MaGEL: A Soft, Transparent Input Device Enabling Deformation Gesture Recognition

Fumika Oguri
Aoyama Gakuin University
Sagamihara, Japan
oguri.fumika@x-lab.team

Katsutoshi Masai
Kyushu University
Fukuoka, Fukuoka, Japan
masai@ait.kyushu-u.ac.jp

Yuta Sugiura
Keio University
Yokohama, Japan
sugiura@keio.jp

Yuichi Itoh
College of Science and Engineering
Aoyama Gakuin University
Sagamihara, Japan
itoh@it.aoyama.ac.jp

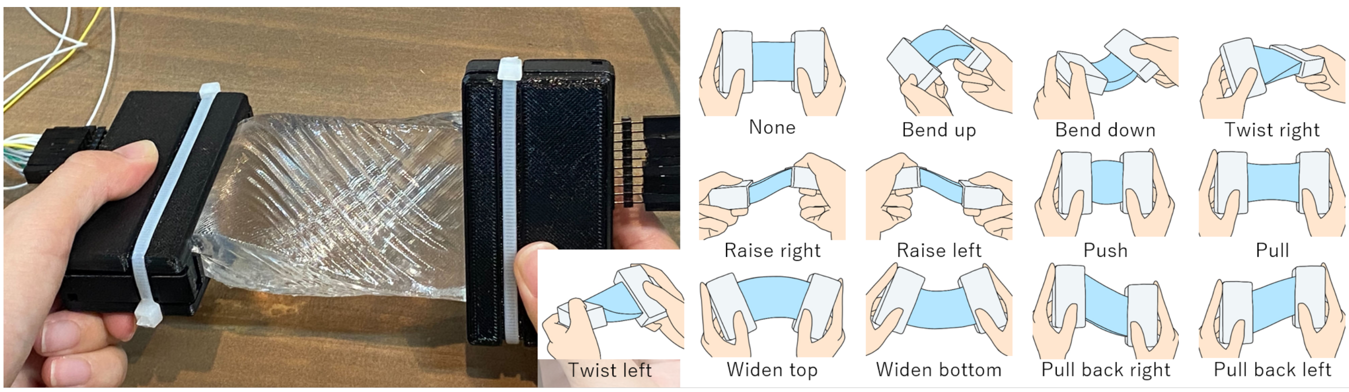

Figure 1: Implemented MaGEL (left) and Gestures (right). MaGEL identifies the deformation gestures and can be used for input device.

Abstract

We propose MaGEL, a soft-input device that utilizes light intensity to detect and interpret user deformation interactions. Unlike traditional rigid input devices, MaGEL enables three-dimensional interactions such as twisting, bending, and pulling. Additionally, MaGEL incorporates elastic haptic feedback, providing users with tactile sensations that reflect the tension or resistance of their interactions. These factors realize intuitive and natural user interaction experiences, and users can employ familiar physical gestures as input. For example, bending the device may simulate turning the page of a book, or stretching it may zoom in on an image. The device consists of a transparent urethane resin gel with LED lights and phototransistors on both sides. When the device gel deforms, the intensity of the light passing through the gel undergoes a specific change due to the deformation. The system analyzes these changes using machine learning to identify the user gestures. We evaluated the optimal configuration and number of LEDs and phototransistors to classify the deformation accurately. We acquired data for 13 types of deformation gestures from 14 participants. The results showed that a combination of four LEDs and ten phototransistors enabled MaGEL to identify 13 types of deformation gestures with an accuracy of 94.1 %. Using MaGEL, we provide novel interactive experiences, such as game controllers that employ bending, pulling, or twisting to mimic natural gaming motions.

CCS Concepts

• **Human-centered computing → Haptic devices**; *Gestural input*.

Keywords

User Interface, Input Device, Tangible, Double-Handed Interaction, flexible controller

IUI '25, Cagliari, Italy

ACM ISBN 979-8-4007-1306-4/25/03
https://doi.org/10.1145/3708359.3712100

ACM Reference Format:
Fumika Oguri, Katsutoshi Masai, Yuta Sugiura, and Yuichi Itoh. 2025. MaGEL: A Soft, Transparent Input Device Enabling Deformation Gesture Recognition. In *30th International Conference on Intelligent User Interfaces (IUI '25), March 24–27, 2025, Cagliari, Italy*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3708359.3712100

1 Introduction

To enable more intuitive interactions, input interfaces have progressed beyond traditional methods, such as mouse, keyboard, and touch panel. Gestural input, an example of three-dimensional interaction, has become particularly relevant in virtual reality (VR) applications. Depth-sensing technologies, such as Kinect and Leap Motion, and gyroscopic sensors embedded in devices are commonly used to capture hand and body movements. While gesture-based input in a virtual space offers greater freedom of movement, the lack of physical feedback poses challenges for precise control and fine adjustments. Additionally, prolonged use of such systems may cause user fatigue. Thus, incorporating physical objects offers physical haptic feedback, potentially enhancing the sense of control and the overall interaction experience. In physical feedback systems, users can perceive the extent of their movements through physical sensations or enable the accurate determination of object contact, facilitating precise control through minimal movement. Moreover, unlike gesture-based input methods, haptic interfaces allow quantification of the applied force for an input action.

In this context, tangible user interfaces (TUIs), which leverage physical objects for user interactions, have attracted substantial attention. TUIs offer the advantage of allowing users to manipulate digital information intuitively, thereby providing both visual and tactile feedback concurrently. However, most TUIs are constructed from rigid materials such as plastic and metal, limiting their ability to support deformation interactions, such as pulling, bending, and twisting. The development of input devices from soft materials can enable three-dimensional interactions. This capability allows for more intuitive interactions, such as bending a device like a book to turn pages or stretching it to zoom into an image, potentially leading to enhanced user engagement and usability.

In this study, we propose a novel input device, MaGEL(Fig. 1), that utilizes soft materials for deformation-based input with elastic haptic feedback, providing users with tactile sensations that reflect the tension or resistance of their interactions. By utilizing soft materials, interaction with more natural and physical sensations is possible, and users can operate the device intuitively through physical sensations. This has the potential to create new experiences in fields, such as gaming and music production.

MaGEL consists of a soft gel, LEDs, and phototransistors. The gel was made of transparent urethane resin and LEDs and phototransistors were placed on both sides of the gel. When the device is deformed, it alters the intensity of light passing through the gel. By applying machine-learning techniques to analyze the changes in phototransistor values, MaGEL can detect the deformation of the device and estimate user gestures.

The contributions of this study are as follows:

(1) We developed a soft-input controller, MaGEL, which identifies the three-dimensional gestures using LEDs, phototransistors, and machine learning.
(2) We conducted an experiment to determine the optimal number and arrangement of LEDs and phototransistors for gesture identification in MaGEL and the blinking pattern of the LEDs.
(3) The results of the evaluation experiments showed that the system identified 13 gestures, with 94.1 % accuracy for within-individual identification and an 85.1 % accuracy rate for between-individual identification.

2 Related work

2.1 Tangible User Interface

Extensive research has been conducted on tangible objects that directly manipulate physical objects to enhance intuitive operation[7]. Weiss et al. proposed an FTIR device that utilizes a rear camera to capture the diffuse reflection of light projected onto the back of a transparent panel at the interface between the panel and device[26]. The displayed content changes based on the widget's position and interactions. Hwang et al. proposed MagGetz, which represents an approach that leverages a smartphone's built-in magnetic sensor to receive input from physical controls such as buttons and sliders[6]. This method enables tangible input without requiring an external power source, as it relies solely on smartphone sensors and magnets. SurfaceIO, as proposed by Ding et al., is a methodology designed to enhance tactile feedback for touch-based interactions by incorporating subtle surface irregularities into objects[2]. This approach facilitates touch and slide interactions through vibration, and despite the presence of fine surface textures, enables tactile operation without visual dependency. Thus, direct manipulation of interface elements enhances the user experience by providing immediate feedback on operations and facilitating intuitive interactions. Users can readily associate operational methods with the visual cues presented by knobs, buttons, and other interface components.However, these devices are all constructed from rigid materials such as plastic or metal, and they lack the flexibility characteristic.

2.2 Soft Input Devices

Input devices made of soft materials such as silicone, cloth, or skin rather than rigid materials such as plastic and metal are being used to extend input expression by incorporating natural, intuitive gestures and soft haptic feedback[1]. Those soft input devices are being evaluated for their potential implementation in various applications, including controllers for home appliances, haptic touch interfaces on LCD displays, music production, gaming, and numerous other fields.

Harrison et al. proposed a method for creating three-dimensional buttons by deforming a soft membrane using air pressure[4]. MimicTile, designed for mobile devices, employs shape memory alloys (SMAs) to impart smooth, lifelike movements to metal components, and enable deformation-based input[15]. Gummi is a flexible display technology that enables deformation-based inputs while simultaneously presenting visual information[20]. This functionality facilitates features, such as screen magnification, through bending gestures. Mazursky et al. proposed MagnetIO, a soft patch with magnets embedded in soft silicone that can be attached to curved surfaces[14]. It provides force feedback through the attachment of a coil to the user's finger. He et al. proposed a method to recognize squeezing interaction by attaching a tube to an air puff and acquiring air vibrations with an acoustic sensor[5]. Kildal et al. developed a deformable controller and established guidelines for its future potential and design considerations[12]. In addition, they conducted evaluations of gesture identification systems and their controller applications[11]. Shorey et al. demonstrated that incorporating deformable inputs into action control schemes enhanced the overall user experience[21]. Weinberg et al. proposed a methodology for music creation that utilizes a soft, spherical device that responds to squeezing and pulling actions[25]. Furthermore, they engineered a flexible controller incorporating a rubber bridge, enabling four distinct types of deformation inputs, such as bending and twisting, in addition to the conventional button operations. While these game controllers recognize four types of deformations, including bending and twisting, our proposed MaGEL system is capable of recognizing a diverse range of deformation gestures, enabling its use in a wider range of applications. Thus, soft-input devices demonstrate potential for application across diverse domains. This study proposes a methodology for detecting deformations in soft materials and utilizing such deformations as input mechanisms.

2.3 Methodologies for Capturing Soft Object Deformation

Several methodologies have been proposed for obtaining the deformation of soft objects using camera-based systems or position sensors to ascertain the deformation position of an elastic body. GelForce proposed by Kamiyama et al[10]., Forcetile proposed by Kakehi et al[9]., and DeforMe proposed by Punpongsanon et al[17]. are all which employ a technique involving the embedding of markers within an elastic body and subsequently tracking their positions via camera. Sato et al. proposed an approach that utilizes a polarized camera to capture the position and displacement of elastic body deformation on an LCD display, leveraging the photoelastic properties of the material[19]. Reed introduced a method that involves embedding multiple wireless locators within clay to estimate its shape based on their positions[18]. Luo et al. proposed a method of weaving conductive threads into fabric to extend it as a flexible input surface[13] Although these methods effectively capture the deformation of soft materials and can serve as input devices, they are limited by occlusion issues inherent to camera-based recognition systems. Furthermore, their functionality is constrained by the range of a stationary camera or receiver.

Harrison et al. have proposed a method of acquiring pushing and shearing forces utilizing a 3D stick[3]. Similarly, Tsuchida et al. have proposed a method of acquiring pushing force strength and direction using magnets and springs[24]. While these studies acquire the strength of force intensity alongside two-dimensional input, but they are merely an extension of input onto a plane and cannot be considered three-dimensional input.

Slyper et al. introduced a technique for detecting deformation in soft materials by measuring the contact between electrodes embedded in silicone resin[22]. This method allows for the detection of deformation through the contact or separation of electrodes within slits as the silicone resin is bent or stretched. Kadowaki et al. developed an embedded flexible tactile sensor utilizing infrared LEDs[8]. This sensor detects interactions by measuring the changes in the amount of reflected infrared light when an object is deformed. The aforementioned investigations facilitated tactile input through the integration of soft materials with electrodes and structural frameworks. In contrast, the device proposed in this study incorporates distinct gel and sensor components, enabling the facile replacement of the gel to accommodate various input modalities.

A methodology employing photoreflectors was proposed to capture object movement and density variations. Infrared LEDs and phototransistors are cost-effective compared with bending sensors, high-performance IMUs with minimal drift, and camera-based image recognition. Sugiura et al. developed a system that estimates gestures by measuring cotton cushion density using a photoreflector[23]. Ogata et al. introduced SenSkin, an infrared sensor band worn around the wrist to detect skin movement on the arm, thereby extending the touch panel functionality by utilizing the skin as a soft input device[16]. While the cotton-based approach is suitable for furniture, such as sofas and stuffed toys, there are concerns regarding potential shifts in cotton positioning over time, which may affect sensor placement. Skin-based methods are limited to body surfaces. Consequently, this study proposes a novel approach utilizing a soft transparent gel that leverages the reflective and refractive properties of infrared light within the gel to employ device deformation as the input. By utilizing a moldable gel, it is feasible to create deformable input devices of various shapes tailored to specific applications, extending beyond the rectangular form proposed in this study.

3 MaGEL:Soft-Input Device

We propose MaGEL, a soft-input device designed to recognize gestures, such as twisting, bending, and pulling. By utilizing the elastic properties of soft materials, MaGEL provides an intuitive and immersive user experience. In this section, we first explain the deformation identification principle, followed by an explanation of the device configuration and each component.

3.1 Sensing Principle

The infrared light emitted by the infrared light-emitting diode (LED) on one side of the device propagates through the transparent gel, reflects off its surface, and the phototransistors placed on opposite sides detect it using the phototransistor (Fig. 2). When a user interacts with the device, the deformation of the gel modulates the infrared light transmission, resulting in corresponding voltage variations in the phototransistor. The MaGEL system uses machine learning to interpret these voltage changes, allowing it to accurately estimate different types of deformations and user gestures.

Fig. 3 provides a visual representation of light transmission through the transparent gel captured using an infrared camera. This figure shows that the appearance of light undergoes observable changes as the device experiences deformation. The camera was positioned on one side of the device, with infrared LEDs situated on the opposite side. By fixing relative positions of the gel, LED, and phototransistor, it is feasible to detect alterations in LED light using the phototransistor and estimate the gel's deformation.

3.2 Hardware

The device comprises a transparent gel, two grips for securing sensors to the gel, infrared LEDs, and infrared-sensing phototransistors for detecting gel deformation. Fig. 4 illustrates the implemented device.

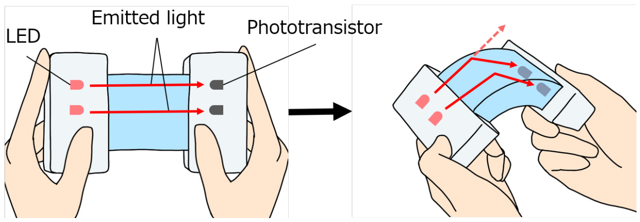

Figure 2: Principle: phototransistor detects the change of light intensity.

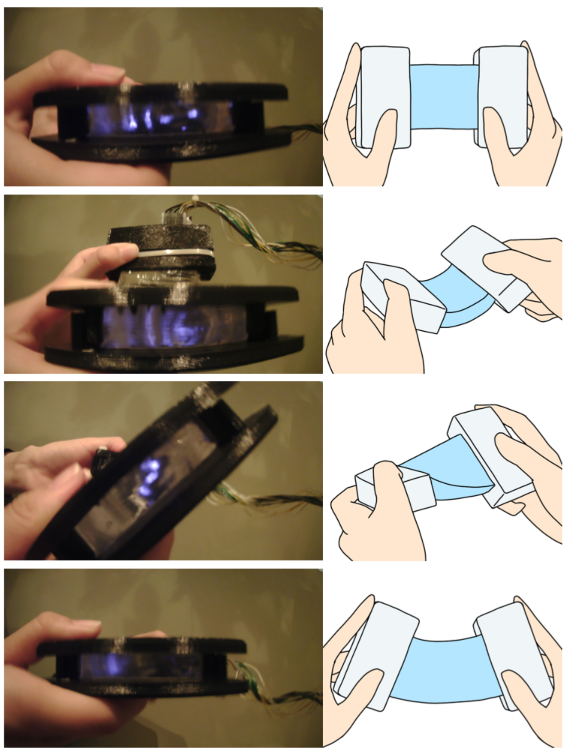

Figure 3: Appearance of infrared LED light when the device is deformed.

The device's body was constructed from transparent soft urethane resin with an Asker hardness of 0, allowing for easy manual deformation to input various gestures. Gesture-based input, characterized by hand or body movements in space, presents limitations owing to the absence of tactile feedback. This deficiency results in imprecise control, and may lead to user fatigue caused by extensive motion. The incorporation of an elastic gel provides tactile feedback to the gesture, thereby enhancing the user's ability to perceive the applied force and improve overall control. The gel is formed by combining two liquids, the main agent and hardener, and in this study, we molded it into a $50 \times 100 \times 20$ mm rectangular shape, and reinforced its surface using a transparent coating agent. Although a rectangular gel was used in this study, the principle can be applied to devices other than controllers using molds with different shapes.

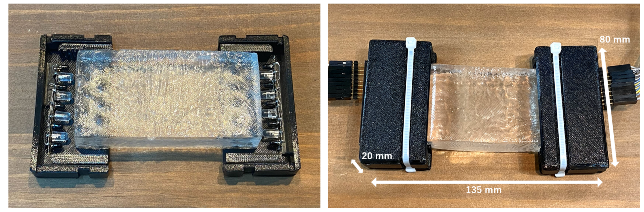

Figure 4: Implemented device.

Infrared LEDs and infrared-receiving phototransistors with a peak wavelength of 940 nm were used. This infrared spectrum choice was intended to minimize interference from ambient indoor light when measuring light transmission through the gel, which could affect measurements using visible light components. We conducted a investigation to determine the optimal positions and number of LEDs and PTr on the device in evaluation section.

To ensure the relative positioning of the gel, LED, and phototransistor, we fabricated grips using a 3D printer with polyethylene terephthalate (PET) and fixed them on both sides of the device. Each grip comprised three parts: upper, lower, and side. The upper and lower parts featured protrusions designed to secure a transparent gel, and the LED and phototransistor were mounted on the side parts. These three parts were fastened with cable ties to maintain fixed positions of the transparent gel, LEDs, and phototransistors.

The device dimensions were $135 \times 80 \times 20$ mm. This measurement was derived from the dimensions of the flat surface area of the Nintendo Switch Pro Controller, excluding the handles, which are devices designed for bimanual operation and grip.

A microcomputer (Arduino MEGA 2560) controlled the infrared LED and the phototransistor. The phototransistor's voltage is acquired by the Arduino as a 1024-step analog value and transmitted via serial communication to a personal computer (Windows 11 Pro OS, Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz 2.59 GHz). The PC processes the received value to estimate the deformation of the device using the following software requirements.

3.3 Software

The LEDs incorporated in the device were individually controlled using the aforementioned microcomputer. Serial communication between the microcomputer and personal computer was established at a baud rate of 115,200 bps. The gesture recognition algorithm was implemented using Python version 3.12.0.

The proposed system is designed to classify multiple deformations. The voltage magnitude applied to the phototransistors was used as a feature, and for each feature dimension of the training data, the measured data were standardized to achieve a mean of 0 and a variance of 1 Machine learning techniques are employed to classify the gestures based on the deformation of a device, which corresponds to phototransistor values. We conducted a investigation to evaluate and compare various machine learning classifiers with the objective of selecting the classifier that demonstrates the highest recognition accuracy.

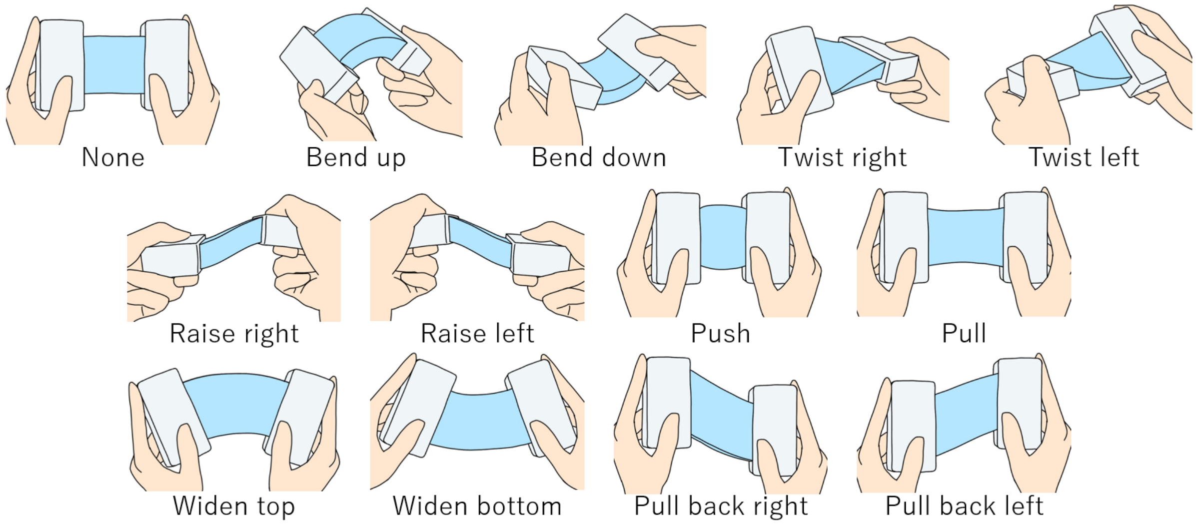

Figure 5: Gestures to identify.

4 Evaluation experiment

Evaluation experiments were conducted to confirm the performance of the proposed system in detecting soft gel deformation. Gesture data were collected from 14 participants for analysis. The initial phase of the study focused on identifying the most suitable machine learning classifier and the positions of the components to optimize accuracy. Following this, intra- and inter-person identification tests were performed under these conditions to assess the accuracy of MaGEL gesture recognition. Lastly, the response speed of gesture recognition was also evaluated to confirm the real-time performance of the MaGEL.

4.1 Gesture Data Collection from Multiple Users

4.1.1 Overview. In this experiment, we aimed to identify 13 distinct types of deformation gestures using the MaGEL device. Fig. 5 shows each of the deformation gestures. The selected gestures encompass rotation and translation along the three axes in three-dimensional space. We chose six rotational gestures, including reverse rotations, and six translational gestures, including reverse directions, in addition to a neutral (no deformation) position, resulting in 13 gestures. These 13 gestures were selected to leverage the soft material's ability to bend, twist, and stretch, offering a comprehensive range of deformations that allow natural and intuitive inputs.

Each side of the grips in the experimental device incorporated four LEDs and five phototransistors arranged in a linear configuration. In total, the device utilized eight LEDs and ten phototransistors. This configuration is the maximum number of LEDs and phototransistors that can be accommodated without electrode contact, considering the vertical width of the transparent gel. The LEDs and phototransistors were positioned in an alternating pattern, with each component assigned sequential numbers. Fig. 6 shows the layout of these sensors.

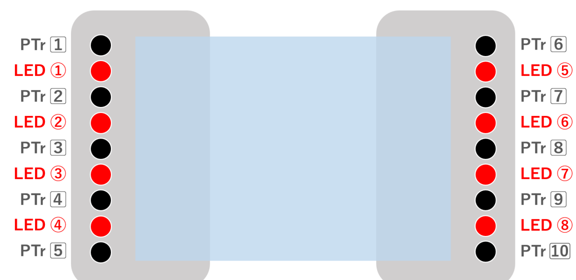

Figure 6: The Layout of LEDs and phototransistors(PTrs) for Experiment.

The voltage values from the ten phototransistors were acquired for each of the 255 lighting patterns, encompassing all possible combinations of the eight LEDs being illuminated or extinguished, except for configurations where none of the LEDs were activated. Microcomputers regulate the activation and deactivation of LEDs.

4.1.2 Procedure. The participants were instructed to deform the device into 13 distinct gestures, during which the phototransistor values were recorded. After participants signed the consent form, they were presented with images of 13 gestures and given time to practice gestures. When the experiment started, the experimenter displayed one of the 13 gesture images, illustrated in Fig. 5, on a monitor in a randomized sequence. The participants were required to deform the device in accordance with the displayed image and maintain the deformation for 6 seconds, during which time the voltage values from the phototransistor were acquired. This process was repeated for all 13 gesture types. The experiment consisted of ten sets, with each set comprising 13 gestures. To mitigate order effects, the presentation sequence of the images was randomized for each set using a computer-generated random number algorithm to ensure that each gesture type appeared once per set. This experiment passed the university's ethical review(H22-008).

Fourteen participants(10 males, 4 females; mean age 22.5 ± 1.3 years) participated in the investigation. The maximum duration of the experiment for one person was 40 minutes. In total, 1,820 data points (14 participants × 13 gestures × 10 sets) were collected. Each data point has 2,550 features (10 PTrs and 255 LED patterns).

Table 1: Gesture Recognition Accuracy of Four Classifiers

Classifier	Accuracy	Tuning Parameters
SVM	0.932	C, gamma, kernel
Random Forest	0.896	n estimators, max depth, min samples split
Neural Network	0.834	hidden layer sizes, activation, solver, alpha, learning rate
XG Boost	0.832	max depth, min child weight, eta, reg alpha, reg lambda

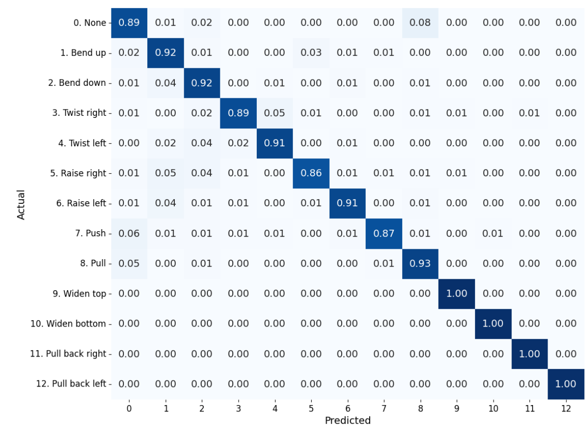

Figure 7: Confusion matrix of intra-personal gesture classification with 255 LED lighting patterns × 10 PTr.

4.2 Selection of Classifiers for Machine Learning

To determine the most suitable classifier for deformation gesture recognition, we conducted a comparative analysis of gesture classification accuracy using four classifiers: a support vector machine (SVM), random forest, neural network, and extreme gradient boosting (XGBoost). The classifiers were implemented using Python libraries scikit-learn and xgboost. Utilizing the data collected in Section 4.1, we employed leave-one-out cross-validation for each participant, designating one of the 10 sets as test data. The evaluation was based on the average intrapersonal recognition accuracy of 14 participants. Prior to implementation, the classifiers were optimized using a grid search. Table 1. lists the tuned parameters and the recognition accuracies of the 13 gestures for each classifier.

For this study, we selected SVM with the following parameters: $C = 0.1, gamma = 1, andkernel =`linear.'$

4.3 Arrangement of LEDs and Phototransistors used in the Device

A reduction in the number of sensors employed in a device offers significant advantages in terms of processing efficiency and cost-effectiveness. However, an excessively limited number of sensors may compromise the classification accuracy of deformation gestures. Consequently, this investigation utilized the data collected in Section 4.1 to determine the optimal number and position of LEDs and phototransistors (PTr) to be incorporated into the device as well as the appropriate LED lighting patterns.

Considering the accuracy using phototransistor values for all LED illumination patterns as the benchmark, we investigated the optimal number and configuration of features that would yield comparable performance with fewer sensors. The Support Vector Machine (SVM) model outlined in Section 4.1 was employed as the classification algorithm. To assess classification accuracy, the mean intra-personal identification accuracy across the 14 participants was utilized. When incorporating all 2550 features comprising 255 lighting patterns with 10 PTrs as feature inputs, the classification accuracy for 13 gestures reached 93.2 %.Fig. 7 shows the confusion matrix.

From the 10 PTrs, 1023 distinct combinations of usage patterns were considered as potential feature sets, excluding instances where none were utilized. The classification accuracy for 13 gestures was evaluated using 260,865 LED-PTr patterns derived from a combination of 255 LED lighting patterns and 1023 PTr usage patterns. The analysis revealed a trend towards higher accuracy when fewer LEDs were simultaneously illuminated and a greater number of PTrs were employed. The LED and PTr combination pattern, yielding the highest accuracy, consisted of LED1 illuminated in conjunction with phototransistors 1-3 and 5-10. This configuration resulted in a 13-gesture classification accuracy of 72.4 %.

To enhance the classification accuracy, we expanded our approach beyond a simultaneous LED lighting pattern to incorporate a sequential blinking pattern. This pattern involved activating an additional LED after the deactivation of the previous one and recording of the phototransistor value. Utilizing all ten PTrs, we employed a sequential forward method to identify the optimal combination of 255 LED lighting patterns that would yield the highest classification accuracy. The search process continued until an accuracy of 93.2 % was achieved, which corresponds to the accuracy obtained when utilizing all 2550 features.Ultimately, a classification accuracy of 94.1 % was attained through sequential activation of LED1, 4, 5, and 8 in isolation.

Based on these results, we selected LED1, LED4, LED5, and LED8, positioned at the four corners of the device, in conjunction with all phototransistors (PTr1-10) for further analysis considering the trade-off between measurement time and accuracy. We blinked four LEDs sequentially and recorded the values of the ten phototransistors for each LED. Consequently, a total of 40 features (4 LEDs × 10 PTrs) were employed for the machine learning analysis.

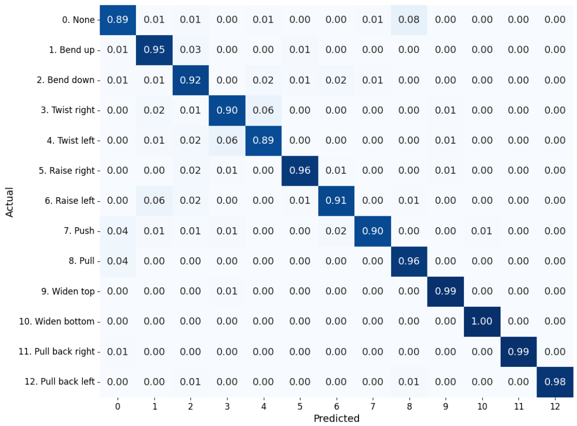

Figure 8: Confusion matrix of intra-personal gesture classification.

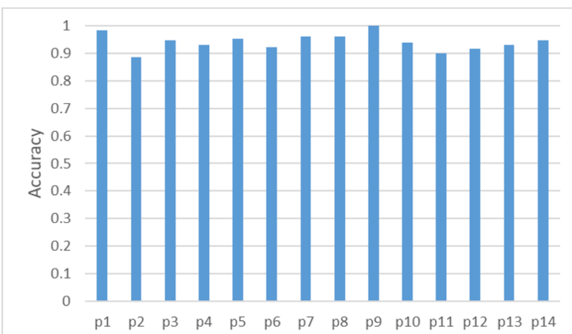

Figure 9: Classification accuracy for each participant.

4.4 Gesture Classification by the Proposed System

In this section, we use the device configuration selected in the previous section to perform intra-person gesture classification using data from the same participant, and inter-person gesture classification using data from other participants. The device used in this experiment was equipped with four LEDs and 10 phototransistors, as described in Section 4.3. Therefore, from the data acquired in Section 4.1, the values of the 10 phototransistors that measure the intensity of light emitted from each of the four LEDs were utilized, resulting in 4 LEDs × 10 PTrs = 40 features for classification. The SVM determined in Section 4.2 is used as the machine learning classifier, and the features were standardized to have a mean of 0 and a variance of 1.

4.4.1 Intra-personal Classification. To evaluate the gesture classification accuracy for individual users, the intra-person classification accuracy was assessed. For each of the 14 participants (p1-p14), leave-one-out cross-validation was conducted using one of the 10 datasets as the test data.The average intra-person classification accuracy across all the participants was 94.1 ± 3.0 %. Fig. 8 shows the confusion matrix and illustrates that while many gestures were accurately classified, some errors occurred. Specifically, "Pull" was

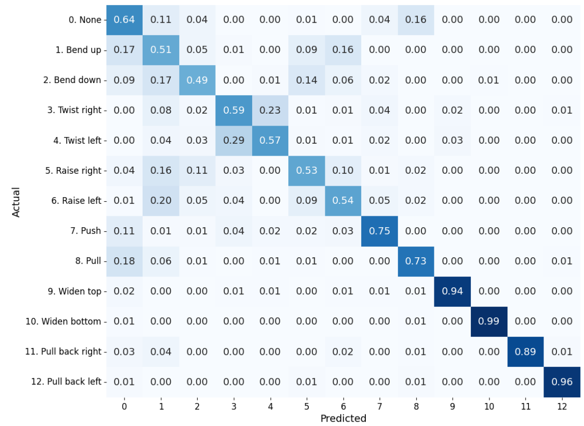

Figure 10: Confusion matrix of inter-people gesture classification.

sometimes confused with "None," "Raise left" with "Bend up," and "Twist left" and "Twist right" were sometimes confused. Fig. 9 shows the inter-person classification accuracy of each participant. Accuracy varied across participants, with p9 achieving the highest gesture identification accuracy at 100.0 %, and p2 the lowest at 88.5 %.

4.4.2 Inter-person Classification. Considering a case in which multiple users share a single device, we examined the classification accuracy across different users. For each participant, data from 13 other people were used as training data to evaluate intra-person classification accuracy.

The mean inter-person classification accuracy across all participants was 72.6 ± 1.5 %. Fig. 10 shows this confusion matrix. In addition to the previously observed confusion among None, Pull, Push, and Twist in intrapersonal classification, confusion between Raise and Bend was noted in the inter-person context.

A plausible explanation for the lower accuracy of inter-person classification compared with intra-person classification may be attributed to the variation in experimental timing across participants. The experiment was performed indoors with the blinds closed and under the same lighting, but the differing environmental light conditions between participants who conducted the experiment during the daytime and those who performed it at night may have affected the classification accuracy. Therefore, to mitigate the influence of environmental light, we implemented a centering approach. The values obtained when each LED was turned on were centered on those obtained when no LEDs were activated.

This adjustment resulted in an increase in the average inter-individual discrimination accuracy of the 14 participants to 85.1 %. Fig. 11 shows the confusion matrix after the feature values were centered. Fig. 12 shows the inter-individual discrimination accuracy for each participant after centering. The centering process reduced the confusion between None and Pull/Push, as well as between Bend and Raise, thereby improving the overall accuracy. However, confusion between bend-up and bend-down, twist-right and twist-left, and raise-right and raise-left remained. We performed the same

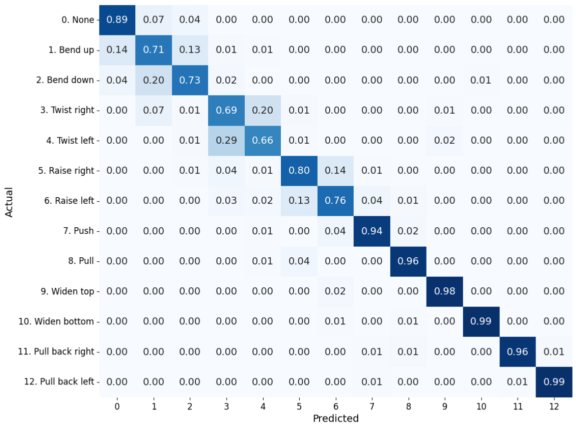

Figure 11: Confusion matrix of inter-people gesture classification.

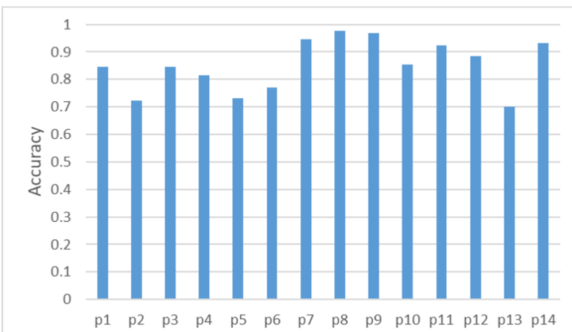

Figure 12: Classification accuracy for each participant.

centering process for intra-person classification, but the accuracy did not improve.

4.4.3 Accuracy and Number of Training Sets. To evaluate the trade-off between accuracy and training data size, we assessed the performance of the model using varying amounts of training data. Fig. 13 shows the relationship between the number of sets used as training data and the classification accuracy for the 13 gestures. The vertical axis represents the mean classification accuracy across 14 participants. The sets used for training and testing were randomly selected for each participant, and one set was selected for the test set regardless of the number of training sets. The results demonstrate a positive correlation between the number of training sets and classification accuracy. The graph shows that the average accuracy exceeded 90 % after training for more than six sets.

4.5 Processing Speed of Real-time Classification

To investigate the response time of the real-time classification system, a classification model was created in advance, and the time required for classification was measured. The duration for reading data via serial communication, standardizing the data, and classifying gestures were assessed for 100 frames, and the average time was calculated. The experimental setup was consistent with

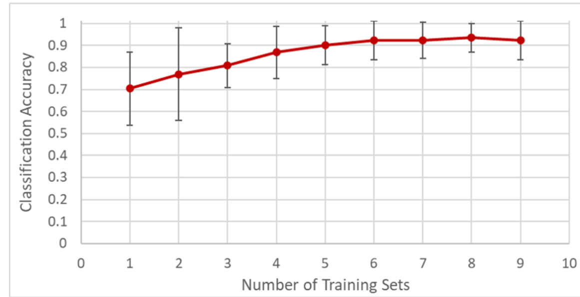

Figure 13: Accuracy and Number of Training Sets.

the hardware requirements outlined in Section 3.2. For intrapersonal classification, an SVM model, selected in Section 4.2, was constructed from 10 sets of data collected using four LEDs and 10 phototransistors and employed for gesture classification. The processing time for real-time recognition averaged 29.5 ms/frame over 100 frames. This corresponds to a frame rate of 33.9 Hz, suggesting that the proposed system can identify gestures at a processing speed suitable for real-time applications.

5 Discussion and Future Work

5.1 Discussion of experimental results

The evaluation experiments confirmed that MaGEL demonstrated high accuracy in identifying 13 types of deformation gestures for both intra- and inter-person classification. Specifically, the system achieved an accuracy of 94.1 % for intra-person classification and 85.1 % for inter-person classification. This performance was attained using four LEDs positioned at the four corners of the device and ten phototransistors.

Within the classified gestures, confusion occurred in some gestures. These included "Pull" and "None," "Twist right" and "Twist left," and "Raise left" and "Bend up" in the intra-person classification. Pull and None were likely to be misclassified when the participants applied a weak force during deformation. This was because the gel deformation was minimal, making it difficult to distinguish between the two. "Raise left" was confused with "Bend up" when the opposite side of the grip was turned inward when lifting the grip, resulting in a shape similar to "Bend up." In inter-person classification, "Twist right" and "Twist left," "Raise left" and "Raise right," and "Bend up" and "Bend down" were confused sometime. This may be because it was difficult for the light from the LED to reach the phototransistor on the opposite side when twisted to the left or right, making it difficult to distinguish between them. Based on the results of inter-individual identification, it is possible to identify deformable gestures by training on other users' data, but it is preferable to train on the same user's data because each user has his or her own movement habits.

In the evaluation of the processing speed for real-time gesture identification, 33.9 Hz, an average response time from data acquisition to gesture identification, was achieved using the machine learning model. This indicates that the MaGEL has sufficient real-time responsiveness and can be used as an input device. These

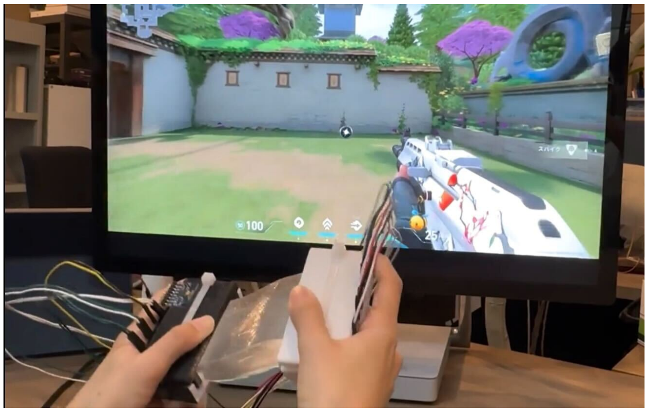

Figure 14: Playing an FPS game using MaGEL as an input device.

results indicate that MaGEL is an input device that can identify soft gel deformations in real-time with high accuracy.

5.2 Application Example

An example application of MaGEL is a game controller that enables deformation input. This application is designed for first-person shooter (FPS) games and allows players to input commands by deforming the soft controller, thereby providing a gaming experience that is not possible with a conventional rigid controller. Fig. 14 shows an image of playing an FPS game using MaGEL as an input device.

As part of a preliminary investigation, we examined the input gestures required for such a flexible game controller and obtained the deformation gestures from seven participants. Table 2 outlines the correspondence between the game operations and deformation gestures employed. For the deformation of bending one of the grips, the post-deformation shape of the gel exhibits similarity, which reduces the accuracy of identification when relying only on the MaGEL system. Therefore, an inertial measurement unit (IMU) was integrated into one of the grips, enabling the system to distinguish the deformation more effectively. The participants proposed deformation gestures that were inspired by real-world body movements. Their feedback included suggestions such as "synchronizing device movement with the sound of footsteps," "mimicking the action of pulling a trigger," and "developing a movement control that incorporates bodily motions."

Overall feedback was positive, confirming the usefulness of the MaGEL-based application. Many of the participants commented that "it was fun and enjoyable," "Having the actual object was great," and "can be controlled by feeling," "the soft gel was cute," and "I would like to use such a controller if one existed." However, a participant who was unfamiliar with FPS games reported difficulties owing to the complexity of the operations. Some participants also commented that they would like to avoid moving their hands too much for permanent operations, such as moving, considering the operability of the game. In this implementation, deformable inputs were used for all operations to confirm the usefulness of the

Table 2: Game control operations and corresponding deformable inputs

Game Operations	Deformation
Move forward	Move grips up and down.
Move left	Bend the right grip upwards.
Move right	Bend the left grip upwards.
Shoot	Move grips back and forth.

deformable controllers. Therefore, it is necessary to match operations and inputs considering hand fatigue and operability when developing actual games.

5.3 Future Challenges

5.3.1 Estimation of Gesture Intensity. The MaGEL proposed in this study uses a transparent gel to estimate device deformation, enabling deformation gestures to serve as input. In the evaluation experiments described in Section 4, we focus only on gesture classification. However, if the intensity of gestures could be estimated through regression analysis, or if multiple gestures, such as Push and Twist, were executed simultaneously, the system would have the capability to recognize a broader range of gestures. In addition to gestures that can be estimated by this system, conventional input methods can also be used by placing buttons on hard cases or gel parts.

5.3.2 Various Shapes of Gel. Gesture identification is feasible, even with alterations to the gel's shape, provided that the light transmission from the LED to the phototransistor varies with gel deformation. The MaGEL device exhibits a simple structure comprising a transparent gel, LED, and phototransistor. In this study, the gel was molded into various shapes, enabling flexible design customization.

The gel shape can be modified to enhance gesture recognition accuracy or to guide user interactions based on its form. The higher accuracy observed in the identification of the pull-back gesture, in comparison to Raise and Bend, suggests that using multiple rows of phototransistors, rather than a single row, may provide a more precise detection of movements perpendicular to the device. Although thin rectangular gels were employed in this experiment, thicker gels could accommodate two rows of phototransistors. In another configuration, a narrow tubular gel would likely exhibit a more pronounced deformation when the user manipulates the left and right handles towards themselves, potentially facilitating improved identification. Other gel configurations, such as an accordion shape or the inclusion of slits, can encourage the pulling of gestures from users. In addition, the risk of misclassification can be minimized by selecting gestures that are less prone to confusion during application implementation. If the device design incorporates shapes that allow its operation using various body parts, including arms and legs, it extends its accessibility to individuals with physical disabilities. In addition, unlike conventional button-based interfaces, it operates through a gripping motion, making it accessible to users who experience difficulties with precise finger movements. This design approach enhances the device's usability for a broader range of individuals, including those with limited dexterity in their hands.

The transparency of the gel, which covers a significant portion of the device, presents opportunities for utilizing the device's internal components and rear surface for gesture identification. For instance, the application of a gel sheet to a smartphone can enable operations such as pinching and shifting. Additionally, it could function as a magnifying tool, where the user holds the gel up to the screen and pulls it to zoom in the desired area.

However, the optimal number and arrangement of LEDs and phototransistors, as discussed in Section 4.3, were designed for the rectangular gel and the 13 gestures selected in this study. While the principle of gesture identification can be applied to gels of various shapes, the optimal arrangement should be re-evaluated based on the specific shape and intended gestures for each application.

5.3.3 Gel Durability. In this study, urethane resin was selected because of its softness; however, it exhibited degradation over time and demonstrated a tendency to tear. Additionally, under the application of a significant force, such as when a player is absorbed in a game, the gel component is prone to cracking at points of contact with the protrusions on the grip. It is necessary to examine its deformation capacity, specifically evaluating the intensity and number of deformations that it can withstand without tearing.

While not embedding the LED/phototransistor within the gel facilitates easier replacement of the deteriorated gel, the need for replacement persists. Given that the proposed method is compatible with other transparent and flexible materials beyond urethane resins, it would be advantageous to explore alternative materials that are more resistant to degradation.

6 Conclusion

In this study, we proposed a soft-input device, MaGEL, to realize a more intuitive deformation input. MaGEL comprises a soft and transparent gel, infrared LEDs, and phototransistors. When users perform deformation gestures, the intensity of the LED light passing through the transparent gel changes. The system detects changes in phototransistor values and identifies the deformation of the device through machine learning.

In the evaluation experiment, we confirmed the device configuration that is suitable for identifying 13 types of gestures, the recognition accuracy of the 13 types of gestures, and the real-time responsiveness of the system. The device configuration was such that four LEDs were placed at each corner of the gel, and the phototransistors were placed evenly on both sides of the gel. By sequentially turning on these four LEDs, the gestures can be identified with the highest accuracy. Ten datasets for 13 gestures were obtained from 14 participants, and each gesture could be identified with a classification accuracy of 94.1 % for intra-personal identification and 85.1 % for inter-personal identification. In addition, when gestures were identified in real time from the obtained data, it was confirmed that the gestures could be identified with a response speed of 33.9 Hz.

These results show that MaGEL has sufficient gesture identification accuracy and response speed to function effectively as a real-time input device. As a prospect, we plan to refine the system to estimate not only gesture classification but also deformation intensity through regression analysis.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number JP24H00745.

References

[1] Alberto Boem and Giovanni Maria Troiano. 2019. Non-Rigid HCI: A Review of Deformable Interfaces and Input. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 885–906. https://doi.org/10.1145/3322276.3322347

[2] Yuran Ding, Craig Shultz, and Chris Harrison. 2023. Surface I/O: Creating Devices with Functional Surface Geometry for Haptics and User Input. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 423, 22 pages. https://doi.org/10.1145/3544548.3581037

[3] Chris Harrison and Scott Hudson. 2012. Using shear as a supplemental two-dimensional input channel for rich touchscreen interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 3149–3152. https://doi.org/10.1145/2207676.2208730

[4] Chris Harrison and Scott E Hudson. 2009. Providing dynamically changeable physical buttons on a visual display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 299–308.

[5] Liang He, Gierad Laput, Eric Brockmeyer, and Jon E. Froehlich. 2017. SqueezaPulse: Adding Interactive Input to Fabricated Objects Using Corrugated Tubes and Air Pulses. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction* (Yokohama, Japan) *(TEI '17)*. Association for Computing Machinery, New York, NY, USA, 341–350. https://doi.org/10.1145/3024969.3024976

[6] Sungjae Hwang, Myungwook Ahn, and Kwang-yun Wohn. 2013. MagGetz: customizable passive tangible controllers on and around conventional mobile devices. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 411–416. https://doi.org/10.1145/2501988.2501991

[7] H Ishii. 1997. Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms. In *Conference on Human Factors in Computing Systems CHI/ACM Press*.

[8] Asuka Kadowaki, Tomoaki Yoshikai, Marika Hayashi, and Masayuki Inaba. 2009. Development of soft sensor exterior embedded with multi-axis deformable tactile sensor system. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*. 1093–1098. https://doi.org/10.1109/ROMAN.2009.5326073

[9] Yasuaki Kakehi, Kensei Jo, Katsunori Sato, Kouta Minamizawa, Hideaki Nii, Naoki Kawakami, Takeshi Naemura, and Susumu Tachi. 2008. ForceTile: Tabletop Tangible Interface with Vision-Based Force Distribution Sensing. In *ACM SIGGRAPH 2008 New Tech Demos* (Los Angeles, California) *(SIGGRAPH '08)*. Association for Computing Machinery, New York, NY, USA, Article 17, 1 pages. https://doi.org/10.1145/1401615.1401632

[10] Kazuto Kamiyama, Hiroyuki Kajimoto, Kevin Vlack, Naoki Kawakami, Terukazu Mizota, and Susumu Tachi. 2004. Gelforce. In *ACM SIGGRAPH 2004 Emerging technologies*. 5.

[11] Johan Kildal and Marion Boberg. 2013. Feel the Action: Dynamic Tactile Cues in the Interaction with Deformable Uis. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 1563–1568. https://doi.org/10.1145/2468356.2468636

[12] Johan Kildal, Susanna Paasovaara, and Viljakaisa Aaltonen. 2012. Kinetic Device: Designing Interactions with a Deformable Mobile Interface *(CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1871–1876. https://doi.org/10.1145/2212776.2223721

[13] Yiyue Luo, Kui Wu, Tomás Palacios, and Wojciech Matusik. 2021. KnitUI: Fabricating Interactive and Sensing Textiles with Machine Knitting. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 668, 12 pages. https://doi.org/10.1145/3411764.3445780

[14] Alex Mazursky, Shan-Yuan Teng, Romain Nith, and Pedro Lopes. 2021. MagnetIO: Passive yet Interactive Soft Haptic Patches Anywhere. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 213, 15 pages. https://doi.org/10.1145/3411764.3445543

[15] Yusuke Nakagawa, Akiya Kamimura, and Yoichiro Kawaguchi. 2012. MimicTile: A Variable Stiffness Deformable User Interface for Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA,

745–748. https://doi.org/10.1145/2207676.2207782

[16] Masa Ogata, Yuta Sugiura, Yasutoshi Makino, Masahiko Inami, and Michita Imai. 2013. SenSkin: Adapting Skin as a Soft Interface. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 539–544. https://doi.org/10.1145/2501988.2502039

[17] Parinya Punpongsanon, Daisuke Iwai, and Kosuke Sato. 2013. DeforMe: Projection-Based Visualization of Deformable Surfaces Using Invisible Textures. In *SIGGRAPH Asia 2013 Emerging Technologies* (Hong Kong, Hong Kong) *(SA '13)*. Association for Computing Machinery, New York, NY, USA, Article 8, 3 pages. https://doi.org/10.1145/2542284.2542292

[18] Michael Reed. 2009. Prototyping digital clay as an active material. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (Cambridge, United Kingdom) *(TEI '09)*. Association for Computing Machinery, New York, NY, USA, 339–342. https://doi.org/10.1145/1517664.1517733

[19] Toshiki Sato, Haruko Mamiya, Hideki Koike, and Kentaro Fukuchi. 2009. PhotoelasticTouch: Transparent Rubbery Tangible Interface Using an LCD and Photoelasticity. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) *(UIST '09)*. Association for Computing Machinery, New York, NY, USA, 43–50. https://doi.org/10.1145/1622176.1622185

[20] Carsten Schwesig, Ivan Poupyrev, and Eijiro Mori. 2004. Gummi: A Bendable Computer. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 263–270. https://doi.org/10.1145/985692.985726

[21] Paden Shorey and Audrey Girouard. 2017. Bendtroller: An Exploration of In-Game Action Mappings with a Deformable Game Controller. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1447–1458. https://doi.org/10.1145/3025453.3025463

[22] Ronit Slyper, Ivan Poupyrev, and Jessica Hodgins. 2010. Sensing through structure: designing soft silicone sensors. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*. 213–220.

[23] Yuta Sugiura, Gota Kakehi, Anusha Withana, Calista Lee, Daisuke Sakamoto, Maki Sugimoto, Masahiko Inami, and Takeo Igarashi. 2011. Detecting shape deformation of soft objects using directional photoreflectivity measurement. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 509–516.

[24] Taichi Tsuchida, Kazuyuki Fujita, Kaori Ikematsu, Sayan Sarcar, Kazuki Takashima, and Yoshifumi Kitamura. 2022. TetraForce: A Magnetic-Based Interface Enabling Pressure Force and Shear Force Input Applied to Front and Back of a Smartphone. *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 564 (Nov. 2022), 22 pages. https://doi.org/10.1145/3567717

[25] Gili Weinberg, Maggie Orth, and Peter Russo. 2000. The embroidered musical ball: a squeezable instrument for expressive performance. In *CHI'00 extended abstracts on Human factors in computing systems*. 283–284.

[26] Malte Weiss, Julie Wagner, Yvonne Jansen, Roger Jennings, Ramsin Khoshabeh, James D. Hollan, and Jan Borchers. 2009. SLAP Widgets: Bridging the Gap between Virtual and Physical Controls on Tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 481–490. https://doi.org/10.1145/1518701.1518779