Soft modularized robotic arm for safe human–robot interaction based on visual and proprioceptive feedback

Subyeong Ku\textsuperscript{1}, Byung-Hyun Song\textsuperscript{2}, Taejun Park\textsuperscript{2}, Younghoon Lee\textsuperscript{3} and Yong-Lae Park\textsuperscript{2}

Abstract
This study proposes a modularized soft robotic arm with integrated sensing of human touches for physical human–robot interactions. The proposed robotic arm is constructed by connecting multiple soft manipulator modules, each of which consists of three bellow-type soft actuators, pneumatic valves, and an on-board sensing and control circuit. By employing stereolithography three-dimensional (3D) printing technique, the bellow actuator is capable of incorporating embedded organogel channels in the thin wall of its body that are used for detecting human touches. The organogel thus serves as a soft interface for recognizing the intentions of the human operators, enabling the robot to interact with them while generating desired motions of the manipulator. In addition to the touch sensors, each manipulator module has compact, soft string sensors for detecting the displacements of the bellow actuators. When combined with an inertial measurement unit (IMU), the manipulator module has a capability of estimating its own pose or orientation internally. We also propose a localization method that allows us to estimate the location of the manipulator module and to acquire the 3D information of the target point in an uncontrolled environment. Using the feedback information from the internal sensors and camera, we implemented closed-loop control algorithms to carry out tasks of reaching and grasping objects. The manipulator module shows structural robustness and the performance reliability over 5,000 cycles of repeated actuation. It shows a steady-state error and a standard deviation of 0.8 mm and 0.3 mm, respectively, using the proposed localization method and the string sensor data. We demonstrate an application example of human–robot interaction that uses human touches as triggers to pick up and manipulate target objects. The proposed soft robotic arm can be easily installed in a variety of human workspaces, since it has the ability to interact safely with humans, eliminating the need for strict control of the environments for visual perception. We believe that the proposed system has the potential to integrate soft robots into our daily lives.

Keywords
Soft robotics, soft actuators, soft sensors, modularized soft robotic arm, computer vision, deep learning

1. Introduction
Industry 5.0 represents a human-centric smart production environment in which humans occupy a dominant role in the manufacturing process (Xu et al., 2021; Maddikunta et al., 2022; Leng et al., 2022). In this environment, robots serve as assistants to human production activities and creativities, and cooperation between humans and robots is emphasized. Compared to the straightforward, repetitive automated production environment characterized by Industry 4.0, there are increased interactions between humans and robots, and robots are required to perform more complex and diverse tasks than before (Lasi et al., 2014; Frank et al., 2019; Gobakhloo, 2020). Consequently, a focus on safe collaboration between humans and robots and the dexterity of the robots is essential.

For safe human–robot interactions and dexterous manipulation, continuum manipulators or soft robotic arms made of soft bellows actuators have been studied, and bellow-type artificial muscles have been widely used for...
their capabilities of large displacement and bi-directional actuation (Ranzani et al., 2013; Nguyen et al., 2019; Yang and Asbeck, 2020; Jing et al., 2022). In addition, the innate safety of soft materials makes the robots safe even in the case of collisions with humans, and their large degrees of freedom allow for complex and diverse motions for manipulation. However, the bellows that drive the robotic arms are in general made with a manual process by molding and casting or by thermal bonding of fabric and elastomer materials, making the fabrication process complicated and difficult to control the quality of the device.

To simplify the fabrication process and overcome the issue of relatively large manufacturing tolerance, recently advanced 3D-printing technologies have been employed to fabricate the complex structure of the bellow at once (Peele et al., 2015; Patel et al., 2017; O’Bryan et al., 2017; Anver et al., 2017; Hsu et al., 2018; Schaffner et al., 2018; Gonzalez et al., 2022). There have been studies on continuum manipulators composed of 3D-printed bellow actuators (Grzesiak et al., 2011; Falkenhahn et al., 2015; Qi et al., 2015; Sadeghi et al., 2017; Sui et al., 2022).

For closed-loop control of soft manipulators or arms (Figure 1(a) and (b)), wire encoders have been used to measure the states (Mahl et al., 2014; Kapadia et al., 2014; Tamadon et al., 2020). Other studies have combined the kinematics of the robot with the motion data from cameras or motion capture systems (Hannan and Walker, 2005; Kang et al., 2013; Dalvand et al., 2016; Ansari et al., 2017; Del Giudice et al., 2017; Werner et al., 2020). In spite of the simplicity of the system, wire encoders, that directly measure the displacements do not usually provide compact form factors, making it difficult to modularize the robot. In addition, the tension of the string sometimes interrupts the operation of the manipulator. On the other hand, motion capture systems are able to track 3D positions of the robot accurately without any physical interferences. However, they can be applied only in controlled environments with cleared surroundings, resulting in limited applications of pHRI.

With the development of optical equipment, it is possible to obtain 3D images using a depth measurement camera (Intel, 2021). Although the accuracy is not as good as motion capture systems (Breitharth et al., 2021), it has an advantage of simplicity in system construction and reconfiguration, since it requires only a single camera. When combined with deep learning techniques, it is possible to recognize and track target objects in a 3D space in near-real time (Girshick, 2015; Ren et al., 2015; Jocher et al., 2022) (Figure 1(c)).

Studies on soft strain sensors that could overcome the limitations (See Figure S1) of the wire encoders have been conducted (Park et al., 2012; Chossat et al., 2013; Chen et al., 2018). Omnidirectional strain sensor has also been proposed (Zymelka et al., 2017). By embedding microchannels filled with room-temperature liquid conductors in an elastomer matrix, axial strains or displacements can be easily detected based on the change in electrical resistance of the microchannels (Park et al., 2010). Automated processes, such as direct writing, printing, or thin-film deposition have recently been developed for fabrication of liquid-conductor microchannels (Boley et al., 2014; Shin et al., 2020; Park et al., 2022; Kim et al., 2023; Shin et al., 2023).

We demonstrated localization methods using soft strain sensors, and 3D vision with deep learning. By combining these two technologies, we controlled the soft robotic arm. However, it is also required to detect contacts with humans for autonomous and interactive operations. Thus, we directly embedded a triboelectric nanogenerator (TENG) in each bellow actuator as a tactile sensor specialized for detecting human touches.

A TENG converts a mechanical touch to electrical energy based on coupling of contact electrification and electrostatic induction (Wu et al., 2019; Pan and Zhang, 2019), and the electrical energy is used as a sensing signal for interaction between the human and a robot. The TENG sensor has a simple structure that consists of only a pair of conductive and dielectric layers, making it easy to be embedded or integrated into a structure without any major modification of the host system. To provide compliance for the structure, ionic hydrogel was used as the conductive layer. In this layer, ions dissolved in the gel act as charge carriers (Yang and Suo, 2018; Lee et al., 2020a), and hydrogel made of organogel with reduced evaporation was used (Lee et al., 2018; Kim et al., 2019; Lee et al., 2022).

In this study, we propose a modularized soft robotic arm composed of multiple soft manipulator modules connected in series, each consisting of three bellow-type soft actuators, pneumatic valves, soft strain sensors, and an on-board sensing and control circuit. The bellow-type soft actuators were designed to facilitate smooth and large deformations. They incorporate air chambers for actuation and embedded channels on the wall for organogel-based TENG tactile sensing (tactile sensor). Compared to previous studies on soft robotic arms (Gong et al., 2018; Chen et al., 2019), our robot stands out due to its complete
modularity. It also incorporates localization and adjustment algorithms, enabling precise feedback control. Moreover, it offers integrated sensing of human touch for enhanced physical human–robot interactions (pHRI).

For modularity, we designed a custom manifold and a control circuit that integrates all the electronic components. In this way, each manipulator module becomes a complete system as itself and can be operated independently. When connected in series, multiple manipulator modules form and act like a single robotic arm. The modular design also makes it easy to assemble the modules since they have the same components and are made by the same fabrication process. Furthermore, having the manipulator modules of the same shape has an advantage of implementing only a single trained model for detecting their motions.

The manipulator modules are localized and controlled using omnidirectional soft string sensors. The proposed sensor is extremely lightweight and has a compact form factor compared to a commercial wire encoder or a fiber-optic sensor (Liu et al., 2015), which requires a dedicated device for sensing. It was possible to estimate the pose of the robotic arm with low computational power, not requiring additional external devices. In addition, compared to the studies that required controlled environments for shape sensing and control (Camarillo et al., 2008; Li et al., 2020), we propose a vision-based deep learning method for localization suitable for application in even uncontrolled environments. By combining the data from the soft sensor and those from a depth camera with the deep learning model, we were also able to localize and control the position of the end effector in real-time.

Finally, to acquire rich information about contact during the pHRI, we further present a method that can both recognize and localize human contacts using the TENG tactile and the soft string sensors, while existing approaches can acquire only the contact information on the surrounding environment (Funamizu et al., 2016; Yamauchi et al., 2022). Embedded beneath the surface of the soft bellow actuators, the tactile sensors contribute to seamless integration and enhanced functionality, enabling safe collaboration with humans.

2. Design

The manipulator module consists of 3D-printed soft bellow actuators, soft sensors, and an inertial measurement unit (IMU) to measure and estimate the length and the pose of the module, a pneumatic system to deliver air pressure to each module, and a control circuit (Figure 2(a)). All the components were fixed at the base structure, and the soft bellows were connected to the bases through 3D-printed circular rings to prevent stress concentration on the soft part. The assembled manipulator module has a height of 234.5 mm and a weight of 1.1 kg.

2.1. 3D-printed tribo-sensitive soft bellow

When designing the bellow actuator, one of the primary requirements was achieving a gradual response in deformation to the input pressure, without any abrupt changes. Additionally, the bellow design should also incorporate the functions of features for tactile sensing, along with an easy method of fabrication to ensure high throughput. To meet these requirements, we first evaluated different shapes for the bellow design and tested their responses to the input pressure through simulation. We then selected the shape with rounded convex and concave areas for gradual deformation and large contraction range at negative input pressures. This design showed physical robustness against structural damages. More details on the simulation will be discussed in Section 5.

The triboelectric voltage generator was embedded in the bellow as the tactile sensor for recognizing and localizing external contacts. Since the amount of voltage generation is related to the contact area (Lee et al., 2017), we designed a 3D printable bellow structure that secures a large contact area. The outer radius of the bellow was designed to maximize the actuation range while maintaining the printability. The total length of the bellow actuator is 171.5 mm ($l_{\text{total}}$), with an active region of 141.5 mm ($l_{\text{active}}$). Both convex and concave areas of the bellow have the same curvature radius of 8.25 mm ($r_{\text{curve}}$), and four channels were embedded for organogel (Figure 2(b)). Four ports (diameter, $d_{\text{vent}}$) 2.2 mm) were made on the bellow to inject the organogel into the channel at the top. Each channel has vent holes, allowing residual resin after printing to be discharged. The outer wall thickness is 4 mm, and the depth of the embedded channel is 1.5 mm ($t_{\text{channel}}$) (Figure 2(b)-i). The channel depth was selected considering the structural stiffness and the clogging during printing. The minimum width of the organogel channel is 3.8 mm ($w_{\text{min}}$) which is located at the concave part of the bellow. The maximum width of the channel is 8.7 mm ($w_{\text{max}}$) and located at the outer convex part. This part was designed to have as large area as possible, for easy detection of external contacts. Four channels were placed to cover a range of 140° ($\theta_{\text{tactile}}$). The maximum width between two adjacent channels is 7.7 mm.

To increase the stability of printing and robustness of the bellow, cross-shaped structural reinforcements were added inside the concave area, where the bellow actuator has the minimum cross-section area and is prone to structural weakness (Figure 2(b)-ii). A structure for physically constraining the motion of the three actuators to prevent the actuators from buckling is shown in Figure 2(b)-iii. This is necessary because pure bending is impossible as the bellows would buckle during actuation if they were fixed only at the top and the bottom.
2.2. Modular design

The control system was designed to generate the required pneumatic pressure to actuate the bellows within the manipulator module, while also delivering pneumatic power to other manipulator modules. Six proportional solenoid valves (VSO series 11, Parker) were installed to regulate the input pressure, connected using a 3D-printed custom-designed manifold to reduce the weight. Compressed air was then provided through the single-coiled pneumatic tube connecting the preceding and the subsequent modules.

A custom-designed circuit board was installed to control the manipulator module. The board consists of a current driver for the solenoid valve, a circuit for power delivery, and circuits for acquiring and processing the signals from the tactile sensors and string sensors, and a port to communicate with the IMU (WitMotion, WT901). A wireless communication unit (LOLIN D1 mini, WeMos) was installed to make the entire system untethered. To measure the pressure of the bellow, pressure sensors (XGZP6847-040KPG, CFSensor) were installed in the middle (Figure 2(d)).

To achieve a large range of motion through contraction and elongation of the bellow while also maintaining a simple and modular design with a single pneumatic source tube, we utilized ejectors to generate negative pressure from positive pressure (Figure 2(e)).

2.3. Material selection

The bellow was printed with a photopolymer resin (Elastic 50A, Formlabs) with the lowest shore hardness of 50A and the maximum strain of 160% (Formlabs, 2020). This material is soft enough for safe pHRI and also capable of tolerating large enough strains without failures during repeated operations. It also shows a fast shape recovery rates after actuation. The other parts were made of rigid materials (Clear and Black, Formlabs).

The organogel was fabricated using acrylamide (AAm; Sigma, A8887) and N,N-methylenebisacrylamide (MBAAm; Sigma, M7279) as a monomer and a crosslinker, respectively. Ethylene glycol (EG; DAEJUNG, 4026-4105) was used as a liquid constituent, and lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP; Sigma, 900889) was used as a photoinitiator. Lithium chloride (LiCl; DAEJUNG, 5086-4405) was used as ionic charge carriers. Trichlorosilane (Heptadecafluoro-1,1,2,2-tetrahydrodecyl) (HDFS; JSI Silicone Co., H5060.1) was used as a surface perfluorination agent.
The soft string strain sensor was made of highly stretchable elastomer (Ecoflex-0030, Smooth-On). Since the manipulator module has a height ranging from 100 mm to 200 mm, the soft string strain sensor was fabricated with an initial length of 90 mm and prestrained. Ecoflex-0030 was chosen as it easily stretches to the strain of 150% or more, and the stress applied under this range is low (Xavier et al., 2021).

3. Fabrication

3.1. Tribo-sensitive soft bellow

The bellow actuator is made of relatively complex structure with different features, such as curvatures, channels, and connectors. We fabricated this structure using an SLA 3D-printer (Form3, Formlabs) (Figure 3(a)). Since the width of the embedded channel was small, residual resin had to be removed after printing, and it was done by blowing compressed air through the open port of the channel and rinsing the channel using isopropyl alcohol. Afterward, the bellow was cured to complete the post process (Form Cure, Formlabs). Printing the bellow horizontally, as opposed to vertically in relation to the build plate, resulted in an increased printed area of each layer. This orientation also improved the bonding strength between adjacent layers, reducing the likelihood of delamination when the bellow was subjected to stretching (Dwiyati et al., 2019).

3.2. Organogel

The 3D-printed bellow was treated with air plasma under vacuum for 30 s to form hydroxyl terminated on the surface of the bellow (EQ-PCE-3, MTI Corp.) (Figure 3(b)). The surface-activated bellow was immersed in HDFS solution dissolved in hexane with a mixed ratio of 1:300 (Figure 3(c)). Self-assembled monolayer formation was performed for 30 min (Lee et al., 2018, 2020b). The treated surface was then rinsed with hexane for 10 min. Organogel precursor solution, composed of 3.5 M AAm, 1.5 M LiCl, 0.17 mM LAP, and 8 mM MBAAm was injected into the bellow (Figure 3(d)) and then polymerized under 400 nm ultraviolet irradiation in a curing device (Form Cure) for 10 min (Figure 3(e)).

3.3. Soft string strain sensor

To measure the length and to estimate the position of the manipulator module, we developed and used a soft string strain sensor. Since the manipulator module was designed to bend in all directions, the strain sensors were required to show consistent sensor readings for strains in all directions. A liquid-metal trace was directly printed onto a flexible silicone substrate and sealed, and then the substrate was rolled to form a cylindrical shape. The length of the manipulator module was estimated using the resistance change when the module was actuated.

The string sensor was fabricated with the following procedures. First, a thin layer of liquid-state Ecoflex-0030 was spread using an applicator (Elcometer 4340, elcometer®) and cured. Liquid-metal (eGaIn) trace was printed on the cured silicone substrate using a motorized x-y-z stage (Shotmaster 300VX, Musashi), a pneumatic dispensing system (SuperΣ CMIII V2, Musashi), and a laser distance sensor (LK-G32, Keyence) (Figure 4(a)). Then, the printed liquid-metal trace was covered with another layer of Ecoflex-0030, and the signal wires were connected (Figure 4(b)). The sensor was cut and placed on a flexible plastic sheet. This sheet was then rolled into a cylindrical shape (Figure 4(c) and (d)). Liquid-state Ecoflex-0030 was injected into the center of the rolled sheet (Figure 4(e)) and cured. The flexible sheet was then removed to complete the string sensor with a diameter of 4.5 mm and a length of 90 mm. 3D-printed rigid parts were added to both ends of the sensor for mechanical connection with the manipulator module.

4. Modeling

The bellow actuator contracts or expands depending on the air pressure applied to the chamber. Combinations of single-axis motions of three actuators enable linear and bending motions of the manipulator module. We modeled the pressure response of the manipulator module in 3D space and the inverse kinematics for calculating the required internal pressure for the manipulator to reach the desired point in the space.

The general method for solving the linkage system of a rigid body by considering contraction and expansion with the changes in link length, was not suitable for a soft body. The analysis is more complex since continuous linear deformation and bending occur in all the elements constituting the body. One approach to simplify this complex system is to assume the manipulator module has a constant curvature (Webster and Jones, 2010; Yang and Asbeck, 2020).
4.1. Forward kinematics

In the assumption of constant curvature, the pose of the manipulator module is described with respect to the center point of both ends. The posture is determined by the arc lengths $l_1$, $l_2$, and $l_3$ of the three bellows. These lengths are variables in the actuator space ($q = [l_1, l_2, l_3]^T$) and are used to calculate transformation into configuration space.

Factors affecting the bellow length include the internal pressure, geometrical parameters, and the material property of the bellow. According to Hooke’s law, the pressure $p_{bi}$ applied to $i$th bellow causes the length change and the arc length of the $i$th bellow can be expressed as

$$\Delta l_i = \frac{p_{bi} A_{eff}}{k_i},$$

$$l_i = l_i^0 + f(p_{bi}') ,$$

$$f(p_{bi}') = a_0 p_{bi}' + a_1,$$

$$p_{bi}' = \begin{cases} p_{bi} - \min(|p_{bi}|) & \text{if } \text{sgn}(p_{bi}) = 1 - 1 \\ p_{bi} & \text{if } \text{sgn}(p_{bi}) = 0 \end{cases}$$

where $i = 1, 2, 3$, and $\Delta l_i$ is the deformed length of the $i$th bellow, $A_{eff}$ is the effective cross-sectional area of the bellow, assumed to be constant, $l_i$ is the arc length, $l_i^0$ is the initial length of the active region, $k_i$ is the structural stiffness, and $f(p_{bi}')$ is the first-order fitting function that adjusts the arc length between predictions by simulation and Hooke’s law. $a_0$ and $a_1$ are the coefficients of the adjust function and $p_{bi}'$ is the pressure subtracting the minimum absolute gauge pressure among the three bellows, considering the direction of the pressure. $p_{bi}$ is the gauge pressure of the $i$th actuator.

When all three pressures are the same (Figure 5(a)), $f(p_{bi}')$ becomes zero in the equation (3), so only translational motion is considered. In the case of bending, the sequence is divided into two (Figure 5(b)). Translation occurs first due to $p_{bi}$, and then bending occurred due to the $p_{bi}'$ of equation (4). Since the strain energy by the input pressure will be equal to the sum of the axial and the bending strain energies, the different pressures of bellows indicate the bending strain energy. In this case, the axial strain will be reduced, and as a result, the arc length will be increased, as in equation (2). Thus, the function $f(p_{bi}')$ plays a role in the consideration of elastostatics when determining the length of the bellow. The stiffness $k_i$ was determined through simulation based on the bellow geometry and the material properties (Hermann et al., 1997).

The configuration spaces of the manipulator module $l_c$, $\phi$ and $\kappa$ are then determined by the lengths $l_i$ as follows (Webster and Jones, 2010; Yang and Asbeck, 2020)

$$l_c(q) = \frac{l_1 + l_2 + l_3}{3},$$

$$\phi(q) = \tan^{-1}\left(\frac{\sqrt{3}(l_1 + l_2 + l_3)}{3(l_2 - l_1)}\right),$$

$$\kappa(q) = \frac{2\sqrt{l_1^2 + l_2^2 + l_3^2 - l_1 l_2 - l_1 l_3 - l_2 l_3}}{d(l_1 + l_2 + l_3)},$$

$$\theta(q) = l_c(q) \kappa(q),$$

where $l_i$ is the arc length between the center point of the top and the bottom, $\phi$ is the angle formed about $x$ axis when the center point on the top is projected on $x - y$ plane, $\kappa$ is the curvature of the manipulator module, $d$ is the distance between the center point on the top and the center of each bellow, and $\theta$ is the angle formed by the arc segment with the $x - y$ plane.

From the configuration spaces, the top center position is determined through the geometrical relation (Drotman et al., 2018). Our manipulator module has interfaces for

Figure 4. Fabrication process of soft strain sensor. (a) Direct printing of a liquid-metal trace on a silicone substrate. (b) Encapsulation of the printed trace with another layer of elastomer and wire connection. (c) Transferring the flat sensor on a flexible sheet. (d) Rolling up the sensor to a thin cylinder. (e) Filling the center void with liquid-state silicone and curing.

Figure 5. Two actuation modes of the manipulator module: (a) Linear translation and (b) bending in sequence.
connection to the next module at the top and the bottom. These are passive structures that are not affected by the pneumatic actuation, and so can be modeled as

\[
c = 2\rho \sin \left( \frac{\theta}{2} \right),
\]

(9)

\[
x_c = c \sin \left( \frac{\theta}{2} \right) \cos \phi + l_u \sin \theta \cos \phi,
\]

(10)

\[
y_c = c \sin \left( \frac{\theta}{2} \right) \sin \phi + l_u \sin \theta \sin \phi,
\]

(11)

\[
z_c = c \cos \left( \frac{\theta}{2} \right) + l_b + l_u \cos \phi
\]

(12)

where \( c \) is the chord length, \( \rho \) is the radius of curvature, \( l_u \) and \( l_b \) are lengths of passive structures at the top and the bottom, respectively (Figure 6), and \( x_c, y_c, \) and \( z_c \) are the Cartesian coordinates of \( p_c \).

4.2. Inverse kinematics

In this study, inverse kinematics was used to determine the length of each bellow from a given point \( p(x, y, z) \) in the task space, and to find the input pressure \( p_i \) using the length and equations (2) to (4). The inverse kinematics was solved based on the coordinates of the point \( p \) rather than \( p_i \) for convenience of calculation (Figure 6). Alternatively, it is possible to use position \( p_i \) and the circular reference in the calculation. The relationship of the configuration space parameters, the lower passive height \( l_u \) and the coordinates of the given point \( p(x, y, z) \) is as follows (Webster and Jones, 2010; Drotman et al., 2018)

\[
\phi = \begin{cases} \tan^{-1}(y/x) & \text{if } x > 0, y > 0 \\ 2\pi - \tan^{-1}(y/x) & \text{if } x > 0, y < 0 \\ \pi + \tan^{-1}(y/x) & \text{if } x < 0 \end{cases}
\]

(13)

\[
\kappa = \frac{2\sqrt{x^2 + y^2}}{x^2 + y^2 + (z - l_b)^2}
\]

(14)

\[
\theta = \begin{cases} \cos^{-1} \left( 1 - \kappa \sqrt{x^2 + y^2} \right) & \text{if } z > 0 \\ 2\pi - \cos^{-1} \left( 1 - \kappa \sqrt{x^2 + y^2} \right) & \text{if } z \leq 0, \end{cases}
\]

\[
l = \frac{\theta}{\kappa}
\]

(15)

\[
l_i = l - \theta d \cos \left( \frac{2\pi}{3} (i - 1) + \frac{\pi}{2} - \phi \right). i = 1, 2, 3
\]

(16)

From the above equations, the top center point \( p_c(x_c, y_c, z_c) \) is determined as

\[
x_c = x + l_u \sin \theta \cos \phi,
\]

(18)

\[
y_c = y + l_u \sin \theta \sin \phi,
\]

(19)

\[
z_c = z + l_u \cos \theta
\]

(20)

By combining equations (2)–(4), the deformation \( \Delta l_i \) of the \( i \)th bellow and the corresponding internal pressure \( p_{bi} \) are determined as

\[
\Delta l_i = l_i - l_0 - a_{i} p_{bi}' - a_0,
\]

(21)

\[
p_{bi} = \frac{\Delta l_i k_i}{A_{eff}}
\]

(22)

Here, the coefficients \( a_0 \) and \( a_1 \) of the function \( f'(p_{bi}) \) are determined through simulation. To solve the equation (21), the pressure with the lowest absolute value among the three bellows must first be found. We determined this value \( p_{bi}' \) from the result of equation (17). If the calculated \( l_i \) is larger than the length of the initial active region \( l_0, p_{bi}' \) is applied to the shortest bellow \( l_i \). Likewise, if \( l_i \) is smaller than \( l_0, p_{bi}' \) (a negative pressure in this case) is applied to the bellow with the longest \( l_i \). Then, for the shortest or the longest bellow for each case, respectively, equation (21) becomes

\[
\Delta l_i = l_i - l_0 - a_0
\]

(23)

Therefore, equation (22) can be solved to determine the lowest pressure, and this value is used to calculate pressures of the other bellows.

5. Simulation

We conducted finite element analysis (FEA) using the COMSOL Multiphysics® (COMSOL) to find the optimum design of the bellow actuator, structural stiffness of the manipulator module, and the pressure response to various
input pressures. The Neohookian model, suitable for small strain simulation (Xavier et al., 2021), was used to analyze the behavior of soft bodies. The simulation parameter was determined from the stress-stretch curve obtained from a uniaxial tensile test (Figure S2).

5.1. Design optimization

To find an optimal design of the bellow actuator, we created four designs featuring combinations of rectangular and circular curves at both the convex and concave regions, and measured the height during deformation under both positive and negative pressure conditions. We kept the other parameters, such as thickness of the wall, the maximum and the minimum radii of the convex and the concave regions, respectively, the number of repetition, and the height of the bellow. The selection criteria were the gradual rate of deformation and the large contraction range when the actuator was subjected to negative input pressure, as mentioned in Section 2.

We designed and simulated a simple flat-wall bellow, a flat-wall bellow with rounded corners, a curved-wall bellow, and a hybrid-wall bellow of the flat and the curved walls for the outer convex and the inner concave regions, respectively (Figure 7(a)). All bellows were designed to have an overall height of 80 mm, three layers with a height of 20 mm for each layer, and a wall thickness of 5 mm. Depending on the geometry of each layer of the bellows and the shape of the corner, the contraction and the expansion lengths were different, even for the same pressure (Figure 7(b)). We chose the curved-wall bellow for our actuator design, since it showed the largest deformation when vacuums and a smoother change than the hybrid-wall bellow for a given input pressure range (−30 kPa ∼ +30 kPa), meeting our selection criteria.

5.2. Linear translation

We provided the pressure of the same magnitude and direction to the three bellows and simulated the response of linear translation (Figure 7(c)). The response was smaller in the positive deformation range, and the response pattern changed based on a specific threshold in the negative deformation range. Pressure smaller than the threshold showed a linear behavior because the bellow was fully collapsed and behaved like a cylinder. We performed curve fitting based on the results and obtained the stiffness by differentiating the fitting curves with respect to the displacement (Figure 7(d)). The linear fit lines of the structural stiffness $k_i$ of the bellow are expressed as

$$ k_i = \begin{cases} 0.0152(\Delta d) + 0.32 & \text{if } \Delta d > 0 \\ 0.0038(\Delta d) + 0.29 & \text{if } \Delta d \leq 0, \end{cases} $$

where $\Delta d$ is the deformed length.

5.3. Bending

We provided negative pressure to one of the three bellows in a single manipulator module to simulate the bending response of the manipulator module since the system showed a higher response to the negative pressure than to the positive pressure from the stiffness result.

Since estimation of the bellow length based on Hooke’s law was not accurate in the case of bending, the error was compensated by using the simulation result and the prediction by Hooke’s law. Figure 7(e) schematically shows the difference in arc length between the simulation and the kinematic model. (f) Arc length difference between model and simulation, and a linearized estimation of the difference. (g) $x$ and $z$ position in the task space during bending of the module from input pressure of 0 to −35 kPa. (h) Bending angle response to input pressure.
simulation increased as the pressure decreased, we im-
plemented the adjustment function to compensate for the
difference. A linear estimation that connects the maximum
difference and zero (i.e., the initial difference) was used,
which was sufficient in preventing overfitting and solving
the inverse kinematics. The coefficients $a_1$ and $a_0$ of the
adjustment function $f$ were $-0.325$ and $0$, respectively.

The model without adjustment is represented by the gray
triangular markers in Figure 7(g). Given the same pressure
input, the model prediction demonstrates a positional dif-
ference compared to the simulation and the actual mea-
surement. This is the most noticeable at the end point of the
curve at the minimum vacuum pressure, where the model
without adjustment reached a greater $x$ position and a
smaller $z$ position. However, by applying the adjustment to
the deformation of the bellows, a more accurate position
estimation was possible, as shown with the black circular
markers. The bending angle response corresponding to the
result Figure 7(g) can be seen in Figure 7(h). Similarly,
without adjustment to the model, bending angle estimation
was larger than the actual angle, the difference increasing
towards smaller vacuum pressure. The results of the sim-
ulation for linear translation and bending can be seen in
Supplementary Movie S1.

6. System configuration

The proposed soft robotic arm consists of one or more
manipulator modules (Figure 2(a)), and its control hardware
is composed of a sensing and actuation system, a pneumatic
system, and a depth camera (Figure 8(a)). All the compo-
nents except for the depth camera that is connected to the
host PC, such as the string sensor, the tactile sensor, the
IMU, and the pressure sensor are connected to the controller
on the manipulator module. Compressed air generated by
the pump is delivered to the solenoid valves through the
pressure regulator and the manifold. With this hardware
setup, the internal pressure of the bellows is controlled.

We propose two methods of localization of the manip-
ulator module for estimating and controlling its pose. One is
to use the depth camera and a trained deep learning model
(YOLOv5) to measure the pose of the manipulator module
on the host PC (Fig. 8(b)). This external measurement with
the depth camera has the advantage of providing additional
information on the surrounding environment. However, it
requires installation of both the camera and the host PC.
Moreover, it cannot be used in an environment with oc-
cclusions. The other method is to use the string sensor and the
IMU (Fig. 8(c)). The advantage of this method is that the
robotic arm can be fully independent without any assistance
from external devices.

We also propose a method for detecting and dis-
tinguishing the types of contacts made to the robotic arm,
which allows the robot to perform additional tasks when the
robot is in contact with humans using the signals from the
string sensor, the pressure sensor, and the tactile sensor.

7. Manipulator module localization

We calculate the 3D position of the manipulator module
through the forward kinematics calculation based on the on-
board pressure sensor readings. However, an error exists
between the actual position and the model prediction.
Therefore, we used two sensors to localize and control the
position of the manipulator module.

7.1. String strain sensor-based localization

To estimate the center position of the top, three position
vectors of the top plate of the manipulator module are re-
quired. Therefore, three string sensors were placed between
the bellows. Experiments were first conducted to charac-
terize the string sensors. The string sensor showed low
hysteresis for the displacement up to 100 mm (Figure 9(a)).
The maximum difference with hysteresis was 18 mV, as
shown in the subset of the figure, and the corresponding
position difference was 2.6 mm.
The sensor also demonstrated durability and consistent signal readings over 2,000 cycles of 125% strain (Figure 9(b)). The result of quadratic polynomial fitting of the sensor data showed the $R^2$ value of 0.99.

Different experiments were also conducted to verify the omnidirectionality of the string sensor. An industrial robot arm (UR5e, Universal Robots) was used to stretch the sensor in various directions (Figure 9(c)). The sensor was stretched up to 200 mm with the stretching direction varying from $-60^\circ$ to $+60^\circ$ from the vertical, incremented by $30^\circ$. The robot’s end-effector was driven along the dotted lines from the initial point (i.e., red dotted line) to the blue target point in each direction for a total length of 200 mm (Supplementary Movie S4). The sensor responses in all directions were nearly identical (Figure 9(d)-top). This differs from the test results of a planar-type liquid-metal sensor, a commonly used design of soft strain sensor in various applications. The planar design usually shows inconsistency in the responses for different directions, as shown in Figure 9(d)-bottom. Furthermore, this experiment was also conducted with a planar-shaped sensor with an even smaller width (the same length for the liquid-metal trace) and the results can be found in the Figure S3.

Algorithm 1 Projected gradient descent with backtracking line search

**Parameters:** initial step size of gradient direction $\alpha$, return $\beta$, quadratic cost function $f$, termination threshold $c_0$, step size $\eta$

**Input:** goal position vector $\vec{x}_g$, projection function $P_c$

**Output:** position vector $\vec{x}$

$\vec{x} \leftarrow \vec{x}_g$, $\eta \leftarrow 1$, $\alpha \leftarrow 0.5$, $\beta \leftarrow 0.12$

while $(f(\vec{x}) > c_0)$ do
  while $(f(\vec{x}) - \alpha \eta \|\nabla f(\vec{x})\|^2 - f(\vec{x} - \eta \nabla f(\vec{x})) < 0)$ do
    $\eta \leftarrow \beta \cdot \eta$
  end while
  $\vec{x} \leftarrow P_c(\vec{x} - \eta \cdot \nabla f(\vec{x}))$
  $\eta \leftarrow 1$
end while

return $\vec{x}$

When assembling the string sensor with the manipulator module, errors may occur due to the different measurement environment. We adjust this using the initialization structure depicted in Figure S4. We localized the top center point using the measured lengths from the string sensors and the Euler angle measured by the IMU (Figure 9(e)). From the string sensors, we obtained the following equations:

\[
\left\| \vec{p}_{1,1}\vec{p}_{2,1} \right\|^2 = l_{1,1}^2, \tag{25}
\]

\[
\left\| \vec{p}_{1,2}\vec{p}_{2,2} \right\|^2 = l_{2,2}^2, \tag{26}
\]

\[
\left\| \vec{p}_{1,3}\vec{p}_{2,3} \right\|^2 = l_{3,3}^2. \tag{27}
\]

where $p_{1,i}$ and $p_{2,i}$ ($i = 1, 2, 3$) are the start and end positions of the string sensors. We also obtained the following equations from the IMU data and the geometric relationship of the manipulator module.

\[
R = R_z(\psi)R_y(\theta)R_x(\phi), \tag{28}
\]

\[
R \cdot \vec{O}_{p_{1,1}} = [a_1, \beta_1, \gamma_1]^T = \vec{O}_{p_{2,1}}, \tag{29}
\]

\[
R \cdot \vec{O}_{p_{1,2}} = [a_2, \beta_2, \gamma_2]^T = \vec{O}_{p_{2,2}}, \tag{30}
\]

\[
R \cdot \vec{O}_{p_{1,3}} = [a_3, \beta_3, \gamma_3]^T = \vec{O}_{p_{2,3}}. \tag{31}
\]

Figure 9. Experimental result of soft string sensor. (a) Normalized sensor signal during a cycle of stretching and releasing. The blue and the red curves in the inset plot represent extension and recovery, respectively. (b) Sensor signal over 2,000 cycles. The inset plot is the result of the 280-th cycle. (c) The experimental setup for directional response of the string sensor. (d) Signal response to stretching in various directions of the string sensor and a planar sensor. (e) Schematic of manipulator module localization using string sensors. (f) Measurement and localization result.
Figure 9(f). Cost calculation result with and without the localization result using the string sensor can be seen in Figure S5. Since these systems of nonlinear equations can only be solved for ideal cases, instead, we found a solution that minimizes the error of the following cost function.

\[
 f(\vec{x}) = \sum_{i=1}^{3} \left( \| \vec{x} + R \cdot \overrightarrow{O_1p_{i,1}} - \overrightarrow{O_1p_{i,1}} \| - l_i \right)^2, \tag{35}
\]

where \( \vec{x} \) is the position vector of the center of the top. We used projected gradient descent (PGD) with a backtrack line search algorithm Armijo (1966), as shown in Algorithm 1.

The termination parameter \( c_e \) was set as 1, so the iteration is terminated when the cost function value becomes smaller than \( c_e \). We iteratively found the step size \( \eta \) by backtracking line search. After multiplying the tangent by the direction \( \alpha \), we repeated multiplying \( \eta \) by return \( \beta \) until the difference between \( f(\vec{x}) - \alpha \cdot \eta \| \nabla f(\vec{x}) \| \) and \( f(\vec{x} - \eta \cdot \nabla f(\vec{x})) \) became positive. Through this process, the step size \( \eta \) was adaptively adjusted, and the cost function could be updated at every step. Considering the limited workspace of the manipulator module, the solution is projected if the calculated position exceeds the range of motion of the manipulator module. Since the edge device with low computing power is used, this adaptation helps to reduce the number of iterations. The localization result using the string sensor can be seen in Figure 9(f). Cost calculation result with and without the PGD can be seen in Figure S5.

7.2. Vision-based localization

In addition to string sensors, the manipulator module was localized externally with a depth camera (RealSense L515, Intel®). Colored landmarks were attached to the top and the bottom plates, and the length of each bellow was calculated using the measured depth information and a transformation matrix. While the color detection mechanism is relatively simple, similar color objects or noises also affect the performance of the depth camera. To solve this issue, the surrounding must be cleared, which poses more constraints to the operating environment.

Instead, we utilized an one-stage object detection model named YOLOv5 (Jocher et al., 2022) (Figure 10(a)) to find a region of interest which contains the manipulator module in the image. Train images were obtained from various distances from the manipulator module with different orientations. Images shown in Figure S6 are not part of the training dataset, yet the model was able to detect the module even in complex environments. The captured images were rotated 90°, 180°, and 270° through image processing to diversify the dataset. Demonstration of detecting the manipulator module under various conditions can be seen in Supplementary Movie S2. The trained YOLOv5 model can achieve 60 frames per second on a 720 × 480 image and 30 frames per second on a 1280 × 720 image with a graphics processing unit (Geforce 1060, NVIDIA), which allows sufficiently fast processing for close-to-real-time control. Figure 10(b) compares the performance of the depth camera with and without object detection: the color detection (left) showed difficulty in localizing the manipulator module due to the various surrounding objects with similar colors, while object detection (right) performed better in localization despite these noises. In addition, even if a gripper was mounted at the end of the manipulator module, the trained model was able to detect it (See Figure S6), and the performance of the trained model is summarized in Supplementary Table S1.

With the depth information obtained from the depth camera and the trained model, the top center point of the manipulator module can be calculated using vectors determined by the IMU and other design parameters. The depth camera provides two position vectors \( \overrightarrow{OC_{LM1}} \) and \( \overrightarrow{OC_{LM2}} \), which connect the landmarks \( LM_1 \) and \( LM_2 \) on the manipulator module from the origin of the camera frame \( \{ C \} \) (Figure 10(c)). Also, we can define the position vectors from the center points of the top and the bottom to each bellow as \( \overrightarrow{O_1A_1}, \overrightarrow{O_1B_1}, \overrightarrow{O_1C_1}, \overrightarrow{O_1D_1}, \overrightarrow{O_2A_2}, \overrightarrow{O_2B_2}, \overrightarrow{O_2C_2}, \overrightarrow{O_2D_2} \), and \( \overrightarrow{LM_1} \) and \( \overrightarrow{LM_2} \) which connect the points \( A_2, B_2, C_2, \) and \( D_2 \) with the origin of the local frame \( \{ 2 \} \). Then, the vectors \( \overrightarrow{OC_{LM1}} \) and \( \overrightarrow{OC_{LM2}} \) can be described with respect to local frame \( \{ 1 \} \) as

\[
\begin{align*}
\overrightarrow{OC_{LM1}} &= \mathbf{R}^T \overrightarrow{OC_{LM1}}, \\
\overrightarrow{OC_{LM2}} &= \mathbf{R}^T \overrightarrow{OC_{LM2}},
\end{align*}
\]

where \( \mathbf{R} \) is the rotation matrix that transforms the local frame \( \{ 1 \} \) into the camera frame \( \{ C \} \). This matrix is obtained from the initial setup. From equations (36) and (37), we get vectors \( \overrightarrow{OC_{O_1}}, \overrightarrow{OC_{O_2}} \) as

\[
\begin{align*}
\overrightarrow{OC_{O_1}} &= \mathbf{R} \overrightarrow{OC_{LM1}} - \overrightarrow{OC_{LM1}}, \\
\overrightarrow{OC_{O_2}} &= \mathbf{R} \overrightarrow{OC_{LM2}} - \overrightarrow{OC_{LM2}},
\end{align*}
\]

where \( \mathbf{R} \) is the rotation matrix that transforms the local frame \( \{ 1 \} \) into the local frame \( \{ 2 \} \). This rotation matrix is derived from the roll, pitch, and yaw angles measured by the IMU. Finally, we obtain the vectors \( \overrightarrow{O_1O_2}, \overrightarrow{O_1A_2}, \overrightarrow{O_1B_2}, \overrightarrow{O_1C_2} \) as

\[
\begin{align*}
\overrightarrow{O_1O_2} &= \mathbf{R} \overrightarrow{O_1O_2}, \\
\overrightarrow{O_1A_2} &= \mathbf{R} \overrightarrow{O_1A_2}, \\
\overrightarrow{O_1B_2} &= \mathbf{R} \overrightarrow{O_1B_2}, \\
\overrightarrow{O_1C_2} &= \mathbf{R} \overrightarrow{O_1C_2}.
\end{align*}
\]
Localization result using equations (44)–(46) can be seen in Figure 10(d). Regardless of the direction of the camera or the magnitude of the pressure, it is possible to localize the manipulator module and estimate the length of the bellows actuators in real-time as shown in Supplementary Movie S3.

8. Experiment

8.1. Pressure response

An experiment was conducted to verify the workspace and repeatability of the manipulator module. First, we measured the bending response to different input pressures for each bellows actuator (Figure 11(a) and (b)). The maximum bending angle of 70.4° was achieved with an input pressure of −40 kPa at \( p_{b1} \) and +30 kPa at \( p_{b2} \) and \( p_{b3} \). The bending angle can be further increased by increasing the positive pressure, but we limited the maximum pressure to +30 kPa for safety (Position 1). When only the positive pressure of \( p_{b2} \) and \( p_{b3} \) were removed while maintaining the negative pressure at \( p_{b1} \), the bending angle decreased to 49.2° (Position 2). The dash-dotted line shows the bending responses when only the negative pressure was applied. Figure 11(c) shows the maximum and the minimum lengths achieved by expansion and contraction, respectively. The manipulator module reaches 102 mm, and expanded to 204 mm when the input pressures of −40 kPa and +50 kPa, respectively, were applied to the three actuators. However, as mentioned previously, we limited the maximum input pressure to +30 kPa for safety, with the corresponding length of 198 mm.

To verify the workspace of the manipulator module in a 3D space, we tracked its top center position using a motion capture system (OptiTrack) with all possible combinations of pressure inputs ranging from −40 kPa to 30 kPa with a 10 kPa increment (Figure 11(d)). The initial position is shown with a red dot in the figure. The manipulator showed larger displacement in contraction than in expansion. The maximum velocity of the manipulator was also measured using the same experimental setup. The maximum speeds during linear translation were 21.7 mm/s and 36.4 mm/s when 40 kPa and −30 kPa were applied to all three bellows, respectively. The maximum speeds for bending were 54.6 mm/s and 61.0 mm/s when −30 kPa was applied only to a single bellows and when 40 kPa was additionally applied to the remaining two bellows, respectively.

Next, the pressure of the manipulator module was controlled by a proportional-derivative (PD) controller with \( P \) and \( D \) gains \( K_P \) and \( K_D \) of 0.013 and 0.5, respectively. We first tested step input responses (Figure 12(a)). The goal
pressure was increased with an increment of 10 kPa in the range −40 kPa to +20 kPa, which was selected to prevent any damages to the bellows. As a result, when the difference between the input and the initial value increased, so did the settling time.

We also experimented the effect of the control limit (Figure 12(b)). The flow rate of the solenoid valve was controlled by the input voltage, and we limited the change of the input voltage to 10 mV, which was determined empirically; voltage changes smaller than 10 mV required too long settling times for effective control of the manipulator module compared to the voltage change larger than 10 mV. On the other hand, voltage change larger than 10 mV resulted in an overshoot with positive pressure inputs. As a result, we selected a control limit of 10 mV. In addition, we conducted an experiment to determine the resolution of the manipulator module. The response was recorded while the input pressure was changed by 0.1 kPa or 0.05 kPa. By controlling each solenoid valve, the desired pressure was achieved (Figure 12(c)-left). While the difference in bending angle in response to the pressure change of 0.05 kPa is indiscernible, as shown with the brown and green curves in Figure 12(c)-right, it becomes clear for changes of 0.1 kPa, as shown with the black-blue and the red-brown curves. Consequently, our conclusion for the resolution is 0.1 kPa.

We conducted an experiment to verify the repeatability of the manipulator module. The manipulator module was tested under repeated actuations and length changes (Figure 12(d)) for more than 5,000 cycles. All three bellows were loaded and unloaded with an input pressure of −40 kPa pressure. The inset in Figure 12(d) shows the length changes during one cycle. The minimum height of the manipulator module in each cycle was plotted with a red square marker, and a linear line was fitted over all markers to verify the consistency. The slope of the fitted line was $-5 \times 10^{-6}$, which suggests that over 100,000 cycles, the manipulator module shows the change of −0.5 mm in height, and we were able to verify the consistent performance of the actuators and the manipulator module.

We conducted a frequency sweep test on the manipulator module. The pressure response was tested with sinusoidal inputs with various periods ranging from 3.96 s to 1.98 s (Figure 12(e)). The shortest period the manipulator module was able to follow was 2.64 s, as it can be observed that the pressure failed to follow the input with a period of 1.98 s. Figure 12(f) shows the pressure response and the corresponding response of the bending angle with a sinusoidal input wave with a period of 3.96 s. The manipulator module responded to even small fluctuations in pressure, with a delay of approximately 130 msec. after the minimum or the maximum pressure was measured.

Finally, the payload of the manipulator module was determined through experiment. Since the payload of soft actuators is dependent on the orientation of the actuator as well as the input pressure, we defined the payload as the mass that causes a 10% increase in the contraction length of the bellow, in a suspended orientation. For −40 kPa, which is the limit of the onboard pneumatic system, the manipulator was able to withstand a weight up to 2 kg while the contraction length changes from 101 mm to 111 mm. When a lower input pressure of −80 kPa was applied to the bellows from the external pneumatic source, the manipulator module was able to lift 4 kg of weight with only 5% increase in the contraction length. Considering the manipulator’s own weight (1.1 kg), we are able to connect up to three modules and form a functioning manipulator. However, it is expected that the additional payload of the modules will degrade the manipulator performance, such as reduced workspace of the preceding modules or increased settling time.

### 8.2. Contact recognition

We tested contact sensing for safe interaction with surroundings including humans. The manipulator module first detects contacts through the air pressure sensor. When a contact is made, the internal volume of the bellow changes instantaneously, resulting in a change in the internal pressure. We applied the density-based local outlier factor
(LOF) algorithm for contact recognition (Breunig et al., 2000; Alghushairy et al., 2020). Based on the algorithm, the lower the local density of a given point, the more likely it is to be an outlier of the cluster. For example, when a weak contact causes a pressure change of 0.1 kPa, the LOF value would be \( \frac{3.02}{C_0} \), which is clearly different from the average of the cluster. To quantify the LOF, the followings are defined.

\[ k/N_k(p) \text{ is the } k\text{th smallest distance from the given point } p. \]

\[ N_k(p) \text{ is the } k\text{-nearest neighbor set of data points whose distance from the point } p \text{ is less than } k - \text{dist}(p) \text{ is denoted as } N_k(p). \]

\[ \text{The reachability distance is defined as: } \]

\[ r - \text{dist}_k(p, q) = \max\{\text{dist}(p, q), k - \text{dist}(q)\} \quad (47) \]

where \( q \) is another data point in the k-nearest neighbors \( N_k(p) \). The local reachability density \( \text{ldr}(p) \) is the reciprocal of the average reachability density of given point \( p \) and its k-nearest neighbors. That is,
\[ \text{lof}(p) = \frac{1}{\sum_{q \in N_k(p)} \text{dist}_k(p, q) / |N_k(p)|} \]

where \(|N_k(p)|\) is the number of elements in the set. Finally, the local outlier factor lof\(\text{lof}(p)\) is defined as

\[ \text{lof}(p) = \frac{\sum_{q \in N_k(p)} \text{lof}(q) / |N_k(p)|}{\text{lof}(p)}, \]

which is the ratio of \(\text{lof}(p)\) to the average of \(\text{lof}(q)\) at point \(q\) belonging to the k-nearest neighbors of point \(p\).

Table 1. Calculation of LOF means and standard deviations according to the pressure and the number of samples.

<table>
<thead>
<tr>
<th>Pressure (# data)</th>
<th>(-10) (40)</th>
<th>(-10) (20)</th>
<th>(-20) (40)</th>
<th>(-30) (40)</th>
<th>(-40) (40)</th>
<th>10 (40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
</tr>
<tr>
<td>Std-dev</td>
<td>0.005</td>
<td>0.01</td>
<td>0.006</td>
<td>0.006</td>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Ku et al. 15

This was confirmed by the LOF calculations for multiple different pressures in Table 1. The mean values differ only in the fourth decimal place, and the standard deviation is determined by the number of data points, not the pressures. This suggests that adjusting the dataset by the difference \(p_g - p_0\) does not affect the performance of the LOF algorithm.

The manipulator module’s capability of touch sensing was tested using the data collected for \(-10\) kPa. The pressure of all three bellows was set to \(-10\) kPa, and we touched each bellow one by one and measured its pressure. The pressure response and the touch sensing results can be found in Figure 13(a). Three spikes in pressure level were detected (Figure 13(a)-left), corresponding to the three instances of touch (Figure 13(a)-right). We then applied different negative pressures to each bellow to generate bending and repeated the experiment. It was possible to detect touches even in the bending states, by detecting the pressure changes (Figure 13(b)). Finally, we conducted an experiment to detect contacts in the positive pressure region using the data collected from the negative pressure. The contact was detected even if different pressures were applied to the bellows (Figure 13(c)).

Since all bellows are mechanically connected, contact in one bellow affects the other bellows as well. Triboelectric touch sensing mechanism is an attractive approach in that it ensures decoupling of the sensing signals from the three bellows. The simplicity in both the structure and working mechanism of triboelectric sensing allows for independent sensing of the three actuators (See Figure S7).

To quantitatively evaluate the triboelectric sensing capability, we conducted an experiment using a film-type sample (See Section 3 for detail). Figure 13(d) and (e) show voltage generation when the sample was touched by copper and finger under a contact frequency of 5 Hz. More than 10 V of noticeable voltage was generated. Based on the results, we injected organogel into the bellow and measured the voltage generation caused by different external contacts (Figure 13(h)). The test was conducted without actuation, and the results are shown in Figure 13-(f). The measured intensity was decoupled, and touch was identified on each bellow, even in the case of multi-touch. The same experiment was also conducted while the manipulator module was actuated. When the configuration of each bellow changed by actuation (Figure 13(g)), the measured baseline intensity also changed, but the sensor was still able to distinguish touch and its position. Supplementary Movie S4 demonstrates the performance of touch sensing. From these results, we mapped different functions to each bellow that can operate through touch which will be discussed in Section 8.4.

8.3. Control of manipulator module

We controlled the manipulator module based on forward and inverse kinematics, using measurements from the string sensors, the IMU, and the depth measurement camera. We set goal positions in a 3D space and calculated the target length of each bellow based on inverse kinematics. We then implemented feedback control without calculating the input pressures corresponding to the target lengths of the actuators to reduce the computational load on the processor. Here, a controller with P and D gains of 0.010 and 0.5, respectively, was used to control the lengths of the bellows.

We first controlled the manipulator module using only a depth camera. The three bellows of the manipulator module are controlled to reach the target lengths in the actuator space (Figure 14(a)). The corresponding results in the task space show that the manipulator module can be controlled to reach the target positions (Figure 14(b)). Figure 14(c) shows the tracking results projected onto the \(x - y\) plane of the top center point of the manipulator module. These results validated the constant curvature model of our manipulator.

The manipulator module can also be controlled using the string sensors and the IMU. We tested the accuracy of the
control for repeated motions by programming the manipulator module to repeatedly move between points (0, 0, 162) and (0, 30, 140), using feedback control based on the measurements from the string sensor and the IMU. Figure 14(d) and (e) each show that the bellows reached the target lengths in the actuator space, and the target positions in the task space. The manipulator module was also simultaneously tracked by the depth camera for comparison. The measurement accuracy of the camera can be found in Figure S9. For the vision-based control, the steady-state error and its standard deviation were 2.9 mm and 1.3 mm, respectively. For the string sensor-based control, the steady-state error and its standard deviation reduced to 0.8 mm and 0.3 mm, respectively. Even in the steady state of the vision-based control, subtle oscillations were observed due to the deviations in the camera measurements, while the string sensor-based control remained stable. Therefore, we confirmed that the control based on the string sensor provides higher accuracy and stability than the vision-based method. However, the string sensor does have its drawbacks since it is exposed to the surrounding and it may experience signal changes upon contact, resulting in unintended motions. Additionally, the manipulator module and the robotic arm need the information on the surroundings to perform desired tasks, and this information cannot be obtained by the string sensors. We thus employed the method of vision-based localization and control not only to ensure the stability of the robot but also to acquire the information on the environment.

As was done for the vision-based control, we also tested the string sensor-based control for reaching various goal positions in 3D space, also the effects of the settling time. Two experiments were conducted with different time intervals (5 s and 0.05 s) between the goal positions. The longer interval time allowed the manipulator to reach the target positions with the higher accuracy (Figure 14(f) and (g)).

While the experiment discussed above only concerned a single manipulator module alone, additional loads or manipulator modules can be applied or connected. We mounted another manipulator module’s pneumatic parts (Figure 15(a)-left), weighing about 500 g, and tracked the motion of the entire arm. Compared to Figure 14(c),
Figure 15(a)-right shows noticeable difference between the goals and the measured positions when the load or another manipulator module was mounted or connected. This is because the constant curvature model did not consider the change in the kinetics due to the external load. The response of the module with an interval of 1 sec. interval between the goal positions can be seen in Supplementary Movie S5. Therefore, an algorithm correcting the position error was applied to compensate the inaccuracy of the constant curvature model in the presence of an external load. The error $e$ between the measured current position $p'$ and the target position $g_p$ was calculated, and it was determined whether the distance $||e||$ exceeded the threshold $k$ when $t_e > t_s$ (where $t_e$ is elapsed time and $t_s$ is settling time). Within the threshold, the current state of the module is maintained. Otherwise, the target position is adjusted to $g_p'$ by subtracting the product of the error $e$ and the decay rate $\gamma$ from the former target position $g_p$. The decay rate $\gamma$ was set as $0 < \gamma < 1$ to prevent oscillation. The whole process is provided in Algorithm 2. Figure 15(a)-right shows the motion of the module when this algorithm was applied. The manipulator module was given with seven goal positions, and the measurements were recorded with both the depth camera and the string strain sensors. Table 2 summarizes experimental results. In the table err denotes the position error just before the adjustment is applied, and err_adj indicates the position error after the adjustment is made in Figure 15(a)-right. The maximum position errors exceeded 6 mm without the adjustment, but with the adjustment, the robot was able to control its position to the goal positions with an error of less than 1 mm (Figure 15(b)). The coordinates in the time domain are shown in Figure 15(c), and we confirmed that the manipulator successfully reached the goal positions through continuous adjustments. We also tested translation of the manipulator from $(0, 0, 162)$ to $(30, 0, 140)$, with and without the adjustment algorithm (Figure 15(d)) in the vision-based control. Similarly, the final positions of the manipulator module were closer to the goal positions with the adjustment algorithm, reached within one or two adjustments. The insets show the maximum error of 10 mm in the x-position and 4 mm in the z-position.

**Algorithm 2 Goal position adjustment**

**Input:** decay rate $\gamma$, goal position $g_p$, adjusted goal position $g_p'$, error $e$, current position $p$, settling time $t_s$, elapsed time $t_e$, threshold $k$

$k \leftarrow 1$

while ($k < ||e||$) do
  if ($t_e > t_s$) then
    $e' = p' - g_p$
  if ($||e|| < k$) then
    Maintain current goal position
  else
    $g_p' = g_p - \gamma \cdot e'$
    PD control based on $g_p'$
  end if
end if

end while
Next, we assembled and tested a robotic arm formed by connecting two modules. Arbitrary points in a 3D space were given to the manipulator, and the trajectory was tracked using the motion capture system and the string sensors. The results in the $x/y$ plane can be seen in Figure 15(e). The control results by the adjustment algorithm based on the manipulator’s own coordinate system are shown in Supplementary Movie S6. The corresponding measurements in a 3D space are shown in Figure 15(f).

Table 2. Comparison results with and without applying the adjustment (unit: mm).

<table>
<thead>
<tr>
<th>Goal position $(x_g, y_g, z_g)$</th>
<th>err</th>
<th>err$_{adj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (30, 0, 140)</td>
<td>6.6</td>
<td>0.9</td>
</tr>
<tr>
<td>2 (0, 30, 140)</td>
<td>3.9</td>
<td>0.9</td>
</tr>
<tr>
<td>3 (-30, 0, 140)</td>
<td>4.3</td>
<td>0.9</td>
</tr>
<tr>
<td>4 (0, -30, 140)</td>
<td>5.3</td>
<td>0.9</td>
</tr>
<tr>
<td>1 (30, 0, 140)</td>
<td>3.9</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Figure 15. Control results of loaded manipulator: (a) Manipulator module with pneumatic components (left) and control result using the string sensors with the adjustment algorithm (right). (b) Tracked positions near the goals 1 and 2 in (a). The position of the manipulator module followed the arrows by the adjustment and satisfied the positional error criteria at the end. (c) Coordinates of the manipulator module in the time domain. (d) Comparison of the two control methods. (e) Experimental results of reaching the goal positions tracked by the motion capture system and the string sensors and (f) corresponding results in the 3D space. (g) Soft robotic arm made of two manipulator modules: i Initial state, ii. bending of both modules with the same curvature, iii. bending of the upper module only, and vi. bending of both modules in opposite directions.
assembled soft robotic arm is shown in Figure 15(g)-(i) through Figure 15(g)-vi as well as in Supplementary Movie S7. The mean position error of 14 randomly chosen goal positions calculated from the string sensors was 0.78 mm, and the standard deviation was 0.31 mm. The maximum position error was 0.98 mm, which was smaller than the termination condition \( k = 1 \), confirming that the algorithm worked on the manipulator’s own coordinate system. From the measurements of the motion capture system, the mean position error was 1.94 mm, the standard deviation was 0.96 mm, and the maximum error was 3.70 mm. We also tested the ability of the manipulator to reach the goal positions with contacts during actuation, and the result can be seen in Figure S10.

8.4. Application

We tested the proposed manipulator for safe human interactions (Figure 16(a) and (b)). The robotic arm is composed of two manipulator modules and a gripper. We set the human touch signals as triggers to perform three actions: one for closing and opening of the gripper, another for contraction of the upper manipulator module, and other for bending of the entire arm. The string sensor-based localization method was used to control the manipulator. For the bellows numbered 1, 2, and 3 on the upper module (Figure 16(a)), the touch sequence 3-2-1-3 (i.e., gripping an object, contracting the upper manipulator module, bending forward both modules in the forward direction, and releasing the object) was performed. Images of the assembled soft robotic arm during the sequence are shown in Figure 16(b). In the intensity plots, the threshold level was different for each bellow since the baseline intensity changed according to the operation state of the manipulator module, and so the threshold level for each actuator was automatically adjusted after each touch. The process of this actuation is demonstrated in Supplementary Movie S8.

![Figure 16](image-url)
By customizing the end-effector and mapping the specific tasks, this application can be further extended for multi-touch or continuous-touch sequence with specific intervals, which will allow the soft robotic arm to perform more complex tasks.

9. Discussion

In this study, we proposed the design of bellow-type compliant actuators for building a modularized soft robotic arm and methods of localizing and controlling the arm. Although the proposed system showed the required functionality, there is still room for further improvement.

First, the payload of the arm is relatively low due to the compliance of the material. This can be solved by using a stiffer material to fabricate the bellow actuators. One candidate material that is also compatible with the SLA 3D-printer used in this research (Form3, Formlabs), is Flexible 80A that presents higher stiffness and shore hardness but still with flexibility and elasticity (Formlabs, 2020). However, this increased stiffness of the manipulator also requires modification to the tactile sensing channels for higher sensitivity to touch as well as a greater range in the input pressure to achieve the same workspace. Also, each module consisting the manipulator can be fabricated with different structural stiffness. The module near the base that needs to bear the weights of the other module as well as the object for manipulation can be made stiffer than the module near the end effector that bears less payload.

Second, the string sensors may degrade the control performance of the arm since they are externally exposed and susceptible to contacts made to the string sensors directly. By moving them to inside the bellow, we can prevent unwanted perturbations to the sensors and consequently expect more reliable and stable physical interactions between the robot and humans.

Third, although our kinematic model assumes that the effective cross-sectional area of the bellow, $A_{\text{eff}}$ is constant, it may actually change due to the complex deformation of the soft structure, necessitating adoption of a pressure-dependent value of $A_{\text{eff}}$ to further elaborate the kinematic model. This would also allow the model to depend less on the approximated adjustment function, implemented in equation (3) of Section 4, and thereby increase the robustness and the accuracy of the model.

Lastly, the TENG tactile sensors are an effective method to detect human touches. However, they are responsive only to physical contacts but not to any objects in proximity. It is sometimes useful to recognize objects or humans in proximity that approach to the robot. One possible solution is to use the organogel channels as a proximity sensor by measuring the change in capacitance (Navarro et al., 2021; Kwon et al., 2022). In this way, the robot will be able to more proactively respond to any possible contacts or collisions in advance.

10. Conclusions

In this study, we first proposed a bellow-type soft actuator with integrated tactile sensing to build a human-interactive soft robotic arm. An SLA 3D-printing technique enabled easy fabrication of the complex geometry of the actuator. This structure contains a pneumatic chamber for actuation, and embedded channels filled with organogel for detecting human contacts.

We also proposed a modular design of manipulator module composed of three actuators, soft string sensors, and custom-built pneumatic and electronic control hardware. This modular design enabled not only independent operation of a single module but also easy assembly of multiple modules to form a soft robotic arm. Since the arm use the identical modules, the same deep learning model can be applied to each module, saving the training cost and time.

To localize the robotic arm, we first developed and integrated a compact and lightweight omnidirectional soft string strain sensor, that provided information on multidirectional bending and linear motions of the arm. We also implemented a method to localize the robotic arm using a depth measurement camera. An object detection model with deep learning was applied to enhance the performance of this vision-based control.

Finally, we demonstrated control of the soft robotic arm with a total mass of more than 2.5 kg and performed different tasks with human interactions.

We believe the proposed system showed a potential in adopting robots for assistance and interactions in our daily life in the future.

Acknowledgments

The authors would like to thank Dr. Gyowook Shin for his technical support for setting up the simulation environment and Mr. Jaewoong Jung for his suggestions for the soft manipulator control strategy.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by the National Research Foundation (NRF) grant (RS-2023-00208052 and 2021R1A2C2093790), in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant (2021-0-00896), both funded by the Korean Government (MSIT), and in part by SNU-Hojeon Garment Smart Factory Research Center funded by Hojeon Ltd.

ORCID iDs

Subyeong Ku https://orcid.org/0000-0001-5184-3073
Byung-Hyun Song https://orcid.org/0000-0001-5369-8981
Supplemental Material

Supplemental material for this article is available online.

References


Anver HM, Mutlu R and Alici G (2017) 3D printing of a thin-wall soft and monolithic gripper using fused filament fabrication. 2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), Munich, Germany, 03–07 July 2017


Maddikunta PKR, Pham QV, Prabadevi B, et al. (2022) Industry 5.0: a survey on enabling technologies and potential applications. Journal of Industrial Information Integration 26: 100257.


