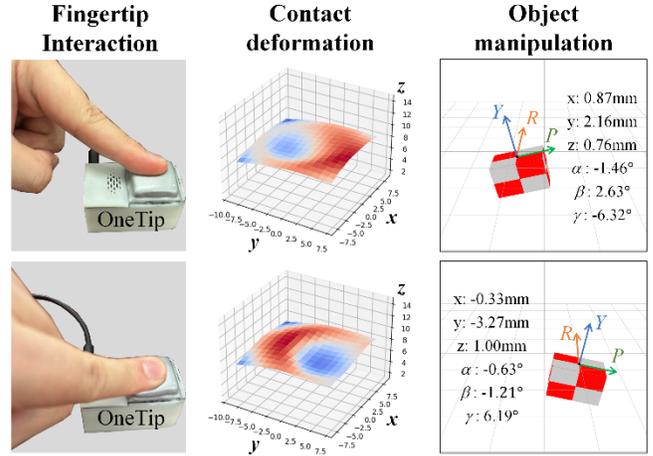


# OneTip Is Enough: A Non-Rigid Input Device for Single-Fingertip Human-Computer Interaction With 6-DOF

\*Mingxuan Li (0000-0003-4159-1831), Yen Hang Zhou (0000-0002-0284-7470), Lunwei Zhang (0000-0001-7800-2821), Tiemin Li (0000-0002-4755-5105), and Yao Jiang (0000-0002-6036-6413)

**Abstract**—With the development of novel display technologies such as 3-D graphical applications, virtual and augmented reality, there are new demands for more efficient and natural multi-degree-of-freedom (DOF) interaction devices. However, traditional rigid touchscreens provide only the 2-D position of the touch point as input information, which cannot meet the needs of high DOF interactions, and many of the existing interaction technologies suffer from the problems of low resolution and lack of information. At the same time, non-rigid interaction interfaces provide deformability for input and have been shown to extend the richness of input vocabulary. This article proposes a new non-rigid input device, OneTip, for single-fingertip human-computer interaction with 6-DOF. In terms of design and manufacture, OneTip employs the bio-inspired design of Skin-On interfaces to mimic the sensitivity of human skin and provide 6-DOF interaction capabilities. In terms of sensing, OneTip uses the visuotactile sensing technique based on the marker displacement method to achieve high-resolution and multi-modal measurements. We propose a novel fingertip pose estimation method based on incipient slip detection, a non-learning algorithm that does not require registration and priori information. Experiments show that OneTip had good 6-D pose estimation accuracy, with RMSEs of translation and rotation not exceeding 0.1mm and 2.6°, respectively, within the linear interval. Extensive experiments were also conducted to explore the application of OneTip in typical virtual manipulation tasks and the possibility of combining it with other interaction devices. This work is intended to serve as a reference for other researchers exploring innovative interaction techniques.



**Index Terms**—Human-computer interaction, non-rigid interaction, force and tactile sensing, visuotactile sensors.

## I. INTRODUCTION

THE development of information visualization and interaction technology is enhancing people’s ability to understand and process 3-D data. Compared to traditional 2-D screens, new visualization methods, such as augmented reality and immersive displays, have expanded the application of 3-D graphics and data models with stereoscopic views, and provided users with beneficial information about medicine, biology, engineering, architecture, and entertainment [1]. Such trends require more effective 3-D input devices to serve 3-D interactive applications fully [2]. However, for typical interactive tasks such as navigation, manipulation, and

selection, the potential number of degrees of freedom (DOF) for 3-D input is six [3], [4]. Compared to the cases of 2-D input with only two or three DOF, 3-D input devices are more complex and have more possibilities in design. Therefore, unlike the mouse that has been established as the standard 2D input device, no input device has proved superior for 3-D interaction tasks so far [5].

Finger-based human-computer interaction has become the most common input mechanism among existing interactive devices. Research mainly explores two approaches: Gesture interaction systems [6] and fingertip touch interfaces [7]. Studies on gesture interaction have focused on obtaining the spatial position and pose of fingers through motion-tracking devices for virtual object operations [8] or emotional communication [9]. Such tasks require the whole finger to move for input operations, making it impossible to achieve intimate and subtle interactions solely relying on fingertips [10]. Touch interfaces have been widely used commercially for measuring finger movement through methods such as capacitive sensing. However, the output of a touch screen is

\*This work was supported in part by the National Natural Science Foundation of China under Grant 52375017, and in part by the iCenter Star of Innovation Program from the Fundamental Industry Training Center, Tsinghua University.

Mingxuan Li, Yen Hang Zhou, Lunwei Zhang, Tiemin Li, and Yao Jiang (corresponding author) are with the Institute of Manufacturing Engineering, Department of Mechanical Engineering, Tsinghua University, Beijing, China. (E-mail: jiangyao@mail.tsinghua.edu.cn).

usually a change in the position of a 2-D touch point. Related research has shown that supplementing additional degrees of freedom, such as finger angles, can enhance the interactive experience [11]. It requires researchers to continuously explore various ways to achieve the full 6-DOF input, which is naturally provided by fingertip pose changes.

The development of non-rigid input technologies [12] (or shape-changing interfaces [13]) provides new solutions for enhanced fingertip interaction. For shape-variable inputs, the physical shape of the contact surface changes dynamically as the user touches it, and can give a richer interaction paradigm than rigid interfaces. Appropriate soft input devices have also been shown to provide better interactivity, feedback, and user experience [14]. However, while existing research has focused on the impact of deformable outputs on effects such as comfort, realism, and emotional communication [15], studies of input performance are relatively limited. Although non-rigid interfaces provide rich expression for touch inputs in many studies, the effective information conveyed for each contact point usually remains 2-D. Therefore, although shape-variable interaction has inspired novel paradigms for finger gestures, the potential for individual fingertips to convey input information has yet to be fully explored.

This article presents OneTip, a novel non-rigid input device for single-fingertip human-computer interaction. Its main advantage lies in supporting 6-DOF input with just one fingertip. OneTip employs a design philosophy based on bio-inspired Skin-On interfaces [16] to facilitate the discoverability of the interaction. Using the marker displacement method (MDM) [17] in visuotactile sensing, we propose an incipient slip-based method for measuring 6-D fingertip pose. OneTip is suitable for performing 3-D operations in a limited touch space, *i.e.*, changing the position and orientation of virtual objects (similar to a joystick), rather than navigating or traveling in a virtual environment on a large scale (similar to a mouse). Experiments show that the maximum RMSE for translation and rotation of the fingertip pose changes is 0.1 mm and 2.6°, respectively, over a linear interval. Finally, the feasibility of OneTip in real-world applications and combination with other interactive devices has been explored through use cases.

## II. RELATED WORK

### A. Touchscreen for Fingertip

Touchscreens are the most common component used in mobile and stationary interactive devices. Tactile interfaces based on capacitive sensing provide a 2-D interaction plane for the fingertips and can convert the disturbances and changes of the electric field generated by touch into capacitive images. Therefore, the 2-D input from the fingertips can be transformed into navigation or travel in virtual environments by establishing a transformation mechanism for obtaining input patterns from capacitive images. More input features have been explored: Harrison *et al.* increased interaction opportunities by categorizing fingertip impact sounds [18]. Watanabe *et al.* estimated manipulation force and finger angle by measuring transmitted light from the fingernail [19]. Boring *et al.* enriched fingertip inputs by measuring the size of the thumb's contact area [20]. Wherein the 3-D finger-in-touch angles (roll, pitch,

and yaw) have received much attention in recent studies [21]-[23]. Such works extracted feature information for estimating fingertip angles from capacitive images by methods such as machine learning. It has been shown that the recovered fingertip angle could effectively complement the interaction degrees of freedom and improve the interaction richness. However, due to the lack of raw information accessed from capacitive screens, similar approaches usually had low resolution and poor accuracy. In addition, none of these works could efficiently estimate the roll angle, which proved to be highly inspirational and ergonomic.

Another solution is to attach more sensors or employ new hardware. Kratz *et al.* [24] and Mayer *et al.* [25] mounted depth cameras on the outside of the touchscreen to estimate yaw and pitch angles by tracking the point cloud of the finger. However, there are limitations in the application scenarios of similar optical tracking-based methods due to the reliance on additional equipment and data processing. Motion tracking-based approaches such as magnetic motion tracking [26] have advantages in terms of localization accuracy and occlusion-free ability. Still, they require additional equipment to be worn on the user's finger. More recent research has explored the feasibility of estimating fingertip pose from fingerprints. A research team from Tsinghua University developed two 3D angle estimation methods based on deep neural networks [11] and fingerprint matching [27], respectively, and have achieved SOTA performance. These works successfully accurately estimated fingertip 3D angles, where the rolling angles were quantitatively measured for the first time. However, such methods rely on large data collection to train models with generalization, otherwise they cannot cover all possible combinations of 3D angles [11]. Even if non-learning frameworks are used, a pre-registration library needs to be constructed for fingerprint matching [27]. Such issues may limit the generality of the method and lead to privacy problems.

According to the above analysis, obtaining more complete fingertip contact degrees of freedom has received extensive attention. Existing studies have mainly focused on 2-D rigid touch interfaces. The insufficient input DOF of the hardware increases the difficulty in the sensing and algorithm design. This article introduces a non-rigid interface to obtain accurate 6-DOF input information. On the one hand, the deformable input mode improves the responsiveness of the touch interface to changes in the 3-D angle of the fingertip. On the other hand, the non-rigid interface can quantitatively measure the depth of the fingertip's press (normal displacement), which has not yet been adequately realized in existing research.

### B. Non-Rigid Input Device

Non-rigid (or deformable) input devices often contain soft and malleable materials that allow the user to interact with the interface by changing its shape, which cannot be supported by rigid interfaces. This article mainly focuses on fingertip-based deformable interaction. In existing studies, deformable interfaces add depth information to 2-D touch inputs and allow users to perform explicit physical operations such as isometric stretching, squeezing, twisting, poking, etc. [12]. For example, Ngyuen *et al.* proposed the SOFTii input system constructing embossed contours for digital design and game control through

sequential modeling [28]. The FoamSense sensor employed conductive ink and a porous soft object to measure the user's compression, bending, twisting, and shearing operations [29]. Weigel *et al.* reported a wearable mini-device, DeformWear, for single-point deformation interaction (e.g., squeezing, compression, and shearing) on a limited surface [30]. The above studies have demonstrated that flexible inputs can provide richer and more expressive user interactions than traditional touch interfaces. In addition, researchers have taken inspiration from human skin to design novel interactive interfaces. Teyssier *et al.* proposed a new bio-driven paradigm, Skin-On Interface, to enhance interactivity by mimicking human skin's sensory, interactive, and sensing properties [16]. Park *et al.* developed a multilayered structure of skin-inspired robotic skin capable of sensing multimodal contact [31]. Such research has emphasized the inspiration of bio-driven approaches for non-rigid input devices and further expanded the possibilities of flexible interaction.

A common problem with the above work is the loss of raw information during measurement. The raw contact information that can be sensed by using electrodes, circuits, and other components is a 2-D response. Ideally, however, we would like to have direct access to the 3-D deformation field of a non-rigid interface. In other words, the potential of flexible interaction has yet to be fully explored. Early studies have introduced structured light techniques to non-rigid input devices. Watanabe *et al.* applied structured light-based triangulation to achieve 3-D sensing of deformable interaction surfaces [32]. The deForm input devices supported 2.5-D touch gesture interactions by measuring 3-D deformation and 2-D texture with a structured light scanner [33]. Follmer *et al.* implemented Jamming user interfaces with a structured light depth sensing system to realize a deForm-like interaction device [34]. Although these methods could acquire 3-D deformation of the input interface, the large-scale and complex imaging elements still limited the usage flexibility. New measurement techniques of 3-D deformation are needed in portable devices.

Considering the above factors, this article employs advanced visuotactile sensing technology to further enhance the performance of non-rigid interaction. Benefiting from the interactivity advantages of non-rigid input devices, visuotactile sensing can acquire dense and complete contact information. Compared with structured light techniques, it can be utilized in more compact devices with better usability.

### C. Visuotactile Sensing Technology

Visuotactile sensing (or vision-based tactile sensing) technology is an effective method for measuring 3-D contact properties and has been widely discussed in recent studies. Such sensors mainly consist of an elastomer with a marker/reflective layer, an optical system and a camera [35]. When an object is in contact with the soft elastomer, the elastomer deformation can be visualized and characterized by the markers or reflected light, which can be captured by the camera. Raw information related to the contact characteristics can be further obtained through image processing and information mapping. Relevant studies include sensors such as GelSight [36], TacTip [37], and GelStereo [38]. Compared with the flexible electronic skin with the principles of capacitance and piezoelectricity, these sensors have the advantages of

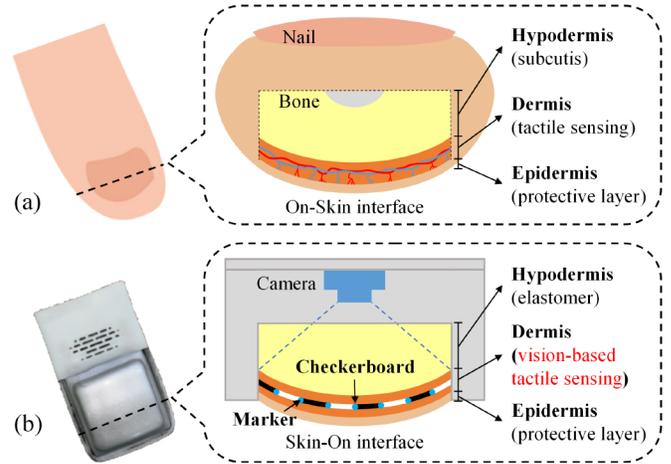


Fig. 1. Skin-on non-rigid interface inspired by human skin. (a) Human skin. (b) Non-rigid interface of OneTip.

multimodality and high resolution, and are suitable for application scenarios with high requirements for sensing performance but low requirements for wearability and power dissipation. Recent visuotactile sensors are still mainly applied for robot grasping and manipulation. Researchers have explored their application in tasks such as dexterous manipulation [39], grasp control [40], and dynamic evaluation [41]. However, attempts for human-computer interaction have been limited and dated [33], [42]. In recent years, with advances in miniaturization, visuotactile sensors have moved away from relying on the large-sized components used in earlier studies and could be shrunk to the size of a human fingertip [43], [44]. This trend hints at their possibility of driving the development of human-computer interaction based on non-rigid interfaces.

Our previous work reported the Tac3D sensor based on the continuous marker pattern (CMP) [45]. It is a visuotactile sensing approach based on the 3-D MDM [17]. Corresponding experiments have demonstrated its advantages of resolution and reliability, and the capability of measuring dense 3-D contact deformation [46]. Based on this approach, we further explore the application of visuotactile sensing in human-computer interaction.

## III. PROTOTYPE AND METHOD

### A. Skin-on Non-rigid Interface

The Skin-On interface proposed by Teyssier *et al.* provides an idea to fabricate interactive devices by mimicking the structure and function of human skin [16]. Human skin has complex microstructures and receptors to realize rich tactile senses. The skin structurally consists of three main layers: the epidermis, the dermis, and the hypodermis [47], as shown in Fig. 1(a). The epidermis provides the contact interface and protects the internal tissues. The dermis contains the majority of sensory receptors and gives the primary tactile feedback. The hypodermis stores a large amount of fat and can provide depth and resistance during contact. According to the investigation by Teyssier *et al.*, when focusing on the interaction performance of the skin and disregarding the visual impact, artificial interfaces

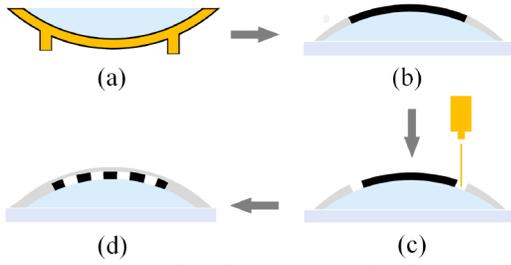


Fig. 2. Fabrication process of the skin-on non-rigid interface. (a) Cast molding to obtain the blank material. (b) Spin coated with black silicone layer as the laser absorption medium. (c) Laser engraving the void patterns and filling the void with colored silicone. (d) Spin coated with the protective layer.

need to replicate human skin functions in terms of texture, strain/thickness, and acuity/sensing [16]. Therefore, the design of OneTip’s non-rigid interface should consider and mimic the sensing properties of human skin in terms of material and structure.

Meanwhile, visuotactile sensing has been shown to have the advantage of mimicking human tactile perception. For example, the TacTip sensor mimics dermal papillae and intermediate ridge structures via pin-like markers to estimate skin surface deformation [16]. Compared to the array-based sensing elements (capacitive sensors) used in the previous Skin-On interface, we chose the visuotactile sensing technique based on the 3-D marker displacement method [17], since it has the advantages of dense deformation measurement, as shown in Fig. 1(b). When the interface deforms due to external contact, the markers in the dermis layer, which provides feature information, will move accordingly. A binocular camera measures the coordinates and displacements of these markers in real time, and the 3-D deformation field of the Skin-On interface can be obtained by interpolation. We shifted the sensing function from the dermis to the posterior end (i.e., the camera that is not directly connected to the non-rigid interface). This change ensures that the sensing layer is less than 1.2 mm thick (to match the thickness of the human dermis) while ensuring the capability to measure dense deformation.

In addition, the continuous marker pattern (CMP) [45] was used in OneTip to provide the texture for the sensing layer. In contrast to discrete markers, the corner points of the black-and-white checkerboard were used as marker points to provide feature information. Teyssier *et al.* used the matrix-layout component arrangement to rely on fewer components and devices to realize the sensing function. Similarly, previous research has demonstrated the usefulness of CMP for ensuring real-time performance and robustness of the sensor under large contact deformations [46]. This characteristic inspires us to explore its applicability in HCI devices fully.

### B. Design and Manufacture of OneTip

Silicone has been shown to be suitable for mimicking the three-layer structure of human skin [16]. To support visuotactile sensing, there are two additional issues to consider for the fabrication of OneTip’s non-rigid interface. First, the hypodermis should provide both simulated kinesthetic compliance and good light transmission. Second, the marker

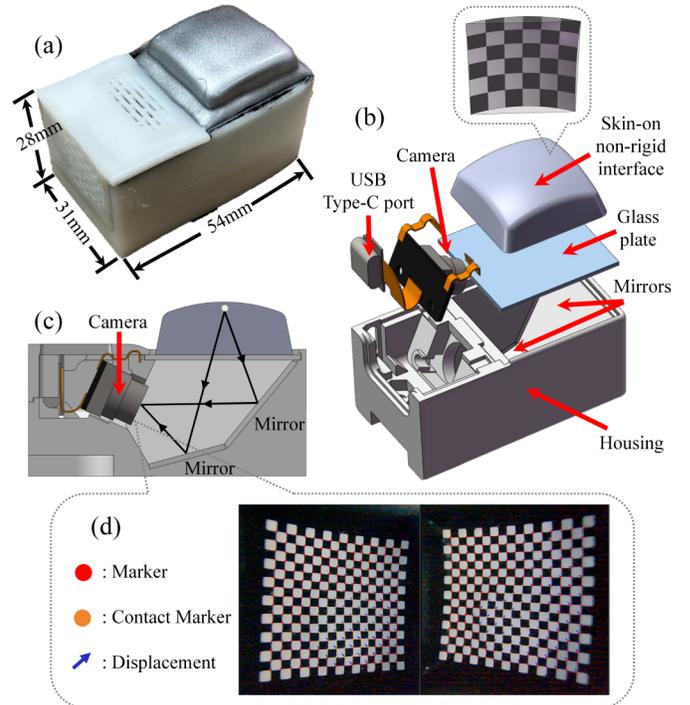


Fig. 3. Design of the OneTip input device. (a) Structure of OneTip. (b) Exploded diagram of OneTip. (c) Virtual binocular vision system used in OneTip. (d) An example of original tactile images.

pattern in the dermis should be easily recognized even when deformed. Based on these, we carried out the following material selection and design exploration [see Fig 2] for the Skin-On interface of OneTip.

#### 1) The hypodermis

The hypodermis is important for reproducing the contact properties of the skin. We selected a transparent organic silicone with a base-to-curing agent ratio 1:1 (XINBANG Co., Ltd., Dongguan, China). This material was characterized by its softness and flexibility close to human fat, and its ability to achieve a hardness suitable for fingertip contact. According to Teyssier *et al.* [16], 5~10 mm was a reasonable range for skin thickness. Considering the needs for sensor compactness and force reconstruction accuracy [48], we compromised by choosing a thickness of about 7 mm. After thoroughly mixing and degassing, the hypodermis was molded by heating at 70°C for 5 hours and casting, as shown in Fig. 2(a).

#### 2) The dermis

The major difficulty of creating the dermis layer is ensuring that the marker pattern can accommodate contact deformation. In this work, we used a continuous fiber laser for fabrication. Since low laser absorption rates of the silicone with bright color may make it difficult for laser processing, a layer of black silicone was used as a laser absorption medium, as shown in Fig. 2(b). On top of this layer, the specified area was etched and filled with white silicone by heated for solidification to form a predefined marker pattern, as shown in Fig. 2(c). Practices have been demonstrated that fiber lasers could achieve higher processing precision than other technologies like water transfer printing. In addition, since the marker pattern was made of the same silica gel as the material of the hypodermic layer, they can deform together without cracking.

### 3) The epidermis

The epidermis layer should provide strong and tough protection for the internal structure and the material and shape that facilitates interaction. This article used AB Platinum silicone with a mixing ratio of 1:1 (SJ3220, SANJING Co., Ltd., Beijing, China). This material possesses flexible and skin-friendly properties as well as good mechanical behaviors. According to Teyssier *et al.*, a Skin-On interface with a large difference from skin color was more likely to be perceived by the user as a device rather than as skin, and a material-free interface could give the user a more comfortable feeling on contact [16]. Considering that OneTip does not have the design requirement to mimic skin in terms of appearance, we only use texture-less silver pigmentation to reduce anthropomorphism. As shown in Fig. 2(d), spin-coating technologies were used to construct the epidermal made of silicone mixed with gray pigment. Spin-coating could ensure a suitable thickness, resulting in a smooth and uniform surface.

The structure of OneTip is shown in Fig. 3(a). Its overall dimensions were 54 mm × 31 mm × 28 mm, and the housing was made of 3D printing. The aforementioned Skin-On non-rigid interface was fixed to the housing, and a camera (1920×1080) was arranged inside the sensor. In addition, components, including an LED light source, metal reflector, and cooling fan, were also fixed inside the sensor [see Fig. 3(b)]. The virtual binocular vision system (VBVS) was also used in OneTip, as shown in Fig. 3(c). This design could achieve a balance between measurement performance and compactness. By cleverly arranging the metal reflectors, the imaging from the marker pattern to the camera can be divided into two optical paths, constituting left and right fields of view on the camera's imaging plane, as shown in Fig 3(d). As a result, OneTip could achieve both simultaneous binocular triggering and a compact structure. Through the standard USB Type-C port, OneTip could be easily connected to external devices for data communication and power supply.

#### C. 3-D Contact Reconstruction

The method for acquiring 3-D contact deformation for OneTip has been introduced in our previous works [45], [46]. It consists of three main steps: recognition, tracking, and reconstruction. First, based on the double-layer circular sampler proposed in [46], the detection of corner features (marker points) can be realized. Next, based on the rigid connection relationship possessed by the features in the continuous marker pattern, the dynamic movement of the marker points can be tracked in real-time. Finally, based on the method for processing binocular tactile images presented in [45], the calculation of the 3-D coordinates of all markers can be realized. The obtained position information can be used as a discrete sampling result of the original contact deformation, and further rely on the interpolation algorithm to reconstruct the 3-D deformation field.

To realize the estimation of the fingertip pose, the contact region should be determined first. The contact region is defined as the region where the contact force is not zero. Considering that the distributed force calculation relies on cumbersome force calibration and is always noisy, we estimate the contact region using the 3-D deformation field, which is simpler and more reliable. The modulus of the deformation vectors of the

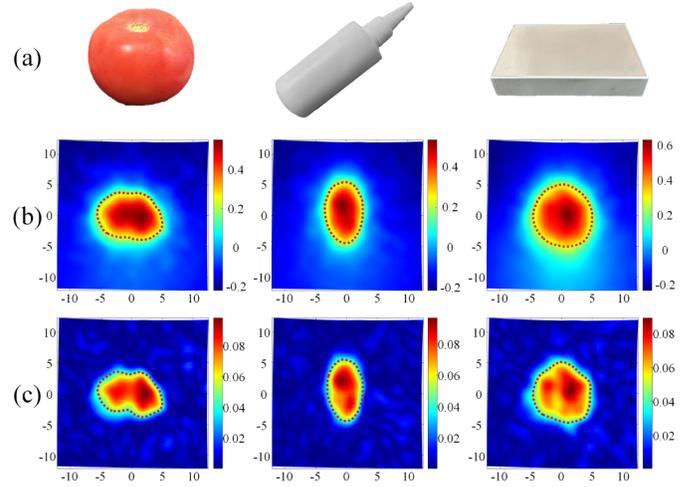


Fig. 4. Contact region estimation for three typical objects. (a) Objects (tomato, plastic bottle, and hard disk). (b), (c) Results based on contact deformation and distributed force [48], respectively. The red dotted line indicates the contact region.

marker arrays is calculated, and the largest of the normal deformation vectors is used as a benchmark. The marker points whose differences are within a certain threshold (50% is chosen in this article) are considered in the contact region. Fig. 4 shows the results of contact region estimation using the 3-D deformation field and the distribution force field obtained using the state-of-the-art force reconstruction algorithm [48], respectively. We used three types of household objects for the evaluation. The results show that the contact regions judged using the 3D deformation field were slightly larger than those based on the distributed force. In addition, the reconstructed deformation field exhibited less noise, which was conducive to reducing the misjudgment under multi-point contact. Considering the effect of noise, we determine that contact occurred when the maximum deformation modulus exceed a certain threshold (0.1 mm is chosen in this article).

#### D. Fingertip Pose Estimation

We further utilize the deformation within the contact region to estimate the fingertip's pose. Previously, there have been studies to measure the 3-D pose of contacting objects based on distributed tactile perception [49]. Assume that the 3-D coordinates of  $n$  marker points within the contact region at the initial moment are  $\mathbf{P}^0 = \{\mathbf{p}_1^0, \mathbf{p}_2^0, \dots, \mathbf{p}_n^0\}$ , and the 3-D coordinates at the  $k$ th moment are  $\mathbf{P}^k = \{\mathbf{p}_1^k, \mathbf{p}_2^k, \dots, \mathbf{p}_n^k\}$ . The pose transformation from the initial moment to the current moment can be realized by the classical ICP point-cloud alignment, which is the process of the least-squares fitting problem between two 3-D point sets:

$$\min_{R, t} E(\mathbf{R}, \mathbf{t}) = \min_{R, t} \frac{1}{n} \sum_{i=1}^n \|\mathbf{p}_i^0 - (\mathbf{R}\mathbf{p}_i^k + \mathbf{t})\|^2. \quad (1)$$

The Euclidean transformations  $R$  and  $t$  denote the rotation and translation matrices between two point-sets, which can be solved by singular value decomposition (SVD).

However, the range of changes in fingertip pose during contact is much greater, and the contact pattern is more complex. A typical gesture input method is scrolling around the

longitudinal axis of the finger, which is more often attempted by users due to its affinity to ergonomics. In this type of contact action, incipient slip or even macroscopic slippage usually occurs on the contact surface between the finger and the non-rigid interface. However, since the point cloud alignment method uses marker points on the non-rigid interface to describe the object's motion, it implicitly assumes that the contact surface is fully sticking with the finger.

Therefore, considering the influence of the slippage, a novel contact pose estimation method is proposed. First, the stick/slip state on the contact surface is distinguished using the method in [50], which detects whether the marker points belong to the stick or slip points. The method can be used for measuring fingertip contact since it was proven to apply to the incipient slip measurement of soft objects. As shown in Fig. 5, the coordinate systems  $\{O_S\}$  and  $\{O_F\}$  are constructed on the contact surfaces of the non-rigid interface and the fingertip, respectively. The points  $p_S$  and  $p_F$  whose positions coincide at the initial moment  $t_0$  are selected in  $\{O_S\}$  and  $\{O_F\}$ , and the position vectors of them denote  $\mathbf{r}_S^{OS}(t_0)$  and  $\mathbf{r}_F^{OF}(t_0)$  (the superscripts denote that they are within the  $\{O_S\}$  and  $\{O_F\}$ , respectively), which satisfy

$$\mathbf{r}_S^{OS}(t_0) = \mathbf{r}_F^{OF}(t_0), \quad (2)$$

$$\mathbf{r}_F^{OS}(t_0) = \mathbf{r}_c^{OS}(t_0) + \mathbf{R}_c^{OS}(t_0) \cdot \mathbf{r}_F^{OF}(t_0), \quad (3)$$

where  $\mathbf{R}_c^{OS}(t_0)$  and  $\mathbf{r}_c^{OS}(t_0)$  denote the rotation and translation matrices of  $\{O_F\}$  with respect to  $\{O_S\}$  at the moment  $t_0$ , respectively. After  $\Delta t$  time (small quantity), let the position vectors of  $p_S$  and  $p_F$  be  $\mathbf{r}_S^{OS}(t_0 + \Delta t)$  and  $\mathbf{r}_F^{OF}(t_0 + \Delta t)$ , and the rotation and translation matrices become  $\mathbf{R}_c^{OS}(t_0 + \Delta t)$  and  $\mathbf{r}_c^{OS}(t_0 + \Delta t)$ . Transforming  $\mathbf{r}_F^{OF}(t_0 + \Delta t)$  into  $\{O_S\}$  gives:

$$\mathbf{r}_F^{OS}(t_0 + \Delta t) = \mathbf{r}_c^{OS}(t_0 + \Delta t) + \mathbf{R}_c^{OS}(t_0 + \Delta t) \cdot \mathbf{r}_F^{OF}(t_0 + \Delta t). \quad (4)$$

At this point, the displacement of  $p_S$  and  $p_F$  in  $\{O_S\}$  can be expressed as

$$\mathbf{u}_S^{OS} = \mathbf{r}_S^{OS}(t_0 + \Delta t) - \mathbf{r}_S^{OS}(t_0), \quad (5)$$

$$\mathbf{u}_F^{OS} = \mathbf{r}_F^{OS}(t_0 + \Delta t) - \mathbf{r}_F^{OS}(t_0). \quad (6)$$

According to Eq. (6),

$$\mathbf{u}_F^{OS} = \Delta \mathbf{r}_c^{OS}(t_0) + \mathbf{R}_c^{OS}(t_0 + \Delta t) \cdot \Delta \mathbf{r}_F^{OF}(t_0) + \Delta \mathbf{R}_c^{OS}(t_0) \cdot \mathbf{r}_F^{OF}(t_0 + \Delta t). \quad (7)$$

Consider that  $p_S$  maybe in the stick or slip states:

1)  $p_S$  is in the stick state, and it can be inferred from the geometric constraint that  $\mathbf{u}_F^{OS} = \mathbf{u}_S^{OS}$ . Therefore,

$$\mathbf{r}_c^{OS}(t_0 + \Delta t) + \mathbf{R}_c^{OS}(t_0 + \Delta t) \cdot \mathbf{r}_F^{OF}(t_0 + \Delta t) - \mathbf{r}_S^{OS}(t_0 + \Delta t) = 0. \quad (8)$$

Substituting Eq. (2) and Eq. (3) into Eq. (8) yields:

$$\Delta \mathbf{r}_c^{OS}(t_0) + \Delta \mathbf{R}_c^{OS}(t_0) \cdot \mathbf{R}_c^{OS}(t_0)^{-1} \cdot [\mathbf{r}_S^{OS}(t_0) - \mathbf{r}_c^{OS}(t_0)] - \Delta \mathbf{r}_S^{OS}(t_0) = 0. \quad (9)$$

2)  $p_S$  is in the slip state, and there is relative slipping between points  $p_S$  and  $p_F$ . According to the characteristics of the slippage, the relative slip should be along the tangential

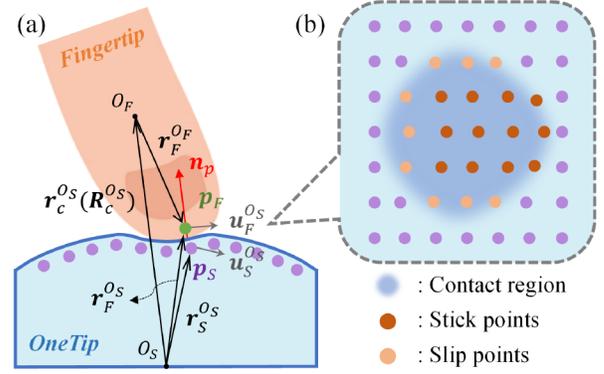


Fig. 5. (a) Contact model between the fingertips and the non-rigid interface. (b) Incipient slip phenomenon and division of stick/slip states of marker points.

direction of the contact surface. Since  $p_S$  is still in contact with the fingertip at this point, it should maintain the same displacement with the fingertip's surface in the normal direction. Considering that  $\Delta t$  is a small amount, the normal direction of the contact geometry at  $p_S$  and  $p_F$  is taken to be the same value  $\mathbf{n}_p$ . According to geometric constraint,  $\mathbf{n}_p \cdot \mathbf{u}_F^{OS} = \mathbf{n}_p \cdot \mathbf{u}_S^{OS}$ . Therefore,

$$\mathbf{n}_p \cdot [\mathbf{r}_c^{OS}(t_0 + \Delta t) + \mathbf{R}_c^{OS}(t_0 + \Delta t) \cdot \mathbf{r}_F^{OF}(t_0 + \Delta t) - \mathbf{r}_S^{OS}(t_0 + \Delta t)] = 0. \quad (10)$$

Similarly,

$$\mathbf{n}_p \cdot [\Delta \mathbf{r}_c^{OS}(t_0) + \Delta \mathbf{R}_c^{OS}(t_0) \cdot \mathbf{R}_c^{OS}(t_0)^{-1} \cdot [\mathbf{r}_S^{OS}(t_0) - \mathbf{r}_c^{OS}(t_0)] - \Delta \mathbf{r}_S^{OS}(t_0)] = 0. \quad (11)$$

We let the rotation matrix and translation matrix of the fingertip concerning the contact interface be  $\mathbf{R}$  and  $\mathbf{t}$ , respectively. Select a point  $p$  on the contact surface, and let the normal direction of the contact geometry at the point be  $\mathbf{n}_p$ . According to Eq. (9) and Eq. (11), the 3-D coordinates of  $p$  at a certain moment  $k$ ,  $\mathbf{p}^k$ , satisfy the relations:

$$\begin{cases} \Delta \mathbf{t}^k + \Delta \mathbf{R}^k \cdot (\mathbf{R}^k)^{-1} \cdot (\mathbf{p}^k - \mathbf{t}^k) - \Delta \mathbf{p}^k = 0, & \text{if } p \text{ is stick point} \\ \mathbf{n}_p \cdot [\Delta \mathbf{t}^k + \Delta \mathbf{R}^k \cdot (\mathbf{R}^k)^{-1} \cdot (\mathbf{p}^k - \mathbf{t}^k) - \Delta \mathbf{p}^k] = 0, & \text{if } p \text{ is slip point} \end{cases} \quad (12)$$

where  $\Delta \mathbf{t}^k$ ,  $\Delta \mathbf{R}^k$  and  $\Delta \mathbf{p}^k$  denote the increments of the corresponding physical quantities at the moment  $k$ . By assuming that the number of stick and slip points in the contact surface are  $n_{\text{stick}}$  and  $n_{\text{slip}}$  with the 3D coordinates of  $\mathbf{P}^k = \{\mathbf{p}_1^k, \mathbf{p}_2^k, \dots, \mathbf{p}_{n_{\text{stick}}}^k\}$  and  $\mathbf{Q}^k = \{\mathbf{q}_1^k, \mathbf{q}_2^k, \dots, \mathbf{q}_{n_{\text{stick}}}^k\}$ , respectively, the transformation of the fingertip's pose from moment  $k-1$  to moment  $k$  can be solved by least-squares fitting as

$$\min_{\mathbf{R}^k, \mathbf{t}^k} E(\mathbf{R}^k, \mathbf{t}^k) = \min_{\mathbf{R}^k, \mathbf{t}^k} \left[ \frac{1}{n_{\text{stick}}} \sum_{i=1}^{n_{\text{stick}}} \|F_i(\mathbf{R}^k, \mathbf{t}^k)\|^2 + \frac{1}{n_{\text{slip}}} \sum_{j=1}^{n_{\text{slip}}} \|G_j(\mathbf{R}^k, \mathbf{t}^k)\|^2 \right], \quad (13)$$

where

$$F_i(\mathbf{R}^k, \mathbf{t}^k) = \mathbf{t}^k - \mathbf{t}^{k-1} + (\mathbf{R}^k - \mathbf{R}^{k-1}) \cdot (\mathbf{R}^{k-1})^{-1} \cdot (\mathbf{p}_i^{k-1} - \mathbf{t}^{k-1}) - \mathbf{p}_i^k + \mathbf{p}_i^{k-1}, \quad (14)$$

and

$$G_j(\mathbf{R}^k, \mathbf{t}^k) = \mathbf{n}_{q_j} \cdot [\mathbf{t}^k - \mathbf{t}^{k-1} + (\mathbf{R}^k - \mathbf{R}^{k-1}) \cdot (\mathbf{R}^{k-1})^{-1} \cdot (\mathbf{q}_i^{k-1} - \mathbf{t}^{k-1}) - \mathbf{q}_i^k + \mathbf{q}_i^{k-1}]. \quad (15)$$

In Eq. (15),  $\mathbf{n}_{q_j}$  denotes the contact geometry normal direction at  $\mathbf{q}_j$ . In Eq. (13),  $\mathbf{n}_{q_j}$ ,  $\mathbf{p}_i^k$  ( $\mathbf{p}_i^{k-1}$ ) and  $\mathbf{q}_i^k$  ( $\mathbf{q}_i^{k-1}$ ) can be measured by OneTip in real-time, and  $\mathbf{R}^{k-1}$  and  $\mathbf{t}^{k-1}$  are derived from the results of the previous round of calculation. Initially,  $\mathbf{R}^0 = \mathbf{I}$  and  $\mathbf{t}^0 = \mathbf{0}$ . Thus, the pose transformation from the initial moment to the current moment can be obtained by continuously computing  $\mathbf{R}^k$  and  $\mathbf{t}^k$  at the current moment according to Eq. (13) from the initial moment.

In addition, there are six unknowns in  $\mathbf{R}^k$  and  $\mathbf{t}^k$  in Eq. (13). For the stick point, each vector equation  $F_i(\mathbf{R}^k, \mathbf{t}^k)$  can provide three constraints, while  $G_j(\mathbf{R}^k, \mathbf{t}^k)$  can only offer one constraint for the slip point. Therefore, to make the solution of the pose transformation based on Eq. (13) feasible, the number of effective stick and slip points within the contact surface should be sufficient to provide at least six constraints. In addition, considering the high pathology of the solution relying only on slip points in practice, at least one effective stick point should exist to provide reliable 3-D position information. Therefore,  $n_{\text{stick}}$  and  $n_{\text{slip}}$  should satisfy

$$\begin{cases} 3n_{\text{stick}} + n_{\text{slip}} \geq 6 \\ n_{\text{stick}} \geq 1 \end{cases}. \quad (16)$$

Compared with Eq. (1), the solution of Eq. (13) takes into account the influence of the stick/slip state on the pose constraints, which can avoid the calculation error caused by the potential assumption of a full sticking. Thus, the proposed method can still ensure the solution accuracy when the incipient slip occurs during the contact. When macroscopic slip occurs, the described method is difficult to meet the constraints of Eq. (16) and may fail. In this case, the displacement of the center of the contact region is used as the output of OneTip instead.

### E. 6-D Signal Input

Based on the above process, we can obtain the pose variation of the fingertip relative to the initial moment (when contact has just occurred) at each frame moment. Considering that the interaction device needs to output an effective 6-D analog quantity and the interaction space typically does not reach singular angles (can avoid gimbal deadlock), the rotation matrix can be transformed into Euler angles. We use a simplified pose accumulation strategy for 6-D signal input. The transformation matrix of each neighboring frame is transformed into 3-DOF Euler angle increments and 3-DOF displacement increments and then directly accumulates to obtain the total input at each moment. At the same time, a moving average filter with a window length of five is introduced to obtain a smooth output. In addition, the outputs are individually multiplied by a fixed gain value, considering the fingertips' suppleness. Ultimately, at each moment  $i$ , a 6-D vector containing the fingertip pose is calculated:

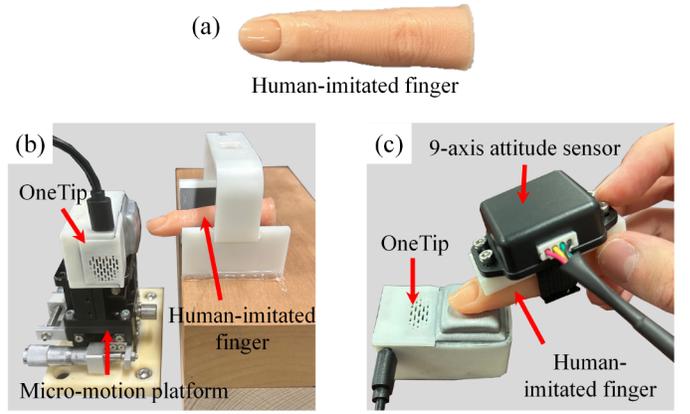


Fig. 6. Experiment Setup. (a) Human-imitated finger. (b) Experimental platform for displacement measurement evaluation. (c) Experimental platform for rotation measurement evaluation.

$$\text{Input}_i = \{\Delta x_i, \Delta y_i, \Delta z_i, \Delta \alpha_i, \Delta \beta_i, \Delta \gamma_i\} \in R^6, \quad (17)$$

where  $\alpha, \beta, \gamma$  correspond to the angles of yaw, pitch, and roll axes, respectively. The video is available [51].

It is worth mentioning that the geometric information of the fingertip is always unknown during the input process, considering generalizability and user privacy. In other words, the result of the described method is to obtain the change of the fingertip's pose rather than the absolute pose. Therefore, the described method does not need to obtain a priori information about the contacting object, thus being able to be applied to different people's fingers. For each contact, the 6-D pose of the fingertip at the beginning of the contract (based on the thresholds defined in Section III-C) is taken as the initial value, and the change in pose is calculated in real time. The initial value is cleared when the current contact ends, and the next contact is waiting.

## IV. EXPERIMENTS

We evaluated the performance of OneTip through specific experiments. First, the effective interaction space of OneTip was assessed based on the collected data generated from real interactions. The estimation accuracy of OneTip in displacement ( $x, y$ , and  $z$  axes) and rotation (yaw, pitch, and roll axes) was evaluated through qualitative and quantitative experiments. Due to the unavoidable joint bending movements of real fingers during the interaction, it is difficult for wearable devices to accurately measure the variation of fingertip orientation. Meanwhile, real fingers are hard to securely fix on the calibration platform for the measurement of position. Therefore, we used a human-imitated finger, which has a similar degree of hardness to that of a human finger, and a built-in artificial bone [see Fig. 6(a)] for the quantitative evaluation experiments. The experimental platforms shown in Fig. 6(b) and Fig. 6(c) were set up to evaluate displacement and rotation estimation, respectively. Finally, the data output frequency (real-time performance) of OneTip was assessed.

### A. Effective Interaction Space Analysis

OneTip's interaction space was analyzed to assess the effective range of 6-DOF pose changes, as well as to estimate

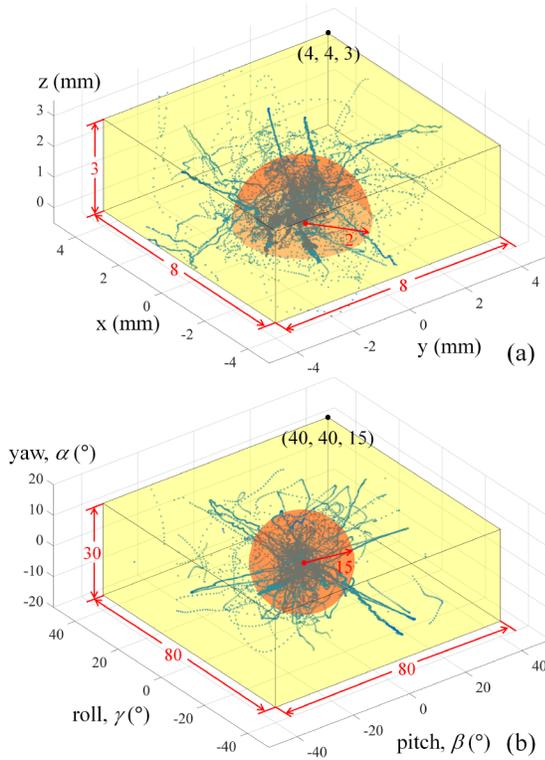


Fig. 7. The range of effective interaction space. (a) Displacement space. (b) Rotation space. The yellow area is the maximum output range, and the orange area indicates the centralized distribution region of the output data.

intervals in which the output data is centrally distributed for usage guidance. During the interaction, the non-rigid interface of OneTip was touched to slowly and steadily change the fingertip pose to ensure that the widest possible range of displacements and angle changes were covered. A total of 29425 sets of valid data were captured.

Fig. 7 shows the collected data's distribution in displacement and rotation space, respectively. Except for a very small number of discrete values, the output data of OneTip is distributed in the range of

$$\begin{cases} \Delta x \in [-4, 4] \\ \Delta y \in [-4, 4] \\ \Delta z \in [0, 3] \\ \Delta \alpha \in [-15, 15] \\ \Delta \beta \in [-40, 40] \\ \Delta \gamma \in [-40, 40] \end{cases} \quad (18)$$

which is regarded as the maximum output range of OneTip (i.e., the yellow area in Fig. 7). Considering that the fingertip movement are relatively conservative in real interactions, we focus on the common output range of OneTip, which is the region of data concentration distribution indicated by the orange area in Fig. 7 (72.8% and 74.9% of all data, respectively). For displacement space, the common output range is roughly a hemisphere with the origin as the midpoint and a radius of 2 mm (non-rigid interfaces have only positive outputs in the z-direction), while the output range for rotation space is roughly a hemisphere with a radius of 15°.

## B. Displacement measurement evaluation

Displacement measurement experiments were conducted using the platform shown in Fig. 6(b). Considering there have not been interaction devices with similar principles and methods as OneTip, only evaluations of its performance were conducted without comparative experiments. By adjusting the 3-axis micro-motion platform with a translation accuracy of 0.02 mm, the displacement of the human-imitated finger in contact with OneTip can be quantitatively controlled. According to the maximum output range obtained in Section IV-A, we compared the displacement estimated by OneTip and the ground truth provided by the micro-motion platform distributed in three directions.

The results are shown in Fig. 8(a). In the z-axis direction (i.e., normal direction), OneTip shows good accuracy and linearity over the effective range of 0~3.5 mm. In contrast, the linear region in the x- and y-axis directions is relatively small, and only when the horizontal displacement is less than about 1.5mm do the measurement results conform to the linear trend. And when exceeding this range, the output of OneTip shows a non-linear trend and the rate of increase of the output displacement increases significantly. This result is mainly due to the effect of macro slip on the method described in Section III-D. Although the described method considers the impact of the incipient slip, the constraints described by Eq. (14) invalidate the approach after a full macro-slip. As the horizontal displacement increases, the slip region gradually expands until complete slip occurs. At this point, the displacement estimation method is changed to a strategy based on the center of the contact area, leading to an increase in the growth ratio of the output. Besides, Fig. 8(b) demonstrates the quantitative assessment results of the accuracy in the linear region for each of the three directions. The results show that the measured displacements are very close to the same straight line within the linear region, and none of the root-mean-square errors exceed 0.1 mm.

## C. Rotation measurement evaluation

Fig. 6(a) shows the experimental setup for rotation measurements. A 9-axis attitude sensor WT901BC with an accuracy of 0.2 degrees was worn on the bionic finger through a wearable standoff made by 3D printing. Similar to the inertial measurement unit used by Duan *et al.* [27], a small sensing device better fits the experimental needs. The yaw, pitch, and roll angles of the human-imitated finger were steadily varied separately to ensure that the range of each angle covered the maximum output range obtained in Section IV-A. At the same time, it was ensured that the contact depth in the z-direction was about 0.5 mm during the process.

The data from OneTip and the three-axis angular information provided by the attitude sensors (as the ground truth) are recorded in real-time, as shown in Fig. 9(a). The results show that OneTip achieves good linear estimation over the maximum output range for all three directions of rotation angles. Qualitative comparisons show that the measurements of the roll and pitch angles are more consistent with the ground truth than the yaw angle. In addition, both the measured yaw and pitch angle are larger than the ground truth when the rotation exceeds the maximum output range. At this point, OneTip exceeded its permissible range and could not give a credible estimation.

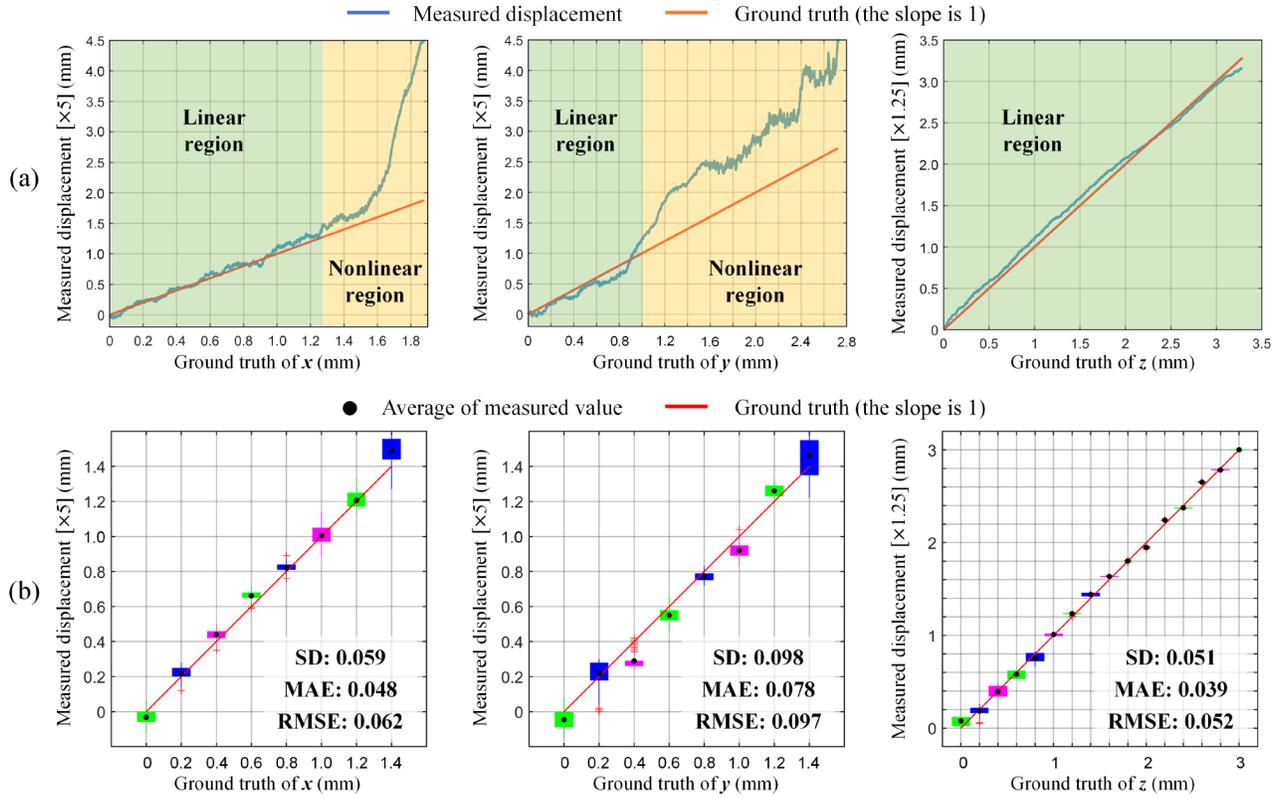


Fig. 8. Displacement measurement evaluation. (a) Qualitative evaluation of displacement measurement. (b) Quantitative assessment of OneTip's displacement estimation in the linear regions. SD: standard deviation, MAE: mean absolute error, RMSE: root-mean-square error.

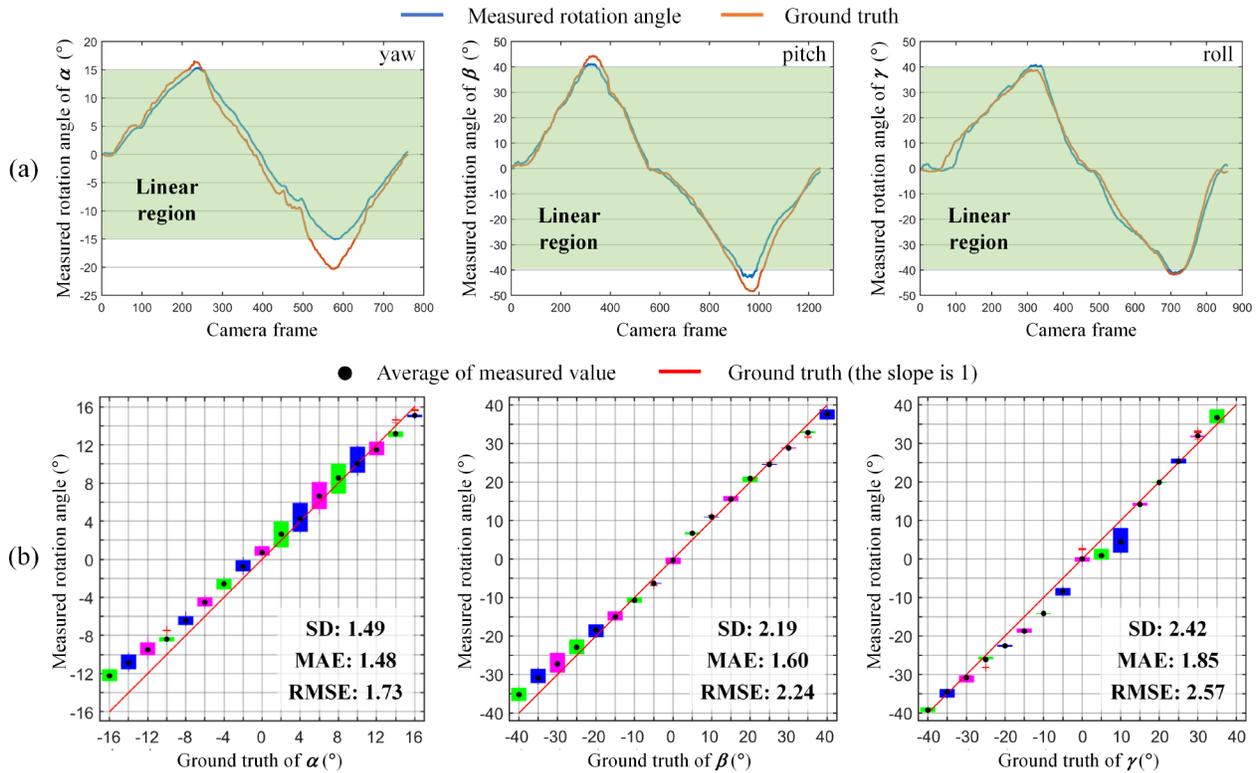


Fig. 9. Rotation measurement evaluation. (a) Qualitative evaluation of rotation measurement in continuous interaction. (b) Quantitative assessment of OneTip's rotation estimation. SD: standard deviation, MAE: mean absolute error, RMSE: root-mean-square error.

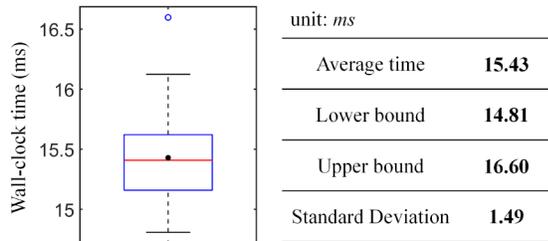


Fig. 10. Real-time measurement performance.

Fig. 9(b) illustrates a quantitative assessment of the measured data in Fig. 9(a). In this case, the data in each boxplot contains measurements attributed to intervals centered on the corresponding integer scale (the length of the intervals is determined by the minimum scale unit). The results show that the root-mean-square error (RMSE) does not exceed  $2.6^\circ$  for both the roll and pitch angles, and the root-mean-square error does not exceed  $1.8^\circ$  for the yaw angles. Since the output range of the roll and pitch axes is much larger, the above results explain that the accuracy of OneTip’s angle estimation for the yaw axis (i.e., normal-to-normal) is lower than that of the other two, which is consistent with the qualitative results.

Due to the difference in sensing principles, device designs, and experimental methods, objectively comparing with existing fingertip interaction devices is impossible. From a subjective point of view, compared to state-of-the-art methods using rigid touchscreens (e.g., fingerprint-based 3-D finger angle estimation [11]), OneTip improves the accuracy by sacrificing the effective interaction space (i.e., the range of variations in fingertip pose). Thus, OneTip is more suitable for manipulating virtual objects (like a joystick) rather than navigating and traveling in a virtual environment (like a mouse).

#### D. Real-time Performance Evaluation

Since the computing efficiency of an input device affects its interaction smoothness and sensitivity, there are concerns about real-time performance. The computing speed of OneTip was evaluated using the average wall-clock time (i.e., the full time it took the computer to complete the application). All programs were implemented in C++/OpenCV and ran on a laptop with a 2.30 GHz Intel i7-12700H processor. The average execution time for 100 consecutive frames of tactile images was calculated over multiple interactions, and the experiment was repeated 30 times.

Fig. 10. shows that the average processing time is 15.43 ms, and the longest processing time does not exceed 16.60 ms. Thus, the output frequency of OneTip is nearly 65 Hz, equivalent to one-fourth of that of a typical gamepad. For general interactive tasks, such as digital design and model manipulation, OneTip has met the corresponding real-time requirements. Further algorithmic and hardware optimizations are needed for tasks with higher payback requirements. For example, the speed of OneTip can be increased using field-programmable gate arrays (FPGAs) and sequential marker patterns with lower information density to meet the desired requirements.

### V. APPLICATION

#### A. 3-D Object Manipulation

The experiments in Section IV demonstrate that OneTip can

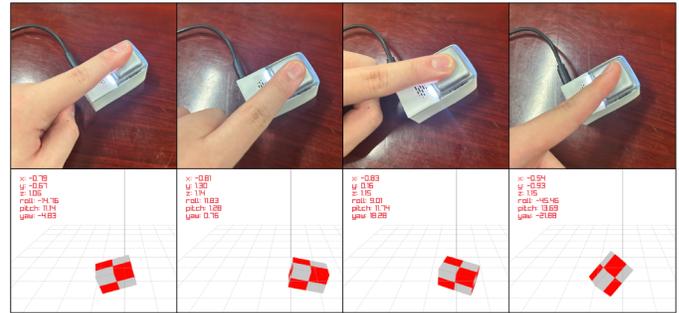


Fig. 11. 3-D model manipulation using OneTip. Video is available [51].

output variation of the 6-DOF fingertip pose and is suitable for performing 3-D human-computer interaction in a limited touch space. Fig. 11 shows an application that utilizes OneTip to manipulate a virtual object. When polling, OneTip outputs a 6-D vector to change the position and orientation of the object. The pose change of the fingertip is mapped to the translation and rotation of the model, and orientation compliance (i.e., the directional correspondence between the magnified rotation and the rotation of the fingertip) is preserved. Such ways of interaction are characterized by the ability to simultaneously rotate and translate virtual objects with only a single finger. Compared to existing 6-DOF devices [4], OneTip occupies a small amount of space and limb movement, making it suitable for private and subtle interactions. In addition, the variation of the fingertip pose corresponds intuitively to that of the model pose. In other words, OneTip provides an interaction model of “what you see is what you get”. However, given the small output range of OneTip, the ability to handle user intentions such as precise navigation, flipping objects, and rapid rotation remains to be explored.

OneTip’s natural interaction revelation allows it to be effectively used in object pose control. We implemented a simple example of controlling a virtual aircraft using OneTip based on the Unity game engine and related free resources<sup>1</sup>. Fig. 12 shows the interactions when OneTip inputs 3-D rotation angles, 3-D translation, and the full 6-D pose variation. The results show that OneTip resembles a joystick device with 6-DOF. Notably, OneTip exhibits certain output coupling during interaction: rotation and translation outputs in the vertical direction may affect each other. For example, when the fingertip translates along the x-axis in the contact interface, avoiding a certain degree of rotational motion around the y-axis is difficult. Since single-fingertip 6-DOF interaction is novel and unfamiliar for new users, it may take several practices before users can become proficient in quantitatively controlling finger movements. Besides, we suggest setting a response threshold to minimize the excess output signal generated by such accompanying motions.

#### B. Combined Input Device

In practice, multi-sensor combinations have proven important for exploration in the design space. We further explore the possible combinations of OneTip with existing interaction devices. As shown in Fig. 13(a), we attempt to combine the OneTip with a typical flying joystick. This

<sup>1</sup> <https://nvjob.github.io/unity/nvjob-dynamic-sky-lite>



Fig. 12. Virtual aircraft flight control using OneTip. (a) Input 3-D orientation variation, and change the rotation angles of the three axes separately. (b) Input 3-D position variation, and change the translation of the three axes separately. (c) Input 6-D pose variation, and changing fingertip pose arbitrarily. Video is available [51].

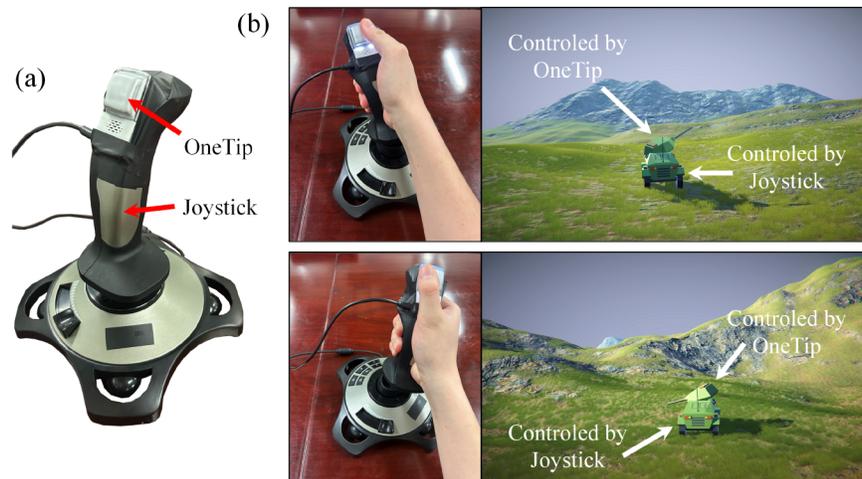


Fig. 13. (a) Combined input device consisting of OneTip and joystick. (b) Virtual tank control using this combined input device, where the joystick controls the tank movement and the OneTip controls the turret rotation. Video is available [51].

combined input device is designed for applications centered on combinatorial object control. While the user controls the rotation of the joystick around three axes through wrist

movements, OneTip provides six more DOFs (or limited to three DOFs) that can be manipulated by the thumb. We built a simple interaction scene in Unity, as shown in Fig. 13(b). In this

scene, the joystick is used to control the traveling and steering of the virtual tank, while OneTip is used to control the 3-DOF rotation of the turret. In addition to controlling the rotation of the virtual object, the thumb press action can also be used as a control message for the turret to fire. Experiments with the device promise good usability for controlling two virtual objects simultaneously with only one hand.

## VI. CONCLUSION

This article describes a novel input device, OneTip, for 6-DOF human-computer interaction with a single fingertip. OneTip is based on the design of non-rigid HCI and Skin-On interfaces to facilitate the discoverability of the interaction actions, and realizes accurate 6-D input based on the visuotactile sensing and the proposed method of fingertip pose measurement. We detail the design, fabrication, and modeling of OneTip, and discuss the experimental evaluation and application exploration. Future work will address the limitations of the OneTip device in terms of the effective interaction space and real-time performance, and expand its application to more human-computer interaction tasks.

Note: This article has been uploaded to TechRxiv as a preprint [51].

## REFERENCES

- [1] B. Bach, R. Sicat, J. Beyer, M. Cordeil, and H. Pfister, "The hologram in my hand: How effective is interactive exploration of 3D visualizations in immersive tangible augmented reality?" *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 457–467, Jan. 2018.
- [2] M. Sundin and M. Fjeld, "Softly elastic 6 DOF input," *Int. J. Hum.-Comput. Interact.*, vol. 25, no. 7, pp. 647–691, Oct. 2009.
- [3] J. F. Lapointe, P. Savard, and N. G. Vinson, "A comparative study of four input devices for desktop virtual walkthroughs," *Comput. Hum. Behav.*, vol. 27, no. 6, pp. 2186–2191, Nov. 2011.
- [4] M. Fiorentino, A. E. Uva, M. D. Fabiano, and G. Monno, "Improving bi-manual 3D input in cad modelling by part rotation optimisation," *Comput.-Aided Des.*, vol. 42, no. 5, pp. 462–470, May 2010.
- [5] X. Wang, L. Besançon, M. Ammi, and T. Isenberg, "Understanding differences between combinations of 2d and 3d input and output devices for 3d data visualization," *Int. J. Hum.-Comput. St.*, p. 102820, Mar. 2022.
- [6] L. Guo, Z. Lu, and L. Yao, "Human-machine interaction sensing technology based on hand gesture recognition: A review," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 4, pp. 300–309, Aug. 2021.
- [7] T. Grosse-Puppenthal *et al.*, "Finding common ground: A survey of capacitive sensing in human-computer interaction," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 3293–3315.
- [8] R. D. Vatavu, "iFAD gestures: Understanding users' gesture input performance with index-finger augmentation devices," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2023, pp. 1–17.
- [9] Y. Li, Q. Meng, T. Yang, Y. Wang, and H. Hou, "Touch gesture and emotion recognition using decomposed spatiotemporal convolutions," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, Feb. 2022.
- [10] L. Chan *et al.*, "FingerPad: Private and subtle interaction using fingertips," in *Proc. 26th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2013, pp. 255–260.
- [11] K. He, Y. Duan, J. Feng, and J. Zhou, "Estimating 3D finger angle via fingerprint image," in *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 1, Article no. 14, Mar. 2022.
- [12] A. Boem and G. M. Troiano, "Non-rigid HCI: A review of deformable interfaces and input," in *Proc. ACM Designing Interactive Syst. Conf.*, Jun. 2019, pp. 885–906.
- [13] J. Alexander *et al.*, "Grand challenges in shape-changing interface research," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2018, pp. 1–14.
- [14] K. Chung and D. Shin, "Effect of elastic touchscreen and input devices with different softness on user task performance and subjective satisfaction," *Int. J. Hum.-Comput. St.*, vol. 83, pp. 12–26, Jun. 2015.
- [15] M. Sturdee and J. Alexander, "Analysis and classification of shape-changing interfaces for design and application-based research," *ACM Comput. Surv.*, vol. 51, no. 1, Article no. 2, Jan. 2018.
- [16] M. Teyssier, G. Bailly, C. Pelachaud, E. Lecolinet, A. Conn, and A. Roudaut, "Skin-on interfaces: A bio-driven approach for artificial skin design to cover interactive devices," in *Proc. 32nd Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2019, pp. 307–322.
- [17] M. Li, L. Zhang, T. Li, and Y. Jiang, "Marker displacement method used in vision-based tactile sensors—from 2D to 3D: A review," *IEEE Sensors J.*, vol. 23, no. 8, pp. 8042–8059, Apr. 2023.
- [18] C. Harrison, J. Schwarz, and S. E. Hudson, "TapSense: Enhancing finger interaction on touch surfaces," in *Proc. 24th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2011, pp. 627–636.
- [19] Y. Watanabe, Y. Makino, K. Sato, and T. Maeno, "Contact force and finger angles estimation for touch panel by detecting transmitted light on fingernail," in *Proc. Haptics: Perception Devices Mobility Commun.*, Jun. 2012, pp. 601–612.
- [20] S. Boring, D. Ledo, X. A. Chen, N. Marquardt, A. Tang, and S. Greenberg, "The fat thumb: Using the thumb's contact size for single-handed mobile interaction," in *Proc. 14th Int. Conf. Human-Comput. Interact. Mob. Dev. Serv.*, Sep. 2012, pp. 39–48.
- [21] R. Xiao, J. Schwarz, and C. Harrison, "Estimating 3D finger angle on commodity touchscreens," in *Proc. Int. Conf. Interact. Tabletops Surfaces*, Nov. 2015, pp. 47–50.
- [22] S. Mayer, H. V. Le, and N. Henze, "Estimating the finger orientation on capacitive touchscreens using convolutional neural networks," in *Proc. ACM Int. Conf. Interact. Surfaces Spaces*, Oct. 2017, pp. 220–229.
- [23] F. Choi, S. Mayer, and C. Harrison, "3D hand pose estimation on conventional capacitive touchscreens," in *Proc. 23rd Int. Conf. Mob. Human-Comput. Interact.*, Oct. 2021, pp. 1–13.
- [24] S. Kratz, P. Chiu, and M. Back, "PointPose: Finger pose estimation for touch input on mobile devices using a depth sensor," in *Proc. ACM Int. Conf. Interact. Tabletops Surfaces*, Oct. 2013, pp. 223–230.
- [25] S. Mayer, M. Mayer, and N. Henze, "Feasibility analysis of detecting the finger orientation with depth cameras," in *Proc. 19th Int. Conf. Human-Comput. Interact. Mob. Dev. Serv.*, Sep. 2017, pp. 1–8.
- [26] J. Yi, J. Liu, C. Zhang, and X. Lu, "Magnetic motion tracking for natural human computer interaction: A review," *IEEE Sensors J.*, vol. 22, no. 23, pp. 22356–22367, Dec. 2022.
- [27] Y. Duan, K. He, J. Feng, J. Lu, and J. Zhou, "Estimating 3D finger pose via 2D-3D fingerprint matching," in *Proc. 27th Int. Conf. Intell. User Interfaces*, Mar. 2022, pp. 459–469.
- [28] V. Ngyuen, P. Kumar, S. H. Yoon, A. Verma, and K. Ramani, "SOFTii: Soft tangible interface for continuous control of virtual objects with pressure-based input," in *Proc. 9th Int. Conf. Tangible Embedded Embodied Interact.*, Jan. 2015, pp. 539–544.
- [29] S. Nakamaru, R. Nakayama, R. Niiyama, and Y. Kakehi, "FoamSense: Design of three dimensional soft sensors with porous materials," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2017, pp. 437–447.
- [30] M. Weigel and J. Steimle, "DeformWear: Deformation input on tiny wearable devices," in *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 2, Article no. 28, Jun. 2017.
- [31] K. Park, H. Yuk, M. Yang, J. Cho, H. Lee, and J. Kim, "A biomimetic elastomeric robot skin using electrical impedance and acoustic tomography for tactile sensing," *Sci. Robot.*, vol. 5, no. 39, Jun. 2022, Art. no. eabm7187.
- [32] Y. Watanabe, A. Cassinelli, T. Komuro, and M. Ishikawa, "The deformable workspace: A membrane between real and virtual space," in *Proc. IEEE Int. Workshop Horizontal Interact. Hum.-Comput. Syst.*, Oct. 2008, pp. 145–152.
- [33] S. Follmer, M. Johnson, E. Adelson, and H. Ishii, "deForm: An interactive malleable surface for capturing 2.5D arbitrary objects, tools and touch," in *Proc. 24th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2011, pp. 527–536.
- [34] S. Follmer, D. Leithinger, A. Olwal, N. Cheng, and H. Ishii, "Jamming user interfaces: Programmable particle stiffness and sensing for malleable and shape-changing devices," in *Proc. 25th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2012, pp. 519–528.
- [35] S. Zhang *et al.*, "Hardware technology of vision-based tactile sensor: A review," *IEEE Sensors J.*, vol. 22, no. 22, pp. 21410–21427, Nov. 2022.
- [36] W. Yuan, S. Dong, and E. H. Adelson, "GelSight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no.

- 12, p. 2762, Nov. 2017.
- [37] N. F. Lepora, "Soft biomimetic optical tactile sensing with the TacTip: A review," *IEEE Sensors J.*, vol. 21, no. 19, pp. 21131–21143, Oct. 2021.
- [38] S. Cui, R. Wang, J. Hu, J. Wei, S. Wang, and Z. Lou, "In-hand object localization using a novel high-resolution visuotactile sensor," *IEEE Trans. Ind. Electron.*, vol. 69, no. 6, pp. 6015–6025, Jun. 2022.
- [39] S. Athar, G. Patel, Z. Xu, Q. Qiu, and Y. She, "VisTac toward a unified multimodal sensing finger for robotic manipulation," *IEEE Sensors J.*, vol. 23, no. 20, pp. 25440–25450, Oct. 2023.
- [40] R. Sui, L. Zhang, Q. Huang, T. Li, and Y. Jiang, "A novel incipient slip degree evaluation method and its application in adaptive control of grasping force," *IEEE Trans. Autom. Sci. Eng.*, doi: 10.1109/TASE.2023.3241325.
- [41] H. J. Huang, X. Guo, and W. Yuan, "Understanding dynamic tactile sensing for liquid property estimation," in *Proc. Robot. Sci. Syst.*, Jun. 2017, pp. 1–11.
- [42] K. Kamiyama, K. Vlack, T. Mizota, H. Kajimoto, K. Kawakami, and S. Tachi, "Vision-based sensor for real-time measuring of surface traction fields," *IEEE Comput. Graph. Appl.*, vol. 25, no. 1, pp. 68–75, Jan. 2005.
- [43] A. Padmanabha, F. Ebert, S. Tian, R. Calandra, C. Finn, and S. Levine, "OmniTact: A multi-directional high-resolution touch sensor," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2020, pp. 618–624.
- [44] N. F. Lepora et al., "Towards integrated tactile sensorimotor control in anthropomorphic soft robotic hands," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2021, pp. 1622–1628.
- [45] M. Li, L. Zhang, T. Li, and Y. Jiang, "Continuous marker patterns for representing contact information in vision-based tactile sensor: Principle, algorithm, and verification," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, Aug. 2022.
- [46] M. Li, Y. H. Zhou, T. Li, and Y. Jiang, "Real-time and robust feature detection of continuous marker pattern for dense 3-d deformation measurement," *Measurement*, vol. 221, p. 113479, Nov. 2023.
- [47] C. Edwards and R. Marks, "Evaluation of biomechanical properties of human skin," *Clinics Dermatol.*, vol. 13, no. 4, pp. 375–380, Jul./Aug. 1995.
- [48] L. Zhang, T. Li, and Y. Jiang, "Improving the force reconstruction performance of vision-based tactile sensors by optimizing the elastic body," *IEEE Robot. Autom. Lett.*, vol. 8, no. 2, pp. 1109–1116, 2023.
- [49] D. Ma, S. Dong, and A. Rodriguez, "Extrinsic contact sensing with relative-motion tracking from distributed tactile measurements," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2021, pp. 11262–11268.
- [50] R. Sui, L. Zhang, T. Li, and Y. Jiang, "Incipient slip detection method for soft objects with vision-based tactile sensor," *Measurement*, vol. 203, p. 111906, 2022.
- [51] M. Li, Y. H. Zhou, L. Zhang, T. Li, and Y. Jiang, "OneTip is enough: A non-rigid input device for single-fingertip human-computer interaction with 6-DOF," Feb. 2024. [Online]. Available: <https://doi.org/10.36227/techrxiv.170775314.44150885/v1>.