

# Deformation Capture via Soft and Stretchable Sensor Arrays

OLIVER GLAUSER, ETH Zurich

DANIELE PANIZZO, New York University

OTMAR HILLIGES, ETH Zurich

OLGA SORKINE-HORNUNG, ETH Zurich



Fig. 1. Left to right: We propose a method for the fabrication of soft and stretchable silicone based capacitive sensor arrays. The sensor provides dense stretch measurements that, together with a data-driven prior, allow for the capture of surface deformations in real-time and without the need for line-of-sight.

We propose a hardware and software pipeline to fabricate flexible wearable sensors and use them to capture deformations without line of sight. Our first contribution is a low-cost fabrication pipeline to embed multiple aligned conductive layers with complex geometries into silicone compounds. Overlapping conductive areas from separate layers form local capacitors that measure dense area changes. Contrary to existing fabrication methods, the proposed technique only requires hardware that is readily available in modern fablabs. While area measurements alone are not enough to reconstruct the full 3D deformation of a surface, they become sufficient when paired with a data-driven prior. A novel semi-automatic tracking algorithm, based on an elastic surface geometry deformation, allows to capture ground-truth data with an optical mocap system, even under heavy occlusions or partially unobservable markers. The resulting dataset is used to train a regressor based on deep neural networks, directly mapping the area readings to global positions of surface vertices. We demonstrate the flexibility and accuracy of the proposed hardware and software in a series of controlled experiments, and design a prototype of wearable wrist, elbow and biceps sensors, which do not require line-of-sight and can be worn below regular clothing.

## 1 INTRODUCTION

Motion capture is an essential tool in many graphics applications, such as character animation for movies and games, sports, biomechanics, VR, and AR. Most commonly, motion capture systems are camera based, either relying on body-worn markers or more recently markerless. Vision based approaches can be highly accurate and in the case of multiview or depth imaging, they can provide dense surface reconstructions. However, such systems rely on extensive infrastructure and are therefore mostly confined to lab and studio use. Other sensing modalities, such as body-worn inertial and magnetic sensors, or resistive and capacitive distance sensors have been explored to provide more mobility, yet these are typically limited to capturing skeletal deformation only.

We introduce a new, practical and affordable approach to deformation sensing and motion capture. Our approach bridges the gap between vision-based and inertial approaches by providing accurate sensing of dense surface deformations while being wearable, and hence practical for scenarios in which stationary cameras are unsuited, for example to capture muscle bulging below clothing.

*Capacitive sensor array.* We propose to leverage a capacitive sensor array, fabricated entirely from soft and stretchable silicone, that is capable of reconstructing its *own* deformations. The sensor array provides dense measurements of area change, which can be leveraged to reconstruct the underlying 3D surface deformation without requiring line-of-sight (see Fig. 2). We furthermore contribute a data driven surface reconstruction technique, allowing for the capture of non-rigid deformations even in challenging conditions, such as under heavy occlusion, at night, outdoors, or for the acquisition of uncommon deformable objects.

Conductive polymers have been leveraged to fabricate resistive bend sensors [Bächer et al. 2016; Rendl et al. 2012], and are the basis of soft capacitive distance sensors, which are now readily available commercially [Par 2018; Str 2018]. Such stretchable capacitive sensors are enticing, since they are thin, durable, and may be embedded in clothing or directly worn on the body. However, so far fabrication has been involved and required specialized equipment, driving up cost. Moreover, such sensors have not been demonstrated to be accurate enough for motion capture and are typically limited to measurement of uniaxial deformation. Please note that capacitive sensing is often considered synonymous with touch sensing [Grosse-Puppendahl et al. 2017; Lee et al. 1985; Rekimoto 2002], in which capacitive coupling effects are leveraged to detect finger contact with a static sensor. In this paper, however, the term is used

in a different sense, referring to the fact that capacitance changes when an electrode undergoes deformations.

*Custom fabrication method.* We introduce a fabrication method for soft and stretchable capacitive deformation sensors, consisting of multiple bonded layers of conductive and non-conductive silicone. Crucially, the method only requires casting silicone and etching conductive traces by a standard laser cutter, and can thus be performed using hardware commonly available in a modern fabrication lab. The precision and accuracy of our sensors is comparable to commercial solutions, and the involved material costs are low. Our approach supports embedding many sensor cells of custom shape in a single thin film. Each cell measures changes of its own area, caused by deformation of the surface it is attached to. The resulting sensor array can be read out at interactive rates.

*Geometric prior.* While providing a rich signal, the area measurements alone are not sufficient to uniquely reconstruct the full 3D sensor shape due to isotropy and lack of direct bending measurement. They are however sufficient when paired with an appropriate geometric prior, if expected deformations involve some amount of non-area preserving stretch.

In addition to the hardware, we propose an effective pipeline to acquire the deformation of the sensor worn by a user, for example wrapped around the wrist or an elbow. We propose a data driven technique based on a neural network regressor to reconstruct the sensor geometry from area measurements. At runtime, the regressor estimates the location of a sparse set of vertices, and the dense deformed surface is computed by a nonlinear elastic deformation method, obtaining a high-resolution reconstruction in real-time (see Fig. 1).

To acquire the necessary training data, we overcome an additional challenge: optical tracking systems struggle with the heavy occlusions and large deformations typical for natural motions of wrists, elbows and other multi-axial joints. Furthermore, when capturing other non-rigidly deforming objects, skeletal priors cannot be leveraged to recover missing markers. We thus introduce a semi-automatic ground truth acquisition technique, enabling capture of the necessary training data in minutes and reducing tedious manual cleanup to a minimum. The approach leverages an elastic simulation of the sensor to disambiguate the marker tracks, deal with unlabeled markers and correctly attribute marker positions to the digital mesh model of the sensor.

*Evaluation.* We demonstrate our sensors in action by acquiring dense deformations of a wrist and lower part of the hand (see Fig. 1), an elbow, an inflating balloon, and muscle bulging. We also capture deformations of flat sensors, both in and out of plane, which shows the precision and localization properties of our capacitive sensor arrays. Finally, we evaluate the prediction accuracy of the learning based prior quantitatively.

## 2 RELATED WORK

Our work relates to several areas of the literature ranging from digital fabrication to motion capture and self-sensing input devices. We briefly review the most important work in these areas.



Fig. 2. An elbow “hidden” below by a jacket. Top: Video frames for comparison. Bottom: With our approach the dense surface deformation is estimated without requiring line of sight.

*Camera-based motion capture.* The acquisition of articulated human motion using cameras is widely used in graphics and other application domains. Commercial solutions require wearing marker suits or gloves and depend on multiple calibrated cameras mounted in the environment. To overcome these constraints, research has proposed marker-less approaches using multiple cameras (cf. [Moeslund et al. 2006]); sometimes these rely on offline [Ballan et al. 2012; Bregler and Malik 1998; Starck and Hilton 2003] and more recently online processing [de Aguiar et al. 2008; Elhayek et al. 2017; Rhodin et al. 2015; Stoll et al. 2011], but always require fixed camera installations. Neumann et al. [2013] capture muscle deformations of a human shoulder and arm with a multi-camera system and derive a data-driven statistical model.

Recent pose estimation methods exploit deep convolutional networks for body-part detection in single, fully unconstrained images [Chen and Yuille 2014; Newell et al. 2016; Tompson et al. 2014; Toshev and Szegedy 2014; Wei et al. 2016]. However, these methods only capture 2D skeletal information. Predicting 3D poses directly from 2D RGB images has been demonstrated using offline methods [Bogo et al. 2016; Tekin et al. 2016; Zhou et al. 2016] and in online settings [Mehta et al. 2017]. Monocular *depth* cameras provide additional information and have been shown to aid robust skeletal tracking [Ganapathi et al. 2012; Ma and Wu 2014; Shotton et al. 2013; Taylor et al. 2016, 2012] and enable dense surface reconstruction even under deformation [Dou et al. 2016; Newcombe et al. 2015; Zollhöfer et al. 2014]. Multiple, specialized structured light scanners can be used to capture high-fidelity dense surface reconstructions of humans [Pons-Moll et al. 2015].

All vision-based approaches struggle with visual clutter, (self-)occlusions and difficult lighting conditions, such as bright sunshine in the case of depth cameras, high contrast or lack of illumination in the case of color cameras. Furthermore, all camera based systems require line-of-sight and often precise calibration, and are therefore not well suited in many scenarios, such as outdoors. Our sensor is a first step in removing these limitations, allowing mobile and self-contained sensing, without line of sight.

*Self-sensing input devices.* An important feature of our method is the capability of measuring the sensor’s own deformation without requiring any external cameras. Such *self-sensing* input devices, usually not designed for motion capture, have been first demonstrated

in the Gummi system [Schwesig et al. 2004], which simulated a handheld, flexible display via two resistive pressure sensors. Other early work used the ShapeTape sensor [Danisch et al. 1999] for input into a 3D modeling application [Balakrishnan et al. 1999]. Metallic strain gauges embedded into flexible 3D printed 1D strips measure the bending and flexing of custom input devices [Chien et al. 2015]. Rendl et al. [2014] use eight transparent printed electrodes on a transparent and flexible 2D display overlay to reconstruct 2.5D bending and flexing of the sheet in real time, but do not allow for stretch. [Bächer et al. 2016] propose an optimization based algorithm to design self-sensing input devices by embedding piezo-resistive polymer traces into flexible 3D printed objects. [Sarwar et al. 2017] use polyacrylamide electrodes embedded in silicone to produce a flexible, transparent 4×4 sensing grid, and [Xu et al. 2016] propose a PDMS based capacitive array; both are limited to detecting touch gestures. Hall effect sensors embedded into hot-pluggable and modular joints can measure joint angles of tangible input devices used for character animation [Glauser et al. 2016; Jacobson et al. 2014]. While demonstrating the rich interactive possibilities afforded by flexible input devices, none of the above approaches are directly suitable for the acquisition of dense non-rigid surface deformation.

*Inertial measurement units (IMUs).* Attaching sensors directly onto the body overcomes the need for line-of-sight and enables use without infrastructure. IMUs are the most prominent type of sensors used for pose estimation. Commercial systems rely on 17 or more IMUs, which fully constrain the pose space, to attain accurate skeletal reconstructions via inverse kinematics [Roetenberg et al. 2007]. Good performance can be achieved with fewer sensors by exploiting data-driven methods [Liu et al. 2011; Schwarz et al. 2009; Tautges et al. 2011] or taking temporal consistency into account, albeit at high computational cost and therefore requiring offline processing [von Marcard et al. 2017]. While IMUs provide mobility and accuracy, they cannot sense dense surface deformations.

*Strain gauges, stretch and bend sensors.* Strain sensors fabricated from stretchable silicone and attached directly to the skin have been proposed to measure rotation angles of individual joints [Lee et al. 2016]. Shyr et al. [2014] propose a textile strain sensor, made from elastic conductive yarn, to acquire bending angles of elbow and knee movements. [Mattmann et al. 2008] and [Lorusso et al. 2004] use strain gauges embedded into garments to classify discrete body postures. [Scilingo et al. 2003] propose polymerized fabric strain sensors and demonstrate use of the sensor in a data glove. Specifically designed for the capture of wrist motion, [Huang et al. 2017] use five dielectric elastomer sensors and achieve an accuracy of 5° for all motion components, highlighting the difficulty of reconstructing joint orientation of complex, multi-axial joints such as the wrist, shoulder or ankle. Bending information can be used to recover articulated skeletal motion, and resistive bend sensors are typically used in VR data gloves. However, these suffer from hysteresis [Bächer et al. 2016]; imprecise placement and sensor slippage can impact accuracy [Kessler et al. 1995]. A soft bend sensor that is insensitive to stretching and mountable directly on the user’s skin is proposed in [Shen et al. 2016], increasing angular accuracy, but it is inherently limited to measuring uni-axial bending.

We propose a wearable, soft and stretchable silicone-based capacitive sensor design, focused on measuring dense area changes, which allows us, in combination with a data-driven reconstruction technique, to accurately capture dense, articulated and non-rigid deformations.

*Fabrication.* Producing capacitive elastomer stretching sensors is challenging, and the mechanical, electrical and thermal properties all depend on the type of material used and the pattern of conductive traces or electrodes. Another challenge is that the silicone is hydrophobic, hence the adhesion of non-silicones is extremely difficult. For an extensive review of various ways to manufacture conductive layers for such sensors or actuators, we refer to [Rosset and Shea 2013]. Composites of carbon black (conductive powder) and silicone are widely used, see e.g. [Araromi et al. 2015; Huang et al. 2017; O’Brien et al. 2014; Rosset et al. 2016]. A large range of fabrication methods for manufacturing conductive trace patterns have been proposed. Most methods rely on the potentially costly fabrication of intermediate tools like screen printing masks [Jeong and Lim 2016; Wessely et al. 2016], molds [Huang et al. 2017; Sarwar et al. 2017] or stencils [Rosset et al. 2016]. To circumvent the adhesion issue, specialized plasma chambers are often required to selectively pre-treat the base layer [Jin et al. 2017]. An alternative procedure, introduced by [Lu et al. 2014], involves patterning conductive PDMS sheets, manually removing excess parts with tweezers, sealing the resulting circuit with PDMS and bonding multiple such circuit layers to form capacitive touch sensors (as demonstrated by [Weigel et al. 2015]). Similar to [Araromi et al. 2015], our process leverages a standard laser cutter to etch away the negative sensor pattern, opening up the possibility to digitally design electrode patterns and produce them with low error tolerance. However, in contrast to prior work, our fabrication method does not require a plasma chamber or manual alignment and gluing of the different layers. Hence it allows for the production of larger sensors with a high alignment quality (see Fig. 8). To the best of our knowledge, we are the first to propose a fabrication method that requires almost no specialized hardware and enables creating large high-resolution multi-layer sensor arrays.

*Capacitive (touch) sensing.* Ever since the introduction of the Theremin [Glinsky 2000], an experimental musical instrument, researchers have explored the use of capacitive sensing in the context of HCI. Most notably, capacitive coupling effects are the basis of early [Beck and Stumpe 1973; Lee et al. 1985] and virtually all modern touchscreen devices [Rekimoto 2002]. Capacitive coupling effects exist naturally between many objects (including humans) and their surroundings, and by measuring the changes in relative values it is possible to recover relative position, proximity and other properties. The seminal works by Smith [1995] and Zimmermann et al. [1995] introduced and categorized the various electric field sensing aspects to the interaction research community and demonstrated applications that went well beyond binary touch detection. Since then capacitive coupling effects have been used to sense touch, detect and discriminate user grip and grasp, detect and track objects on interactive surfaces, track 3D positions and proximity and coarsely classify 3D poses and gestures. We refer to the survey by Grosse-Puppeldahl et al. [2017] for an exhaustive treatment. Notably, flexible and bendable sensors [Gotsch et al. 2016; Han et al.

2014; Poupyrev et al. 2016] and those directly worn on the user’s skin [Kao et al. 2016; Nittala et al. 2018; Weigel et al. 2015] have been proposed. However, virtually all of the above work measures one or a combination of different capacitive coupling effects, that is, the change in capacitance due to a conductive object (such as a finger) approaching an electrode. Our work is fundamentally different in that we do not sense capacitive coupling effects but instead measure changes in the electrodes’ properties themselves: under deformation, the area of the electrode’s plates changes, which in turn changes the capacitance of the plate and hence the charge time of the capacitor. We show how this effect can be leveraged to recover, using appropriate geometric priors, detailed 3D surface deformations, albeit at the cost of requiring a custom read-out scheme.

### 3 OVERVIEW

We present a stretchable silicone elastomer based sensor and its corresponding fabrication procedure. The sensor senses its *own deformation* and estimates the local surface area changes during deformation when wrapped around an object or a body part of interest (e.g., a wrist). The sensor array is fabricated layer onto layer entirely from 2-component silicone elastomer with conductive elements made from the same silicone but mixed with carbon black particles. The conductive layers can be designed to contain custom electrode patterns via etching with a standard laser cutter. This approach avoids the production of masks or molds and makes interlayer alignment very straightforward and precise.

As a further contribution we introduce a silicone-based capacitive area sensor *array*, whereas prior work only demonstrated individual stretch sensing elements, and arrays only to detect dense touch or pressure (e.g., [Block and Bergbreiter 2013; Engel et al. 2006; Lipomi et al. 2011; Nittala et al. 2018; Ponce Wong et al. 2012; Sarwar et al. 2017; Wissman et al. 2013; Woo et al. 2014]).

Our key insight is that such arrays could also be used to attain dense localized area changes, given an appropriate read-out scheme. Our arrays are made by placing electrode strips in two conductive layers, separated by a dielectric, together forming a non-uniform grid of capacitors. Furthermore, we propose a scanning based read-out scheme that does not require individually connected capacitors, which would require a large number of layers or a large portion of the sensor area dedicated to connection leads. Instead, we propose a time-multiplexing procedure to indirectly read out capacitance values, which allows for a drastically simplified routing of electric connections. By integrating all the capacitance readings, we can acquire area changes with a sufficient granularity and accuracy to reconstruct the geometry of an object, given suitable geometric priors. These dense area measurements are therefore combined with a deep learning based regressor to attain 3D position estimates of key points on the surface and an elastic deformation optimization to obtain dense deformation reconstructions.

In the following sections we provide a brief primer on capacitive sensing (Sec. 4.1), detail our sensor design (Sec. 4.2) and detail the fabrication (Sec. 4.3). We then complete our method by introducing our data capture and cleanup, learning and surface reconstruction approaches (Sec. 5).

## 4 SENSOR DESIGN

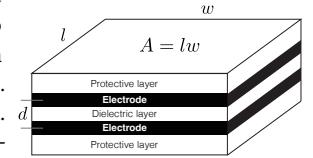
### 4.1 Preliminaries

The capacitance  $C$  (in Farads) of a plate capacitor is given by

$$C = \epsilon_r \epsilon_0 \frac{A}{d} = \epsilon_r \epsilon_0 \frac{lw}{d}, \quad (1)$$

where  $A$  is the area of overlap of the two electrodes (in square meters),  $\epsilon_r$  is the dielectric constant,  $\epsilon_0$  is the electric constant and  $d$  is the separation between the plates (in meters). Assuming a rectangular plate capacitor,  $l$  is its length and  $w$  the width. While originally derived for static plate capacitors, this relationship also holds for capacitors made from silicone elastomers [Atalay et al. 2017; Huang et al. 2017; O’Brien et al. 2014]. To minimize capacitive coupling effects with other objects, capacitors are typically shielded via insulating layers (see inset). Using Eq. (1), and assuming the same Poisson ratio of width and thickness of the sensor ( $d/d^0 = w/w^0$ ), a linear relationship between the ratio of the stretched capacitor’s length  $l$  to the rest pose length  $l^0$ , and the ratio of the capacitance of the stretched capacitor  $C$  to the rest pose capacitance  $C^0$  can be established:

$$\frac{C}{C^0} = \frac{\epsilon_r \epsilon_0 \frac{lw}{d}}{\epsilon_r \epsilon_0 \frac{l^0 w^0}{d^0}} = \frac{l}{l^0} \frac{w}{w^0} \frac{d^0}{d} = \frac{l}{l^0}. \quad (2)$$



Prior work applies this principle to the design of capacitive, uniaxial stretch sensors [Atalay et al. 2017] by continuously measuring a capacitance, which is then transformed to length measurement using Eq. (2). Note that here, an assumption is made that stretch only happens along  $l$ , which typically requires fabricating isolated, individual capacitors (Fig. 3a). Our aim is to create a dense array of sensing elements, for which stretch may occur in multiple directions and hence each sensing element captures changes in area.

*Area changes.* Starting from Eq. (1), and assuming volume conservation ( $V = V^0 \Leftrightarrow A d = A^0 d^0 \Leftrightarrow d^0/d = A/A^0$ ) and constant stretch throughout the entire sensor cell, the ratio of capacitance before and after deformation can be expressed as

$$\frac{C}{C^0} = \frac{\epsilon_r \epsilon_0 \frac{A}{d}}{\epsilon_r \epsilon_0 \frac{A^0}{d^0}} = \frac{A}{A^0} \frac{d^0}{d} = \left( \frac{A}{A^0} \right)^2. \quad (3)$$

Thus, if we know the current capacitance  $C$  of a sensor cell and have recorded its rest pose area  $A^0$  and capacitance  $C^0$ , we can compute the change in area between the rest state and the current configuration as

$$\frac{A}{A^0} = \sqrt{\frac{C}{C^0}}. \quad (4)$$

*Touch vs. pressure vs. stretch.* We note that there are fundamental differences between capacitive sensing of touch, pressure, and stretch. The majority of the HCI literature on capacitive sensing measure capacitive coupling effects (e.g., changes in capacitance due to an approaching finger). Applied pressure can be measured capacitively since the thickness  $d$  is reduced, which leads to a higher capacitance  $C$  (see Eq. (1)). Finally, in our work, both the overlap

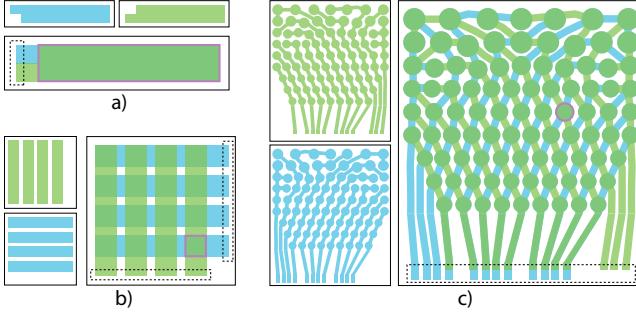


Fig. 3. Various electrode strip patterns, with the bottom layer in blue and the top layer in green. When overlaid, the overlapping regions form *sensor cells*; we highlight one cell in each example in pink. The dashed lines outline the places where the read-out circuit is connected. Example (a) is a classic elastomer strain sensor with 2 leads and 1 sensor cell; (b) is our array concept with 8 leads and 16 sensor cells; (c) depicts our actual prototype sensor, a warped grid that brings all connection leads to the bottom side of the sensor, with 24 leads and 92 sensor cells.

area  $A$  and the thickness  $d$  change due to the deformation of the sensor, requiring a custom read-out scheme (cf. Fig. 5). We now explain how a naive implementation, designed for touch or pressure sensing, must be modified in order to capacitively sense deformation.

#### 4.2 Sensor layout

Dense surface deformation capture requires a sensor that can measure local changes in the surface geometry with high density. This need has to be balanced with the complexity of the electrical design, so that the fabrication remains feasible. Our proposed concept of the sensor array (Fig. 3b), which we call simply *sensor* from now on, strikes this balance with its two-electrode-layers design. The sensor is made of two conductive layers with  $n$  and  $k$  independent electrode strips on each layer, respectively. We call the individual electrodes *strips*, but they may have any shape. Overlapping sections of two electrode strips from separate layers form a local capacitor, which we call a *sensor cell*  $S$ . We lay out the strips in a non-uniform grid arrangement, as shown in Fig. 3c. Each pair of strips from top and bottom layers crosses at most once, amounting to  $s$  sensor cells ( $s \leq kn$ ). This design allows routing all strips to the same side of the sensor, where the silicone-based traces are connected to a PCB for the measurement of capacitances (Fig. 4). However, since sensor cells are daisy-chained, we cannot directly read each one independently. We now derive a read-out scheme that provides the desired localized area measurements.

*Sensor read-out.* As mentioned, our sensor is designed to consist of only two capacitive layers, which renders individual addressing of capacitors difficult without sacrificing sensor surface for complex routing of electrical traces. We experimentally verified that simple scanning schemes common in mutual capacitive touchscreens cannot be applied in the case of geometrically deforming and overlapping capacitor plates and traces, see Fig. 5.



Fig. 4. Left: Our prototype sensor with connector boards. Both conductive layers contain 12 electrode strips each, and the overlaps amount to 92 sensor cells. Right: Using silicone glue, the topology of a flat sensor can be changed to form e.g. a cylinder. See Fig. 14 for our second and larger fabricated sensor.

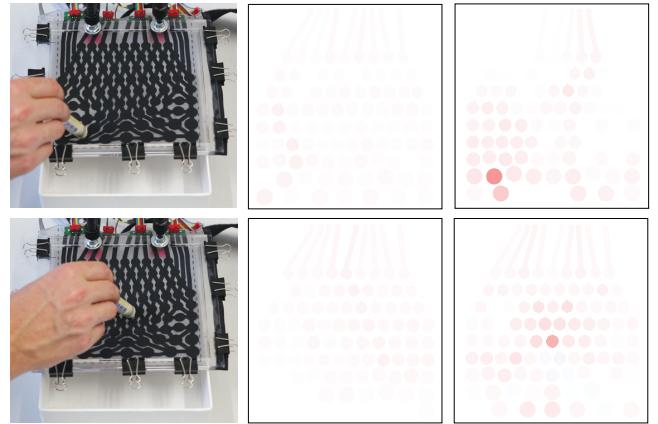


Fig. 5. A naive scanning scheme (mutual-capacitance approach, using charging time to measure capacitance) results in underestimation of the magnitude of stretch, leads to not well-localized measurements, and even gives incorrect readings. Left: Sensor is deformed by poking with a pen. Middle: Change of magnitude per sensor cell, measured by the naive scanning scheme. Right: Change of magnitude per sensor cell, measured by our proposed scheme (see the respective video clip in supplemental material).

We propose a time-multiplexing scheme, in which a voltage is applied to a subset of strips from both layers in turn, and the remaining strips are connected and serve as the second plate of the local capacitor. A simple example of a sensor composed of a  $3 \times 2$  grid of electrode strips, with a total of  $s = kn = 6$  sensor cells, is shown in Fig. 6. For each such measurement, the cells where the combined electrode strips overlap are measured in parallel. The capacitances of these cells add up, leading to a linear relationship between the individual sensor cell capacitances and the measured, combined capacitance. This can be expressed in matrix form:

$$\mathbf{M} \mathbf{C}_c = \mathbf{C}_m .$$

Here,  $\mathbf{M}$  is an  $s \times s$  binary matrix with rows encoding different measurement combinations, so that  $\mathbf{M}$  transforms the vector of sensor cell capacitances  $\mathbf{C}_c$  into the measured capacitances  $\mathbf{C}_m$ . Using our example in Fig. 6 to illustrate the composition of this linear system of equations, the vector  $\mathbf{C}_c$  is

$$\mathbf{C}_c = [C_{1A}, C_{2A}, C_{1B}, C_{2B}, C_{1\Gamma}, C_{2\Gamma}]^\top, \quad (5)$$

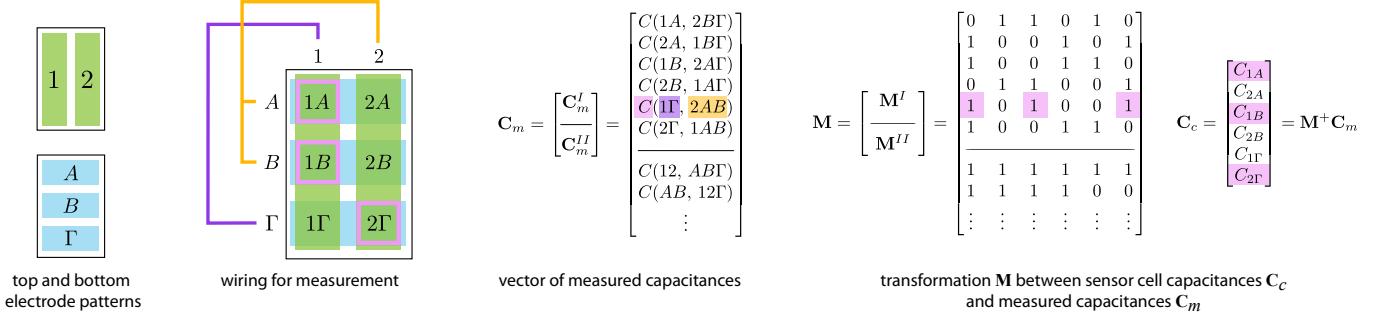


Fig. 6. Measuring capacitance of sensor cells via selective combinations of strips. The measured combination in this example is comprised of strips 1 and  $\Gamma$  as the source electrode, and strips 2, A and B as the ground electrode. The resulting overlaps are highlighted in pink. The measurement contributes the equation  $C(1\Gamma, 2AB) = C_{1A} + C_{1B} + C_{2\Gamma}$  to the linear system that recovers the individual sensor cell capacitances.

where  $C_{1A}$  denotes the sought localized capacitance of sensor cell 1A, and so on. Each row of  $\mathbf{M}$  corresponds to a measurement, where the row elements corresponding to jointly read sensor cells are set to 1 and the remaining elements to 0. In our example (Fig. 6), the highlighted row of  $\mathbf{M}$  corresponds to a measurement where electrodes 1 and  $\Gamma$  are connected to serve as the source electrode, and 2, A, B as the ground electrode. This leads to cells 1A, 1B and 2 $\Gamma$  to form parallel capacitors, and the read-out values are summed.

To reconstruct  $\mathbf{C}_c$  from measurements  $\mathbf{C}_m$ , the matrix  $\mathbf{M}$  needs to be invertible, which is the case if it has  $s$  linearly independent rows. The matrix  $\mathbf{M}^I$  is formed by iteratively connecting one strip from the top and bottom layer as source electrode, with all remaining strips connected as the ground electrode, resulting in the required  $s$  linearly independent rows. We experimentally found that taking additional measurements with all remaining combinations of strips, collected in matrix  $\mathbf{M}^{II}$ , and solving the resulting over-constrained linear system in the least square sense leads to extra robustness:

$$\mathbf{C}_c = \mathbf{M}^+ \mathbf{C}_m. \quad (6)$$

Here,

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^I \\ \hline \mathbf{M}^{II} \end{bmatrix}, \quad \mathbf{C}_m = \begin{bmatrix} \mathbf{C}_m^I \\ \hline \mathbf{C}_m^{II} \end{bmatrix}, \quad (7)$$

where  $\mathbf{C}_m^I, \mathbf{C}_m^{II}$  represent the capacitance readings of the mandatory part  $\mathbf{M}^I$  and the additional measurements  $\mathbf{M}^{II}$ , respectively.

*Non-uniform stretch.* Since our sensor cells have non negligible size (Fig. 4), the uniform stretch assumption may not hold in practice. We therefore model a sensor cell  $S_j$  more accurately by splitting it into several elements (triangles)  $e_i \in S_j$ , each with an individual (uniform) area stretch. Applying Eq. (3) to each element, the capacitance  $C_j$  of the sensor cell becomes

$$\frac{C_j}{C_j^0} = \frac{1}{C_j^0} \sum_{e_i \in S_j} \left( C_j^0 \frac{A_i^0}{A_j^0} \right) \left( \frac{A_i}{A_i^0} \right)^2 = \frac{1}{A_j^0} \sum_{e_i \in S_j} \frac{A_i^2}{A_i^0}, \quad (8)$$

where  $C_i^0 = C_j^0 A_i^0 / A_j^0$  is the rest pose capacitance of element  $e_i$ . This holds because in rest state, the thickness  $d$  is constant, and hence the rest state capacitance is proportional to the area  $A_i^0$ .

### 4.3 Fabrication

We propose a fabrication pipeline, illustrated in Fig. 7, for silicone-based sensors with arbitrarily shaped electrodes.

*Structure.* The sensor consists of two conductive layers with a dielectric layer between them, and it is encased by shielding layers (see inset on the previous page). During fabrication the sensor rests on a flat glass plate to which the silicone elastomer sticks well but the final sensor can be easily detached. We provide the description of the chemical composition of the silicone mixtures in Appendix A. The layers are cast one by one by spreading the silicone using a blade; the correct thickness is ensured by Kapton tape (65  $\mu\text{m}$  thickness) at the borders of the glass plate. After the casting of each layer the sensor is cured for 20 minutes in an oven at 100 °C.

The second, conductive layer (silicone mixed with carbon black) is directly cast onto the shielding layer, and after curing, the desired pattern is etched with a laser cutter. The etching is done with a 100 Watt Trotec Speedy 360 laser cutter. Two rounds of etching are carried out with the following settings: 20 Power, 60 Speed and 500 Pulses/inch. This vaporizes the carbon black to create non-conductive areas between traces, while the underlying silicone-only layer stays intact. The resulting dust can be carefully removed with isopropyl alcohol without damaging the electrodes. The sensor is completed by adding another dielectric, the second capacitive layer (which is also etched and cleaned) and finally another shielding layer. The overall process takes around 3.5 hours (1 h for mixing and casting, 1.5 h for curing and 1 h for laser etching) for producing a sensor of 200×200 mm.

In previous works [Araromi et al. 2015; Lu et al. 2014], the alignment of the different layers of a multilayer sensor had to be done manually. Aligning the layers with high accuracy and without wrinkles can prove a difficult task, especially for larger sensors like ours. With our approach, a high alignment quality is achieved by design, since we directly cast layers onto one another (see the accompanying video from 01:05) and place the base glass plate in the laser cutter aligned with physical stoppers before etching. Fig. 8 shows an alignment experiment.

The thickness of the final sensor is about 500  $\mu\text{m}$ , the conductive

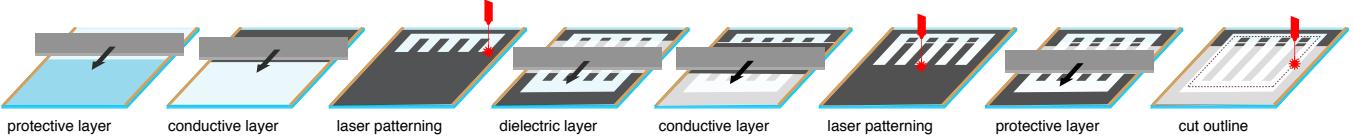


Fig. 7. The proposed fabrication pipeline consists of eight main steps. From left to right: Casting a protective layer; casting a conductive layer; etching the negative electrode strip pattern with a laser cutter; dielectric layer; conductive layer; etching again; protective layer; cutting the desired outline.

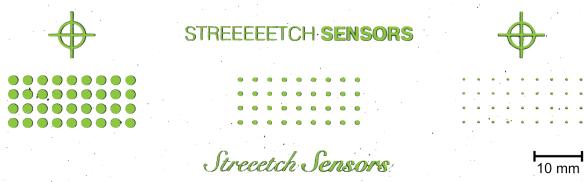


Fig. 8. To demonstrate the alignment quality of our fabrication method, we produced a test pattern with two identical conductive (black) layers. The fabricated pattern was scanned with a flatbed scanner. The scan is overlaid with the digital design (green). Wherever the alignment is perfect, only the green layer is visible.

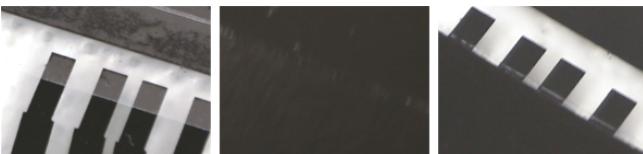
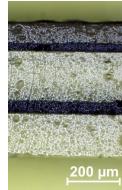


Fig. 9. Left: sensor after casting the dielectric layer, the connector pads are covered by transparent sticky tape. Middle: after casting the second conductive layer. Right: after removal of the sticky tape (before curing in the oven); the connector pads stay exposed.

layers are 45  $\mu\text{m}$  thick each (for the basic protective layer we use 4 layers of offset tape, and for the dielectric layer 2 layers of offset tape). The inset on the right shows a cross section of the sensor layers under a microscope. The sheet resistance of a conductive layer is in the order of 1 kOhm (four-point probe). The stiffness (Young's Modulus) of the pure layered RTV is  $729.6 \pm 13.4$  kPa, with two embedded conductive layers  $979.6 \pm 16.6$  kPa (calculated from three samples each with the setup and method as described in [Hopf et al. 2016]).



**Connectors.** The electrode strips must be connected to our electronic boards for measurement (see Appendix B for details). During fabrication we cover the connectors with sticky tape before casting the remaining layers. The tape is removed before curing the corresponding layer, re-exposing the connectors, see Fig. 9.

**Finalization.** The sensor is cut to the desired outline shape with the laser cutter. The resulting sensor is then pulled off the glass plate, and silicone adhesive can be optionally used to close the sensor to form, for example, a cylinder (Fig. 4) to wrap a wrist or an elbow.

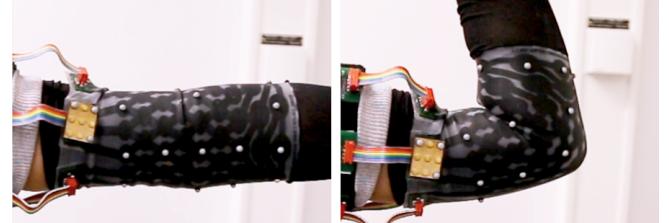


Fig. 10. Our sensor on an elbow. Left: rest pose; right: close to fully bent.

## 5 SURFACE DEFORMATION RECONSTRUCTION

Our sensor is equipped with simple rest state geometry, represented by a triangle mesh  $\mathcal{S} = (\mathcal{V}, \mathcal{F})$ , where  $\mathcal{V}$  is the set of 3D vertex positions and  $\mathcal{F}$  is the connectivity (the set of faces). The connectivity  $\mathcal{F}$  comes from meshing the electrode layout (Sec. 4.2): we represent each sensor cell  $S_j$  with a fan of triangles and mesh the overall layout using Delaunay triangulation using [Shewchuk 1996]. We set the rest state geometry  $\mathcal{V}^0$  to the canonical shape corresponding to the chosen topology: e.g., for the sensor in Fig. 4 (right), we use a circular cylinder of dimensions corresponding to the intrinsic size of the produced sensor. As the sensor is pulled onto a deforming object and capacitance changes are measured, the goal is to reconstruct the deformed geometry  $\mathcal{V}(t)$  for each frame  $t$ , given the measured capacitances  $C_j(t)$  of all sensor cells  $S_j$ .

Through the relation of capacitance to area (Eq. (8)), our sensor provides rich, localized area change measurements at interactive frame rates, but areas alone are not sufficient to define the shape of a general deforming surface in 3D, since area is an intrinsic quantity. We therefore pair these measurements with a data-driven geometric prior, acquired by simultaneously capturing the deformation of the object of interest using our sensor and an optical tracking system, and then training a regressor that maps the capacitance measurements to marker vertex positions.

To this end, we define a sparse set of vertex indices  $\mathcal{M}$  and attach reflective markers onto the corresponding physical locations. To simplify the marker attachment process, the set  $\mathcal{M}$  is a subset of the mesh vertices corresponding to centers of circular sensor cells. The set is chosen to obtain a regular coverage of the cylindrical sensor, allowing a maximal distance of 5 centimeters in-between the individual markers. For all experiments we used a single, fixed marker pattern per sensor layout. Placing the sensor onto the object of interest (Fig. 10), we simultaneously record sensor readings and 3D marker positions tracked by an 8-camera OptiTrack setup [Opt 2018]. Untreated silicone is highly specular, but we found that a matte finish can be attained by densely etching the outer layer on

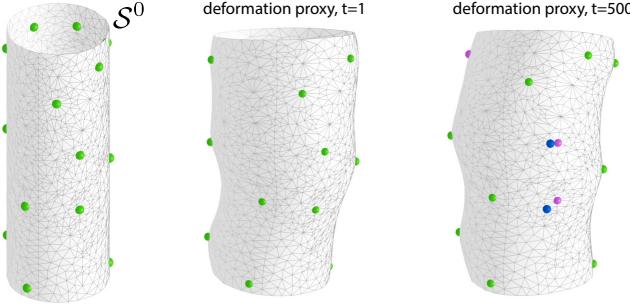


Fig. 11. Left: The rest state sensor mesh  $S^0 = (\mathcal{V}^0, \mathcal{F})$  with marker vertices  $\mathcal{M}$  in green. Middle:  $S^0$  deformed by the marker positions of the first frame in a wrist capture session. Right: the labeled markers in green, two unlabeled marker observations in blue and the two candidate matches in pink; the mesh geometry is estimated by elastically deforming  $S^0$  using the green markers as positional constraints.

the laser cutter (with 60 Power, 100 Speed, and 500 Pulses/inch). The captured and processed data for each frame  $t$  consists of:

- Coordinate frame transformation  $T(t) \in \mathbb{R}^{3 \times 4}$  (a  $3 \times 3$  rotation and a translation, recovered from 3 designated markers);
- Marker positions  $p_i(t) \in \mathbb{R}^3$  w.r.t. the local frame, for each marker vertex  $i \in \mathcal{M}$ ;
- A vector  $C_c(t)$  of capacitance values of all sensor cells, obtained as described in Sec. 4.2.

This data is used to train a regressor  $g_\theta(C_c(t))$  that maps sensor cell capacitance values to marker vertex position estimates  $\hat{p}$ . Given  $g$ , we can employ the sensor at run-time and use the marker positions predicted by  $g$  as positional constraints that guide the deformation of the sensor mesh  $S$ .

### 5.1 Capturing and processing training data

A fundamental challenge with marker based approaches are incorrectly labeled or lost markers, an issue exacerbated in settings like ours, where heavy occlusions and strong non-rigid deformations are combined with the lack of a simple skeletal prior. Fig. 12 provides an illustrative example of tracking 12 wrist-mounted physical markers. The OptiTrack system outputs 165 individual marker observations due to frequent tracking failures (sequence length is 1.5 minutes). This problem quickly becomes unwieldy; in capturing real data we encountered more than 500 marker labels in a dataset of 17000 frames (3 minutes) of 21 physical markers.

Manual cleanup, label merging, and correct attribution would require hours of manual labor and make the acquisition of our deformation prior impractical. We therefore employ a novel semi-automatic marker cleanup and labeling pipeline.

The mocap system outputs a set of marker labels  $\mathcal{I} = \{1, 2, \dots, N\}$ , and for each frame  $t$ , a binary indicator that tells whether the marker was visible in that frame. For each frame  $t$  where marker  $j$  is visible, the system also outputs its 3D position  $p_j(t) \in \mathbb{R}^3$ . We seek an assignment of marker vertices  $i \in \mathcal{M}$  to tracked marker labels  $j \in \mathcal{I}$ , providing a 3D position in each frame  $t$ . Our main insight is to employ a state-of-the-art elastic deformation technique to create a proxy deformation of  $S^0$ , using reliably labeled marker vertices

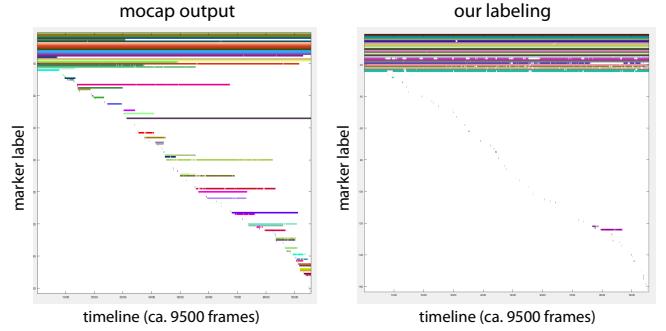


Fig. 12. Marker labeling. For each individual label, we plot horizontal bars spanning the frames where it is visible. Left: Captured markers directly from the mocap system. There are 165 individual labels due to periods of occlusion and subsequent failure to pick up the track, despite the actual number of markers being only 12. Right: sanitized and relabeled markers using our semiautomatic approach. A minority of outliers remain in a few frames; they are discarded from the dataset.

as positional constraints. This allows us to match each unlabeled marker to its closest marker vertex on the proxy.

*Initialization.* Usually the number of tracked labels  $N$  is much larger than the actual number of physical markers, because some markers are temporarily lost, and are then given a new label when they re-enter. We initialize the assignment of marker vertex indices by picking 3 tracked markers in the first frame and manually matching them with their corresponding mesh vertices in our rest pose mesh  $S^0$ . We then rigidly transform  $S^0$  to align it with the tracked data (i.e., put it in the same coordinate system) by solving the Procrustes problem.

We then assign a 3D position to all remaining marker vertices of the mesh by searching for the closest tracked marker position in this frame. This way we obtain  $|\mathcal{M}|$  pairings between marker labels and mesh vertex indices, as typically in the first frame (rest pose) all markers are visible.

*Labeling.* We sort the unassigned tracked markers in chronological order according to the first frame they are visible at.

For each unassigned marker  $j^*$  and for each frame  $t$  where  $j^*$  is visible, we elastically deform  $S^0$  to match the captured geometry in  $t$  by imposing the marker vertices in  $\mathcal{M}$  that already have matched marker positions in frame  $t$  as positional constraints. The output is a set of deformed “proxy” meshes, one for each such frame, which we use to find a match for  $j^*$ . For robustness, we pick the mesh vertex whose average  $L^2$  distance over all frames is the smallest. We accept the match only if this distance is below a threshold  $\tau$  (25 mm in our experiments), otherwise  $j^*$  is marked as an outlier.

Every successful labeling provides an extra positional constraint for the deformations, improving the quality of the proxy (and thus the success rate) for subsequent labeling passes. In our implementation, we use the deformation optimization method by Wang et al. [2015], a state-of-the-art nonlinear elastic deformation technique that expects solely sparse positional constraints as input.

As a post-processing step, we visually inspect the produced assignments via 3D renderings and plots of  $x, y, z$  coordinates over

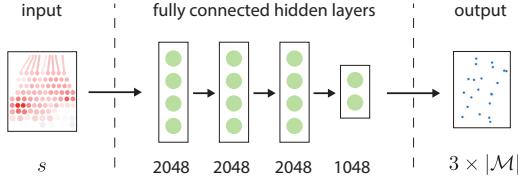


Fig. 13. To train a sensor with  $s$  sensor cells and  $|\mathcal{M}|$  markers, our network takes  $s$  capacitance readings as input and outputs  $|\mathcal{M}|$  vertex position estimates, through three fully connected layers with 2048 units each and one fully connected layer with 1024 units. E.g. for our sensor in Fig. 4 there are 92 inputs and 63 ( $3 \times 21$ ) outputs.

time, to detect incorrect merges. If any are present, we can separate them and rerun the labeling algorithm again. One iteration of this procedure was sufficient for most of our capture sessions.

Our MATLAB implementation takes below 15 minutes per session, allowing us to have a 3 minutes long captured session cleaned in around 10 minutes. Note that we are not guaranteed to find observed 3D positions for each marker vertex of our mesh in each and every frame  $t$ , due to occlusions, outliers and possible failures of our assignment heuristic. We thus discard frames with unassigned markers, which are around 20 % in our acquisition sessions. We encountered one case where too many markers were missing in some frames due to heavy occlusions in the folded elbow, which hampered the regressor training due to insufficient data. We resorted to synthetic 3D data for those frames, taking missing marker vertex positions from the deformation proxy.

## 5.2 Regressor training

We wish to recover dense surface deformations in real time.

To this end, we learn a function  $g_\theta(\mathbf{C}_c)$ , parametrized by a deep neural network, that maps from sensor cell capacitances  $\mathbf{C}_c \in \mathbb{R}^s$  to marker positions  $\hat{\mathbf{p}} \in \mathbb{R}^{3 \times |\mathcal{M}|}$  (in a local frame). We have experimentally verified that nonlinear function approximators such as the fully connected multi-layered neural network used here, perform better than linear models due to the nonlinearities in the mapping from area change to capacitance (Table 1).

Our network architecture, depicted in Fig. 13, takes  $s$  sensor cell capacitance readings as inputs of a linear layer, followed by three fully connected layers with 2048 units each and one fully connected layer with 1024 units. A final linear output layer predicts the marker vertex positions  $\hat{\mathbf{p}}$ . The input and all hidden layers are followed by a ReLu activation function and a BatchNorm layer. Given a training set  $\mathcal{D} = \{(\mathbf{C}_c^i, \mathbf{p}^i)\}$  of  $K$  vectorized ground truth input-output pairs, we perform training via a weight-regularized  $L^2$  loss:

$$\mathcal{L}_{\text{reg}} = \sum_{i=1}^K \|g(\mathbf{C}_c^i) - \mathbf{p}^i\|_2^2 + \lambda \|\theta\|_2^2, \quad (9)$$

where  $\theta$  are the model parameters and  $\lambda$  is a regularization factor.

We implement the network using pyTorch [Paszke et al. 2017] and train it with the ADAM optimizer with a learning rate of  $10^{-4}$ , mini-batch size of 256, regularization  $\lambda = 10^{-5}$  and default values for all other parameters [Kingma and Ba 2014]. All inputs are normalized to be zero-mean unit variance.

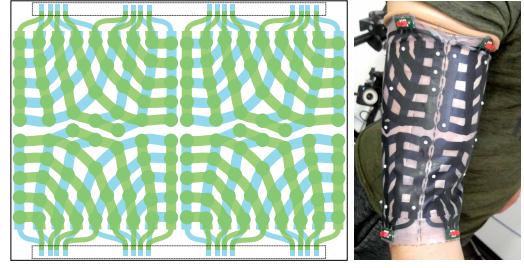


Fig. 14. We fabricated a second larger sensor (300x250 mm) with 144 sensor cells and connectors on two sides. Left: The sensor layout consists of four identical sub-sensors that can be read out in parallel. Right: The produced sensor, glued to form a cylinder and worn on a biceps.

## 5.3 Capturing dense surface deformation at runtime

Once the neural network is trained and the regressor  $g_\theta$  is available, we can deploy our sensor standalone, uncoupled from the optical tracking and estimate the dense surface deformation of an object without line-of-sight. This is illustrated in Fig. 2, where the sensor is worn underneath clothing, rendering vision based approaches infeasible. The regressor provides 3D positions  $\hat{\mathbf{p}} = g_\theta(\mathbf{C}_c)$  of the marker vertices  $\mathcal{M}$  given current sensor measurements  $\mathbf{C}_c$ . We note that the network is able to compensate for inaccuracies in area estimates from capacitive readings (see Fig. 19), which in particular occur under extreme stretch (see Sec. 6.2). To reconstruct the current surface deformation, we deform the rest state mesh  $\mathcal{S}^0$  using the method proposed by Wang et al. [2015], where the marker vertices  $\hat{\mathbf{p}}$  again serve as positional constraints.

## 6 EXPERIMENTS AND RESULTS

To demonstrate the utility of our proposed approach, we evaluate its components in an ablative manner. First, we quantitatively assess the sensor concept and the corresponding fabrication method (Sec. 6.1) and then demonstrate the applications in reconstruction of surface deformations, both qualitatively and quantitatively (Sec. 6.2). Our experiments are performed with two sensor layouts, shown in Figures 4 and 14. The layouts are manually designed, non-uniform grids, with all strips routed to the same side of the sensor, where they are connected to a connector PCB. The first layout is used both in its flat form and as a cylinder.

### 6.1 Sensor characterization

*Distance sensor comparison.* We verify the accuracy of our sensors by fabricating a uni-axial sensor with the same dimensions (15×50 mm) as a commercially available Parker Hannifin industrial sensor [Par 2018]. We stretch both sensors (with a motorized linear stage, see inset) to various lengths and directly compare the readings. The average relative error of the two sensors is comparable (Fig. 15), with a slight but non-significant edge for the Parker Hanafin sensor (0.0085) over ours (0.0096). Overall, we conclude that the accuracy of our



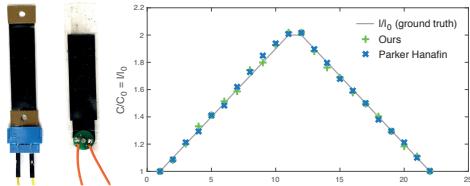


Fig. 15. Left: An industrial sensor by Parker Hanafin and a sensor of the same dimension fabricated by us. Right: Comparison of their accuracy.

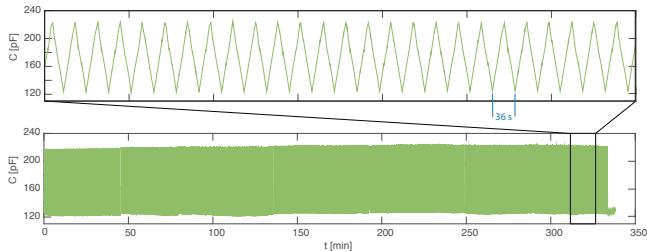


Fig. 16. The uni-axial sensor response stays constant during a cyclic stretch (2x) test of 5 hours and 30 minutes (about 550 cycles).

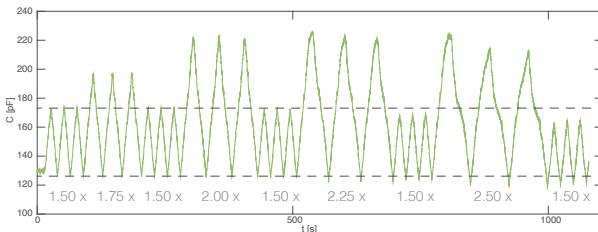


Fig. 17. After a stretch factor of 2.25x, the sensor response when stretched by a factor of 1.5x has changed compared to the first three rounds.

measurements is high and comparable to commercial solutions. We note that there was no observable hysteresis in our experiments.

*Longterm sensor behavior.* In a second set of experiments, we evaluate whether and how the sensor response changes under longterm cyclic stretch and large stretch. For the longterm experiment, the uni-axial sensor is pre-stretched a few times and then continuously stretched and relaxed for 5 h 30 min by a factor of 2x. The sensor response stays constant (see Fig. 16). The maximally allowed stretch before (internal) material damage occurs is found by stretching the sensor a few times to a baseline factor of 1.5x, increasing the maximum stretch factor in each round (see Fig. 17). These experiments show that our fabricated sensor can be stretched without noticeable internal damage by 100% (2x) for at least 5 h 30 min. In our experiments, this stretch factor was never surpassed when capturing body parts.

*2D stretch localization.* To assess the localization capabilities of our sensor layout, we perform a simple experiment, in which we fix a flat sensor to a frame and poke it in different locations. Eq. (4) states that the sensor cells' capacitance changes directly relate

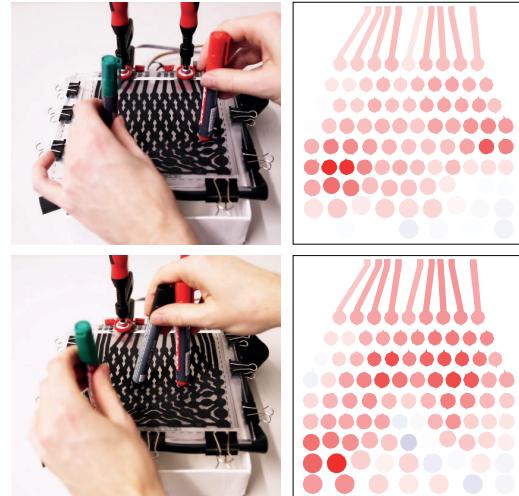


Fig. 18. Left: the sensor is fixed to a frame and poked with pens. Right: Area change magnitude measured per sensor cell.

to area changes. The proposed readout scheme (cf. Sec. 4.2) allows us to measure and localize stretch. Fig. 18 visualizes two example frames extracted from the video in the supplemental material. This capability could be explored in other application scenarios, including detection of touch and pressure.

*2D stretch quantification.* To better understand the accuracy of recovered stretch measurements, we attach clips on strings to a flat sensor, so that we can apply spatially varying tension forces by selectively pulling on the strings. Additionally, we place reflective markers on the sensor, so that we can estimate the actual stretch per sensor cell. Fig. 19 visualizes the results. We report an average relative error of 7.7% when comparing the measured capacitance ratio  $C_c/C_c^0$  with the theoretical capacitance ratio calculated by Eq. (8) per sensor cell from the tracked areas. This error is likely due to our approximate sensor model, which neglects the influence of the (changing) resistance of the electrodes. Close inspection of Fig. 19 reveals that this effect is negligible for our purposes.

## 6.2 Surface deformation capture

*Predictor comparison.* To validate our design choice of parameterizing the regression problem of Eq. (9) with a neural network, we perform a comparison with several alternative models as baseline. Table 1 summarizes the results of a three-way comparison with linear regression and non-linear SVM using an RBF kernel. The neural network achieves the lowest mean and max errors and produces the lowest standard deviation across all datasets used in our experiments.

*Non-skeletal 3D deformation.* To demonstrate the deformation capture abilities of our sensor, we use it to measure the shape of a balloon that is aperiodically inflated (up to a maximum diameter of about 120 mm) and deflated. Despite the apparent simplicity of the setup, the deformation is freeform, and it is not possible to rely on standard geometric priors, such as a skeleton. We captured a

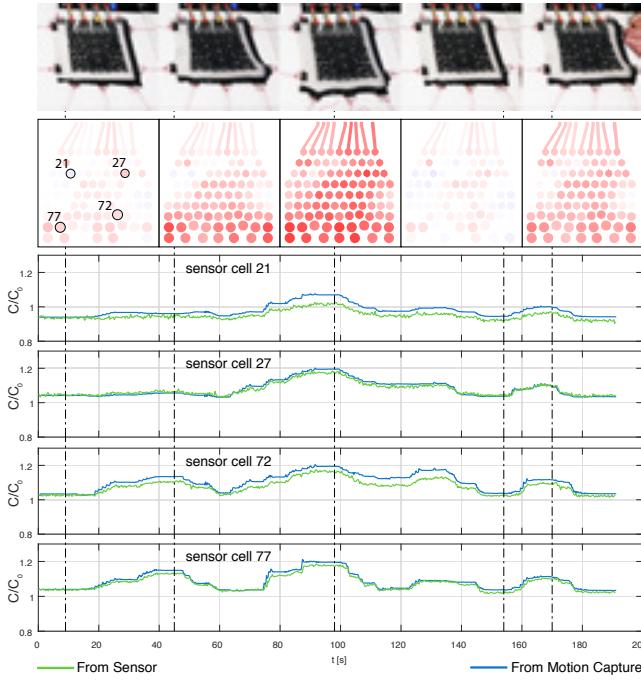


Fig. 19. The sensor is dynamically stretched by selectively pulling on the strings it's attached to. Top: A set of sample frames. Middle: Stretch intensity per cell at the sample frames. Bottom: The relative capacitance of four selected sensor cells over time, comparing ground truth (estimated through mocap markers) in blue and the capacitance change recorded by our sensor in green. The dashed vertical lines show the locations of the sample frames on the timeline.

Table 1. Comparison of prediction accuracy of the chosen DNN regressor (*ours*) with a linear regression model (LR) and a non-linear support vector machine with an RBF kernel (SVM). All errors are in millimeters, lower is better.

Marker error		mean	std	max
Balloon	LR	3.59	1.90	12.84
	SVM	3.22	2.73	25.05
	<i>ours</i>	<b>2.75</b>	<b>1.86</b>	<b>12.85</b>
Biceps	LR	7.64	5.06	53.00
	SVM	6.86	5.18	52.24
	<i>ours</i>	<b>3.85</b>	<b>2.39</b>	<b>25.81</b>
Elbow	LR	7.65	3.31	39.95
	SVM	6.73	5.26	59.79
	<i>ours</i>	<b>3.46</b>	<b>2.48</b>	<b>30.82</b>
Wrist	LR	12.8	4.99	71.89
	SVM	4.36	2.71	44.12
	<i>ours</i>	<b>3.51</b>	<b>2.14</b>	<b>27.22</b>
Forearm	LR	10.64	3.94	52.03
	SVM	4.38	2.30	32.11
	<i>ours</i>	<b>4.02</b>	<b>2.66</b>	<b>38.52</b>

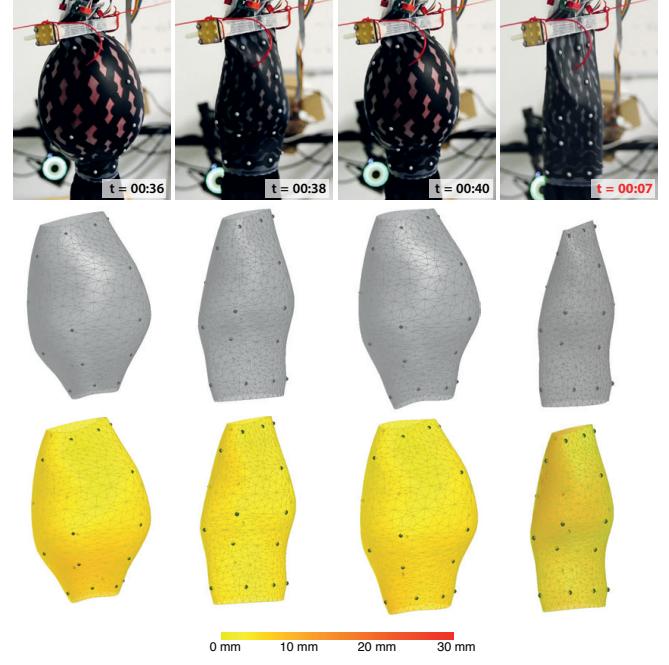


Fig. 20. Four frames of a 1:40 min long balloon capture session. Top: Video frames for comparison. Middle: Mocap ground truth. Bottom: Reconstruction based on the sensor measurements and the trained prior. The rightmost frame corresponds to the frame with the largest individual marker error.

5-minute session with the mocap system (2451 frames), and used the cleaned data to train a regressor (Sec. 5.2). To validate the system, we recorded an additional 1:40 min sequence (946 frames). The errors between our regressor and the mocap output are small, 2.75 mm on average, with a maximum of 12.85 mm (Fig. 20, rightmost column). Note that the maximal resolution of our mocap system, which is used as ground truth for these measurements, is 0.2 mm. Fig. 20 shows four frames extracted from the video in the supplemental material.

As a non-skeletal body part example, we captured a biceps muscle of ca. 36 cm in circumference being flexed, together with a small part of the elbow, using a larger sensor (see Fig. 14). We captured a 6-minute training session with the mocap system (2305 frames) and an additional 2 min test sequence (1224 frames). We report an average marker error of 3.85 mm, with a maximum of 25.81 mm (Fig. 21, rightmost column). Fig. 21 shows four frames (extracted from the video in the supplemental material).

*Uni-axial deformation.* We wrap our sensor around an elbow to capture its movement. This is a challenging scenario due to the strong occlusions when the elbow is fully bent and due to the local non-rigid surface deformation. We use 12 minutes of training data (5369 frames) and a 2-minute test sequence (1329 frames). Our sensor accurately matches the test sequence (Fig. 22) and enables deformation sensing even when worn below clothing (Fig. 2). In this example, the mean error is 3.46 mm and max error is 30.82 mm. In

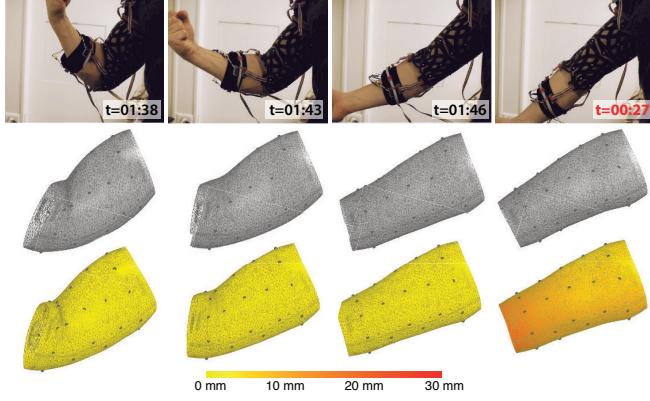


Fig. 21. Four frames of a 2-minute long biceps capture session. Top: Video frames for comparison. Middle: Mocap ground truth. Bottom: Reconstruction based on the sensor measurements and the trained prior. The rightmost frame corresponds to the frame with the largest individual marker error.

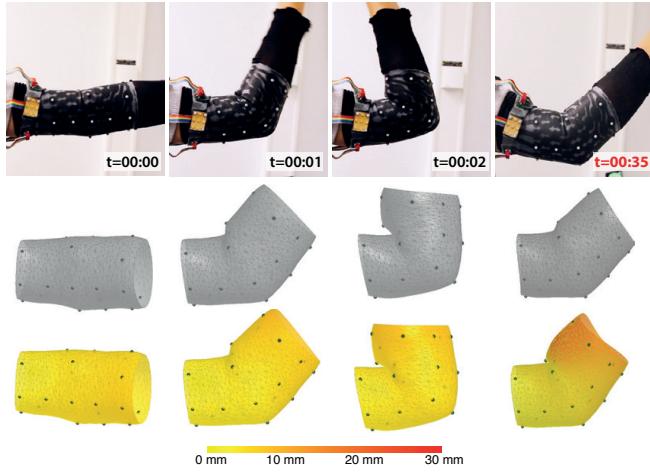


Fig. 22. Four frames of an elbow capture session. Top: Video frames for comparison. Middle: Mocap ground truth. Bottom: Reconstruction based on the sensor measurements and the trained prior. The rightmost frame corresponds to the largest individual marker error.

Fig. 22 we show four frames extracted from the full video sequence (attached in the supplemental material).

**Multi-axial deformation.** Our sensor successfully reconstructs very challenging scenarios, such as a wrist movement containing both a multi-axial skeletal deformation and volume changes when the fingers are splayed. For the wrist example, we trained on a 15-minute session (8799 frames), and tested on a 2:45 minutes session (1774 frames). Even in this case, the errors are low, with a mean of 3.51 mm and max error of 27.22 mm (see Fig. 23).

**Twisting motions.** The sensor also manages to capture the twisting motion of a forearm. For this example the model is trained on a 8-minute session (1846 frames), and evaluated on a 2 minutes session (1320 frames). For such a scenario the errors are slightly higher with

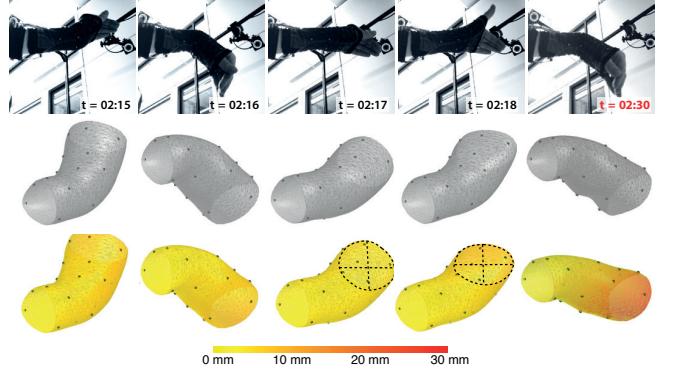


Fig. 23. Four frames from a wrist capture session. Top row: Video frames for comparison. Middle: Mocap ground truth. Bottom: Reconstruction based on the sensor measurements and our trained prior. In the third and fourth frames, note how our sensor correctly senses its shape when the fingers are splayed. The frame corresponding to the largest individual marker error is shown on the right.

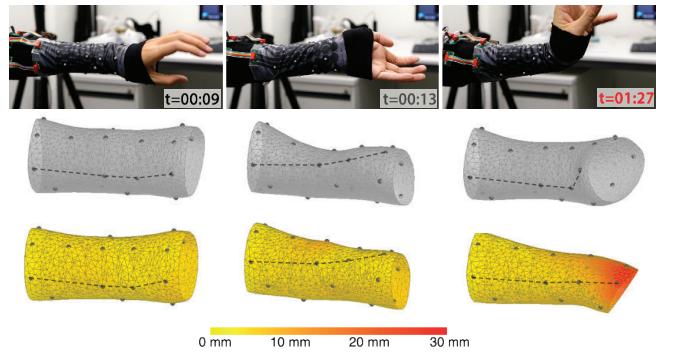


Fig. 24. Three frames from the forearm capture session. Top row: Video frames for comparison. Middle: Mocap ground truth. Bottom: Reconstruction based on the sensor measurements and our trained prior. Be aware that our sensor is only able to capture local stretch occurring below the sensor. The frame with the highest individual error is shown on the right: The sensor fails to correctly predict the bending of the wrist.

a mean error of 4.02 mm and max error of 38.53 mm, (see Fig. 24). The peak in error corresponds to predictions of the markers on the hand when the wrist is fully bent, see Fig. 24 on the left.

**Interpolation behavior.** To demonstrate the robustness of our predictor in test situations with strains deviating from the training data, we artificially reduce the training data of the wrist example, while keeping the test set fixed. We only keep training frames where the angle  $\alpha$  between the arm and the palm is  $\alpha < \gamma$  or  $\alpha > \beta$  ( $\alpha$  is the angle between a line connecting two markers on the arm and another line connecting two markers on the back of the hand). Table 2 shows the remaining number of training frames and the resulting mean and maximum error for a selection of angular limits. The first block (where frames with large angles are removed) shows that the network does not extrapolate well. Note that this is to be

Table 2. Predictor accuracy of the wrist test example with artificially reduced training data. It shows the ability of handling strains in the test data not previously seen during training. The training is reduced to frames with  $\alpha < \gamma$  or  $\alpha > \beta$ , where  $\alpha$  is the angle between the arm and the palm and  $\gamma, \beta$  are angular limits.

$\gamma$	$\beta$	#frames	mean	max
$\alpha_{\min}$	$\alpha_{\max}$	8799	3.51	27.22
60	$\alpha_{\max}$	8668	3.14	28.40
40	$\alpha_{\max}$	7335	3.40	49.20
30	$\alpha_{\max}$	5777	4.07	50.55
20	$\alpha_{\max}$	3229	6.59	76.96
20	30	6251	3.35	26.45
20	40	4693	3.89	31.31
$\alpha_{\min}$	20	5570	3.41	35.76
$\alpha_{\min}$	30	3022	4.67	47.76
$\alpha_{\min}$	40	1464	7.38	52.50

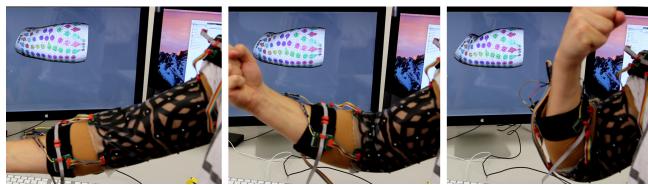


Fig. 25. Three frames from a live capturing session of the biceps.

expected, since most machine learning approaches do not generalize well to situations where the training and test data statistics differ significantly. However, as shown in the middle and the lower block, the method manages to interpolate well, even though there are now training samples at shallow angles. This holds true as long as the training set is large enough. The last row of Table 2 shows the results of exceeding this lower limit in terms of training data size.

**Real-time reconstruction.** To demonstrate the real-time capabilities of our approach, we have implemented a live system in which a user may wear the sensor, and we deform a cylindrical (in rest pose) mesh at interactive rates (approximately 8 Hz). See Figures 1, 25, and the accompanying video for the results. Note that in this setting, the users wear the sensor long after the training data was acquired; when taking the sensor off and putting it on again, one only needs to make sure that the alignment of the sensor and the body part is approximately the same. For the wrist example we quantitatively evaluated this effect of taking the sensor off and putting it on again with an imperfect alignment. For a 2-minute test sequence, the in-session mean error is 4.06 mm (max: 38.28 mm) while the out-of-session mean error is 6.80 mm (max: 47.22 mm).

## 7 LIMITATIONS AND CONCLUDING REMARKS

We proposed a soft and stretchable capacitive sensor array that allows measuring localized area changes. When paired with a learned geometric prior, it can reconstruct complex deformations without line-of-sight.

Our fabrication method and sensor layout open the door to multiple exciting future work venues. The most obvious is combining our area sensor with bend sensors to measure both extrinsic and intrinsic surface geometry, to e.g., also capture isometries. Furthermore it would be compelling to find a way to capture distance changes in such a dense array setting. These extensions would allow to estimate the deformation of general surfaces (like clothing) even if there's no non-area preserving stretching or twisting occurring. Another practical addition would be an assisting mechanism for correct placement of the sensor on the measured object: at present, we simply take a photograph before the training session and peruse it when putting the sensor on again for live session capture.

The acquisition of a large dataset of training sequences with multiple users is necessary to generalize our approach to multiple users, skipping the per-user training session. As with other sensing modalities (e.g., EMG, EEG), additional research into solving the cross-session problem may be required in this setting. Furthermore, the computational design of sensor layouts that are optimized for a specific set of deformations is also an interesting challenge that would directly benefit from the flexibility and simplicity of our fabrication pipeline. Finally, more complex sensor (3D) geometries such as data gloves appointed with our sensor array would enable a number of compelling use cases, such as reconstructing fine-grained hand shape in real-time, sidestepping the various issues (occlusions, lighting) associated with other sensing modalities.

We note that we employ a sparse set of markers as our *ground truth*, and effectively reconstruct this set from our sensor readings. Ideally we would like to have densely captured 3D geometry for training, and match it to denser sensor readings. As discussed in Sec. 2, spatially and temporally dense 3D capture is highly challenging and currently invariably involves some degree of model fitting. A realistic simulator that generates large quantities of high-quality synthetic data could be an alternative. It would be interesting to develop a denser version of our sensor design for more direct, dense geometry measurements. This comes with its own challenges, such as properly housing the electronic boards and a time multiplexing strategy to keep the read-out frame rates interactive; we leave this as future work.

## ACKNOWLEDGMENT

We would like to thank Denis Butscher, Christine de St. Aubin, Raoul Hopf, Manuel Kaufmann, Roi Poranne, Samuel Rosset, Michael Rabinovich, Herbert Shea, Rafael Wampfler, Yifan Wang, Wilhelm Woigk, Shihao Wu and Ji Xiabon for their assistance in the fabrication, with the experiments and for insightful discussions and Seonwook Park, Velko Vechev and Katja Wolff for their help with the video. This work was supported in part by the SNF grant 200021\_162958, the NSF CAREER award IIS-1652515, the NSF grant OAC:1835712, and a gift from Adobe.

## REFERENCES

- 2018. Imerys ENSACO 250G. [http://www.imerys-graphite-and-carbon.com/wordpress/wp-content/uploads/2014/04/Polymer\\_comounds1.pdf](http://www.imerys-graphite-and-carbon.com/wordpress/wp-content/uploads/2014/04/Polymer_comounds1.pdf). (2018). Accessed: 2018-01-18.
- 2018. OptiTrack. <http://optitrack.com/products/prime-13/>. (2018). Accessed: 2018-01-19.

2018. Parker Hannifin EAP Sensor. <http://ph.parker.com/us/en/electroactive-polymer-technology-monitors-movement-and-stretch-eap-sensor-evaluation-kits>. (2018). Accessed: 2018-01-18.
2018. Silbione RTV 4420. [https://silicones.elkem.com/EN/our\\_offer/Product/90060082\\_90060081/SILBIONE-RTV-4420-B-U1](https://silicones.elkem.com/EN/our_offer/Product/90060082_90060081/SILBIONE-RTV-4420-B-U1). (2018). Accessed: 2018-01-18.
2018. STM32 Nucleo-F446RE. <http://www.st.com/en/evaluation-tools/nucleo-f446re.html>. (2018). Accessed: 2018-01-18.
2018. StretchSense. <https://www.stretchesense.com/>. (2018). Accessed: 2018-01-18.
- O. A. Araromi, S. Rosset, and H. R. Shea. 2015. High-Resolution, Large-Area Fabrication of Compliant Electrodes via Laser Ablation for Robust, Stretchable Dielectric Elastomer Actuators and Sensors. *ACS Applied Materials & Interfaces* 7, 32 (2015), 18046–18053. <https://doi.org/10.1021/acsmami.5b04975> PMID: 26197865.
- A. Atalay, V. Sanchez, O. Atalay, D. M. Vogt, F. Haufe, R. J. Wood, and C. J. Walsh. 2017. Batch Fabrication of Customizable Silicone-Textile Composite Capacitive Strain Sensors for Human Motion Tracking. *Advanced Materials Technologies* 2, 9 (2017), 1700136–n/a. <https://doi.org/10.1002/admt.201700136> 1700136.
- M. Bächer, B. Hepp, F. Pece, P. G. Kry, B. Bickel, B. Thomaszewski, and O. Hilliges. 2016. DefSense: Computational Design of Customized Deformable Input Devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3806–3816. <https://doi.org/10.1145/2858036.2858354>
- R. Balakrishnan, G. Fitzmaurice, G. Kurtenbach, and K. Singh. 1999. Exploring Interactive Curve and Surface Manipulation Using a Bend and Twist Sensitive Input Strip. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics (I3D '99)*. ACM, New York, NY, USA, 111–118. <https://doi.org/10.1145/300523.300536>
- L. Ballan, A. Taneja, J. Gall, L. Van Gool, and M. Pollefeys. 2012. Motion capture of hands in action using discriminative salient points. *Computer Vision—ECCV 2012* (2012), 640–653.
- F. Beck and B. Stumpe. 1973. *Two devices for operator interaction in the central control of the new CERN accelerator*. Technical Report. CERN.
- L. Bernardi, R. Hopf, D. Sibilio, A. Ferrari, A. Ehret, and E. Mazza. 2017. On the cyclic deformation behavior, fracture properties and cytotoxicity of silicone-based elastomers for biomedical applications. *Polymer Testing* 60 (2017), 117 – 123. <https://doi.org/10.1016/j.polymertesting.2017.03.018>
- P. D. Block and S. Bergbreiter. 2013. Large area all-elastomer capacitive tactile arrays. In *SENSORS, 2013 IEEE*, 1–4. <https://doi.org/10.1109/ICSENS.2013.6688345>
- F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. J. Black. 2016. Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In *European Conference on Computer Vision*. Springer, 561–578.
- C. Bregler and J. Malik. 1998. Tracking people with twists and exponential maps. In *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*. IEEE, 8–15.
- J. Brunne, S. Kazan, and U. Wallrabe. 2011. In-plane DEAP stack actuators for optical MEMS applications. (2011), 7976–7976 - 10 pages. <https://doi.org/10.1117/12.880232>
- X. Chen and A. L. Yuille. 2014. Articulated pose estimation by a graphical model with image dependent pairwise relations. In *NIPS*. 1736–1744.
- C.-y. Chien, R.-H. Liang, L.-F. Lin, L. Chan, and B.-Y. Chen. 2015. FlexiBend: Enabling Interactivity of Multi-Part, Deformable Fabrications Using Single Shape-Sensing Strip. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 659–663. <https://doi.org/10.1145/2807442.2807456>
- L. A. Danisch, K. Englehart, and A. Trivett. 1999. Spatially continuous six-degrees-of-freedom position and orientation sensor. (1999), 48–56 pages. <https://doi.org/10.1117/12.339112>
- E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun. 2008. Performance Capture from Sparse Multi-view Video. In *ACM SIGGRAPH 2008 Papers (SIGGRAPH '08)*. ACM, New York, NY, USA, Article 98, 10 pages. <https://doi.org/10.1145/1399504.1360697>
- M. Dou, S. Khamis, Y. Degtyarev, P. Davidson, S. R. Fanello, A. Kowdle, S. O. Escalano, C. Rhemann, D. Kim, J. Taylor, P. Kohli, V. Tankovich, and S. Izadi. 2016. Fusion4D: Real-time Performance Capture of Challenging Scenes. *ACM Trans. Graph.* 35, 4, Article 114 (July 2016), 13 pages. <https://doi.org/10.1145/2897824.2925969>
- A. Elhayek, E. de Aguiar, A. Jain, J. Thompson, L. Pishchulin, M. Andriluka, C. Bregler, B. Schiele, and C. Theobalt. 2017. MARCONfåÄTCOnNet-Based MARKer-Less Motion Capture in Outdoor and Indoor Scenes. *IEEE transactions on pattern analysis and machine intelligence* 39, 3 (2017), 501–514.
- J. M. Engel, N. Chen, K. S. Ryu, S. D. Pandya, C. Tucker, Y. Yang, and C. Liu. 2006. Multi-layer Embedment of Conductive and Non-conductive Pdms for All-elastomer Mems.
- V. Ganapathi, C. Plagemann, D. Koller, and S. Thrun. 2012. Real-time human pose tracking from range data. In *European conference on computer vision*. Springer, 738–751.
- O. Glauser, W.-C. Ma, D. Panizzo, A. Jacobson, O. Hilliges, and O. Sorkine-Hornung. 2016. Rig Animation with a Tangible and Modular Input Device. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH)* (Jul 2016).
- A. Glinsky. 2000. *Theremin: ether music and espionage*. University of Illinois Press.
- D. Gotsch, X. Zhang, J. Burstin, and R. Vertegaal. 2016. HoloFlex: A Flexible Holographic Smartphone with Bend Input. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 3675–3678. <https://doi.org/10.1145/2851581.2890258>
- T. Grossé-Puppenthal, C. Holz, G. Cohn, R. Wimmer, O. Bechtold, S. Hodges, M. S. Reynolds, and J. R. Smith. 2017. Finding Common Ground: A Survey of Capacitive Sensing in Human-Computer Interaction. In *Proc. CHI*. ACM, 3293–3315. <https://doi.org/10.1145/3025453.3025808>
- J. Han, J. Gu, and G. Lee. 2014. Trampoline: A Double-sided Elastic Touch Device for Creating Reliefs. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 383–388. <https://doi.org/10.1145/2642918.2647381>
- R. Hopf, L. Bernardi, J. Menze, M. Zündel, E. Mazza, and A. Ehret. 2016. Experimental and theoretical analyses of the age-dependent large-strain behavior of Sylgard 184 (10:1) silicone elastomer. *Journal of the Mechanical Behavior of Biomedical Materials* 60 (2016), 425 – 437. <https://doi.org/10.1016/j.jmbbm.2016.02.022>
- B. Huang, M. Li, T. Mei, D. McCull, S. Qin, Z. Zhao, and J. Zhao. 2017. Wearable Stretch Sensors for Motion Measurement of the Wrist Joint Based on Dielectric Elastomers. *Sensors* 17, 12 (2017). <https://doi.org/10.3390/s17122708>
- A. Jacobson, D. Panizzo, O. Glauser, C. Pradalier, O. Hilliges, and O. Sorkine-Hornung. 2014. Tangible and Modular Input Device for Character Articulation. *ACM Transactions on Graphics (proceedings of ACM SIGGRAPH)* 33, 4 (2014), 82:1–82:12.
- H. Jeong and S. Lim. 2016. A Stretchable Radio-Frequency Strain Sensor Using Screen Printing Technology. *Sensors* 16, 11 (2016). <https://doi.org/10.3390/s16111839>
- H. Jin, S. Jung, J. Kim, S. Heo, J. Lim, W. Park, H. Y. Chu, F. Bien, and K. Park. 2017. Stretchable Dual-Capacitor Multi-Sensor for Touch-Curvature-Pressure-Strain Sensing. *Sci Rep* 7, 1 (Sep 2017), 10854.
- H.-L. C. Kao, C. Holz, A. Roseway, A. Calvo, and C. Schmandt. 2016. DuoSkin: rapidly prototyping on-skin user interfaces using skin-friendly materials. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*. ACM, 16–23.
- G. D. Kessler, L. F. Hodges, and N. Walker. 1995. Evaluation of the CyberGlove as a whole-hand input device. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2, 4 (1995), 263–283.
- D. Kingma and J. Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- H. Lee, J. Cho, and J. Kim. 2016. Printable skin adhesive stretch sensor for measuring multi-axis human joint angles. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 4975–4980. <https://doi.org/10.1109/ICRA.2016.7487705>
- S. Lee, W. Buxton, and K. C. Smith. 1985. A Multi-touch Three Dimensional Touch-sensitive Tablet. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '85)*. ACM, New York, NY, USA, 21–25. <https://doi.org/10.1145/317456.317461>
- D. Lipomi, M. Vosgueritchian, B. Tee, S. L Hellstrom, J. Lee, C. Fox, and Z. Bao. 2011. Skin-like pressure and strain sensors based on transparent elastic films of carbon nanotubes. 6 (10 2011), 788–92.
- H. Liu, X. Wei, J. Chai, I. Ha, and T. Rhee. 2011. Realtime human motion control with a small number of inertial sensors. In *Symposium on Interactive 3D Graphics and Games*. ACM, 133–140.
- F. Lorusso, W. Rocchia, E. P. Scilingo, A. Tognetti, and D. D. Rossi. 2004. Wearable, redundant fabric-based sensor arrays for reconstruction of body segment posture. *IEEE Sensors Journal* 4, 6 (Dec 2004), 807–818. <https://doi.org/10.1109/JSEN.2004.837498>
- T. Lu, L. Finkenauer, J. Wissman, and C. Majidi. 2014. Rapid Prototyping for SoftåÄRMatter Electronics. *Advanced Functional Materials* 24, 22 (2014), 3351–3356. <https://doi.org/10.1002/adfm.201303732> arXiv:<https://onlinelibrary.wiley.com/doi/10.1002/adfm.201303732>
- Z. Ma and E. Wu. 2014. Real-time and robust hand tracking with a single depth camera. *The Visual Computer* 30, 10 (2014), 1133–1144.
- C. Mattmann, F. Clemens, and G. Tröster. 2008. Sensor for Measuring Strain in Textile. *Sensors* 8, 6 (2008), 3719–3732. <https://doi.org/10.3390/s8063719>
- D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H.-P. Seidel, W. Xu, D. Casas, and C. Theobalt. 2017. VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera. *ACM Transactions on Graphics* 36, 4, 14. <https://doi.org/10.1145/3072959.3073596>
- T. B. Moeslund, A. Hilton, and V. Krüger. 2006. A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding* 104, 2 (2006), 90–126.
- T. Neumann, K. Varanasi, N. Hasler, M. Wacker, M. Magnor, and C. Theobalt. 2013. Capture and Statistical Modeling of Arm-Muscle Deformations. *Computer Graphics Forum* 32, 2pt3 (2013), 285–294. <https://doi.org/10.1111/cgf.12048>
- R. A. Newcombe, D. Fox, and S. M. Seitz. 2015. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 343–352.
- A. Newell, K. Yang, and J. Deng. 2016. Stacked hourglass networks for human pose estimation. In *ECCV*, 483–499.
- A. S. Nittala, A. Withana, N. Pourjafarian, and J. Steimle. 2018. Multi-Touch Skin: A Thin and Flexible Multi-Touch Sensor for On-Skin Input. In *Proceedings of the 2018*

- CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 33, 12 pages. <https://doi.org/10.1145/3173574.3173607>
- B. O'Brien, T. Gisby, and I. A. Anderson. 2014. Stretch sensors for human body motion. In *Proc. SPIE*, Vol. 9056, 905618.
- A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. 2017. Automatic differentiation in PyTorch. (2017).
- R. D. Ponce Wong, J. Posner, and V. Santos. 2012. Flexible microfluidic normal force sensor skin for tactile feedback. 179 (06 2012), 62–69.
- G. Pons-Moll, J. Romero, N. Mahmood, and M. J. Black. 2015. Dyna: A Model of Dynamic Human Shape in Motion. *ACM Trans. Graph.* 34, 4, Article 120 (July 2015), 14 pages. <https://doi.org/10.1145/2766993>
- I. Poupyrev, N.-W. Gong, S. Fukuhara, M. E. Karagozler, C. Schwesig, and K. E. Robinson. 2016. Project Jacquard: Interactive Digital Textiles at Scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4216–4227. <https://doi.org/10.1145/2858036.2858176>
- J. Rekimoto. 2002. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 113–120. <https://doi.org/10.1145/503376.503397>
- C. Rendl, P. Greindl, M. Haller, M. Zirkel, B. Stadlober, and P. Hartmann. 2012. PyzoFlex: Printed Piezoelectric Pressure Sensing Foil. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 509–518. <https://doi.org/10.1145/2380116.2380180>
- C. Rendl, D. Kim, S. Fanello, P. Parzer, C. Rhemann, J. Taylor, M. Zirkel, G. Scheipl, T. Rothländer, M. Haller, and S. Izadi. 2014. FlexSense: A Transparent Self-sensing Deformable Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 129–138. <https://doi.org/10.1145/2642918.2647405>
- H. Rhodin, N. Robertini, C. Richardt, H.-P. Seidel, and C. Theobalt. 2015. A versatile scene model with differentiable visibility applied to generative pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, 765–773.
- D. Roetenberg, H. Luinge, and P. Slycke. 2007. Moven: Full 6dof human motion tracking using miniature inertial sensors. *Xsen Technologies*, December 2, 3 (2007), 8.
- S. Rosset, O. A. Araromi, S. Schlatter, and H. R. Shea. 2016. Fabrication Process of Silicone-based Dielectric Elastomer Actuators. *J Vis Exp* 108 (Feb 2016), e53423.
- S. Rosset and H. R. Shea. 2013. Flexible and stretchable electrodes for dielectric elastomer actuators. *Applied Physics A* 110, 2 (01 Feb 2013), 281–307. <https://doi.org/10.1007/s00339-012-7402-8>
- M. S. Sarwar, Y. Dobashi, C. Preston, J. K. M. Wyss, S. Mirabbasi, and J. D. W. Madden. 2017. Bend, stretch, and touch: Locating a finger on an actively deformed transparent sensor array. *Science Advances* 3, 3 (2017). <https://doi.org/10.1126/sciadv.1602200> arXiv:<http://advances.sciencemag.org/content/3/3/e1602200.full.pdf>
- L. Schwarz, D. Mateus, and N. Navab. 2009. Discriminative human full-body pose estimation from wearable inertial sensor data. *Modelling the Physiological Human* (2009), 159–172.
- C. Schwesig, I. Poupyrev, and E. Mori. 2004. Gummi: A Bendable Computer. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 263–270. <https://doi.org/10.1145/985692.985726>
- E. P. Scilingo, F. Lorusso, A. Mazzoldi, and D. D. Rossi. 2003. Strain-sensing fabrics for wearable kinaesthetic-like systems. *IEEE Sensors Journal* 3, 4 (Aug 2003), 460–467. <https://doi.org/10.1109/JSEN.2003.815771>
- Z. Shen, J. Yi, X. Li, M. H. P. Lo, M. Z. Q. Chen, Y. Hu, and Z. Wang. 2016. A soft stretchable bending sensor and data glove applications. *Robotics and Biomimetics* 3, 1 (01 Dec 2016), 22. <https://doi.org/10.1186/s40638-016-0051-1>
- J. R. Shewchuk. 1996. Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In *Applied Computational Geometry: Towards Geometric Engineering*. Lecture Notes in Computer Science, Vol. 1148, 203–222.
- J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore. 2013. Real-time human pose recognition in parts from single depth images. *Commun. ACM* 56, 1 (2013), 116–124.
- T.-W. Shyr, J.-W. Shie, C.-H. Jiang, and J.-J. Li. 2014. A Textile-Based Wearable Sensing Device Designed for Monitoring the Flexion Angle of Elbow and Knee Movements. *Sensors* 14, 3 (2014), 4050–4059.
- J. R. Smith. 1995. *Toward electric field tomography*. Ph.D. Dissertation. Massachusetts Institute of Technology.
- J. Starck and A. Hilton. 2003. Model-based multiple view reconstruction of people. In *null*. IEEE, 915.
- C. Stoll, N. Hasler, J. Gall, H.-P. Seidel, and C. Theobalt. 2011. Fast articulated motion tracking using a sums of gaussians body model. In *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 951–958.
- J. Tautges, A. Zinke, B. Krüger, J. Baumann, A. Weber, T. Helten, M. Müller, H.-P. Seidel, and B. Eberhardt. 2011. Motion reconstruction using sparse accelerometer data. *ACM Transactions on Graphics (TOG)* 30, 3 (2011), 18.
- J. Taylor, L. Bordeaux, T. Cashman, B. Corish, C. Keskin, T. Sharp, E. Soto, D. Sweeney, J. Valentin, B. Luff, A. Topalian, E. Wood, S. Khamis, P. Kohli, S. Izadi, R. Banks, A. Fitzgibbon, and J. Shotton. 2016. Efficient and Precise Interactive Hand Tracking Through Joint, Continuous Optimization of Pose and Correspondences. *ACM Trans. Graph.* 35, 4, Article 143 (July 2016), 12 pages. <https://doi.org/10.1145/2897824.2925965>
- J. Taylor, J. Shotton, T. Sharp, and A. Fitzgibbon. 2012. The vitruvian manifold: Inferring dense correspondences for one-shot human pose estimation. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 103–110.
- B. Tekin, P. Márquez-Neila, M. Salzmann, and P. Fua. 2016. Fusing 2D Uncertainty and 3D Cues for Monocular Body Pose Estimation. *arXiv preprint arXiv:1611.05708* (2016).
- J. J. Tompson, A. Jain, Y. LeCun, and C. Bregler. 2014. Joint training of a convolutional network and a graphical model for human pose estimation. In *NIPS*, 1799–1807.
- A. Toshev and C. Szegedy. 2014. Deeppose: Human pose estimation via deep neural networks. In *CVPR*, 1653–1660.
- T. von Marcard, B. Rosenhahn, M. J. Black, and G. Pons-Moll. 2017. Sparse Inertial Poser: Automatic 3D Human Pose Estimation from Sparse IMUs. *Comput. Graph. Forum* 36, 2 (May 2017), 349–360. <https://doi.org/10.1111/cgf.13131>
- Y. Wang, A. Jacobson, J. Barbić, and L. Kavan. 2015. Linear Subspace Design for Real-time Shape Deformation. *ACM Trans. Graph.* 34, 4, Article 57 (July 2015), 11 pages. <https://doi.org/10.1145/2766952>
- S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh. 2016. Convolutional pose machines. In *CVPR*, 4724–4732.
- M. Weigel, T. Lu, G. Bailly, A. Oulasvirta, C. Majidi, and J. Steimle. 2015. iSkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2991–3000. <https://doi.org/10.1145/2702123.2702391>
- M. Wessely, T. Tsandilas, and W. E. Mackay. 2016. Stretchis: Fabricating Highly Stretchable User Interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 697–704. <https://doi.org/10.1145/2984511.2984521>
- J. Wissman, T. Lu, and C. Majidi. 2013. Soft-matter electronics with stencil lithography. *2013 IEEE SENSORS* (2013), 1–4.
- S.-J. Woo, J.-H. Kong, D.-G. Kim, and J.-M. Kim. 2014. A thin all-elastomeric capacitive pressure sensor array based on micro-contact printed elastic conductors. *J. Mater. Chem. C* 2 (2014), 4415–4422, Issue 22. <https://doi.org/10.1039/C4TC00392F>
- D. Xu, A. Tairych, and I. A. Anderson. 2016. Stretch not flex: programmable rubber keyboard. *Smart Materials and Structures* 25, 1 (2016), 015012. <http://stacks.iop.org/0964-1726/25/i=1/a=015012>
- X. Zhou, M. Zhu, S. Leonards, K. G. Derpanis, and K. Daniilidis. 2016. Sparseness meets deepness: 3D human pose estimation from monocular video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4966–4975.
- T. G. Zimmerman, J. R. Smith, J. A. Paradiso, D. Allport, and N. Gershenson. 1995. Applying Electric Field Sensing to Human-computer Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 280–287. <https://doi.org/10.1145/223904.223940>
- M. Zollhöfer, M. Nießner, S. Izadi, C. Rehmann, C. Zach, M. Fisher, C. Wu, A. Fitzgibbon, C. Loop, C. Theobalt, et al. 2014. Real-time non-rigid reconstruction using an RGB-D camera. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 156.

## A SILICONE MIXTURES

We used the following mixtures for the three types of silicone layers:

**Protective layer:** Silbione RTV 4420 [Sil 2018] component A (weight ratio=1.0) and Toluol (1.0) are mixed, then Silbione RTV 4420 (1.0) component B is added.

**Conductive layer:** Silbione RTV 4420 component A (1.0) and Toluol (2.0) are mixed, then Silbione RTV 4420 (1.0) component B is added. In a separate container, Imerys Enasco 250 P [Ens 2018] conductive carbon black (0.2) is mixed with isopropyl alcohol (2.0) by slowly adding the isopropyl alcohol while stirring. Then both compositions are combined and mixed for about 3 minutes. The 2-component silicone Silbione RTV 4420 was chosen due to its tear behavior as evaluated in [Bernardi et al. 2017] and the Imerys Enasco 250 P carbon black as suggested in [Brunne et al. 2011].

**Dielectric layer:** Same as the protective layer.

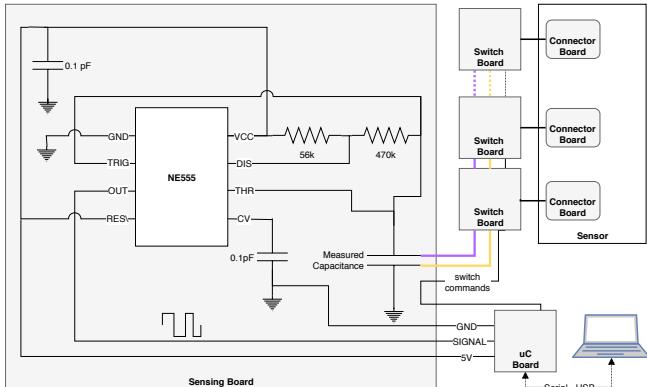


Fig. 26. Our modular setup consists of two parts. Left: The capacitance sensing circuit is implemented with a NE555 timer IC, resulting in a square SIGNAL of the charging time that is read by the uC and sent to the computer. Right: The uC board and the switch boards go through all combinations, dynamically connecting the current set of source electrode strips (purple) and ground electrode strips (yellow); see Sec. 4.2 and Fig. 6 for details.

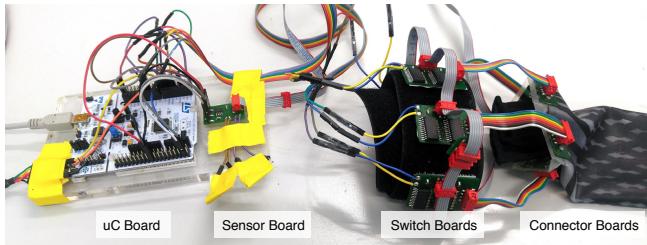


Fig. 27. Our custom modular measurement setup with the four types of boards. Up to 8 switch boards (and according connector boards) can be daisy-chained.

## B MEASUREMENT SETUP

In our setup capacitance is indirectly measured by timing the charging of a capacitor until a predefined voltage level, since the charging time is linearly proportional to the capacitance. However, our setting is more challenging, since we have to dynamically reconnect the electrodes following the measurement protocol described in Sec. 4.2. For this purpose, we design a modular measuring system (Fig. 26 right and Fig. 27), composed of three kinds of custom boards: the *connector board*, which is directly placed in contact with the sensor, the *switch board*, which is connected to the connector board by a set of flexible wires and the *sensing board* that contains the electronics needed to measure the charging times and send them to the connected computer. The connector boards are placed on the sensor on the exposed sensor pads that are shown in Fig. 9, supported by a PET foil and screwed into an acrylic counter-holder. The PET foil acts as intermediary from stretchable (silicone sensor), through flexible (PET), to fully rigid (*connector board*). The switch boards enable switching through the sensor combinations and they can be daisy-chained to allow for a wide variety of sensor layouts. The switching is controlled from the *uC board*: A STM32 microcontroller on a

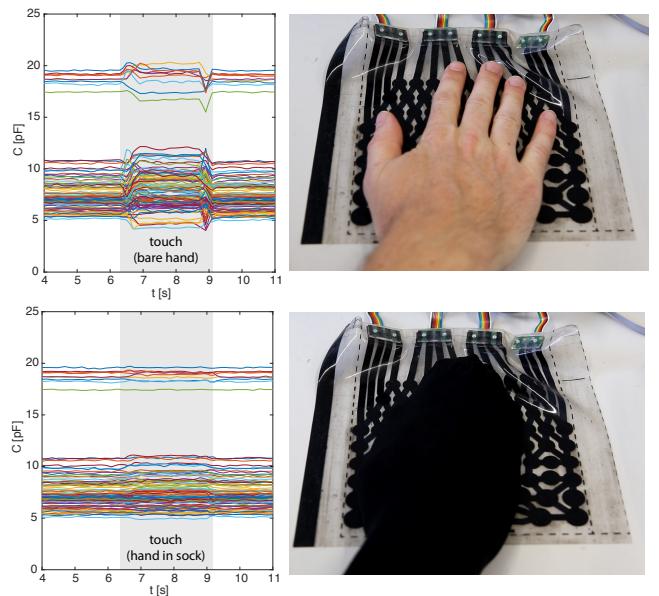


Fig. 28. This experiment demonstrates the effect of the nylon sock, worn below the sensor. Top: If the sensor is touched without the sock, the influence of the body capacitance creates clear spikes in the capacitance measured per sensor cell. Bottom: If the nylon sock is worn the same effect is minimal.

NUCLEO-F446RE board [STM 2018]. The microcontroller continuously transmits the charging time measurements to the computer via a USB-serial connection.

The capacitance measuring circuit (Fig. 26 left) is implemented using a NE555 timer IC. It outputs a square wave SIGNAL with a frequency  $f$  which is converted to capacitance by  $C = 1/(f \cdot (R_1 + 2R_2) \cdot \ln(2))$ , where  $R_1$  and  $R_2$  are the charging resistors. The larger these charging resistors are, the slower the capacitors are charged and dis-charged and the longer it takes for a complete measuring round (going through all sets of combined electrodes as shown in Fig. 6) and get the local capacitance changes updated. Note that our model neglects the influence of the resistance of the electrodes themselves. The full resistance for the longest electrode strip is about 50 kOhm. We experimentally found that setting  $R_1 = 470$  kOhm and  $R_2 = 47$  kOhm is a good compromise that produces sufficient accuracy while still supporting an interactive frame rate of 8 Hz. The parasitic capacitance of the circuit has to be subtracted from all the capacitance measurements. This can be simply done by continuously measuring the capacitance between two unconnected connector board pads. A nylon sock is worn below the sensor when capturing human body part deformation. As demonstrated in Fig. 28, it shields the in silicone embedded capacitor array from body capacitance and lowers the friction between the sensor and the skin, to e.g. pull a cylindrical sensor over a wrist with much less effort.