

RESEARCH ARTICLE

Open Access



Dexterous object manipulation by a multi-fingered robotic hand with visual-tactile fingertip sensors

Seung-hyun Choi¹ and Kenji Tahara^{2*}

Abstract

In this paper, a novel visual-tactile sensor is proposed; additionally, an object manipulation method for a multi-fingered robotic hand grasping an object is proposed by detecting a contact position using the visual-tactile sensor. The visual-tactile sensor is composed of a hemispheric fingertip made of soft silicone with a hollow interior and a general USB camera located inside the fingertip to detect the displacement of the many point markers embedded in the silicone. The deformation of each point marker due to a contact force is measured, and a contact position is estimated reliably through a novel method of creating virtual points to determine the point clouds. The aim is to demonstrate both the estimation performance of the new visual-tactile sensor and its usefulness in a grasping and manipulation task. By using the contact position obtained from the proposed sensor and the position of each fingertip obtained from kinematics, the position and orientation of a grasped object are estimated and controlled. The effectiveness of the method is illustrated through numerical simulation and its practical use is demonstrated through grasping and manipulating experiments.

Keywords: Visual-tactile sensor, Multi-fingered hand, Grasping, Manipulation

Introduction

A human hand can detect a contact position and force. This tactile sensing skill helps to recognize an external environment and makes it possible to manipulate objects inside one's hand. To make a multi-fingered robotic hand perform a human-like object grasping and manipulation motion, many types of information such as each joint angle, contact force, or contact position should be acquired through an encoder, a force sensor, or a tactile sensor. There have been many attempts to realize a human-like grasping and manipulation task using such sensors [1–5]. Among them, the function of tactile and force sensors for determining the contact force and position must be important because these make it possible to detect the grasping force, position, and orientation of

a grasped object, as well as finger slippage by knowing the contact position. Additionally, these sensors contribute to change the position and orientation of a grasped object by in-hand manipulation.

There are various methods to obtain the contact force and position of each fingertip. Many present tactile sensors use electrical resistance, electromagnetics, piezoelectricity, ultrasonic, optics, and strain gauges. However, they need to attach many sensor arrays to the surface of each fingertip to determine a contact position, and these tend to be fragile by impact and cost [6–8]. A large number of sensor arrays also increases the number of wires and it may be difficult to satisfy performance requirements or sampling rate due to a computational burden for processing a large amount of sensor information [8]. However, recently several tactile sensors using a visual sensor, which is called a visual-tactile sensor hereinafter, have been proposed [9–12]. A visual-tactile sensor is mainly based on image processing and therefore they are easily affected by the ambient light condition and require

*Correspondence: tahara@ieee.org

² Department of Mechanical Engineering, Faculty of Engineering, Kyushu University, 744 Moto'oka, Nishi-ku, Fukuoka 8190395, Japan
Full list of author information is available at the end of the article

relatively high programming skills. However, a visual sensor does not need to physically contact an object and thus, it is robust for impulsive contact breakdown, and it can build a low if a generally available USB camera is utilized. Several studies have observed deformation of a soft finger by placing a point cloud inside the finger to observe the contact position and force using a visual sensor [10–12]. These studies demonstrated the effectiveness of a visual-tactile sensor. However, they mainly evaluated the performance of the sensor itself and they did not mention how to use the obtained information in a grasping and manipulation control and how it works in grasping and manipulation tasks. Yamaguchi et al. [13] evaluated a cutting force of an object by a robotic gripper with a visual-tactile sensor. From the viewpoint of grasping and manipulation control, there is a study that demonstrated reliable grasping by increasing the contact area at an initial contact position [14]. Although this study has been applied to a robotic gripper using a visual-tactile sensor to grasp an object reliably, it has not been applied to a method for controlling the position and orientation of a grasped object by a multi-fingered robotic hand.

However, several works which use conventional tactile sensors have demonstrated that they reliably control the position and orientation of the grasped object based on the sensor information. When changing the position and orientation of a grasped object in the hand, humans can realize it with small motions of each finger through a rolling contact between an object surface and each fingertip. Several studies have achieved reliable grasping and manipulation through the rolling constraint [16, 17]. Tahara et al. have proposed a blind grasping and manipulation controller capable of controlling the position and orientation of a grasped object without using a force or tactile sensor [16]. They introduced a virtual object frame that is defined by the position of each fingertip of a three-fingered robotic hand to express the position and orientation of a grasped object virtually instead of a real object position and orientation. The fingertip of the robotic hand is made of flexible hemispheric silicone and by utilizing its rolling contact, the position and orientation of the object can be changed with a little motion of each finger. This method is quite advantageous because when controlling a grasped object, force and tactile sensors are unnecessary. However, the controlled position and orientation is not accurate because there is no external sensor. To achieve more accurate position and orientation control, it is necessary to use a force or tactile sensor that can detect the real contact force and position of each fingertip.

In this study, a new visual-tactile fingertip sensor, contact position, and force estimation method are proposed by using the detected information, and a new

object grasping and manipulation controller based on the blind grasping and manipulation controller is designed. First, the developed new visual-tactile sensor is introduced, and its performance is demonstrated through several experiments. Next, the new object grasping and manipulation controller, which uses the detected contact position information, is designed. Subsequently, the effectiveness of the proposed controller is evaluated through numerical simulations by comparing the proposed controller with the conventional blind grasping and manipulation controller.

Methods

Visual-tactile sensor design

The proposed visual-tactile sensor equipped on a fingertip is shown in Fig. 1.

Several point markers for detecting fingertip deformation are embedded inside the hemispheric silicone fingertip. The position of each point marker is arranged at a constant angle along the spherical surface to show deformation motion when the shape of the silicone fingertip contacts an object. Each point marker comprises 3-mm beads without holes, which are colored to be easy to distinguish. A mold for making silicone fingertip is made using a 3D printer facilitate creating a complex-shaped mold, as shown in Fig. 2.

A camera used in the sensor is a general USB-type camera capable of normal 30 FPS, which is readily available and low cost. The configuration of the sensor is like an endoscopic camera and it possesses a brightness-adjustable LED. Therefore, the influence of ambient light can be controlled to some extent. It is reasonable to use a faster camera to quickly detect information on the fingertip through the movement of the markers. However, this study used a 30 FPS USB camera on purpose because of its cost and implementation of LED. The fixed camera base is made by a 3D printer and designed to disassemble the silicone finger for easy repair as shown in Fig. 1. The base does not require any adhesive or screws.

Contact position and force estimation through the displacement of point markers

Several studies have developed visual-tactile sensors, in which multiple small markers are arranged inside a soft material, and these mainly measured contact force and position by detecting the displacement of the markers from their initial positions when an external force acted on the fingertip [10–13]. In most cases, the shapes of the contact area were flat and utilized a square arrangement of the markers in the soft material. However, these studies did not consider a rolling contact. In our proposed visual-tactile sensor, each point marker is arranged radially according to the hemispheric shape.

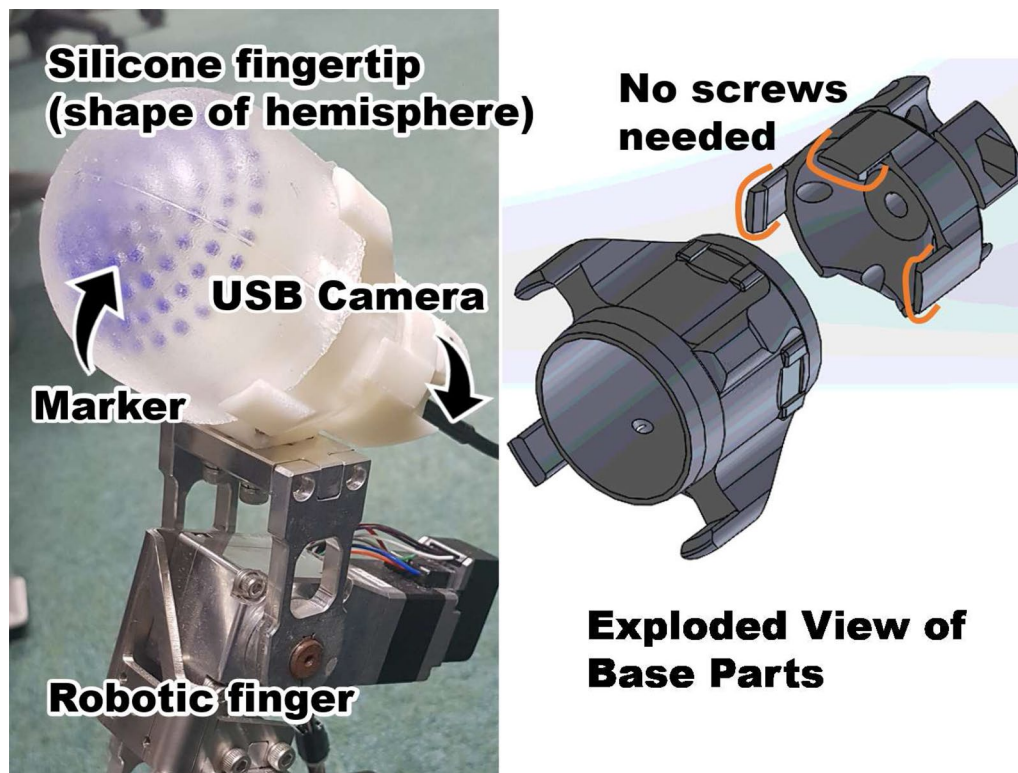


Fig. 1 Visual-tactile sensor and the exploded view of base parts

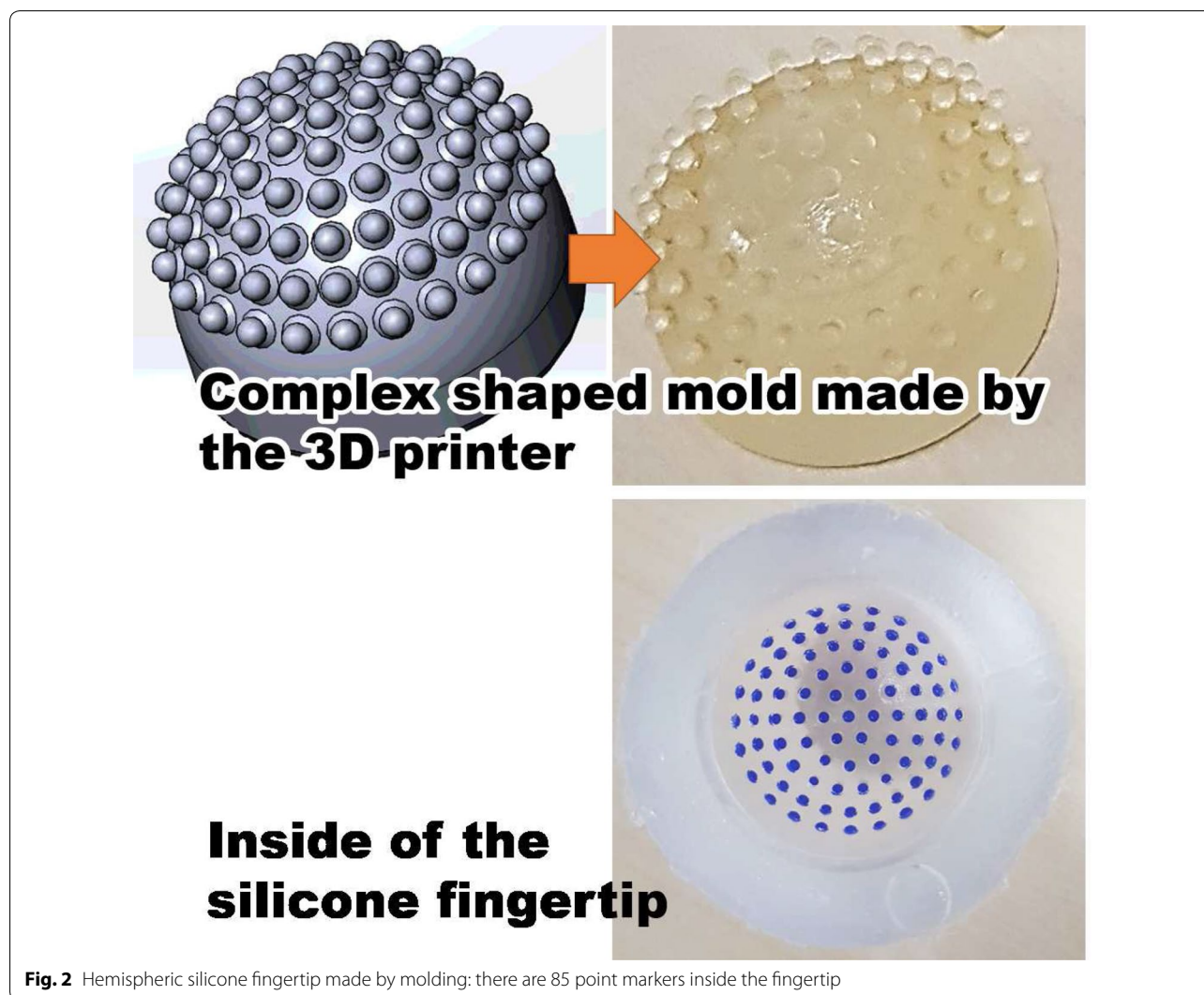
The radial arrangement is reasonable for a hemispheric fingertip even though the displacement of a point marker becomes more complex than the standard flat shape. Similar studies related to the visual-tactile force sensor have been performed (e.g. [14, 15]). However, the advantage of our proposed tactile sensor compared with similar studies is to use the position information of the virtual point markers instead of that of measured point markers. This provides an advantage that the labeling of each point marker is performed robustly against the change in the external lighting environment. When labeling each point marker continuously, incorrect labeling sometimes occurs induced by the loss of the detection of each point marker. Using virtual point markers instead of the detected point markers delivers more reliable labeling even if point markers are misplaced in real-time. Moreover, the most advantageous point of the study compared with other related studies is in addition to developing the sensor itself, this study also proposes an object grasping and attitude control method. Many similar studies focus on the specification and performance of the sensor itself and they hardly mention how to use it in a manipulation task. However, to utilize such a tactile sensor in practice, it is also important to know how to use it in a manipulation

task. In this study, the proposed sensor and controller are designed simultaneously, and these are eventually integrated as a system. Namely, it can be said that the specification of the proposed tactile sensor is demanded by the proposed controller or vice versa. Its effectiveness is evaluated in both numerical simulation and experiments using a prototype of the multi-fingered robotic hand with proposed tactile fingertip sensors. The displacement of each point marker arranges can be observed in real-time at 30 FPS through the USB camera.

We will give a new estimation method for a contact position and force using a geometric relation between each point marker in the next section.

Changes in the inside markers due to external forces

The shape of the fingertip is changed by external forces, which causes the position of the markers to move. The location where the external force is acting on the surface of the fingertip can be identified by comparing the initial and the present states of the changing markers. This deformation is obtained by the USB camera when the actual external force is applied as shown in the following Fig. 3.



Delaunay triangulation and virtual point markers implementation

To estimate a contact position and force, the camera recognizes markers placed on the hemisphere of the fingertip, which means that the three-dimensional location of each marker is projected onto a two-dimensional visual plane and the markers can be classified into several meaningful groups. The Delaunay triangulation and the Voronoi space division are well-known methods that can divide space to be several small parts. However, when using the Voronoi space division, because the edge area of the observed markers becomes infinite size, it is necessary to set a separate closing marker as shown in Fig. 4e. In this study, the Delaunay triangulation was used to divide the markers to be several meaningful parts. The Delaunay triangulation method divides a point cloud into several triangles by

dividing the circumcircle of triangles without including the other vertices of the triangle. By dividing the observed point cloud into several triangles, it is possible to measure the change in contact position or the change in force by comparing the area of each triangle.

When an external force acts on the silicon finger, the labeling order changes depending on the moving position of each marker as shown in Fig. 4. In such a case, the present divided area cannot directly be compared with the previously collected area because the labeling order is different. To overcome this problem, the observed markers are fitted into ellipses using the least-squares method in the LabView fitting ellipse function. From the obtained fitted ellipses, information about the major and minor axes and the center point can also be obtained. Subsequently, the virtual point markers on the ellipse can be generated using the following Eq. (1):

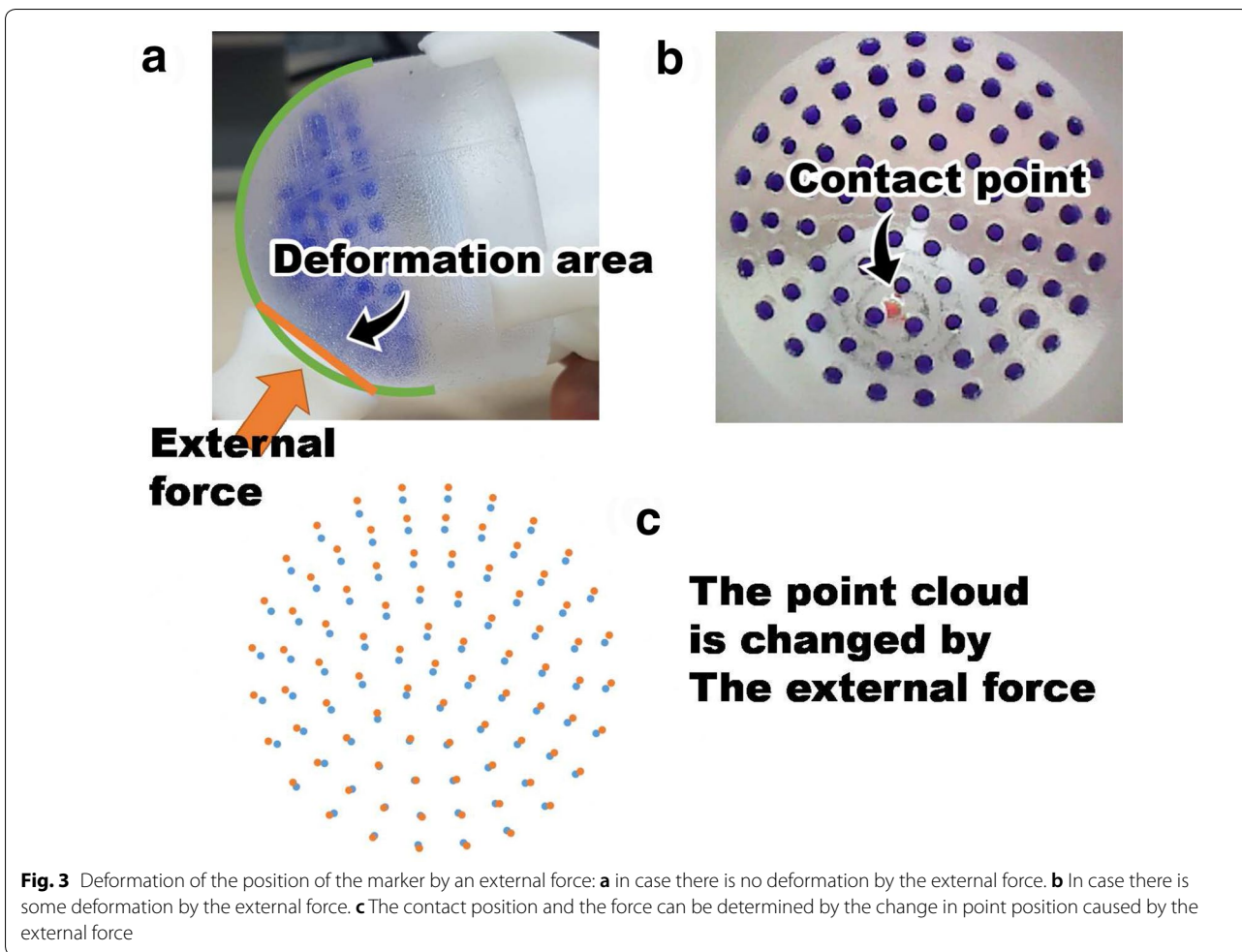


Fig. 3 Deformation of the position of the marker by an external force: **a** in case there is no deformation by the external force. **b** In case there is some deformation by the external force. **c** The contact position and the force can be determined by the change in point position caused by the external force

$$\theta_n = \theta_1 + (n - 1)2\pi/t_n \tag{1a}$$

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} \cos \theta_n & 0 \\ 0 & \sin \theta_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix} \tag{1b}$$

where $[x_n, y_n]^T$ denotes the position of virtual point marker located on the ellipse, θ_n denotes the internal angle of the n -th triangle made by $[x_n, y_n]^T$, $[x_{n+1}, y_{n+1}]^T$, and $[x_c, y_c]^T$ where $[x_c, y_c]^T$ denotes the center position of the fitted ellipse. The total number of sampled points on the ellipse is denoted by t_n , and a and b denote the major and minor axes of the fitted ellipses, respectively. The labeling order on virtual point markers is robust even if the lighting environment changes. The process for this method is visually shown in Fig. 5.

Another advantage of using the virtual point markers is that it is robust to the change in the external environment. The outer layer of the fingertip sensor is made of silicone and is thus, affected by the change in the light condition of the external environment as shown in Fig. 6.

The positions of several markers may not be detected due to the external environment. As each triangle changes according to the position of the markers and if it changes frequently, the force or contact position estimation may become unstable and inaccurate. However, the virtual point markers are created from a fitted ellipse and thus, the Delaunay triangulation can be performed reliably even though the positions of several markers are unknown.

Relocation of the area change and estimation of the contact point position

As we mentioned before, the issue of labeling order of each divided part is solved by the introduction of virtual point markers. However, the virtual point markers are always uniformly distributed in an area. This may cause another issue that if the area size ratio of each triangle is almost identical before and after an external force acts on it, the change in the triangle area cannot be detected even if the shape of the ellipse is changed by an external force as shown in Fig. 7a.

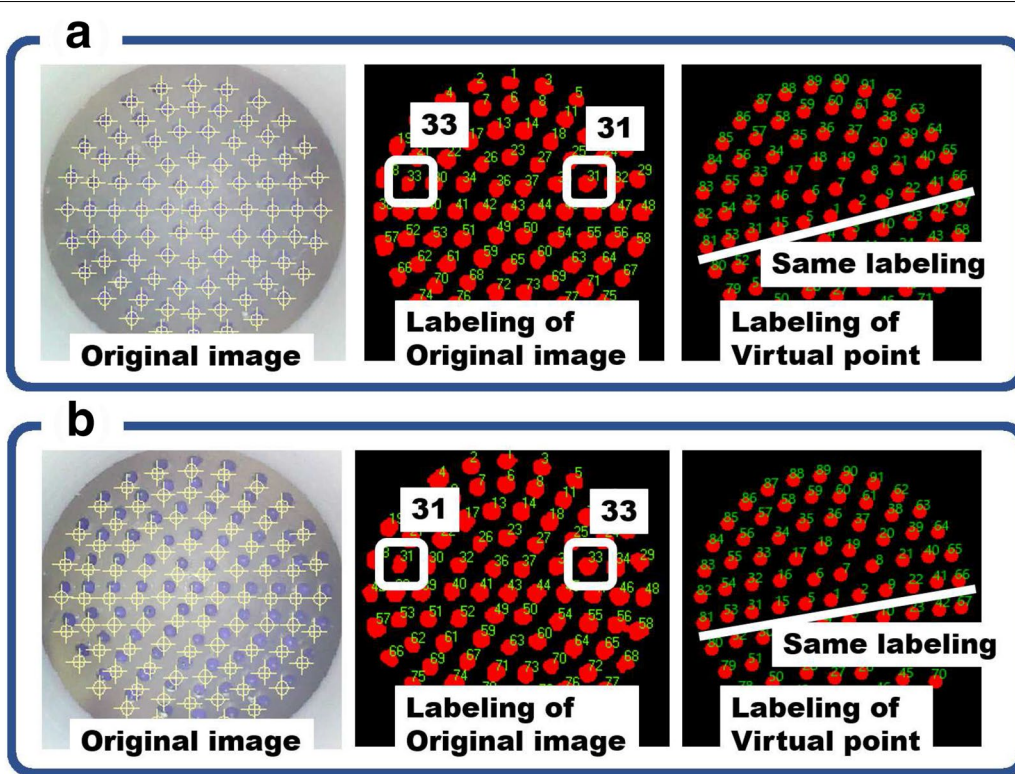


Fig. 4 Labeling by external forces and virtual points: **a** In case there is no deformation by the external force. **b** In case there is some deformation by the external force. Deformation by external forces may change the labeling order of observed points. However, it is possible to maintain the labeling order by using virtual points

To overcome this issue, instead of comparing just each triangle, we compare the triangle area averages including the virtual point markers as in Fig. 7b. The reason for taking the average is that at an edge, especially around the upper edge of the fingertip sensor, the number in the triangle exists only two triangles compared with the center part. A difference in the number of triangles can lead to a considerable difference in the total size area of the sum of all triangles. By utilizing the area average, it is possible to compare the center points containing around six triangles with the edge points containing at least two triangles. By using the average, this undesirable estimation can be avoided to a great extent.

When the contact point position is estimated by the change in triangle area around the marker, the spatial resolution depends on the number of markers. In this study, the proposed visual-tactile sensor has 85 point markers, which are not very dense and thus the spatial resolution is not high. To compensate for the low spatial resolution, the contact point position P_c is estimated using the following equation:

$$A = \frac{BC}{D} \tag{2}$$

where

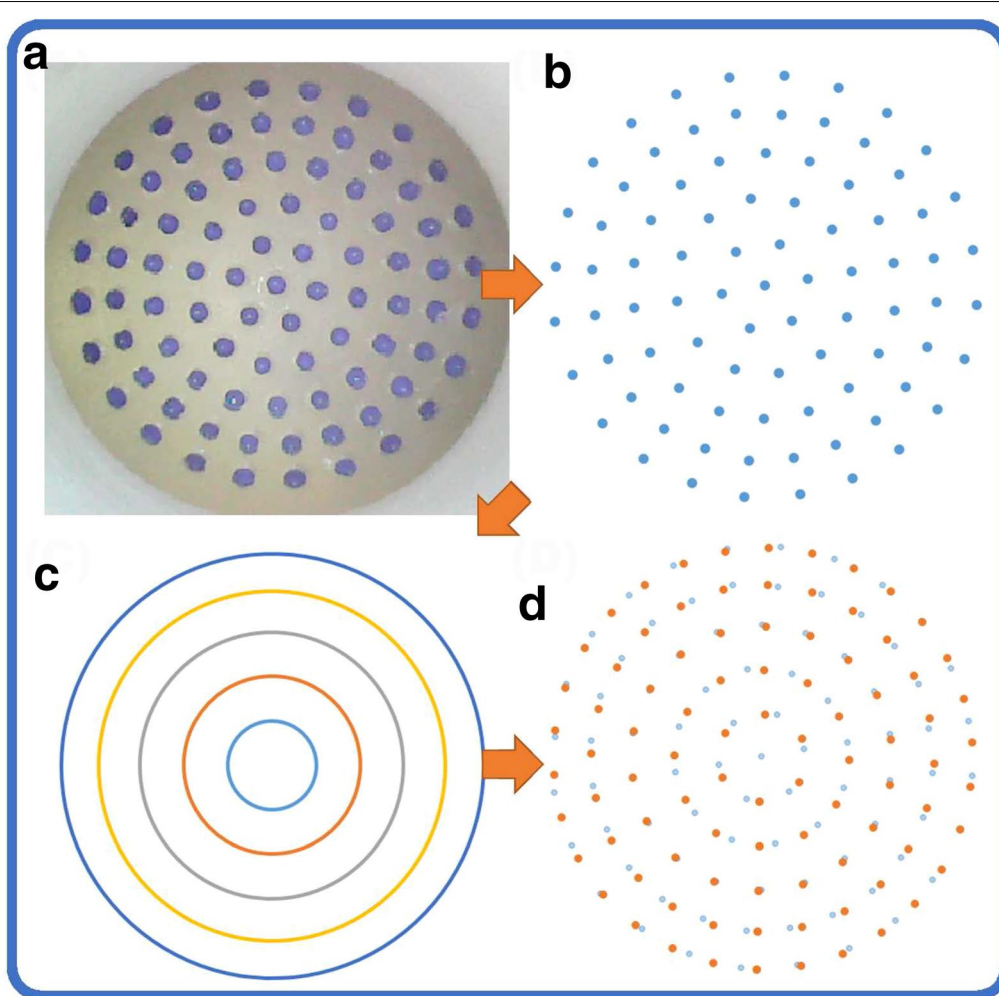
$$A = \begin{bmatrix} P_{cx} \\ P_{cy} \end{bmatrix} \in \mathbb{R}^2$$

$$B = \begin{bmatrix} vp_{2x}, & \dots, & vp_{7x} \\ vp_{2y}, & \dots, & vp_{7y} \end{bmatrix} \in \mathbb{R}^{2 \times 6}$$

$$C = \begin{bmatrix} 1 \\ vp_{1a} - vp_{2a} \\ \vdots \\ 1 \\ vp_{1a} - vp_{7a} \end{bmatrix} \in \mathbb{R}^6$$

$$D = \frac{1}{vp_{1a} - vp_{2a}} + \dots + \frac{1}{vp_{1a} - vp_{7a}}$$

where vp_{1a} denotes the area size of the center point that is thought to be in contact, $vp_{2a} \sim vp_{7a}$ denote the peripheral points of the area size considered to be in contact, and $vp_{2x} \sim vp_{7x}$ and $vp_{2y} \sim vp_{7y}$ stand for x - and y -position of each point. This calculation enables continuous retrieval of the contact position even though the spatial resolution is low. Additionally, this choice makes the sensor easy to assemble and reduces computational costs.



The Process of creating virtual points

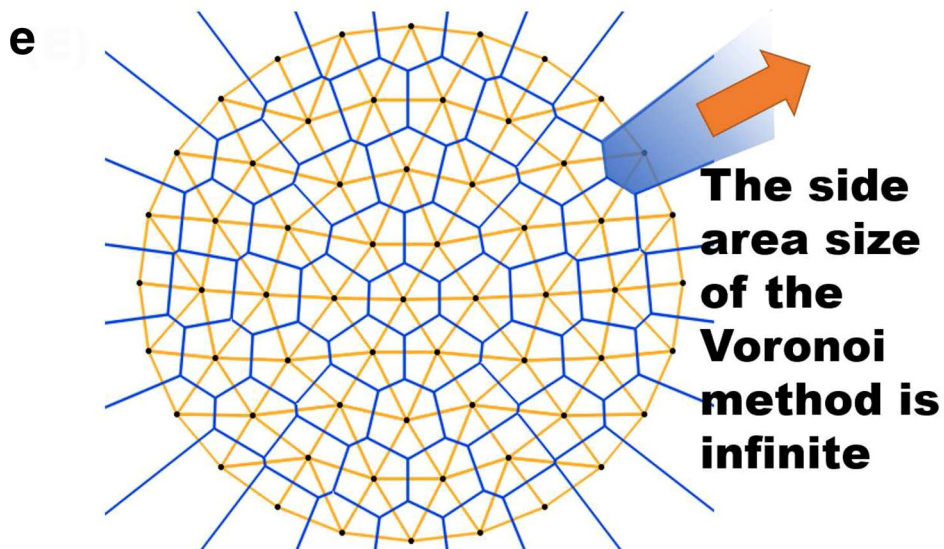


Fig. 5 Process for creating a new virtual point: **a** Original image received from a USB camera. **b** Point detected after image processing. **c** Ellipse fitting based on the detected point. **d** New virtual points created based on the fitted ellipse (Orange: newly created virtual point, Blue: the original point). **e** Comparing spatial partitioning results (Blue: the Voronoi space division, Orange: the Delaunay triangulation)

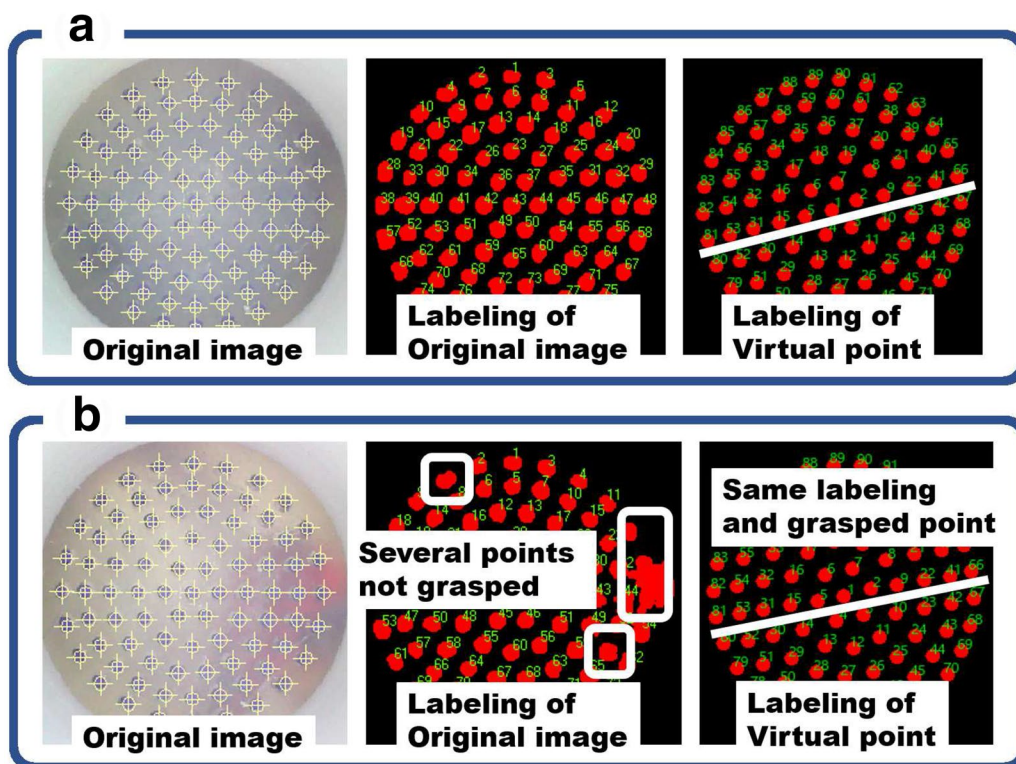


Fig. 6 Labeling by external environment and virtual points: **a** In case there is no influence on the external environment. **b** In case the external environment affects the marker. Markers may not be observed if the external environment is different. These changes in the environment can also be observed using virtual points

Results

Experiments on the comparison between the estimated and actual contact positions

It is necessary to evaluate how much estimation error is in the contact position estimated by the proposed sensor. In this study, transparent silicone rubber is used for the outer layer of the visual-tactile sensor. It is possible to determine the contact position by marking the contact point as shown in Fig. 8a. To evaluate the performance of the sensor according to the contact position, the measured area is sectioned into five parts as shown in Fig. 8b. The experimental device using the linear guide for accurate contact angle division and force measurement is shown in Fig. 9. The estimation error is evaluated by comparing the actual contact position with the estimated contact position by Eq. (2). A total of 30 trials in which an external force is applied to each area is performed, and the error of each trial is averaged.

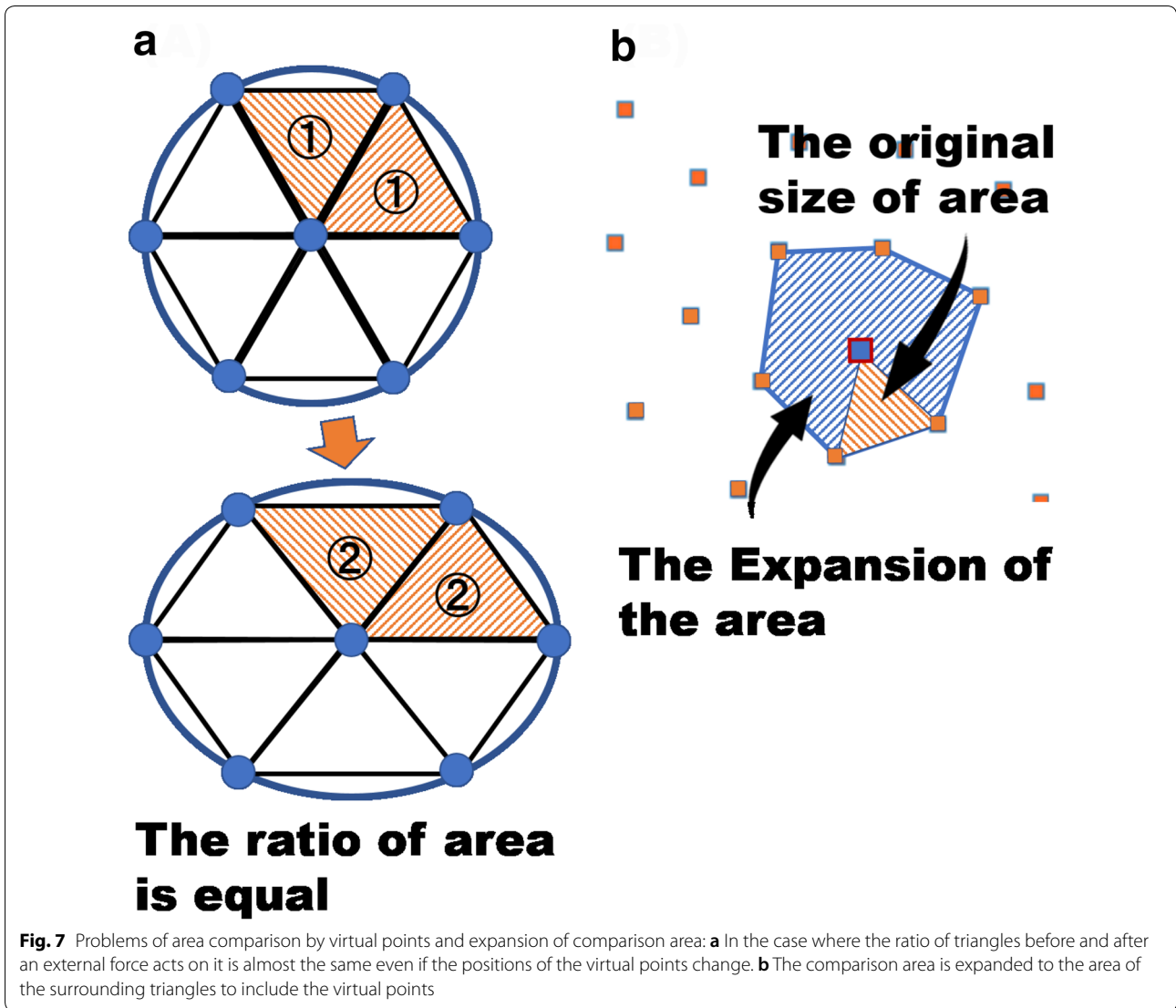
It can be seen from Table 1 that the overall error average is 1.475 (mm). The error was smaller on the side, which is thought to be because of the nature of the contact position improving as the average variation in the area is more abundant on the side. The result of the error

is small enough compared with the diameter of the proposed visual-tactile sensor 60 (mm).

Contact force estimation

In this study, we model a contact force acting on the fingertip such that the change in the contact area when the silicone is deformed due to an external force induces a spring-like force.

It can be seen from Fig. 10 that the change in the triangle area decreases when the contact position approaches the upper sensor edge. This indicates that the triangle area size is different according to where the external force acts even if the applied force is constant. Namely, the relationship between the applied contact force and the change in the triangle area is not linear and it changes depending on the position. To address the non-linearity, we first numerically analyzed the relationship between the displacement of the point markers and the contact force using a finite element method (FEM) offline. The numerical analysis of the fingertip deformation due to a static contact force according to contact angles is shown in Fig. 11. In this analysis, we assume that the upper end of the hemispherical fingertip is fixed, and a rectangular steel plate is pushed up by an external force from the



bottom. The hemispherical fingertip is deformed by the external force applied by the steel plate, and the position of each inner point is changed depending on the relationship between the external force and the contact angle. From FEM, we determined that the distance between each side point spreads linearly when an external force acts on the fingertip as shown in Table 2.

Experiment for determining the revision of the distance between each point marker

The correction function to revise the non-linearity can be determined through numerical analysis as shown in Fig. 11. As a result, we can obtain an almost linearized relationship between the contact position and force using the correction function as shown in Fig. 12.

When the contact position approaches both upper edges, the spread returns to its original position. In the

case of distance 4 and 5, the distance between point markers on both sides can be kept constant even if the contact position is approaching the boundary. Therefore, instead of using the change in the triangle area, an external force can be estimated by applying the correction function to the distance between the point-markers on both edge sides as shown in Eq. (3).

$$\begin{cases} P_{\text{dist}} = \|P_{\text{nc}+1} - P_{\text{nc}}\| + \|P_{\text{nc}-1} - P_{\text{nc}}\| \\ F_{\text{est}} = k_{\text{est}}P_{\text{dist}}, \end{cases} \quad (3)$$

where P_{nc} denotes the nearest estimated point around the location of ⑤ on Fig. 12, $P_{\text{nc}+1}$ and $P_{\text{nc}-1}$ are the points at the front and back of the P_{nc} , F_{est} denotes the estimated force and K_{est} denotes a stiffness coefficient depending on the position and material of the fingertip.

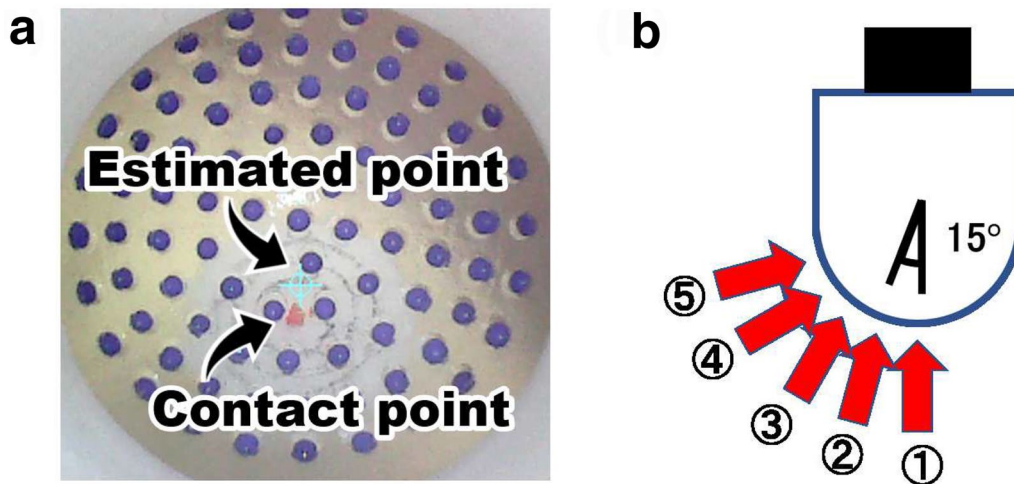


Fig. 8 Estimation point and contact point comparison and the condition of the angle: **a** Estimated contact point and an actual contact point: both the contact point detection and estimation are executed in real-time. **b** Area distinction according to a contact position in the experiment

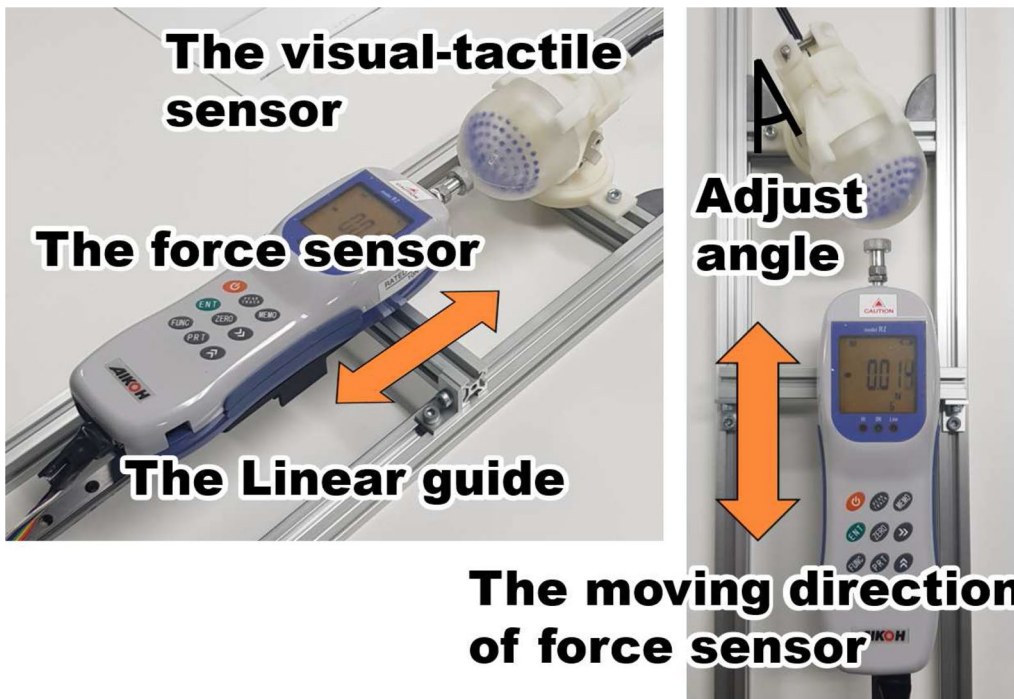


Fig. 9 Experiment device to check the contact position of visual tactile sensor

Table 1 Error between the estimated and the actual contact point positions

	①	②	③	④	⑤
Unit: (mm)					
Error	1.906	1.683	1.369	1.267	1.152

Simulation

In this section, numerical simulations of grasping an object and its orientation control are conducted using a newly redesigned virtual object frame. The object grasping and manipulation controller use the virtual frame, which was proposed by Tahara et al. [16], is an externally

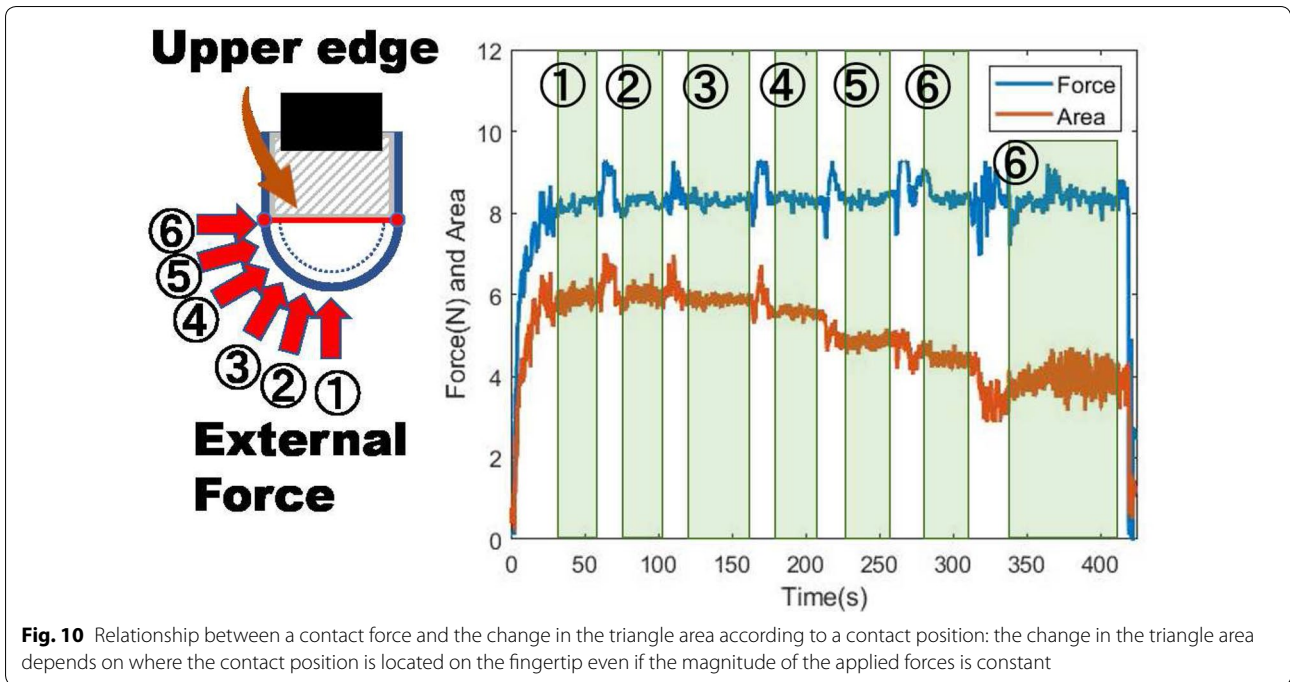


Fig. 10 Relationship between a contact force and the change in the triangle area according to a contact position: the change in the triangle area depends on where the contact position is located on the fingertip even if the magnitude of the applied forces is constant

sensorless controller. It is robust because there is no effect from noise and time delay of sensing information, while the accuracy depends on an initial contact position and object shape. Kawamura et al. [17] has also proposed an object grasping and orientation control method using visual information obtained from a camera to detect the orientation of a grasped object. In Kawamura’s method, the desired virtual object frame is updated by the visual information from a camera as an external sensor. However, generally to detect the orientation of a grasped object in enough accuracy is still difficult even using a very high speed and dense spatial resolution camera. In this study, unlike Kawamura’s method, the virtual object frame is composed using the contact position information obtained from the proposed sensor. It is easy to compose the desired virtual object frame using the contact position information from the proposed sensor. Additionally, our proposed visual-tactile sensor can estimate the contact force if necessary, while Kawamura’s method cannot. Even if the sensing information is not accurate enough, our proposed controller is robust to sensing error, noise, and time delay because the contact position information is not directly used in a feedback controller and it is only used for composing and updating the virtual object frame.

Dynamics of the object-fingers system

The multi-fingered robotic hands of our research have a hemispherical shape and soft contact surface. The

robot hands and grasping object have a nonholonomic rolling constraint when the robot hands hold a grasping object. The objects and the multi-fingered robot hands are modeled and define the dynamics. In addition, the conditions of the contact surface of the soft fingertips are defined to satisfy the physical laws of the real world. The soft fingertip robot hands can be illustrated in three dimensions as shown in Fig 13. We assume that the robotic hand has three fingers in which each finger has five Degree Of Freedom (DOF)s and the grasped object is a triangular prism, and $i(= 1, 2, 3)$ be the index of each finger. Therefore, this robotic hand has a total of fifteen DOFs. The dynamics of the object-finger system have already been modeled in [16, 17]. In this simulation, we use the same dynamics, which can be given as follows:

For the multi-fingered hand:

$$\begin{aligned}
 H(q)\ddot{q} + \left\{ \frac{1}{2}\dot{H}(q) + S_q(q, \dot{q}) \right\} \dot{q} + \sum_{i=1}^3 \frac{\partial T_i^T}{\partial \dot{q}} \\
 + \sum_{i=1}^3 \left(J_i^T C_{iY} f_i + X_{iq}^T \lambda_{iX} + Z_{iq}^T \lambda_{iZ} \right) + g(q) = u
 \end{aligned}
 \tag{4}$$

For the grasped object:

$$M\ddot{x} + \sum_{i=1}^3 \left(-f_i C_{iY} + X_{ix}^T \lambda_{iX} + Z_{ix}^T \lambda_{iZ} \right) = 0
 \tag{5}$$

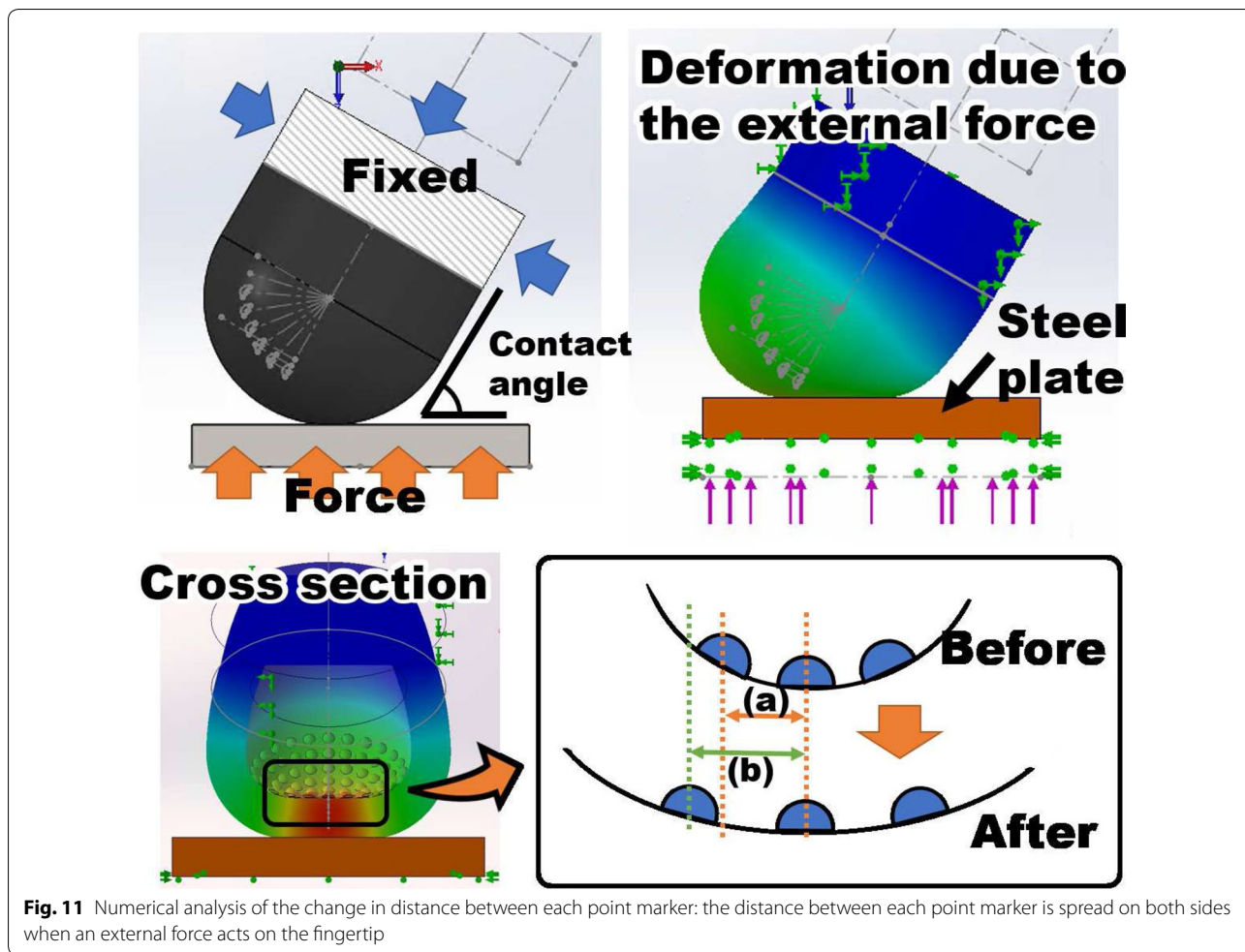


Fig. 11 Numerical analysis of the change in distance between each point marker: the distance between each point marker is spread on both sides when an external force acts on the fingertip

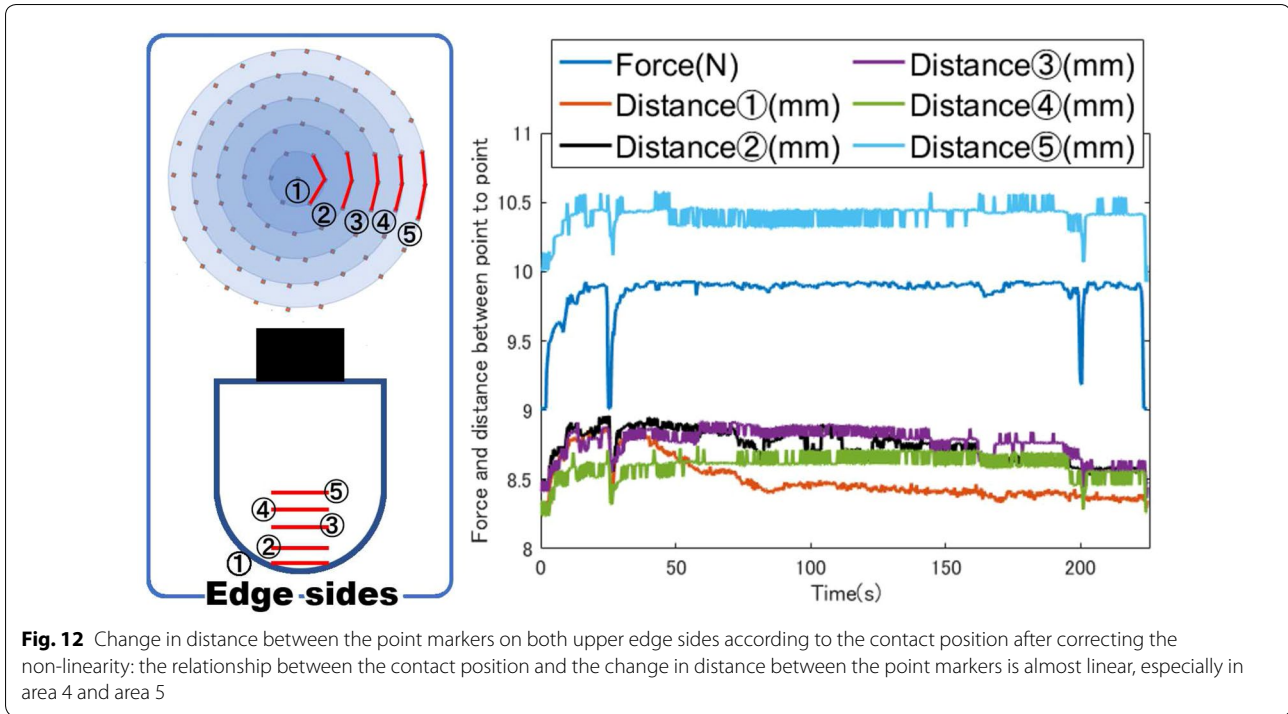
$$\begin{aligned}
 & I\dot{\omega} + \left\{ \frac{1}{2}\dot{I} + S_{\omega} \right\} \omega - \sum_{i=1}^3 \{C_{iY} \times (x - x_i)\} f_i + \sum_{i=1}^3 \frac{\partial T_i}{\partial \omega} \\
 & + \sum_{i=1}^3 (X_{i\omega}^T \lambda_{iX} + Z_{i\omega}^T \lambda_{iZ}) = 0
 \end{aligned} \tag{6}$$

Table 2 Distance of both side point markers in the case where the static external force is applied to the different positions on the fingertip and the ratio between the distance of before and after contacts: the contact angle is defined in Fig. 8b

	Distance before contact Fig. 11a (mm)	Distance after contact Fig. 11b (mm)	Ratio after/ before Fig. 11b/a
Contact angle 90°	7.69	12.09	1.572
Contact angle 75°	8.65	13.83	1.598
Contact angle 60°	8.51	13.59	1.596
Contact angle 45°	8.09	13.02	1.609
Contact angle 30°	9.30	15	1.612

where $H(q) \in \mathbb{R}^{15 \times 15}$ denotes an inertia matrix of all the fingers; $M \in \mathbb{R}^{3 \times 3}$ is the mass of the object; $I \in \mathbb{R}^{t \times 3}$ is the inertia tensor of the object; $q = [q_{11}, \dots, q_{15}, q_{21}, \dots, q_{25}, q_{31}, \dots, q_{35}]^T \in \mathbb{R}^{15}$ stands for the joint angle vector for all fingers; $\dot{q} \in \mathbb{R}^{15}$ and $\ddot{q} \in \mathbb{R}^{15}$ are the joint angular velocity and acceleration vector, respectively; $x \in \mathbb{R}^3$, $\dot{x} \in \mathbb{R}^3$, and $\ddot{x} \in \mathbb{R}^3$ are the position, velocity, and acceleration vector of the object, respectively; and $\omega \in \mathbb{R}^3$ and $\dot{\omega} \in \mathbb{R}^3$ are the angular velocity and acceleration vector of the object, respectively. Furthermore, $S_q \in \mathbb{R}^{15 \times 15}$ and $S_{\omega} \in \mathbb{R}^{3 \times 3}$ denote skew-symmetric matrices including Coriolis and centrifugal forces for the robotic hand and the object, respectively. $\frac{\partial T_i}{\partial \dot{q}}$ and $\frac{\partial T_i}{\partial \dot{\omega}}$ are the viscosity between the contact surface of the object and each fingertip. These terms affect the torsional motion of the fingertip on the object surface. The energy dissipation function of the torsional viscosity of the fingertip is modeled as follows:

$$T_i = \frac{b}{2} \left\| C_{iY}^T (\omega - \omega_i) \right\|^2 \tag{7}$$



where ω_i indicates an angular velocity vector for each fingertip, which can also be expressed by the joint angular velocity vector \dot{q} , and b is a viscosity coefficient that depends on the contact area and the fingertip material. The contact position of each fingertip on the object surface is denoted by $x_i \in \mathbb{R}^3$ and $C_i = [C_{iX}, C_{iY}, C_{iZ}] \in SO(3)$ denotes a rotational matrix to express the orientation of each contact surface on the object from Cartesian coordinates. In the dynamics, two constraints should be considered. One is a contact condition in the normal direction on the contact surface. A constraint force f_i induced by the contact condition implies a grasping force. We assume that there is only one contact point on the fingertip. Moreover, the contact force changes according to the deformation of the soft fingertip as shown in Fig. 14. The relationship between the displacement of the fingertip and its reproducing force is defined as Eq. (8), which was proposed by Arimoto et al. [18].

$$\begin{cases} \dot{f}_i = \bar{f}_i + \xi \Delta \dot{r}_i \\ \bar{f}_i = k \Delta r_i^2 \end{cases} \quad (8)$$

where ξ denotes a positive damping coefficient and k denotes the positive elastic coefficient, both of which depend on the soft fingertip material, and Δr_i denotes the deformation displacement of the i -th soft fingertip. The other is a rolling constraint condition in the tangential direction on the contact surface. The rolling constraint forces λ_{iX} and λ_{iZ} imply tangential contact forces, and X_{iq} , X_{ix} , $X_{i\omega}$,

Z_{iq} , Z_{ix} and $Z_{i\omega}$ are the rolling constraint Jacobians, and $u \in \mathbb{R}^{15}$ denotes an input torque vector for each joint. The details of these constraint Jacobians can be found in [16, 17].

Design of the control input

The object orientation controller using the contact position information and stable grasping controller based on Tahara's and Kawamura's controllers are designed here. These two controllers are eventually combined into one controller, which is given as follows:

$$u = u_{sg} + u_{at}, \quad (9)$$

where u_{sg} denotes a torque control input to each actuator for reliable grasping and u_{at} denotes a torque control input to each actuator for the object orientation. The control input u_{sg} generates a grasping force at the center of each fingertip to approach the fingertips to each other as shown in Fig. 15a. It is given as follows:

$$u_{sg} = \frac{f_d}{\sum_{i=1}^3 r_i} \sum_{j=1}^3 J_j^T (x_{cp} - x_{cj}) - B\dot{q} + g(q) \quad (10)$$

$$x_{cp} = \frac{1}{3} \sum_{i=1}^3 x_{ci} \quad (11)$$

where $B \in \mathbb{R}^{15 \times 15}$ stands for the positive damping coefficient matrix, $J \in \mathbb{R}^{15 \times 3}$ stands for the Jacobian matrix,

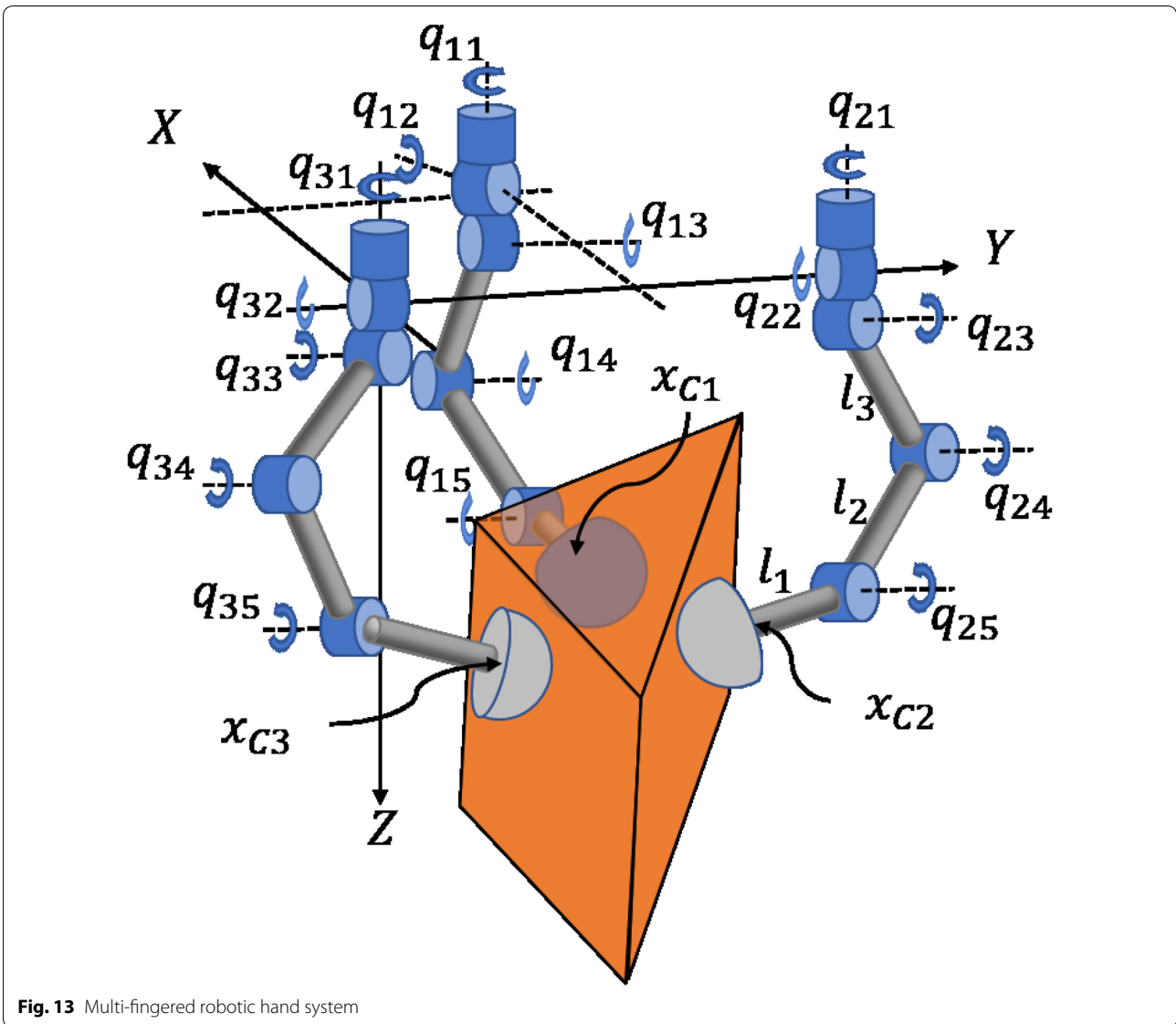


Fig. 13 Multi-fingered robotic hand system

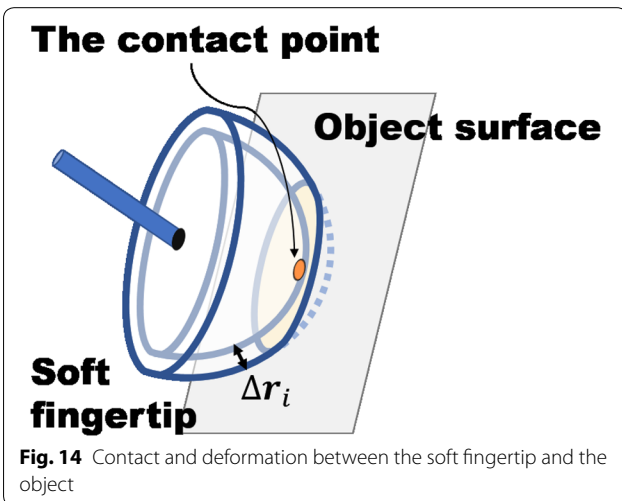


Fig. 14 Contact and deformation between the soft fingertip and the object

$g(q) \in \mathbb{R}^{15}$ denotes the gravity compensation term, f_d signifies a nominal desired grasping force, and r is the radius of each hemispheric fingertip. x_i is the position of each fingertip.

The virtual frame which expresses the orientation of the virtual object is designed as follows:

$$R_{vir} = [r_{xvir} \quad r_{yvir} \quad r_{zvir}] \tag{12}$$

where

$$\begin{bmatrix} r_{zvir} \\ r_{yvir} \\ r_{xvir} \end{bmatrix} = \begin{bmatrix} \frac{(x_{c1}-x_{c3}) \times (x_{c2}-x_{c3})}{\|(x_{c1}-x_{c3}) \times (x_{c2}-x_{c3})\|} \\ \frac{x_{c2}-x_{c3}}{\|x_{c2}-x_{c3}\|} \\ \frac{r_{yvir} \times r_{zvir}}{\|r_{yvir} \times r_{zvir}\|} \end{bmatrix}$$

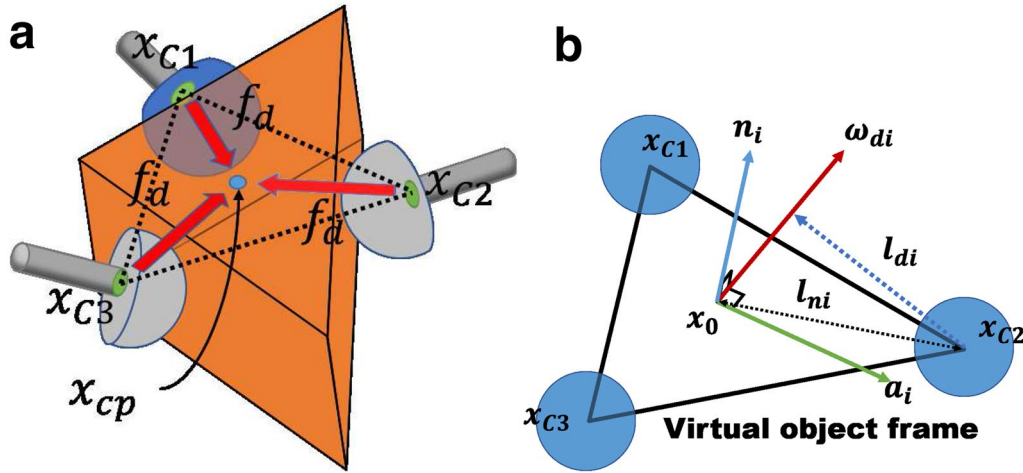


Fig. 15 Concept and component of the object reliably grasping and the object orientation controller: **a** The reliable grasping controller forces each fingertip to the center of the triangle created by the fingertip position. **b** Each finger applies torque around one rotational axis in the object orientation control

Based on Tahara's and Kawamura's controllers [16, 17], a new object grasping and orientation controller is designed as follows:

$$\mathbf{u}_{at} = J_i^T F_{di} \quad (13)$$

where \mathbf{u}_{at} denotes the newly proposed object orientation control input. The orientation of the object is regulated by adding the computed desired contact force F_{di} to the object from each fingertip through the Jacobian of each finger. The desired contact force F_{di} is given as follows:

$$F_{di} = l_{di} \times \omega_{di} \quad (14)$$

where

$$\begin{aligned} l_{di} &= \left(\mathbf{a}_i^T \mathbf{l}_{ni} \right) \mathbf{a}_i \\ \mathbf{l}_{ni} &= \mathbf{x}_0 - \mathbf{x}_{ci} \\ \omega_{di} &= \frac{f_d}{3r} K_{ri} (\mathbf{r}_{x_{vir}} \times \mathbf{r}_{x_{d_{vir}}} + \mathbf{r}_{y_{vir}} \times \mathbf{r}_{y_{d_{vir}}} + \mathbf{r}_{z_{vir}} \times \mathbf{r}_{z_{d_{vir}}}) \\ \mathbf{x}_0 &= \frac{1}{3} (\mathbf{x}_{c1} + \mathbf{x}_{c2} + \mathbf{x}_{c3}) \\ \mathbf{n}_i &= \frac{\omega_{di} \times (\mathbf{x}_{ci} - \mathbf{x}_0)}{\|\omega_{di} \times (\mathbf{x}_{ci} - \mathbf{x}_0)\|} \\ \mathbf{a}_i &= \frac{\mathbf{n}_i \times \omega_{di}}{\|\mathbf{n}_i \times \omega_{di}\|}, \end{aligned}$$

where $\mathbf{r}_{x_{vir}}$, $\mathbf{r}_{y_{vir}}$, and $\mathbf{r}_{z_{vir}}$ each denote the column vectors of \mathbf{R}_{vir} as shown in Eq. (12), and the physical meaning of each vector composing \mathbf{f} the object orientation controller is illustrated in Fig. 15b. The desired rotational axis ω_{di} of the grasped object is updated in real-time by the

desired object orientation and the present virtual object frame. The distance between the rotational axis and the contact position of each fingertip l_{di} and the desired contact force F_{di} to generate the desired torque around ω_{di} are computed.

Orientation control using contact position information with a time delay

Figure 16 shows the coordinates of a robotic hand and a grasped object and the contact position used in the simulations. We assume that the robotic hand has three fingers in which each finger has five DOFs and the grasped object is a triangular prism. It is known that there is a considerable time delay caused by visual image acquisition from a visual sensor and its processing cost. The object contact position observed using the visual-tactile sensor uses a camera sensor of up to 30 FPS. In this study, the time delay can be measured, and its range is 20–30 FPS. The measured visual information by the USB camera is processed between 33 and 50 (ms) in LabVIEW. Namely, the detection of the object orientation can be performed within 50 (ms) at the worst. To overcome the time delay, Kawamura's method is introduced, which controls the virtual object position and orientation (not the real object position and orientation) and updates them according to sensing information, including considerable time delay.

Let t_{delay} be the time delay. The virtual object frame $\mathbf{R}_{vir}(t - t_{\text{delay}})$ that consists of the position of the center of each hemispheric fingertip by each joint angle from encoder embedded on each joint [16], the

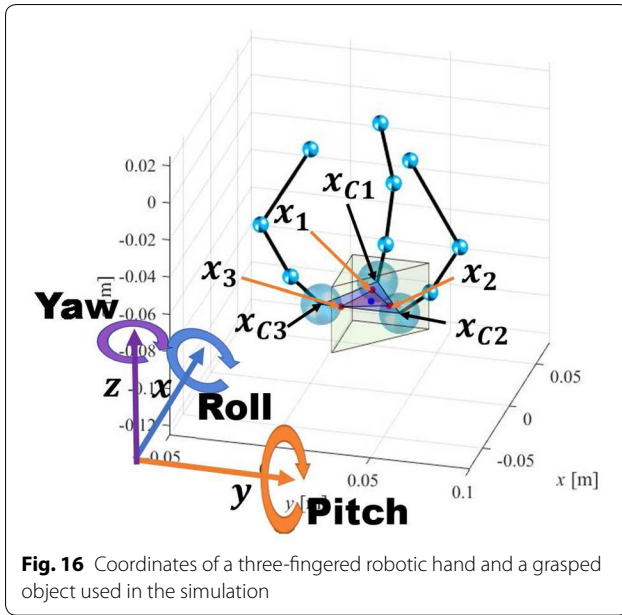


Fig. 16 Coordinates of a three-fingered robotic hand and a grasped object used in the simulation

newly proposed virtual object frame $\mathbf{R}_c(t - t_{\text{delay}})$ that consists of the contact position of each fingertip from the proposed visual-tactile sensor, and the desired virtual object frame $\mathbf{R}_{d_{\text{vir}}}$ that consists of $\mathbf{R}_{\text{vir}}(t - t_{\text{delay}})$, $\mathbf{R}_c(t - t_{\text{delay}})$, and the desired object orientation \mathbf{R}_d are defined as follows:

$$\begin{aligned} \mathbf{R}_{d_{\text{vir}}} &= [\mathbf{r}_{x_{d_{\text{vir}}}} \ \mathbf{r}_{y_{d_{\text{vir}}}} \ \mathbf{r}_{z_{d_{\text{vir}}}}] \\ &= \mathbf{R}_d \mathbf{R}_c^T(t - t_{\text{delay}}) \mathbf{R}_{\text{vir}}(t - t_{\text{delay}}) \end{aligned} \quad (15)$$

where

$$\begin{aligned} \mathbf{R}_{\text{vir}}(t - t_{\text{delay}}) &= [\mathbf{r}_{x_{\text{vir}}} \ \mathbf{r}_{y_{\text{vir}}} \ \mathbf{r}_{z_{\text{vir}}}] \\ \mathbf{R}_c(t - t_{\text{delay}}) &= [\mathbf{r}_{c_x} \ \mathbf{r}_{c_y} \ \mathbf{r}_{c_z}] \end{aligned}$$

Each component of \mathbf{R}_{vir} and \mathbf{R}_c is defined as follows:

$$\begin{cases} \mathbf{r}_{c_z} = \frac{(\mathbf{x}_{c1} - \mathbf{x}_1) \times (\mathbf{x}_{c2} - \mathbf{x}_2) + (\mathbf{x}_{c2} - \mathbf{x}_2) \times (\mathbf{x}_{c3} - \mathbf{x}_3) + (\mathbf{x}_{c3} - \mathbf{x}_3) \times (\mathbf{x}_{c1} - \mathbf{x}_1)}{\|(\mathbf{x}_{c1} - \mathbf{x}_1) \times (\mathbf{x}_{c2} - \mathbf{x}_2) + (\mathbf{x}_{c2} - \mathbf{x}_2) \times (\mathbf{x}_{c3} - \mathbf{x}_3) + (\mathbf{x}_{c3} - \mathbf{x}_3) \times (\mathbf{x}_{c1} - \mathbf{x}_1)\|} \\ \mathbf{r}_{c_y} = \frac{(\mathbf{x}_{c1} - \mathbf{x}_1) \times \mathbf{r}_{c_z}}{\|(\mathbf{x}_{c1} - \mathbf{x}_1) \times \mathbf{r}_{c_z}\|} \\ \mathbf{r}_{c_x} = \frac{\mathbf{r}_{c_y} \times \mathbf{r}_{c_z}}{\|\mathbf{r}_{c_y} \times \mathbf{r}_{c_z}\|} \end{cases} \quad (16)$$

where x_{ci} stands for the position of the center of each hemispheric fingertip that can be obtained from kinematics and encoders embedded on each joint and x_i represents each contact point position on the object surface. The desired virtual object frame $\mathbf{R}_{d_{\text{vir}}}$ as shown in Eq. (15) is updated when new information of the contact position is obtained on-line.

Initial condition of numerical simulation and results

The orientation of the object is regulated by adding the computed desired contact force F_{di} to the object from

each fingertip through the Jacobian of each finger as shown in Eq. (13). The total control input torque for each finger u_i is composed of the summation of the orientation control input and the blind grasping control input that can realize the reliable object grasping already used in Tahara's and Kawamura's methods [16, 17]. The condition of the numerical simulations is shown in Table 3.

The numerical simulation results of the object grasping and orientation control when using our proposed method and when using Tahara's method are shown in Fig. 17.

In our controller, the desired and present virtual frame are expressed as unit vectors of the rotational matrix, but it is not easy to ascertain the orientation intuitively. In the simulation results, the orientation of the object is expressed using the roll-pitch-yaw angle expression. It should be noted that the roll-pitch-yaw angle expression is only used for illustrating the result, and never in control input. It can be seen from these figures that the desired orientation can be realized using the proposed method and its error is clearly reduced compared with Tahara's method. There is an orientation error in Tahara's method because that controller does not use any external sensing information including the contact position, and therefore, the virtual object frame is an approximated frame to express the orientation of the object giving a gap between the actual and the virtual object orientation. Instead, our new controller uses the information of the contact position, which reduces the error. However, there still exists a small error even when our proposed controller is used, which is induced by the design of the virtual object frame as shown in Eq. (12). In this design, we never used the information of the object shape. If the information of the object shape can be used in the controller, the controller would be designed based on a prior knowledge of the object and thus, one of the advantages of our controller, which requires no prior knowledge, would vanish. In this regard, we can confirm that our proposed controller is advantageous compared with Tahara's method.

Experiments

In this section, several grasping and manipulation experiments were performed using a triple-fingered robotic hand in which each finger had three DOFs and the proposed visual-tactile sensors. Due to the difference in degrees of freedom, the movement of the experimental setup of the robot hand was restricted compared with the simulated model. Therefore, in the experiments, in-hand manipulation was performed within a kinematic region in which the number of DOFs would not seriously affect its performance.

Table 3 Condition of the simulations

	Figure 17a, b		Figure 17c, d		Figure 17e, f	
	Initial condition	Desired condition	Initial condition	Desired condition	Initial condition	Desired condition
Roll	0	30	0	0	0	0
Pitch	0	0	0	30	0	0
Yaw	0	0	0	0	0	30
Unit : (deg) (°)						
Notation Value Unit						
Nominal desired grasping force and gains						
Grasping force f_d	1					(N)
Gain for the object orientation control K_r	diag[3.5, 3.5, 3.5]					
Time delay t_{delay}	50					(ms)
Damping coefficient of joint1 B_1	diag[0.07, 0.07, 0.05, 0.04, 0.04] ⁻²					(Ns m/rad)
Damping coefficient of joint2 B_2	diag[0.07, 0.07, 0.05, 0.04, 0.04] ⁻²					(Ns m/rad)
Damping coefficient of joint3 B_3	diag[0.07, 0.07, 0.05, 0.04, 0.04] ⁻²					(Ns m/rad)
Parameters of the robotic hand in simulation						
Length of link1 l_1	0.02					(m)
Length of link2 l_2	0.03					(m)
Length of link3 l_3	0.05					(m)
Mass of link1 m_1	0.02					(kg)
Mass of link2 m_2	0.03					(kg)
Mass of link3 m_3	0.05					(kg)
Inertia of link1 I_1	diag(0.1042, 0.1042, 0.0063) × 10 ⁻⁴					(kg m ²)
Inertia of link2 I_2	diag(0.2250, 0.2250, 0.0375) × 10 ⁻⁵					(kg m ²)
Inertia of link3 I_3	diag(0.6667, 0.6667, 0.2500) × 10 ⁻⁶					(kg m ²)
Radius of fingertip r_f	0.010					(m)
Mass of object M	0.050					(kg)
Inertia of object I	diag(0.0716, 0.1127, 0.0823) × 10 ⁻⁴					(kg m ²)

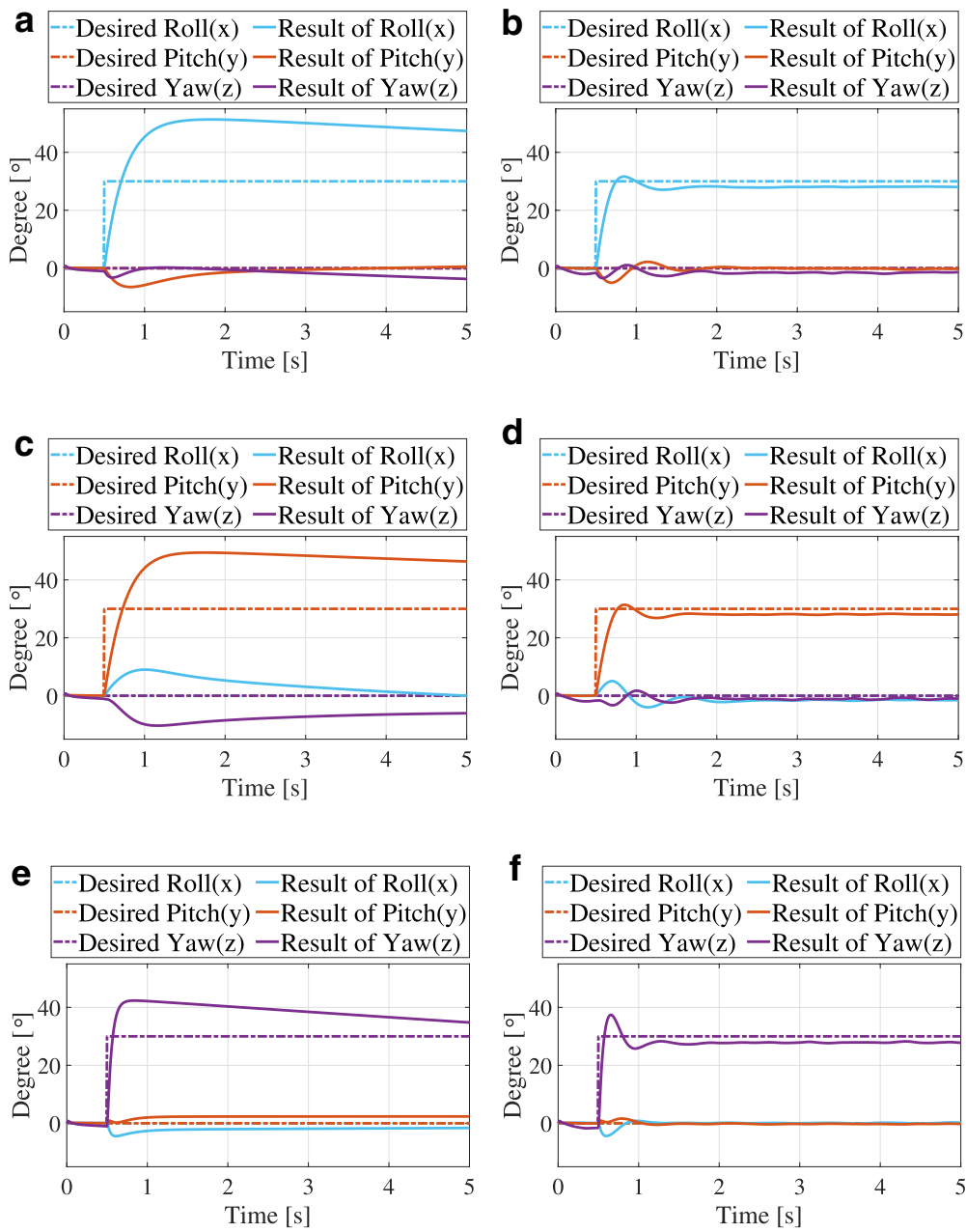


Fig. 17 Simulation results of the object orientation control: (a, c, e) indicate the actual object orientation using Tahara’s method (without using the visual-tactile sensor), (b, d, f) indicate the actual object orientation using the proposed method (using the visual-tactile sensor)

In the experiments, to demonstrate the effectiveness of the proposed method, an external sensor to detect the position and orientation of the object must be used to measure a ground-truth value without the use of proposed visual-tactile sensors.

As shown in Fig. 18, the robotic hand was fixed to a fixture on the ground and the ArUco marker [19] is placed on the grasped object. The position and orientation of the

ArUco marker could be determined through an external USB camera. Even though some errors remained in the detection of object position and orientation by the camera, the attitude of the grasped object is compared using only ArUco markers; thus the measurement through the camera is regarded as ground-truth in the experiment. The overall system of the experimental setup is shown in Fig. 19. The LabVIEW software by National Instruments

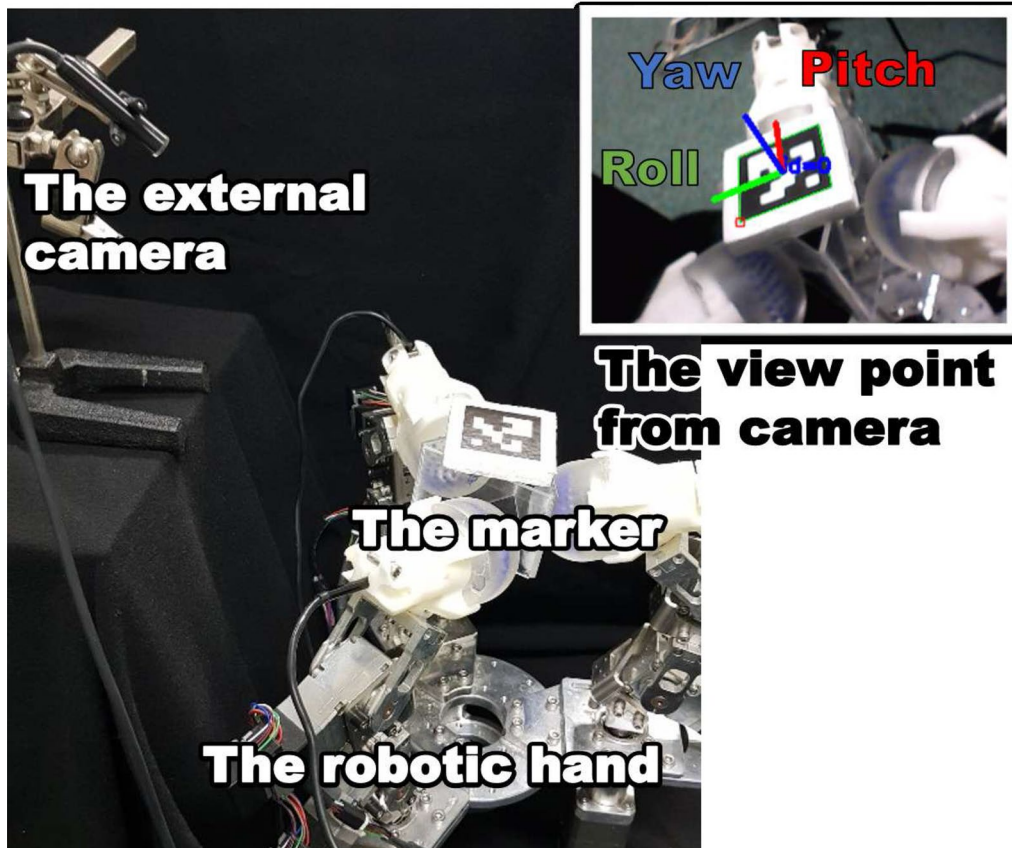


Fig. 18 Experimental setup of the triple-fingered robotic hand with a grasped object and external camera to detect the position and orientation of the object to evaluate the accuracy of the proposed control method

was used to process the visual information from the visual-tactile sensor. The control signal was output to the motor driver by the Compact Rio system by National Instruments. The position and orientation of the grasped object used as a ground-truth were detected by the external camera and LabView. Table 4 shows the initial conditions in the experiment. The experimental results of the object orientation when using our proposed method and when using Tahara's method are shown in Fig. 20. In these figures, the orientation of the object was expressed by roll-pitch-yaw angle for convenience. It should be noted that roll-pitch-yaw angle expression is not used in the controller.

We see from Fig. 20a, b that the proposed method is more accurate in the pitch and yaw directions than Tahara's sensorless control method even though there is a time delay. Figure 20c–e show the yaw angle of the object. In the yaw angle, the accuracy of proposed method is almost like Tahara's sensorless method. It is suggested that in the case of the yaw angle, the

proposed method can achieve the desired angle eventually by suppressing the desired angle. This is because the limit value in the yaw direction is almost up to 30 (deg) due to the mechanical restriction, and in the yaw direction, it is expected to be insensitive to change because it is not much different from the Tahara's virtual frame. On the other hand, the accuracy of Tahara's method is better than the proposed method according to some initial situations because it acts like a feed-forward controller and thus is not affected by the accuracy of feedback information. We can see from Fig. 20e that at the beginning phase, there is an error in the proposed method, but it gradually decreases because the desired virtual object frame is updated by the controller. However, it is evident from these figures that the proposed method can realize a certain level of accuracy regardless of the initial conditions.

Through these experimental results, we can conclude that the proposed method is more accurate than Tahara's method irrespective of initial conditions.

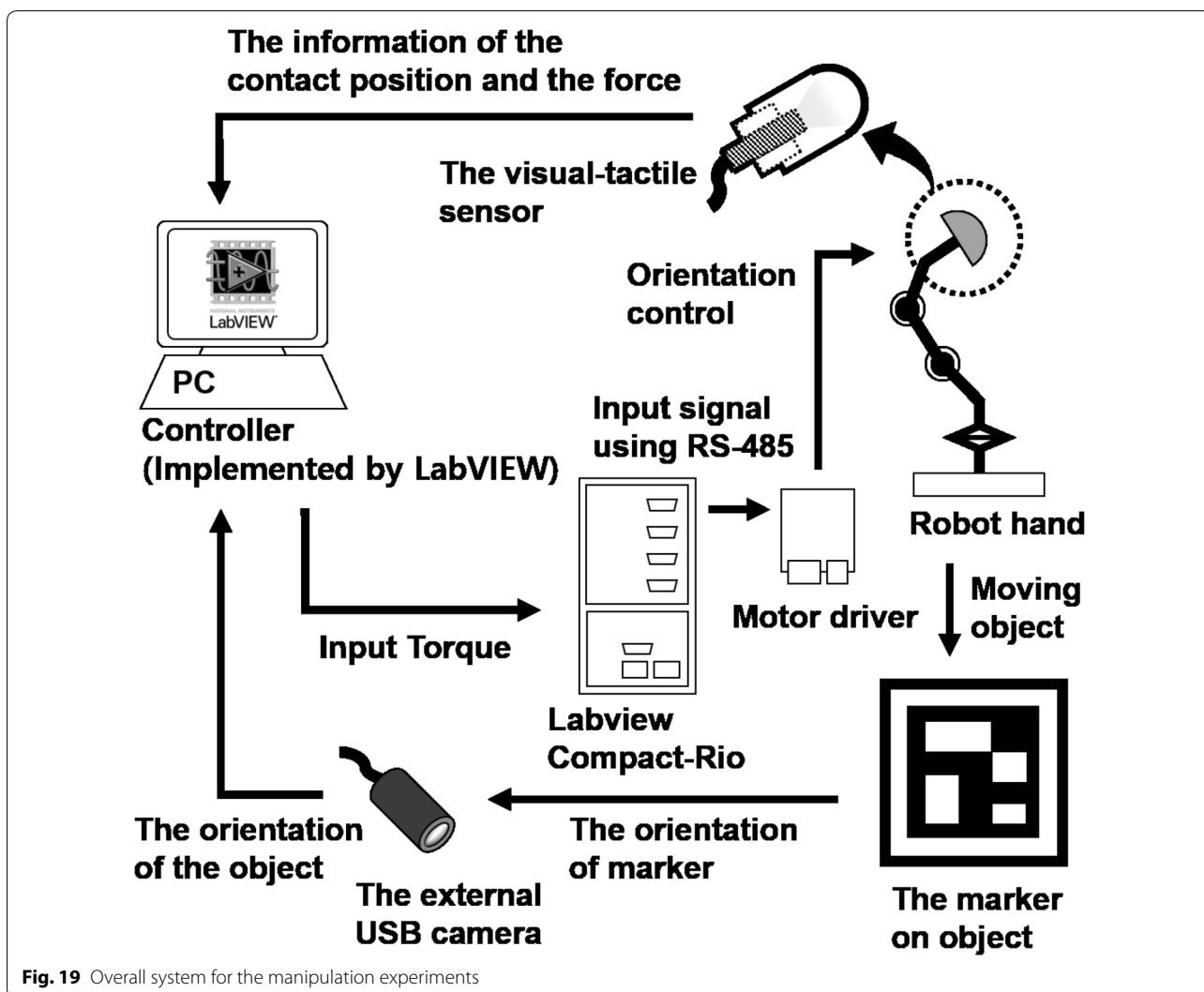


Fig. 19 Overall system for the manipulation experiments

Conclusions

This study proposed a new visual-tactile sensor for a multi-fingered robotic hand and an estimation method for the contact point and force of the sensor that is robust for an external environment. The usefulness of this method was confirmed by comparing the estimated position with the observed contact positions through

several experiments. The new orientation controller was proposed based on the information of contact point positions obtained from the proposed sensor, and its advantages were shown through numerical simulations and experiments. These results confirmed that the proposed control method is more accurate than previous methods and is, therefore, robust for the initial conditions.

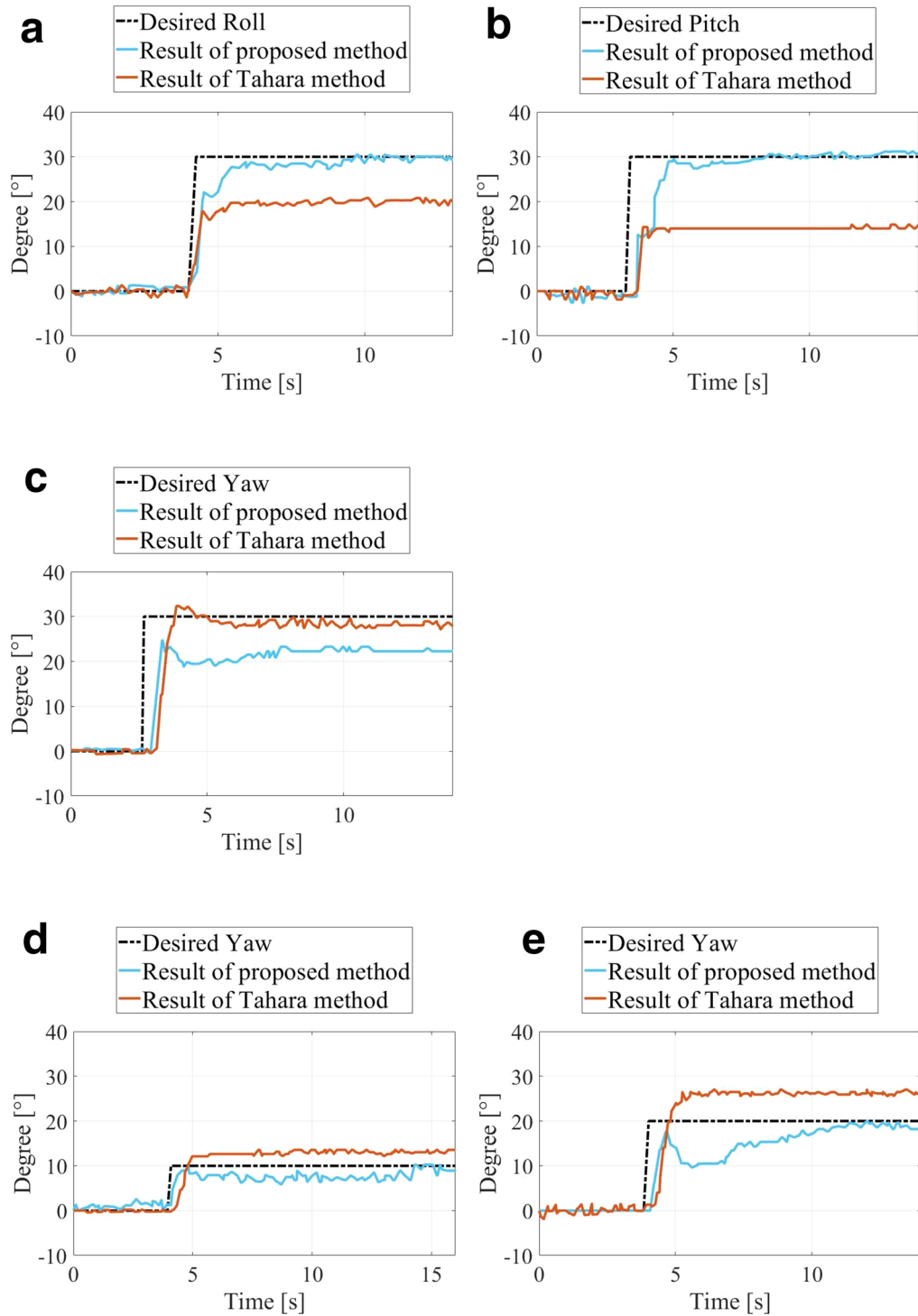


Fig. 20 Experimental results of the object orientation control: **a–c** comparison of the object orientation when the desired **a** roll, **b** pitch, and **c** yaw angles are 30 (deg), **d** comparison of the yaw angle of the object when the desired angle is 10 (deg), **e** comparison of the yaw angle of the object when the desired angle is 20 (deg)

Table 4 Parameters of the prototype of the robotic hand used in the experiments

	Notation	Value	Unit
Radius of fingertip	r_f	0.030	(m)
Initial grasping force	f_d	4	(N)
Length of link1	l_1	0.052	(m)
Length of link2	l_2	0.05	(m)
Length of link3	l_3	0.085	(m)
1nd arm degree	q_1	(0, 0.348, 1.221)	(rad)
2nd arm degree	q_2	(0, 0.348, 1.221)	(rad)
3nd arm degree	q_3	(0, 0.348, 1.221)	(rad)

Specification of motor

Type	Stepping motor (AZM24AK, Oriental Motor Co. Ltd.)
Resolution	0.36°/Pulse
Rotor inertia	9.2×10^{-7} kg m ²
Holding torque	0.095 (N m)
Power input voltage	DC24V
Power input current	1.6A
Gear ratio	20:1

Acknowledgements

This work was supported by the Cabinet Office (CAO), Cross-ministerial Strategic Innovation Promotion Program (SIP), "An intelligent knowledge processing infrastructure, integrating physical and virtual domains" (funding agency: NEDO).

Authors' contributions

All the contents included in the study, such as the idea, method, theory, simulation, and experiments were created and performed by the authors. All authors read and approved the final manuscript.

Funding

This work was supported by the Cabinet Office (CAO), Cross-ministerial Strategic Innovation Promotion Program (SIP), "An intelligent knowledge processing infrastructure, integrating physical and virtual domains" (funding agency: NEDO).

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Mechanical Engineering, Graduate School of Engineering, Kyushu University, 744 Moto'oka, Nishi-ku, Fukuoka 8190395, Japan. ² Department of Mechanical Engineering, Faculty of Engineering, Kyushu University, 744 Moto'oka, Nishi-ku, Fukuoka 8190395, Japan.

Received: 9 October 2019 Accepted: 28 February 2020

Published online: 12 March 2020

References

- Kawasaki H, Komatsu T, Uchiyama K (2002) Dexterous anthropomorphic robot hand with distributed tactile sensor: Gifu hand II. *IEEE/ASME Trans Mech* 7(3):296–303
- Liu H, Greco J, Song X, Bimbo J, Seneviratne L, Althoefer K (2012) Tactile image based contact shape recognition using neural network. In:

- Proceedings of the IEEE international conference on multisensor fusion and integration for intelligent systems, Hamburg, Germany, Sep, pp 138–143
- Jara CA, Pomares J, Candelas FA, Torres F (2014) Control framework for dexterous manipulation using dynamic visual servoing and tactile sensors' feedback. *Sensors* 14(1):1787–1804
 - Kampmann P, Kirchner F (2014) Integration of fiber-optic sensor arrays into a multi-modal tactile sensor processing system for robotic end-effectors. *Sensors* 14(4):6854–6876
 - Yoshikawa T (2010) Multifingered robot hands: control for grasping and manipulation. *Ann Rev Control* 34(2):199–208
 - Takahashi T, Tsuboi T, Kishida T, Kawanami Y, Shimizu S, Iribe M, Fukushima T, Fujita M (2008) Adaptive grasping by multi fingered hand with tactile sensor based on robust force and position control. In: Proceedings of the IEEE international conference on robotics and automation, Pasadena, CA, May, pp 264–271
 - Yusoff H, Abdullah S C, Ohka M (2010) Development of optical three-axis tactile sensor and its application to robotic hand for dexterous manipulation tasks. In: Proceedings of the 4th Asia international conference on mathematical/analytical modelling and computer simulation, Bornea, Malaysia, May, pp 624–629
 - Schmitz A, Maiolino P, Maggiali M, Natale L, Cannata G, Metta G (2011) Methods and technologies for the implementation of large-scale robot tactile sensors. *IEEE Trans Robot* 27(3):389–400
 - Ueda J, Ishida Y, Kondo M, Ogasawara T (2005) Development of the NAIST-hand with vision-based tactile fingertip sensor. In: Proceedings of the IEEE international conference on robotics and automation, Barcelona, Spain, Apr, pp 2332–2337
 - Kamiyama K, Kajimoto H, Kawakami N, Tachi S (2004) Evaluation of a vision-based tactile sensor. In: Proceedings of the IEEE international conference on robotics and automation, New Orleans, LA, Apr, pp 1542–1547
 - Ito Y, Kim Y, Obinata G (2011) Robust slippage degree estimation based on reference update of vision-based tactile sensor. *IEEE Sens J* 11(9):2037–2047
 - Lepora N F, Ward-Cherrier B (2015) Superresolution with an optical tactile sensor. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems, Hamburg, Germany, Sep, pp 2686–2691
 - Yamaguchi A, Atkeson C G (2016) Combining finger vision and optical tactile sensing: reducing and handling errors while cutting vegetables. In: Proceedings of the IEEE-RAS international conference on humanoid robots, Cancun, Mexico, Nov, pp 1045–1051
 - Pestell N, Cramphorn L, Papadopoulos F, Lepora NF (2019) A sense of touch for the shadow modular grasper. *IEEE Robot Autom Lett* 4(2):2220–2226
 - Corradi T, Hall P, Irvani P (2017) Object recognition combining vision and touch. *Robot Biomim* 4(1):2
 - Tahara K, Arimoto S, Yoshida M (2010) Dynamic object manipulation using a virtual frame by a triple soft-fingered robotic hand. In: Proceedings of the IEEE international conference on robotics and automation, Anchorage, AK, May, pp 4322–4327
 - Kawamura A, Tahara K, Kurazume R, Hasegawa T (2012) Robust visual servoing for object manipulation with large time-delays of visual information. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems, Vilamoura, Portugal, Oct, pp 4797–4803
 - Arimoto S, Nguyen PTA, Han H-Y, Dougeri Z (2000) Dynamics and control of a set of dual fingers with soft tips. *Robotica* 18(1):71–80
 - Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas F J, Marín-Jiménez M J (2014) Automatic generation and detection of highly reliable fiducial markers under occlusion. In: IEEE conference on computer vision and pattern recognition, vol. 47, no. 6. Córdoba, Spain, Jun, pp 2280–2292

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.