 0.025 s, stereo registration ~ 0.037 s, and both 3-D shape and force reconstruction completed within 0.0002 s, with MATLAB code.

B. Assessment of the Accuracy of 3-D Reconstruction

The performance of stereo-based 3-D reconstruction was verified by evaluating the measurement error of radial displacement of free nodes on a path. Radial direction was found to play the most important role (see Section VI), and it was challenging to determine the true value of Z -deflection. Specifically, the robot arm was controlled, allowing the probe to create radial displacements ΔR of -5 and -10 mm every ten nodes on the seventh path (Fig. ??). Ten trials were performed at an inner air pressure of zero. Fig. ?? shows the measured radial deflections of ten cells $j = 1, 2, \dots, 10$ compared with baseline robot motion ΔR . Because of geometric constraints on the skin, two nodes located near two ends $j = 1, 10$ could only be displaced by -5 mm. The experiment results of this testing showed that the absolute errors were below 0.7 mm, corresponding to a full-scale error $\sim 5\%$ FS (with FS 15 mm). This error may also derive from several parts of the system (e.g., poor fabrication and installation tolerances with soft materials). These findings indicate that the proposed vision system is both efficient (see Fig. ??) and accurate in estimating deformation of the 3-D skin.

C. Young's Modulus Calibration

The elastic Young's modulus E of the skin material was estimated from a compression test performed under quasi-static conditions. Air was provided at different levels of pressure $P \in \mathcal{P} = \{0.5, 1.0, 1.5\}$ kPa for a total of 30 trials. During inflation, a program automatically recorded data at these pressures, enabling estimation of parameter E by minimizing the difference between experimentally determined deformation and deformation predicted by the linear FE model as

$$\min_E \left(\sum_{P \in \mathcal{P}} \left\| \mathbf{D}_{\text{measured}} - \frac{E^*}{E} \mathbf{D}_{\text{FEM}^*} \right\|^2 \right) \quad (35)$$

$$\text{s.t. } E = E^* \frac{\sum_{P \in \mathcal{P}} \|\mathbf{D}_{\text{FEM}^*}\|^2}{\sum_{P \in \mathcal{P}} \mathbf{D}_{\text{measured}} \cdot \mathbf{D}_{\text{FEM}^*}} \quad (36)$$

where $\mathbf{D}_{\text{measured}}$ and $\mathbf{D}_{\text{FEM}^*}$ are the nodal displacement vectors obtained from experimental data and the FE model (with assigned Young's modulus $E^* = 1$ MPa), respectively.

Because the sum of the squares errors in (??) is a quadratic function of $\frac{1}{E}$, E can be easily derived as in (??). Based on the symmetric shape of the FE model under pressure (i.e., similar curvature of every path), the experimentally determined displacement of all paths was averaged to calculate (??) on a path. Fig. ?? shows the experimental displacements for ten nodes on a path along the skin, together with the theoretical curvature (lines) of the FE model in response to varying levels

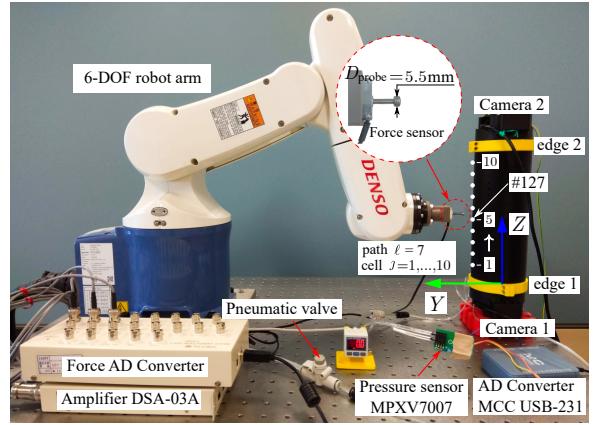


Fig. 11. The experimental platform with all related equipment for measurement. A robot arm was equipped to push the elastic skin through a probe attached to the end-effector through a 3-axes force sensor.

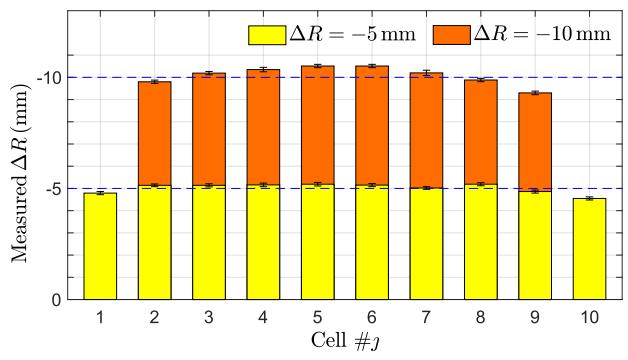


Fig. 12. Measured displacements of ten nodes on a path were compared with true deflections created by robot motion.

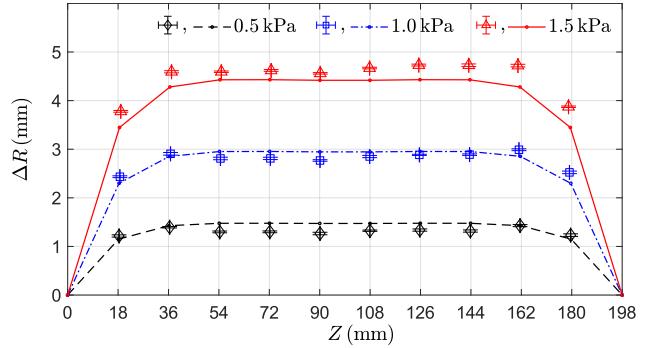


Fig. 13. Comparison of experimental data (solids) and calibrated FE model (lines) of skin curvature in response to applied inner pressures of 0.5, 1, and 1.5 kPa, estimating Young's modulus of the silicone material (Ecoflex 00-50).

of inner pressure. The experimental data fit those of the FE model well with an estimated E value of 0.1 MPa and an average RMSE for nodal displacements of 0.38 mm.

D. Evaluation of FE Model-based Tactile Force Sensing

1) *Single-Point Contact*: The finite element model of the skin (??) enables TacLINK to estimate the external acting force distribution on the whole skin body (??) (i.e., $\mathbf{F}_i^{\text{tac}}, \forall i \in \mathcal{M}$). This experiment evaluated the tactile force-sensing ability and assessed the characteristics of TacLINK when interacting at a

single point under varying inner air pressures. The robot arm was manipulated so that the probe pushing the skin at node #127 along the Y-axis (see Fig. ??) created two desired radial displacements $\Delta R \in \{-5, -10\}$ mm in different conditions of inner air pressure $P \in \{0, 0.5, 1.0, 1.5\}$ kPa for ten trials. Note that although the ideal probe should be narrow enough to represent a point load, a small probe may cause undesired deformation in the thickness direction at the contact point. To mitigate these issues, a cylindrical probe of diameter $D_{\text{probe}} = 5.5$ mm was employed. The tactile force of TacLINK was also simulated by setting an additional boundary condition for (??) on the Y-axis as $\mathbf{D}_{127,Y} = \Delta R$. Fig. ?? shows the comparison between the probe force and the tactile forces at node #127 obtained from TacLINK ($\mathbf{F}_{127}^{\text{tac}}$) and simulation ($\mathbf{F}_{127}^{\text{ext}}$) at each level of deflection ΔR and inner air pressure P .

In the absence of inner pressure (0 kPa), the resultant forces matched well at every deflection. However, when pressure was applied, the nodal tactile force of TacLINK increased linearly but was significantly lower than the probe force. This difference may be due to the probe not being small enough to represent an ideal point load contact scenario in simulation (see Fig. ??). The inner air pressure acting on the inner side of the skin gave rise to tension stresses (circumferential $\sigma_{\hat{y}}$ and longitudinal $\sigma_{\hat{x}}$ stresses, see Fig. ??(c)) in its wall, resulting in a highly increased probe force to move the probe with contact area $S_{\text{probe}} = \frac{\pi}{4} D_{\text{probe}}^2 = 23.8$ mm² that caused greater deformed contact area surrounding the assigned node. The mesh size ($2a = 18$ mm, $2b = 9.5$ mm) was not dense enough to describe actual deformation in the contact area caused by the probe. The computed tactile force distribution of TacLINK representing the probe force would also be distributed on the adjacent nodes, and the largest nodal tactile force should be at node #127 shown in Fig. ??(a). Consequently, single point force contact scenario using a probe attached in a loadcell is rather challenging (i.e., create an ideal point load and perceive the entire applied force by a nodal tactile force), because the probe has a finite area and the skin is soft, etc. The obtained results indicate that TacLINK is capable of determining the distribution of applied forces.

Varying the probe forces revealed actual reaction of artificial skin e.g., the maximum probe forces per unit area ($\equiv \frac{F_{\text{probe}}}{S_{\text{probe}}}$) at 0 kPa and 1.5 kPa were 0.84 N/cm² and 5.6 N/cm², respectively. As the inner pressure increased, so did the slope of probe force, thus by controlling the inner pressure we could alter remarkably the nominal stiffness of the skin and change the interactive interface of TacLINK.

2) *Multiple Large-Areas Contact*: Because the aim of TacLINK is to sense contacts with humans, we performed an experiment in which a human used two fingers to randomly touch different areas of the artificial skin. Fig. ??(b) shows the robust reconstruction of both 3-D deformation and external acting forces in large and multi-contact regions. Tactile forces were distributed only in contact regions; but were absent from other regions, even those that showed deformation. Generally, under conditions of physical interaction in any position, TacLINK registers detailed 3-D deformation of the skin surface caused by external forces, the location and intensity of which are determined by tactile force distribution. It is noteworthy that

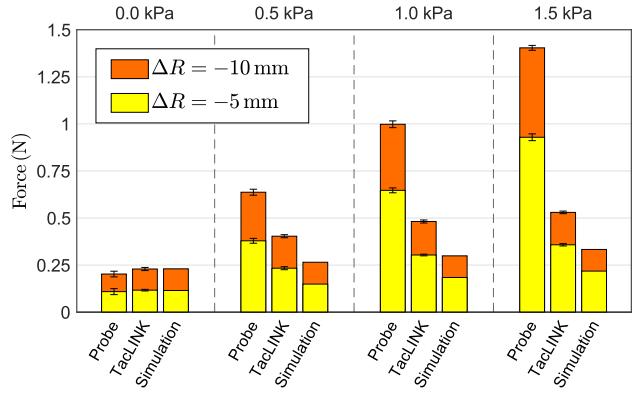


Fig. 14. Probe forces applied yielding radial deflections ΔR at node #127 compared with the nodal tactile forces at node #127 obtained from TacLINK and simulation under different inner air pressures.

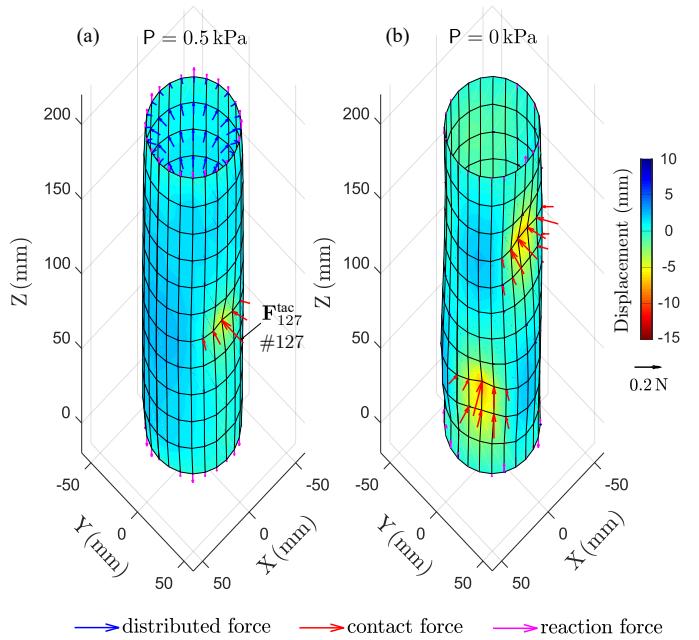


Fig. 15. Experimental results of tactile force reconstruction on the three-dimensional body of the skin surface. (a) Test with single-point (node) contact at node #127. (b) Performance with multiple large-area contacts.

previous vision-based tactile force techniques with markers, such as machine learning ?, and experimental calibration ?, ?, could estimate the total force applied not determine the distribution of applied forces in scenarios such as under large area or multiple point contacts. Also, tactile skin designs with discrete sensing elements ?, ? cannot detect touch in areas not occupied by such elements, however, using the FEM to reconstruct forces over a continuous area, the proposed sensing system can simultaneously perceive tactile forces throughout the skin surface (see supplementary video).

VIII. DISCUSSION

A. Limitations

Using only two co-axial cameras on a simple structure, the proposed system using FEM can efficiently construct the distribution of forces acting on the elongated surface of soft artificial

skin. However, during tracking the skin surface, presence of reflected bright regions or occlusion due to multiple contacts on the same path, may impede image processing and affect the performance of the visual-based system. Since FEM was originally formulated based on the principle of virtual work in elasticity (??), the performance of force computation highly depends on the element model (strains $\hat{\epsilon}$ and stresses $\hat{\sigma}$) and mesh size ($2a, 2b$) representing the strain energy of actual skin deformation (especially in contact regions). Thus, in this paper, force reconstruction might not well represent force intensity, especially over large contact areas with inner air pressure applied. Further design improvements and investigations are needed to overcome these limitations.

B. Technical Issues

The structural rigidity of TacLINK is ensured by a hollow tube with transparent acrylic materials. Although its working load limit can be reduced compared with metals, we can determine suitable dimensions of the tube for specific applications. If M_d is the working load design moment (combined bending and torsion), and $[\sigma]$ represents permissible stress, then the strength condition to choose inner d and outer D diameters of the tube can be expressed as ?

$$W = \frac{\pi}{32} \frac{D^4 - d^4}{D} \geq \frac{M_d}{[\sigma]}, \quad \text{with } [\sigma] = \frac{\sigma_u}{k}, \quad (37)$$

where W , σ_u , and k correspond to section modulus of the tube, ultimate strength of materials, and safety factor, respectively. In this paper, with the parameters $D = 30$ mm, $d = 24$ mm, $\sigma_u = 87$ MPa, and $k = 1.5$, the working load of TacLINK is $M_d \leq 90$ Nm.

For commercialization, the hardware of TacLINK should be compact and capable of high speed operation, such as field-programmable gate array (FPGA), its firmware or software is responsible for processing all tasks regenerating tactile data in real-time and transmitting via a serial bus system, on which a network of TacLINK can share to communicate with user applications. Also, in practice, although running system cables discreetly may be required, a connecting cable can be embedded inside TacLINK.

IX. CONCLUSION

This study describes the developed TacLINK for robotic links with large-scale tactile skin. We showed its feasibility of design and robustness in tactile reconstruction. The system is low cost, affording simple scalable structure, and real-time performance. Especially, TacLINK is not only equipped with sensing ability but can react changing its form and stiffness during interaction with surroundings. Deployment of TacLINK with highly deformable skin and tactile force feedback will enable robots to interact safely and intelligently.

The proposed vision-based force sensing technique can be utilized on shell structures with different shapes, where a set of cameras can be installed to track the displacement and FEM provides the structural stiffness matrix, that enables implementation of this technique on other parts of robots, such as fingers, legs, chests, and heads, and even in robotic prosthetics for

humans, paving the way for development of large-scale tactile sensing devices. Future work will develop the control strategies deploying tactile feedback in various interaction tasks, such as safety control, dexterous manipulation, human-robot interaction, and interpreting tactile feelings.

APPENDIX FINITE ELEMENT EQUATIONS

A four-node rectangular shell was adopted to discretize the artificial skin. The displacement field within an element can be interpolated by: (hereafter, subscripts $i, j = 1, 2, 3, 4$ denote the local nodal indices)

$$\hat{\mathbf{u}} = [u_{\hat{x}} \quad u_{\hat{y}} \quad u_z \quad \theta_{\hat{x}} \quad \theta_{\hat{y}}]^T = \sum_{i=1}^4 \mathbf{N}_i \hat{\mathbf{d}}_i, \quad (38)$$

where $\mathbf{N}_i = \text{diag}(N_i, N_i, N_i, N_i, N_i)$ is a matrix of shape functions $N_i = \frac{1}{4}(1 + \xi_i \xi)(1 + \eta_i \eta)$, with $\xi, \eta \in [-1, 1]$, $\xi_{1,4} = -1, \xi_{2,3} = 1, \eta_{1,2} = -1, \eta_{3,4} = 1$ (see Fig. ??(c)). As the rectangular element $\hat{x} = a\xi$ and $\hat{y} = b\eta$, the Cartesian derivatives of the shape function are: $\partial N_i / \partial \hat{x} = \frac{1}{4a} \xi_i (1 + \eta_i \eta)$, and $\partial N_i / \partial \hat{y} = \frac{1}{4b} \eta_i (1 + \xi_i \xi)$. The essential strains $\hat{\epsilon}$ and stresses $\hat{\sigma}$ of flat shell ? can be derived from displacement field (??) are shown below:

$$\hat{\epsilon} = \begin{bmatrix} \hat{\epsilon}_m \\ \hat{\epsilon}_b \\ \hat{\epsilon}_s \end{bmatrix} = \begin{bmatrix} \epsilon_{\hat{x}} \\ \epsilon_{\hat{y}} \\ \gamma_{\hat{x}\hat{y}} \\ \kappa_{\hat{x}} \\ \kappa_{\hat{y}} \\ \kappa_{\hat{x}\hat{y}} \\ \gamma_{\hat{x}\hat{z}} \\ \gamma_{\hat{y}\hat{z}} \end{bmatrix} = \begin{bmatrix} \frac{\partial u_{\hat{x}}}{\partial \hat{x}} \\ \frac{\partial u_{\hat{y}}}{\partial \hat{y}} \\ \frac{\partial u_{\hat{x}}}{\partial \hat{y}} + \frac{\partial u_{\hat{y}}}{\partial \hat{x}} \\ \frac{\partial \theta_{\hat{x}}}{\partial \hat{x}} \\ \frac{\partial \theta_{\hat{y}}}{\partial \hat{y}} \\ \frac{\partial \theta_{\hat{x}}}{\partial \hat{y}} + \frac{\partial \theta_{\hat{y}}}{\partial \hat{x}} \\ \frac{\partial u_z}{\partial \hat{x}} - \theta_{\hat{x}} \\ \frac{\partial u_z}{\partial \hat{y}} - \theta_{\hat{y}} \end{bmatrix} = \sum_{i=1}^4 \mathbf{B}_i \hat{\mathbf{d}}_i, \quad (39)$$

$$\hat{\sigma} = \mathbf{C} \hat{\epsilon}, \quad (40)$$

where the subscripts m , b , and s are the relevant membrane, bending and transverse shear vectors/matrices, respectively. The strain matrices \mathbf{B}_i are:

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{B}_{m_i} \\ \mathbf{B}_{b_i} \\ \mathbf{B}_{s_i} \end{bmatrix} = \begin{bmatrix} \frac{\partial N_i}{\partial \hat{x}} & 0 & 0 & 0 & 0 \\ 0 & \frac{\partial N_i}{\partial \hat{y}} & 0 & 0 & 0 \\ \frac{\partial N_i}{\partial \hat{y}} & \frac{\partial N_i}{\partial \hat{x}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial N_i}{\partial \hat{x}} & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial N_i}{\partial \hat{y}} \\ 0 & 0 & 0 & \frac{\partial N_i}{\partial \hat{y}} & \frac{\partial N_i}{\partial \hat{x}} \\ 0 & 0 & -\frac{\partial N_i}{\partial \hat{x}} & -N_i & 0 \\ 0 & 0 & \frac{\partial N_i}{\partial \hat{y}} & 0 & -N_i \end{bmatrix} \quad (41)$$

The constitutive matrices of isotropic materials are:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_m & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_s \end{bmatrix}, \quad \mathbf{C}_s = \frac{\kappa Et}{2(1+\nu)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{C}_m = \frac{Et}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix}, \quad \mathbf{C}_b = \frac{t^2}{12} \mathbf{C}_m, \quad (42)$$

where E , ν and t are Young's modulus, Poisson's ratio and shell thickness, respectively. For isotropic materials, the shear

correction factor κ equals 5/6. Substituting (??)–(??) into (??) results in (??) (see ?). The submatrices $\hat{\mathbf{k}}_{ij} \in \mathbb{R}^{5 \times 5}$ of the local stiffness matrix $\hat{\mathbf{k}}$ linking nodes i and j have the form:

$$\hat{\mathbf{k}}_{ij} = \int_{\Omega} \mathbf{B}_i^T \mathbf{C} \mathbf{B}_j d\Omega = ab \int_{-1}^1 \int_{-1}^1 \mathbf{B}_i^T \mathbf{C} \mathbf{B}_j d\xi d\eta. \quad (43)$$

Equation (??) can be approximated by numerical integration as indicated in ?. The distributed force $\hat{\mathbf{f}}_{\text{pressured},i}$ acts on node i of element due to the surface load $\hat{\mathbf{s}}$ is:

$$\hat{\mathbf{f}}_{\text{pressured},i} = \int_{\Omega} \mathbf{N}_i \hat{\mathbf{s}} d\Omega = [0 \ 0 \ abP \ 0 \ 0]^\top. \quad (44)$$

The relationship between displacements in the local and global coordinates is described by the transformation matrix $\mathbf{L}_i^{(e)}$ as

$$\hat{\mathbf{d}}_i^{(e)} = \mathbf{L}_i^{(e)} \mathbf{d}_i^{(e)}, \quad \mathbf{d}_i^{(e)} = [\mathbf{L}_i^{(e)}]^\top \hat{\mathbf{d}}_i^{(e)}. \quad (45)$$

Also, the stiffness matrix and force vector can be written as

$$\mathbf{k}_{ij}^{(e)} = [\mathbf{L}_i^{(e)}]^\top \hat{\mathbf{k}}_{ij}^{(e)} \mathbf{L}_j^{(e)}, \quad \mathbf{f}_i^{(e)} = [\mathbf{L}_i^{(e)}]^\top \hat{\mathbf{f}}_i^{(e)}. \quad (46)$$

As the flat element, the transformation matrices $\mathbf{L}_i^{(e)}$ of the element e are identical for 4-nodes, formulated as

$$\mathbf{L}_i^{(e)}(\phi) = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ -s_\phi & c_\phi & 0 & 0 & 0 & 0 \\ c_\phi & s_\phi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & s_\phi & -c_\phi & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad (47)$$

where s_ϕ and c_ϕ denote $\sin \phi$ and $\cos \phi$, respectively, with angle ϕ of an element e at path ℓ is $\phi = (\ell - \frac{1}{2})\gamma$, see Fig. ??(c).

REFERENCES

- [1] M. Soni, R. Dahiya, “Soft eSkin: distributed touch sensing with harmonized energy and computing,” *Phil. Trans. R. Soc. A*, vol. 378, no. 2164, p. 20190156, 2019.
- [2] B. Shih, D. Shah *et al.*, “Electronic skins and machine learning for intelligent soft robots,” *Sci. Robot.*, vol. 5, no. 41, 2020.
- [3] Zang, Y., Zhang, F., Di, C.A. and Zhu, D., “Advances of flexible pressure sensors toward artificial intelligence and health care applications,” *Mater. Horizons*, vol. 2, no. 2, pp.140-156, 2015.
- [4] D. Kwon, T. Lee *et al.*, “Highly sensitive, flexible, and wearable pressure sensor based on a giant piezocapacitive effect of three-dimensional micro-porous elastomeric dielectric layer,” *ACS Appl. Mater. Interfaces*, vol. 8, no. 26, pp. 16922-16931, 2016.
- [5] Y. Kim, A. Chortos *et al.*, “A bioinspired flexible organic artificial afferent nerve,” *Science*, vol. 360, no. 6392, pp. 998-1003, 2018.
- [6] W. W. Lee, Y. J. Tan *et al.*, “A neuro-inspired artificial peripheral nervous system for scalable electronic skins,” *Sci Robot.*, vol. 4, no. 32, 2019.
- [7] Sundaram, S., Kellnhofer *et al.*, “Learning the signatures of the human grasp using a scalable tactile glove,” *Nature*, vol. 569, pp. 698-702, 2019.
- [8] M. Amjadi, A. Pichitpajongkit, S. Lee, S. Ryu, and I. Park, “Highly stretchable and sensitive strain sensor based on silver nanowireelastomer nanocomposite,” *ACS Nano*, vol. 8, no. 5, pp. 5154-5163, 2014.
- [9] Wang, S., Xu, J. *et al.*, “Skin electronics from scalable fabrication of an intrinsically stretchable transistor array,” *Nature*, vol. 555, pp. 83-88, 2018.
- [10] Ravinder S. Dahiya, Giorgio Metta *et al.*, “Tactile sensing-from humans to humanoids,” *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 1-20, 2010.
- [11] K. Shimonomura, “Tactile image sensors employing camera: A review,” *Sensors*, vol. 19, no. 18, 2019.
- [12] Alexander C. Abad, A. Ranasinghe, Visuotactile Sensors with Emphasis on GelSight Sensor: A Review, *IEEE Sensors J.*, vol. 20, no. 14, pp. 7628-7638, 2019.
- [13] W. Yuan *et al.*, “GelSight: High-resolution robot tactile sensors for estimating geometry and force,” *Sensors*, vol. 17, no. 12, pp. 2762, 2017.
- [14] Sato, K., Kamiyama, K., Kawakami, N., Tachi, S. “Finger-shaped GelForce: Sensor for measuring surface traction fields for robotic hand,” *IEEE Trans. Haptics*, vol. 3, pp. 37-47, 2010.
- [15] B. Ward-Cherrier *et al.*, “The TacTip family: soft optical tactile sensors with 3D-Printed biomimetic morphologies,” *Soft Robot.*, vol. 5, no 2, pp. 216-227, 2018.
- [16] Y. Ito, Y. Kim, G. Obinata, “Multi-axis force measurement based on vision-based fluid-type hemispherical tactile sensor,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2013, pp. 4729-4734.
- [17] M. Greminger, B. Nelson, “Vision-based force measurement,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 4, pp. 290-298, 2004.
- [18] A. N. Reddy *et al.*, “Miniature compliant grippers with vision-based force-sensing,” *IEEE Trans. Robot.*, vol. 26, no. 5, pp. 867-877, 2010.
- [19] A. I. Aviles, S. M. Alsaleh, J. K. Hahn, A. Casals, “Towards retrieving force feedback in robotic-assisted surgery: A supervised neuro-recurrent-vision approach,” *IEEE Trans. Haptics*, vol. 10, no. 3, pp. 431-443, 2017.
- [20] Z. Dequidt, J. Dequidt, C. Duriez, “Vision-based sensing of external forces acting on soft robots using finite element method,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1529-1536, 2018.
- [21] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale *et al.*, “Methods and technologies for the implementation of large-scale robot tactile sensors,” *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 389-400, 2011.
- [22] Philipp Mittendorfer, Gordon Cheng, “Humanoid multimodal tactile sensing modules,” *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 401-410, 2011.
- [23] Quentin Leboutet, Emmanuel Dean-Leon *et al.*, “Tactile-based whole-body compliance with force propagation for mobile manipulators,” *IEEE Trans. Robot.*, vol. 35, no. 2, pp. 330-342, 2019.
- [24] G. Cheng, E. Dean-Leon, F. Bergner *et al.*, “A comprehensive realization of robot skin: Sensors, sensing, control, and applications,” *Proc. IEEE*, vol. 107, no. 10, pp. 2034-2051, 2019.
- [25] Duong, L.V., Asahina, R., Wang, J., Ho, V.A. “Development of a vision-based soft tactile muscularis,” in *Proc. IEEE International Conference on Soft Robotics*, 2019, pp. 343-348.
- [26] S. Das, N. Ahuja, “Performance analysis of stereo vergence and focus as depth cues for active vision,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 12, pp. 1213-1219, 1995.
- [27] Tatsuya Sakuma *et al.*, “A universal gripper using optical sensing to acquire tactile information and membrane deformation,” in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2018, pp. 6431-6436.
- [28] A. Myronenko *et al.*, “Point set registration: coherent point drift,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262-2275, 2010.
- [29] A. Vandini *et al.*, “Unified tracking and shape estimation for concentric tube robots,” *IEEE Trans. Robot.*, vol. 33, no. 4, pp. 901-915, 2017.
- [30] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330-1334, 2000.
- [31] A. W. Fitzgibbon, “Simultaneous linear estimation of multiple view geometry and lens distortion”, *Comput. Vis. Pattern Recognit.*, vol. 1, pp. 125-132, 2001.
- [32] E. Oñate, “Structural analysis with the finite element method linear statics, in: Beams, Plates and Shells of Lecture Notes on Numerical Methods in Engineering and Sciences,” vol. 2, Springer, Berlin, 2013.
- [33] Young W. Kwon , Hyochoong Bang, “The Finite Element Method Using MATLAB,” CRC Press, Inc., Boca Raton, FL, 2000.
- [34] N. M. Belyaev, “Strength of materials,” MIR Publishers, Moscow, pp. 404-406, 1979.



Lac Van Duong was born in Vietnam, in February 1991. He received the Engineering degree (with the excellent degree) in Mechatronics from Hanoi University of Science and Technology (HUST), in 2014, and the M.S. degree in 2016. He is now a PhD student at Japan Advanced Institutes of Science and Technology (JAIST). His research interests include the applications of Control Systems, Robotics, and Mechatronics. He is recipient of 2014 Honda Young Engineer and Scientist’s Award.



Van Anh Ho obtained the PhD degree on Robotics at Ritsumeikan University, Japan in 2012. He had completed the JSPS Post doctoral fellow in 2013 before joining Advanced Research Center Mitsubishi Electric Corp. in Japan. From 2015 to 2017 he worked as Assistant Professor at Ryukoku University in Kyoto where he led a laboratory on soft haptics. From 2017 he joined Japan Advanced Institute of Science and Technology (JAIST) for setting up a laboratory on Soft Robotics. His current research interests are soft robotics, soft haptic interaction, tactile sensing, grasping and manipulation, bio-inspired robots. He is recipient of 2019 IEEE Nagoya Chapter Young Researcher Award, Best Paper Finalists at IEEE SII (2016) and IEEE RoboSoft (2020). He is member of IEEE, RSJ.