3D Upper Body Reconstruction with Sparse Soft Sensors

Zhiyong Chen[1], Ronghui Wu[1], Shihui Guo*[1], Xiangyang Liu[1], Hongbo Fu[2], Xiaogang Jin[3], and Minghong Liao[1]

[1]Xiamen University
[2]City University of Hong Kong
[3]Zhejiang University

Abstract

3D reconstruction of human body has wide applications for example for customized design of clothes and digital avatar production. Existing vision-based systems for 3D body reconstruction require users to wear minimal or extreme-tight clothes in front of cameras, and thus suffer from privacy problems. In this work we explore a novel solution based on a sparse number of soft sensors on a standard garment, and use it for capturing 3D upper body shape. We utilize the maximal stretching range by modeling the nonlinear performance profile for individual sensors. The body shape can be dynamically reconstructed by analyzing the relationship between mesh deformation and sensor reading, with a learning-based approach. The wearability and flexibility of our prototype allow its use in indoor/outdoor environments and for long-term breath monitoring. Our prototype has been extensively evaluated by multiple users with different body sizes and the same user for multiple days. The results show that our garment prototype is comfortable to wear, and achieves the state-of-the-art reconstruction performance with the advantages in privacy projection and application scenarios.

1 Introduction

3D human body reconstruction is the task of recovering the 3D geometry of a real human. It has wide applications for example in producing customized clothes for the textile industry and generating personalized avatars in 3D telepresence and interactive media. Most of the existing solutions (e.g.,[1,2,3,4]) for 3D body reconstructions are vision-based, and require the use of RGB/RGBD/laser cameras. Although such vision-based solutions already achieve reasonably high reconstruction accuracy, they suffer from several limitations. First, due to the incapability of laser/RGB/RGBD cameras in penetrating textiles, the existing solutions often require users to wear minimal or tight clothes in front of a camera. This procedure not only requires additional efforts and but also causes privacy concerns. Additionally, the requirement of camera setup also limits their use in arbitrary (in particular outdoor) environments.

We tackle this problem by building a fully-wearable system, using a sparse network of soft sensors on a standard garment. Benefited from the recent advances in material science and sensor fabrication, the sensors are yarn-like, ultra-light-weight and highly stretchable. The value of sensor

*Corresponding author: guoshihui@xmu.edu.cn

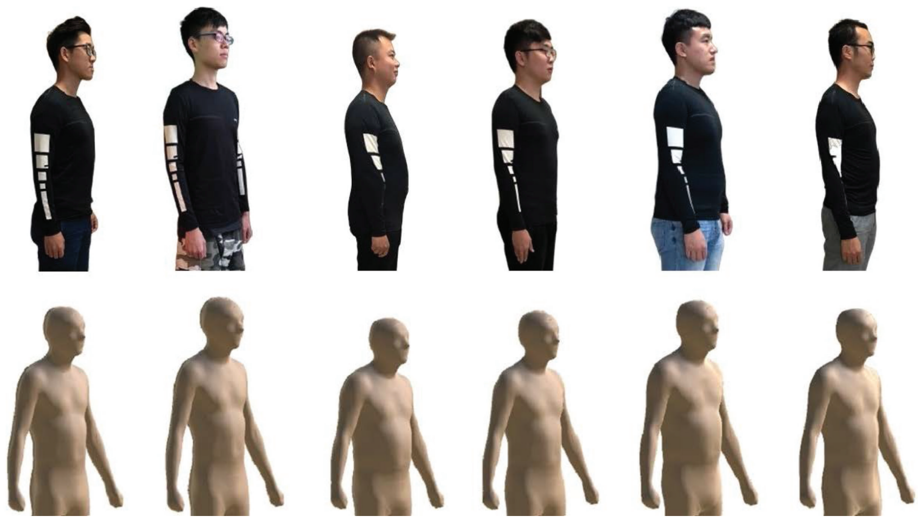


Figure 1: Upper body reconstruction (Bottom) using our wearable garment prototype (Top) with a sparse set of soft sensors.

resistance changes as it is being stretched, leading to a sensing functionality. With such intelligent sensors, our garment is capable of reconstructing 3D human body shape without causing discomfort for users wearing it. We perform 3D body reconstruction by leveraging a learning-based model to analyze the relationship between sensor reading and mesh deformation embedded in a large database of 3D human models. More specifically, we make the following contributions:

- We developed a working prototype, a wearable garment with a small number of soft sensors, and show its applications 3D upper body reconstruction and dynamic breath monitoring. Our prototype is fully stretchable and wearable, ultimately allowing users to wear it for an extended period without interfering with their daily tasks.

- Our method fully leveraged the nonlinear performance profile for individual sensors in order to utilize the maximal stretching range of sensor profile. A Long Short-Term Memory model is trained to take the sensor signal as input and accurately predicts the body girths. The body girth is then translated into the displaced vertex positions for the 3D model, thus creating a personalized human shape.

- We conducted comprehensive usability studies to evaluate our system, including testing its ability to handle various body sizes, repeated attempts of dressing for the same individual, and long-duration dressing in everyday use. The comparison between our method, existing vision-based methods and ground truth for 3D upper body reconstruction shows that our method achieves state-of-the-art performance with the advantages in privacy protection, application scenarios and user comfortableness.

2 Related Work

2.1 Smart Clothes with Soft Sensors

Smart clothes feature soft electronics and interconnections woven into the fabrics. To achieve the ultimate goal of fabricating smart clothes, researchers explored the relevant domains of scalable textile materials, design software, and machine knitting techniques. Project Jacquard[5] presented novel interactive textile materials and the corresponding manufacturing technologies for large-scale production. In addition to material fabrication, researchers also developed a variety of design software[6,7] and machine knitting techniques.[8,9] These recent works foresee future electronic systems to be an integral part of our everyday outfits. Before fabricating smart clothes for consumers, obtaining the information of their body shape is critical in order to guarantee the correct size and body-fitting. However, we found that there still lacks a convenient method to reconstruct human body shape for non-professional customers. Our work addresses this problem by building a sparse network of sensors on a garment and reconstructing human body shape in a user-friendly approach.

The emergence of smart clothes attributes to the rapid development of soft sensors in recent years. Compared with existing wearable sensors (e.g., inertial measurement units), soft sensors demonstrate advantages in terms of flexibility and comfortableness. These features are particularly important for implementing wearable devices. The composition of a soft sensor generally consists of two parts: one for the core conductive material and the other for the flexible support material.[10] The selected materials and fabrication methods are critical factors in determining the performance of a soft sensor. The stretchable sensor in our work chooses polyurethane fiber as the supporting material and silver-plated polyamide yarn as the conductive material.[11] Polyurethane is widely used in the textile industry and recognized for its stretchability and air permeability. The silver-plated polyamide yarn is helically wrapped around the polyurethane core fiber. This yarn-like sensor seamlessly fits onto the standard garment and introduces minimal discomfort to users.

Soft sensors often exhibit nonlinear time-variant behaviors, which make it difficult to accurately monitor their states.[12] Existing works explored the application of deep neural networks (DNNs) to interpret the information of strain sensors and monitor body kinematics. The choices of DNNs include convolutional neural network (CNN),[13] recurrent neural network[14] and long short-term memory model (LSTM).[12,15] A semi-supervised approach[14] achieved a higher performance with a smaller calibration dataset compared to supervised methods. CNN is popular for tasks of image understanding (e.g., object recognition) and suitable for processing the signals of a sensor array. As the CNN input, the resistance or capacitive value of a sensor is equivalent to the color value of an image pixel. Our system uses only a sparse network of five sensors. This sparsity imposes challenges for the down-sampling operation of CNN, indicating that CNN is not appropriate for our task. Since LSTM is known for its capability in processing temporal data, our work chooses LSTM to tackle the challenges of sensor hysteresis and nonlinearity, and to dynamically predict the body girths when users are breathing.

The majority of existing works in smart clothes focus on posture monitoring,[15,13,14] contact sensing[16,17] and gesture classification.[18,19,20] The monitored body parts include fingers,[20] ankle,[21] lower body[13,14] and full body.[15] Monitoring joint rotation can be accomplished by tracking an individual joint with a single sensor, which is placed at the exact location with maximal deformation.[15] To capture the contact with external objects, pressure sensors are placed at specific body parts on the clothes.[17] Different from the purposes of existing methods, our work aims to reconstruct the 3D shape of human upper body. However, the reconstruction of 3D body shape requires the analytic understanding of human shape as a whole model. Our work tackles this challenge by deriving the underlying pattern of human shape with a deep-learning approach.

2.2 Human Shape Reconstruction

Human shape reconstruction has been a long-standing problem in the domains of computer vision and graphics. The Skinned Multi-Person Linear model (SMPL) is a skinned vertex-based model that accurately represents a wide variety of body shapes in natural human poses.[2] The model parameters are learned from the captured data, including the rest pose template, blend weights, pose-dependent blend shapes, identity-dependent blend shapes, and a regressor from vertices to joint locations. Researchers learned the model of soft-tissue deformations from examples using a high-resolution 4D capture system and a method that accurately registers a template mesh to sequences of 3D scans.[22] Vision-based methods have made significant advances in terms of accuracy and time cost, but still face challenges when applying to the outdoor environment where the problems of visual occlusion, over-exposure or lack of illumination may frequently occur. These methods also require setting up specialized camera systems, which are not feasible for non-professional users. The images captured by RGB cameras when users are wearing minimal or tight clothes often cause the privacy concern of image leakage. The goal of our work is to alleviate these limitations and allow long-term wearing and mobility.

Recent works to reconstruct human body shape aimed to tackle the issue of privacy and reconstruct human shape when users are wearing the clothes.[23,24,25] The choice of sensor input includes a monocular video,[23,26] a depth camera,[25] laser scan sequences,[24] a single color image.[27] These methods are generally data-driven and parameterize the human shape and/or motion based on a template shape (e.g., the aforementioned SMPL model). This strategy is capable of producing detailed 3D mesh results, while requiring estimation only of a small number of parameters, making it friendly for direct network prediction. However, these vision-based methods regard clothes as the blocking factor to access the information of human shape. Unlike these methods, our method takes advantage of the clothes as the perfect medium to reconstruct the human body shape. From the aspect of methodology, we parameterized the human shape with a small set of characteristic girths and built a regression model to map the sensor signal to the 3D human mesh.

3 Methodology Overview

Our work builds a sparse network of soft sensors on a garment and reconstructs the 3D human upper body of a user wearing this garment based on the collected sensor reading. The workflow of our method is illustrated in Figure 2. We divide the complete task as two sub-tasks: 1) mapping from sensor signals to body girths; 2) mesh reconstruction from girth prediction. We observe the nonlinear sensor resistance-length relationship and model it by obtaining a large collection of sensor profiles. The pattern analysis of sensor profiles leads to the accurate prediction of the stretched length. We further build a regression model between the characteristic body girths and the vertex displacement of 3D mesh. The predicted length is translated into the vertex displacement position and completes the 3D body reconstruction with the learned model.

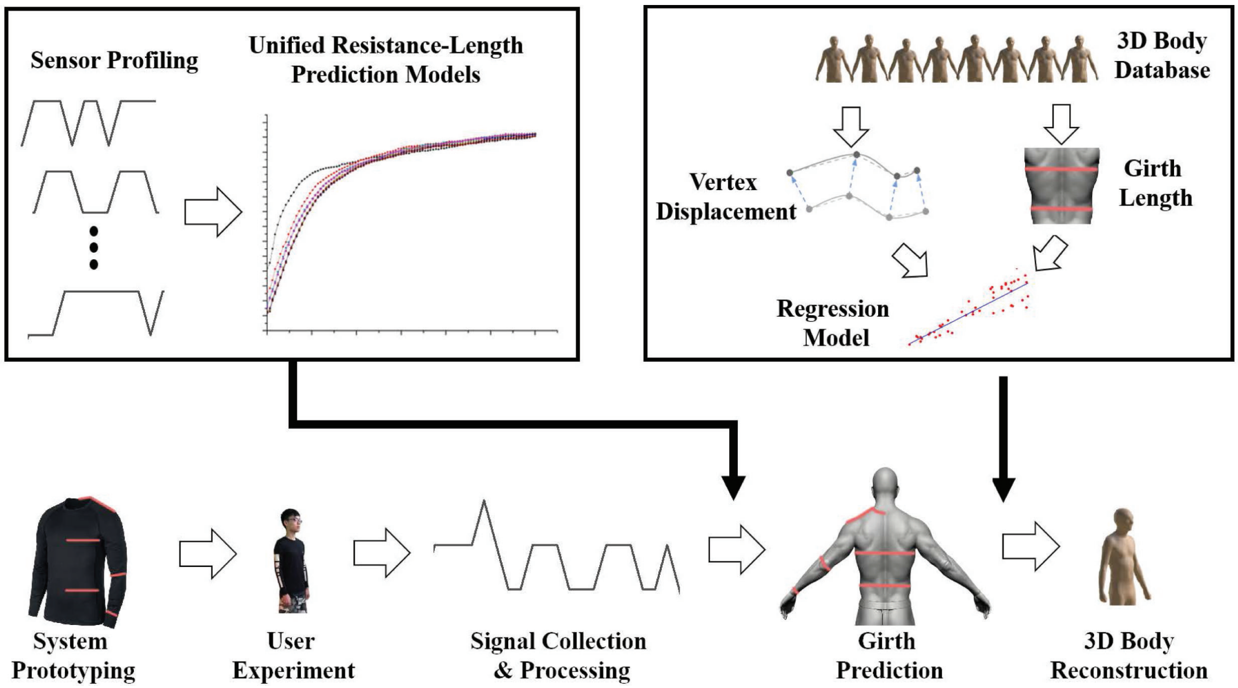


Figure 2: The pipeline of our work.

4 Hardware Development

4.1 Sensor Background

The fabricated sensor in our work is shown as Figure 3(a). The stretchable soft sensor in our work chooses polyurethane fiber as the supporting material and silver-plated polyamide yarn as the conductive material[11] (Figure 3(b)). The silver-plated polyamide yarn is helically wrapped around the polyurethane core fiber. Polyurethane is an ideal material for textile fabrication given its characteristic stretchability and air permeability. Figure 4 shows the relationship of resistance value-sensor length after 1000 cycles of 30% stretching. The results do not show a significant effect of plastic deformation. A previous work[28] showed that polyurethane can stretch up to 300% and could be cycled nearly 300,000 times under 40% stretch without noticeable breakage.

The silkworm fiber is processed with the technique of meso-functionalization,[29] by coating with sensing materials (Ag nanowires in our case). The sensor is further coated with the protection or dielectric layers to minimize the effect of direct contact with the human body. Silkworm is an ideal material for wearable sensors, given its biocompatible and biodegradable properties. It offers the advantages of comfortableness and air permeability, leading to the potential of long-term wearability. For detailed information on sensor material and fabrication, please refer to.[11]

When the sensor is stretched, the resistance increases due to an increasing distance between the wrapped silkworm fibers. Figure 4(b) shows the relationship between the resistance variance rate and the sensor's stretch length. The horizontal 'Length' axis refer to the distance between grips in the stretching apparatus. The static length between the grippers is 10cm in this case. This figure explains the challenges when dealing the sensor signal: hysteresis and nonlinearity. The solid curves

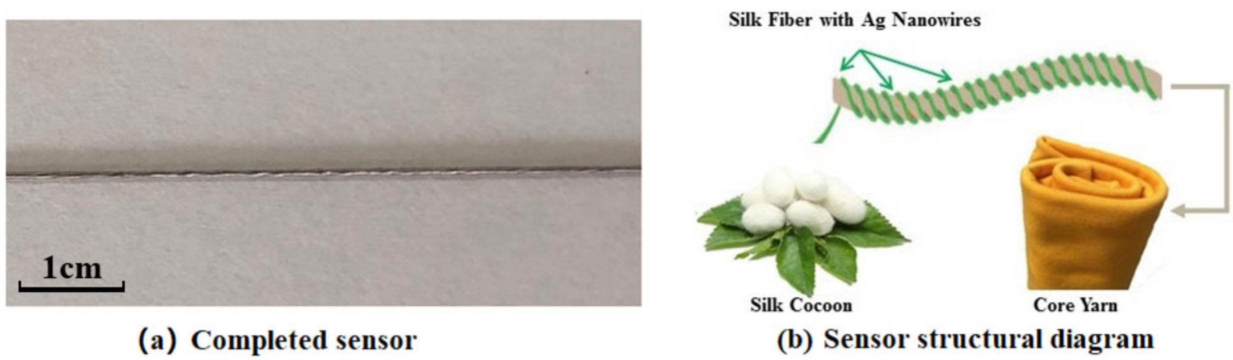


Figure 3: (a) Close-up view of the completed sensor. (b) Illustration of the sensor components.

are the average resistance value and sensor length when the sensor is being stretched (red) and released (blue). The shadowed areas demonstrate the resistance value range at different attempts with varying stretching speeds. The difference between the curves of stretch and release stages show the characteristic challenge of hysteresis. This inspires the use of the long short-term memory (LSTM) model to obtain an accurate length prediction of sensors at the chest and waist during dynamic breathing. The resistance-length curve shows exceptional merits of high sensitivity and approximately linear performance when the sensor is being stretched within 10%. The sensitivity decreases to a smaller value with approximately linear performance for the stretching range of [10%, 30%]. When the stretch length exceeds the threshold (30%), the resistance value changes with a very small ratio. In our work, we thus use the stretch range within 30% and ignore the scenarios with larger stretching. We also embed the capability of modeling this nonlinear relationship with the proposed LSTM model.

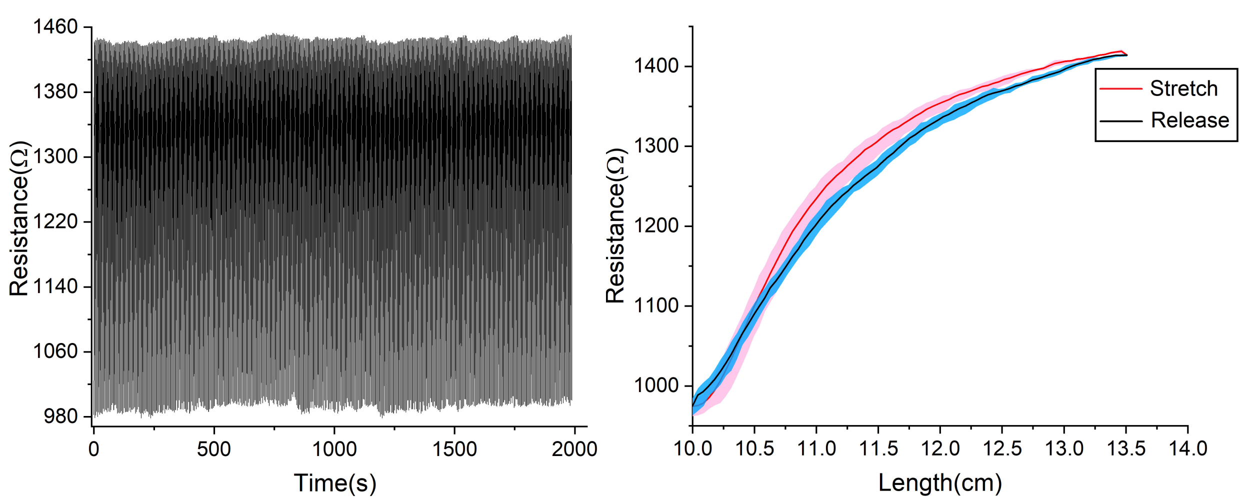


Figure 4: (a) Sensor resistance values for 1000 cycles of stretching. (b) Sensor resistance values in both the extension and release stages.

4.2 Sensor Placement on a Garment

Since it is intuitive to predict the length of our sensors based on their resistance values, we reckon that our problem is similar to the measurement of the body girths for cloth tailoring. Therefore, we consider using the measurement positions of cloth design as a reference for our sensor placement. To this end, we discussed extensively with professional tailors and experts in 3D reconstruction. Finally we decided to follow the 3D Measurement Standard published by the International Organization for Standardization.[30,31] It defines the anatomical landmarks on the human body used to measure its 3D shape.

As a balance between the system complexity and accuracy, we choose five girths for 3D upper body reconstruction, as illustrated in Figure 5:

- Sensor 1: this sensor measures half of the belly girth covering the lowest ribs and the navel. The two ends of this sensor connect from the navel point on the front and the spine area on the back.
- Sensor 2: this sensor measures the body girth circulating the mesosternale. The sensor starts from the middle chest on the front and terminates at the spine area on the back.
- Sensor 3: this sensor measures half of the shoulder width. It starts from the cervical vertebra, and ends at the acromion (the shoulder joint).
- Sensor 4: this sensor circumvents the elbow and measures the elbow size.
- Sensor 5: this sensor circumvents the wrist and measures the wrist size.

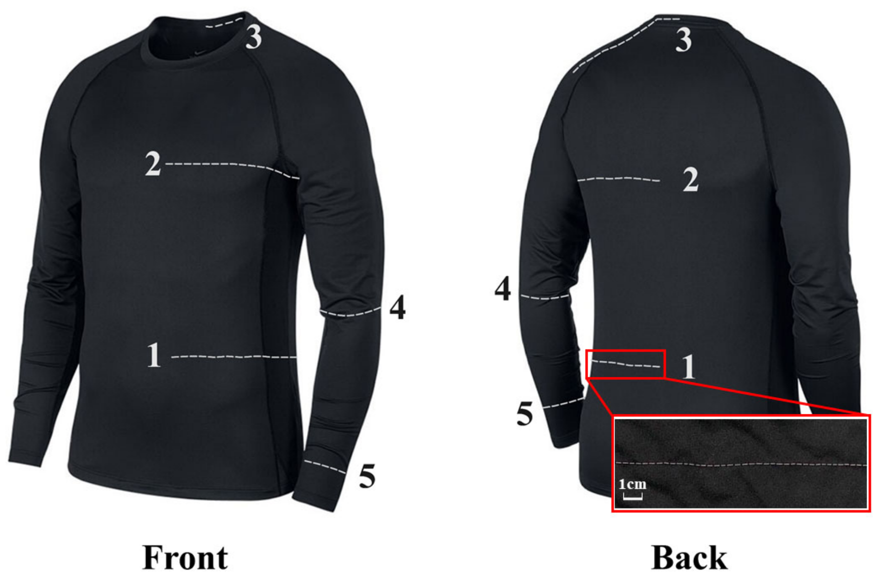


Figure 5: The placement of sensors on a garment according to anthropometry. The bottom right image shows one sensor after being sewed to the garment.

Throughout the experiments, the garments we used were tight sportswear, and each part was tight-fitting to the body. The composition of the garment fabric includes 80% polyester and 20% polyurethane.

We manually sewed sensors to the marked positions with the flat stitching. Each sensor, like a yarn, is sewed by following the needle through the fabric from the front to the back and then from the back to the front. This creates running stitches. One sensor on the garment is illustrated in the bottom right of Figure 5. The length of each stitch segment was small (<1cm), so that the slipping of the sensor on the garment was negligible.

4.3 Circuit Board Development

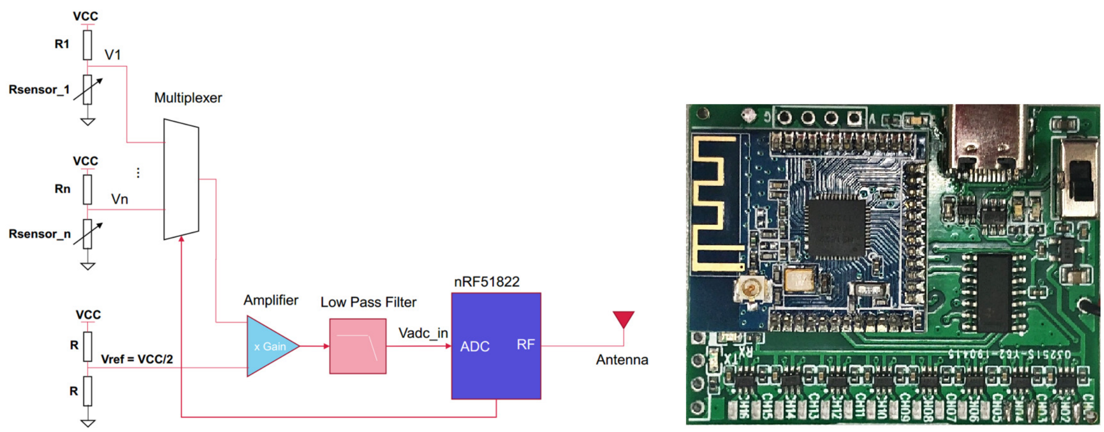


Figure 6: The design diagram (left) and the real circuit board (right) to collect the sensor resistance signals.

We design a circuit board (Figure 6) to collect the sensor resistance signals. The circuit supports a maximum number of 16 channels. The sampling channel is selected by a multiplexing voltage divider at a frequency of 20 Hz. The two electrodes of each sensor are connected to the two welding spots of each channel on the spot. One spot is grounded, and the voltage of the other spot is measured with a Wheatstone bridge circuit. We measure the difference between the voltage across each sensor and the reference voltage $V_{ref} = V_{CC}/2$, where V_{CC} (Volt Current Condenser) represents the access voltage of the circuit. $V_{CC} = 3.3$V. Voltage measurements at both ends of the sensor are processed by a low-pass filter with a bandwidth of 300Hz. The input voltage to the analog–digital conversion $Vadc_{in}$ is defined as (ignoring the effect of the low-pass filter):

$$Vadc_{in} = (\frac{V_{CC} * Rsensor_i}{R_i + Rsensor_i} - V_{CC}/2) * Gain. \quad (1)$$

where $Rsensor_i$ indicates the resistance of the i^{th} soft sensor. R_i denotes the divider resistor with a similar resistance value to $Rsensor_i$, and $Gain$=50 denotes the magnification factor of the amplifier unit.

The circuit board is of size 3.5cm x 4.5cm, and is attached to the garment at the hip-level position on the left part of the body. The signal of each sensor is transmitted to the server/mobile phone with the implementation based on low-energy Bluetooth (Nrf51822). The single sensor measurement of voltage at the two ends is converted to a digital signal. The original voltage range (0-5V) is now encoded in the range of [0, 1023]. The battery capacity is 600mAh and lasts for 30 hours for non-stop use. It is chargeable via a mini-USB port.

5 Shape Reconstruction

5.1 Signal De-noising

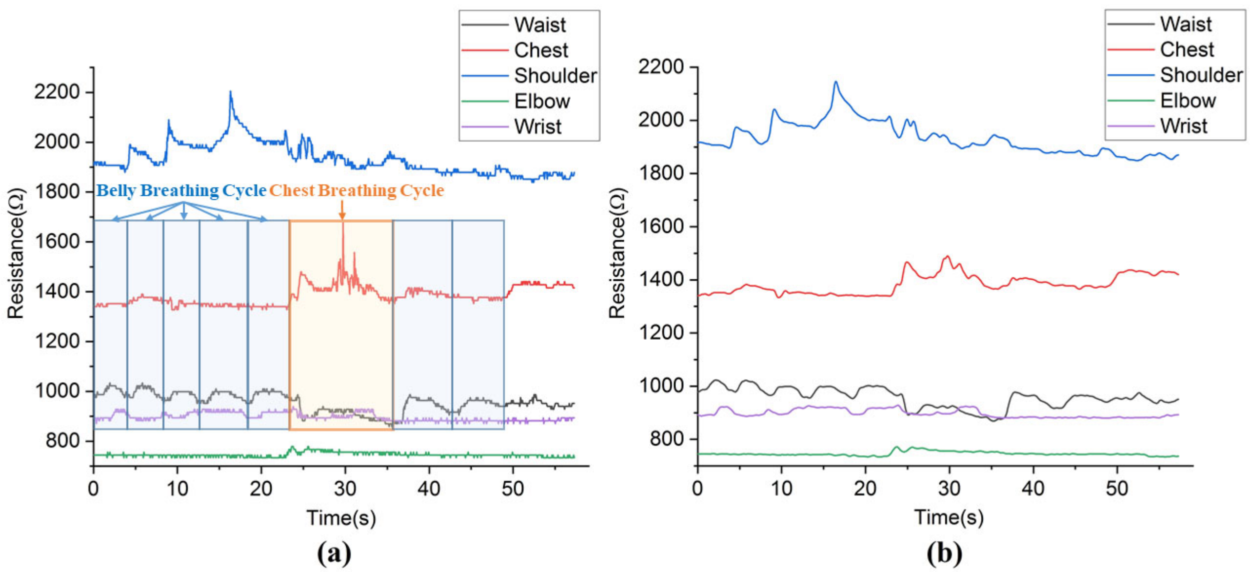


Figure 7: Sensor resistance values before and after smoothing. (a) The original signals. (b) The de-noised signals by the Gaussian smoothing method.

The original sensor signals are mixed with noise (Figure 7(a)). We use the method of Gaussian de-noising to effectively suppress noise and smooth the signals. The principle of action is similar to the averaging filter, which takes the average of the points of each signal in the filter window as the output.

5.2 Mapping from Sensor Signals to Body Girths

Accurate prediction of body girths from sensor signals is a challenging task. From the previous experimental results(Figure 4(b)), it can be seen that the relationship between resistance and tension of the sensor is non-linear, while different segments of the curve exhibit varying levels of linearity. This is particularly true for segments in small ([0, 5%]) and large ([10%, 30%]) stretching ranges. It is worth pointing out that a sensor may have the issue of hysteresis, indicating that the sensor corresponds to different resistance-length curves when it is being stretched or released. Therefore, we propose the use of Long Short-Term Memory (LSTM)[32] to accomplish the task of girth prediction of the waist and chest under breathing condition based on the sensor signals. And other body parts are measured under static conditions, we directly obtain the girths using Second-Order Polynomial Regression.

LSTM is an artificial recurrent neural network (RNN), which can efficiently deal with temporal data. Our network model has 3 LSTM layers, each of 64 hidden units, with 1 softmax layer as the output. The input of the network is a vector:

$$\vec{S} = \{S_{t-(N_p-1)\delta t}, \cdots, S_{t-\delta t}, S_t\}, \tag{2}$$

where S_t is the sensor resistance at a specific time t, δt is the time step for sensor signal reading, and N_P is the number of sample points. In our current implementation, δt and N_p are set to 0.1 second and 50, respectively. The output of the LSTM network is the estimated sensor length. By training the network model, we approach an accurate prediction of the sensor length considering the latent characteristics of nonlinearity and hysteresis.

To build the training dataset, we collected 20 sensors and stretched each sensor for 1000 times. For each stretching attempt, the sensor starts from its static length, and is stretched until its elongation reaches 30% and released to its original length. The stretching rate is dynamically and randomly adjusted. The sensor is stretched by a controlled mechanical motor, therefore we can compute the current sensor length given the historical stretching rate. The resistance value and the length are simultaneously measured and recorded at a fixed time-step of 0.15 seconds. A complete stretching cycle is composed of around 180 sample points (90 for either stretching or releasing stages). We divide the collected recording sequences into segments of a fixed duration ($\delta t \times N_p$). For each segment, the vector of sensor resistance is re-sampled to make its length as consistent of N_p. The predicted output of the network is designed to minimize its deviation from the measured sensor length. We define the loss function as the squared sum of the two, and use the Adam Optimizer with the learning rate of 0.001 and the batch size of 500.

5.3 Mesh Reconstruction from Girth Prediction

We use the CAESAR human body models[33] to calculate the girth of each position which we choose to measure on each 3D body mesh, including ankle, knee, thigh, waist, chest, shoulder, elbow, wrist and height. By calculating the sum of the cross section of each key position and the intersection line of each triangular mesh on each 3D human body mesh, we can get the girth. The height can be obtained by directly calculating the height of the model. Inspired by,[34] we compute the deformation of each triangle facet, and then learn a linear regression between the anthropometric parameters and the deformation of each triangle of each human body mesh.

First, we denote the deformation of each facet in each body mesh as a 3×3 transformation matrix $\mathbf{D}$. Let $\mathbf{v}_i$ and $\tilde{\mathbf{v}}_i$, $i \in 1...3$, be the undeformed and deformed vertexes of the i-th triangle, respectively. To establish how the space perpendicular to the triangle deforms and fully determine the affine transformation, we compute a fourth undeformed vertex as:

$$D = \begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{bmatrix}$$

Let v_i and $\tilde{v}_i$, $i \in 1...3$, be the undeformed and deformed vertexes of the triangle, respectively. To establish how the space perpendicular to the triangle deforms and fully determine the affine transformation, we compute a fourth undeformed vertex as:

$$v_4 = v_1 + (v_2 - v_1) \times (v_3 - v_1) \,/\sqrt{|\,(v_2 - v_1) \times (v_3 - v_1)\,|}$$

We denote the matrix of anthropometric parameters for n body meshes as:

$$G = \begin{bmatrix} p_{1,1} & \cdots & p_{1,9} \\ \vdots & \cdots & \vdots \\ p_{n,1} & \cdots & p_{n,9} \end{bmatrix}$$

A closed form expression for $\mathbf{D}$ is then given by $\mathbf{D} = \tilde{\mathbf{V}}\mathbf{V}^{-1}$, where $\mathbf{V} = [\mathbf{v}_2 - \mathbf{v}_1 \;\; \mathbf{v}_3 - \mathbf{v}_1 \;\; \mathbf{v}_4 - \mathbf{v}_1]$ and $\tilde{\mathbf{V}} = [\tilde{\mathbf{v}}_2 - \tilde{\mathbf{v}}_1 \;\; \tilde{\mathbf{v}}_3 - \tilde{\mathbf{v}}_1 \;\; \tilde{\mathbf{v}}_4 - \tilde{\mathbf{v}}_1]$.

We denote the matrix of anthropometric parameters for n body meshes as $\mathbf{G} = (p_{ij}) \in \mathcal{R}^{n\times 9}$, where $p_{i,j}$ means the j-th ($j \in 1, \cdots, 9$) parameter of body i ($i \in 1, \cdots, n$, n is the number of human body meshes). Then we perform a linear regression between $\mathbf{D}$ and $\mathbf{G}$ of each facet of body mesh.

The regression model can take an input of nine new anthropometric values and produce the deformation $\mathbf{D_k}$ ($k \in 1, \cdots, m$, m is the number of triangles in a body mesh) of each triangle on the new body mesh. Let $\mathbf{N}$ denote the triangular deformation of the new body mesh: $\mathbf{N} = [\mathbf{D}_1 \;\; \mathbf{D}_2 \;\; \cdots \;\; \mathbf{D}_m]^T$. The deformation of the triangles informs the position of each vertex by the following equation:

$$\mathbf{A}^T\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}^T\mathbf{N},$$

where $\tilde{\mathbf{x}}$ represents the vertex positions of our final body mesh. The matrix A is derived from the construction of $\mathbf{V}$.[34] The above system is essentially a sparse linear system and can be solved efficiently.

6 Results

6.1 Implementation and Performance

Our algorithm has been implemented in the Python environment. All source codes and datasets will be released to the public. We tested our algorithm on a standard PC (CPU: Intel i7 9700, GPU: RTX1080Ti, RAM:16G). The offline training of the LSTM-based mapping between sensor resistance and body girths costed 4.6 hours. The offline learning between the body girths and the 3D human body mesh costed 1.3 hours. Fortunately, these two processes need to be performed only once. Predicting the girth with a value of sensor resistance took 0.013 seconds on average. Given a set of girth values as input, the trained model produced the corresponding body mesh within 0.5 seconds on average. In total, it took less than one second to recover a 3D human body mesh from the acquisition of the resistance signal.

6.2 Mapping from Sensor Signal to Length

We compare our method and other polynomial regression (PR) methods (First-Order PR/Third-Order PR/Fifth-Order PR) for the purpose of mapping sensor signals to length. The results are shown in Figure 8. The LSTM method produces smaller error in terms of the predicted sensor length with respect to the ground truth, compared with other methods.

Figure 8 shows that when the sensor is stretched or released, our LSTM model can predict the sensor length with minimal error. A higher degree of error appears when the sensor state transitions from stretching to releasing.

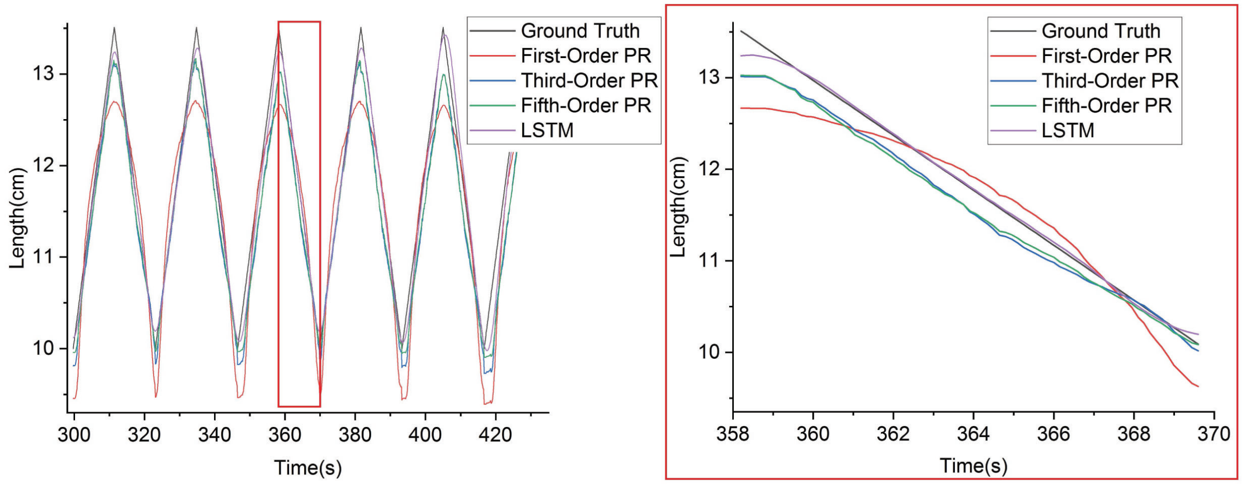


Figure 8: Results of sensor length prediction via our method and comparison methods.

6.3 User Experiment 1

This experiment evaluated our system both quantitatively and qualitatively. We demonstrated the reconstruction accuracy among a group of users and also compared with vision-based methods. User surveys revealed that our system received high scores in terms of comfortableness, convenience and accuracy.

Participants. 25 participants (20 male and 5 female) were recruited in this experiment. They were all students and faculties in a local university. They joined the experiment for free. The average and standard deviation of their age, height and weight were 24.3±4.2, 172.0±6.3 cm, 66.2±10.6 kg and 26.2±2.2, 162.3±5.4 cm, 53.9±5.2 kg for the male and female groups, respectively.

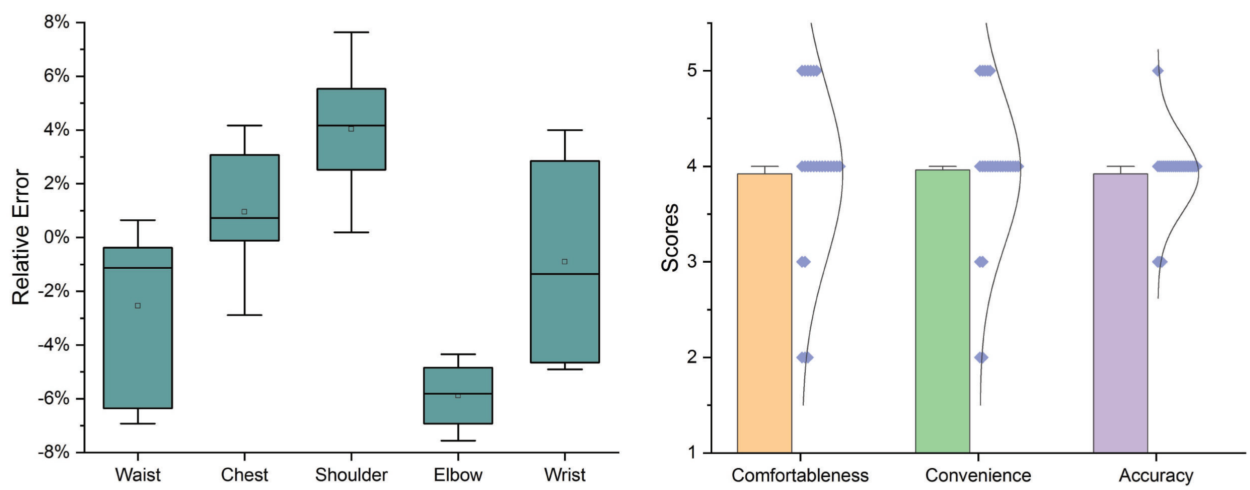
Procedures. The procedure was divided into pre-experiment, main-experiment, meta-experiment and post-experiment.

- *Pre-experiment*: The participants were first informed of the experiment purpose and signed their written agreement to join this study. They first filled in a pre-experiment questionnaire to inform their age, height, weight, and cloth size (XS/S/M/L/XL). This stage costed around 5 minutes.

- *Main-experiment*: They were instructed to put on our garment prototype with only underwear. Unlike vision-based methods, our technique does not suffer from the privacy issue. The participants adjusted the garment to fit their body. The complete measuring process includes three steps: 1) The arm part is stretched so that the feature point aligns with the wrist bone. 2) The sensor No. 1 is aligned with the dot on the belly. 3) The stretched parts are released and the garment returns to its normal mode. For Steps 1) and 2), the participants were instructed to maintain the posture for a duration of 3 seconds. We continuously recorded the sensor reading throughout the whole process and the time for individual steps. The collected data of sensor signals from above steps were considered simultaneously to reconstruct the 3D model, from which we obtained the body girths. This stage costed around 10-15 minutes.

- *Meta-experiment*: After this, the body girths of each participant were measured by a human instructor with a soft ruler. We also recorded the time cost for this process of manual measuring. On average it took 3-5 minutes for each participant. We randomly chose three subjects to conduct the comparative experiment against vision-based methods for the task of human shape reconstruction. the chosen participants were then captured with a RGBD camera (Kinect V2) when standing on a rotating platform in the T posture. The captured depth and color images were fed into two state-of-the-art methods (RGB,[23] RGBD[35]) to produce the 3D reconstruction results for the purpose of comparison. The time cost for the participants involved in the comparative study varied significantly (see the details in the following section on time cost analysis).

- *Post-experiment*: Finally the participants were shown with the reconstructed models by our method (and two vision-based methods for the participants chosen for the comparative study) and visually evaluated reconstruction accuracy. They filled in a 5-scale Likert questionnaire to rate their perception of comfortableness, convenience, and accuracy for the three methods. The participants were also interviewed to provide their subjective comments to explain their rating. This stage costed around 10-15 minutes.



(a) Error distribution (b) Survey results

Figure 9: Results from Experiment 1. (a) Quantitative evaluation of the reconstruction errors for five body girths. (b) Qualitative evaluation using comfortableness, convenience and accuracy scores from the post-study survey.

Quantitative Analysis: Girth Prediction. Figure 9 (a) plots the error distribution of our reconstruction results. The average and standard deviation for waist, chest, shoulder, elbow and wrist were -2.55%±2.96%, 0.95%±2.30%,4.03%±2.33%,-5.50%±1.31%,-0.91%±3.84%. Interestingly, we

SID	Unit(mm)	waist	error	relative error	chest	error	relative error	shoulder	error	relative error
Subject1	**Ground Truth**	743			858			1060		
	RGB	1006	263	35.40%	1051.5	193.5	22.55%	810.6	-249.4	-23.53%
	RGBD	880.8	137.8	18.55%	**857.8**	**-0.2**	**-0.02%**	885.25	-174.75	-16.49%
	Ours	**803.04**	**60.04**	**8.08%**	884.24	26.24	3.06%	**1014.21**	**-45.79**	**-4.32%**
Subject2	**Ground Truth**	824			927			1022		
	RGB	894.6	70.6	8.57%	1078.3	151.3	16.32%	885.3	-136.7	-13.38%
	RGBD	956.85	132.85	16.12%	903.06	-23.94	-2.58%	878.87	-143.13	-14.00%
	Ours	**851.32**	**27.32**	**3.32%**	**929.5**	**2.5**	**0.27%**	**1044.4**	**22.4**	**2.19%**
Subject3	**Ground Truth**	877			984			1120		
	RGB	1145.5	268.5	30.62%	1107.8	123.8	12.58%	920	-200	-17.86%
	RGBD	1015.6	138.6	15.80%	971.5	-12.5	-1.27%	983.1	-136.9	-12.22%
	Ours	**871.23**	**-5.77**	**-0.66%**	**982.07**	**-1.93**	**-0.20%**	**1108.8**	**-11.2**	**-1.00%**

Table 1: Comparison of the reconstruction errors between our method, and the vision-based methods using RGB[23] and RGBD[35] sensors.

found that for the waist, elbow and wrist parts, the measurements tended to under-estimate the body girths, while the measurements tended to over-estimate the chest and shoulder parts. The under-estimation might be potentially caused by the minor displacement of the sensors, with relative to the exact positions. The over-estimation could be caused by the breathing or other subtle movement. From the comparison of the final reconstruction results and the actual values of various parts of the human body, we can see that most of the results were still relatively accurate. However, it can also be seen that the errors of some parts were relatively big, especially for the shoulder part. Measuring the shoulder girth is particularly challenging as we do not actively adjust the sensor position to ensure its accurate position. This opens up research problems for future directions.

Table 1 presents the ground truth of the human body girths and their corresponding predictions by three methods. The results show that our method outperformed vision-based methods using RGB and RGBD sensors in all body girths and for all subjects. This confirms the usability of our method as an alternative solution to existing vision-based systems. In addition, we inherently resolve the privacy concern of users by avoiding using the user-facing cameras.

Qualitative Analysis with Questionnaire Feedback. The comfortableness, convenience and accuracy scores were 3.92±0.91, 3.96±0.79 and 3.92±0.40, respectively (Figure 9 (b)). The participants mentioned that it was convenient to use our prototype. One said that, "this is like a normal garment, and I cannot feel much difference after wearing it". However there were a couple of participants who mentioned that, "it takes some caution to wear the garment". This is possibly due to the user awareness of the circuit board, which most users took special caution and avoided large movements. One participant who had previous experience of vision-based 3D systems mentioned that our tool solved her concern of the picture leakage, while still obtaining satisfactory modeling results.

Quantitative Analysis: Comparison of Time Cost. The time cost to put on our garment was 2.06±0.25 minutes on average. As mentioned previously, the algorithm reconstructs a 3D body mesh in less than one second after receiving the sensor signals. Therefore, the reconstruction process can be regarded as real time, since users can view their 3D body shape before they finish taking off the garment. They can also observe their dynamic shape changing when inhaling and exhaling. In contrast, the reconstruction using the RGBD camera costed 3.26±0.98 on average, and it easily failed when moving the RGBD camera at a fast speed. The reconstruction using the RGB camera costed 18.63±5.69 on average, depending on the rotation speed of users. However, the reconstruction took more than 5 hours for posture estimation and 10 minutes for mesh reconstruction. The reconstruction with RGB images requires iterative optimization to fit the mesh with the extracted human masks and thus is time-consuming. This comparison confirms the advantages of our method in efficiency during user interaction.

6.4 User Experiment 2 - Cross-session Consistency

A common challenge for wearable systems is to consistently maintain high accuracy across worn sessions, each of which is defined as one attempt in which a user puts on and then takes off the wearable system. For each session, the location of sensor placement may vary slightly, since the garment cannot be exactly worn in the same configuration for repeated attempts. Therefore we conducted further experiments to evaluate the cross-session consistency of our system.

Participants. Five participants (students and faculties in the university) were recruited in this experiment. They were different from those in Experiment 1. The average and standard deviation of their age, height and weight were 22.4±1.6, 173.0±4.8 cm, 71.9±15.3 kg, respectively. They joined this study for free.

Procedures. The participants first underwent the same pre-experiment procedure as Experiment 1 (Sec. 6.3) to receive the experiment instructions, sign their written agreement and fill in the pre-experiment questionnaire. Each participant was invited to wear the same garment system for 10 sessions. And they took off and then put back on the garment between each session. For each session, we repeated the same main-procedure as in Experiment 1. Since this study was specifically designed to verify the cross-session consistency, we did not conduct the meta-experiment to collect data for comparison studies or post-experiment to collect user preferences and subjective comments.

Findings. We compared the sensor readings for individuals across different sessions. For five subjects, the distribution of the maximal (inhaling) resistance value was 5.79±0.03, 6.69±0.05, 5.47±0.03, 6.83±0.05 and 6.65±0.05 (Unit: kΩ). In addition, the distribution of the minimal (exhaling) resistance value was 5.13±0.02, 5.76±0.04, 5.29±0.03, 5.84±0.05 and 5.48±0.03 (Unit: kΩ). For the same individual, the standard deviation was 3.81%. Repositioning the sensors between different instances of putting on the shirt would cause the variation between different trials, even on the same wearer. While as shown in Figure 5, the manually sewing is another cause of variation between different trials. The metrics show that the the cross-session data for the same individual is focused within a small range. This confirms the robustness of our system to reconstruct 3D human body across different sessions.

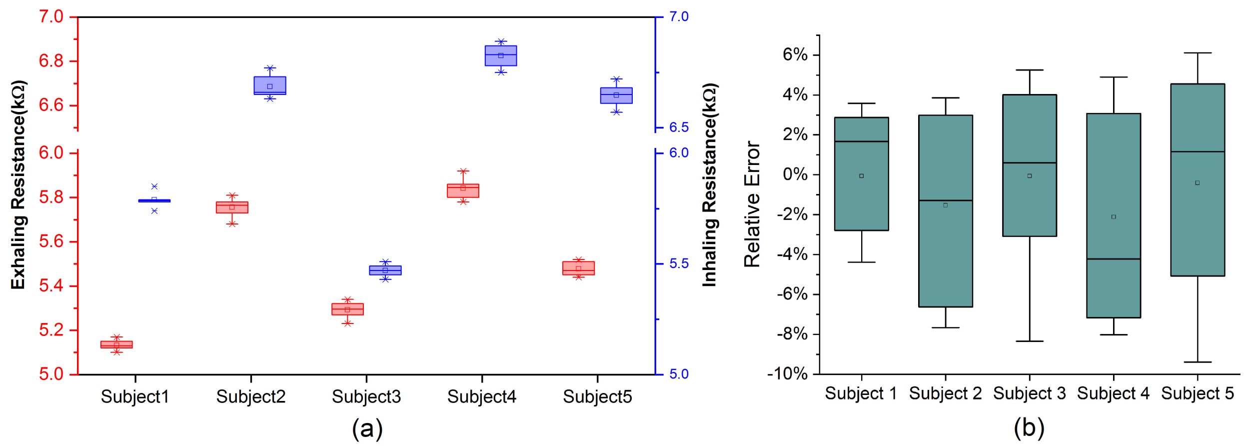


Figure 10: (a) Results from Experiment 2 to validate the cross-session consistency. Five participants repeatedly wore the garment for 10 sessions. This plot shows the maximal and minimal values of the sensor resistance when inhaling (in blue) and exhaling (in red), respectively. (b) Quantitative evaluation of the reconstruction errors for waist girths of five subjects.

6.5 User Experiment 3 - Long-term Wearability

We conducted another experiment to evaluate the performance of our system when a user wears for a long period. The purposes are at least two folds: 1) the usability of users wearing for a long duration. 2) the sensor consistency over a long duration.

Participants. We recruited 1 subject (the student in the university), different from the previous experiments, to join this experiment. His age, height and weight are 25, 165 cm, 65 kg, respectively.

Procedures. The participant received the experiment instruction, signed the written agreement and filled in the pre-experiment questionnaire, as similarly done in Experiments 1 and 2. The subject was asked to wear the garment system continuously from 8 am to 18 pm everyday for 7 days. During the experiment, he conducted his daily routines, including office work and home activities. The battery life was sufficiently large so there was no interruption for battery charging during this experiment. We recorded the maximum and minimum values for every 10 seconds to evaluate the consistency of our sensor. At the end of the everyday session, the participant filled in a 5-scale Likert questionnaire to rate his perception of comfortableness of our prototype. After receiving the daily questionnaire, we also conducted a short semi-structured interview with the participant and collected his feedback.

Quantitative Analysis. For a consecutive of seven days, the distribution of minimal (exhaling) resistance values was 4.62 ± 0.06, 4.64 ± 0.09, 4.57 ± 0.04, 4.61 ± 0.04, 4.66 ± 0.04, 4.67 ± 0.02 and 4.69 ± 0.01 (Unit: kΩ), respectively. Correspondingly, the distribution of maximal (inhaling) resistance values was 5.52 ± 0.07, 5.52 ± 0.11, 5.55 ± 0.08, 5.59 ± 0.06, 5.59 ± 0.06, 5.61 ± 0.05 and 5.67 ± 0.01 (Unit: kΩ), respectively. Figure 11 shows the statistical results for Sensor 1. Figure 11 shows a slight upward drift in the sensor's resistance over time. When we fixed a sensor to the garment

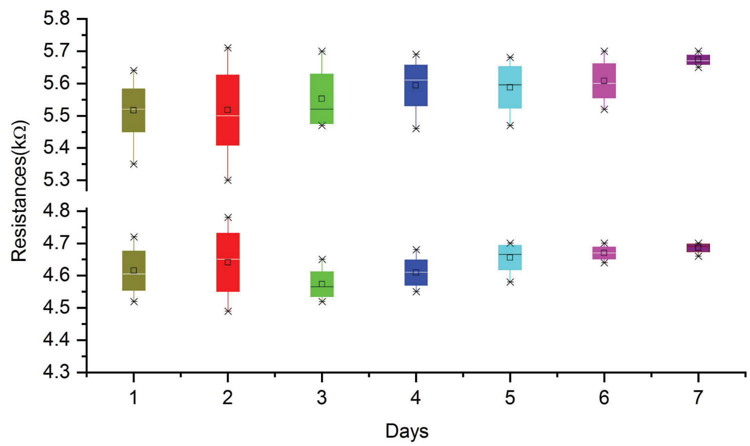


Figure 11: Results from Experiment 3 to validate the long-term wearability by asking a participant to wear our system for a consecutive period of seven days. This plot shows the maximal and minimal resistance value distributions of Sensor 1 for seven days.

(Figure 5), the sewing interval was relatively small (<1cm), and the slipping of the sensor on the garment was almost negligible. The main reason for this slight upward drift of the sensor resistance is attributed to the sewing structure of the sensor. The technique of flat stitching essentially divides the sensor into segments of running stitches, and connects these segments at cross-points between the sensor and the garment. When the sensor stretches and returns to its original length, the contact at these cross-points imposes additional friction preventing the sensor from returning to its original length.

Qualitative Analysis with Questionnaire Feedback. At the beginning of the experiment (Day 1), the participant explicitly expressed the discomfort experience of wearing our garment. The main factor mentioned by the participant is that the wiring and the circuit board for signal collection constantly raised his awareness about the wearable electronics. As a user, he was concerned about whether large movements could lead to the physical damage to the electronics. He mentioned particularly about the try-on and take-off procedures, which involve entangling and stretching wires. However, as the experiment continued, the system proved to be robust and he felt more comfortable with the garment. By the last day of the experiment, he felt fully comfortable with the garment and mentioned that "it is acceptable to try on the cloth as an everyday task".

6.6 Dynamic Capture of Breathing

We explicitly annotated the curves in Figure 7 by dividing them into cycles of belly and chest breathing. The participant mainly used the approach of belly breathing, so the waist-part sensor (the black curve) demonstrated cyclic variation (highlighted with blue-shadowed frames). During the time interval highlighted by the yellow-shadowed frame, the participant was taking a deep chest breathing. Signal variations of the shoulder-part sensor were caused by the secondary movement of the shoulder when the participant was taking breaths.

We segmented a short sequence (two breathing cycles) from Experiment 3 and reconstructed

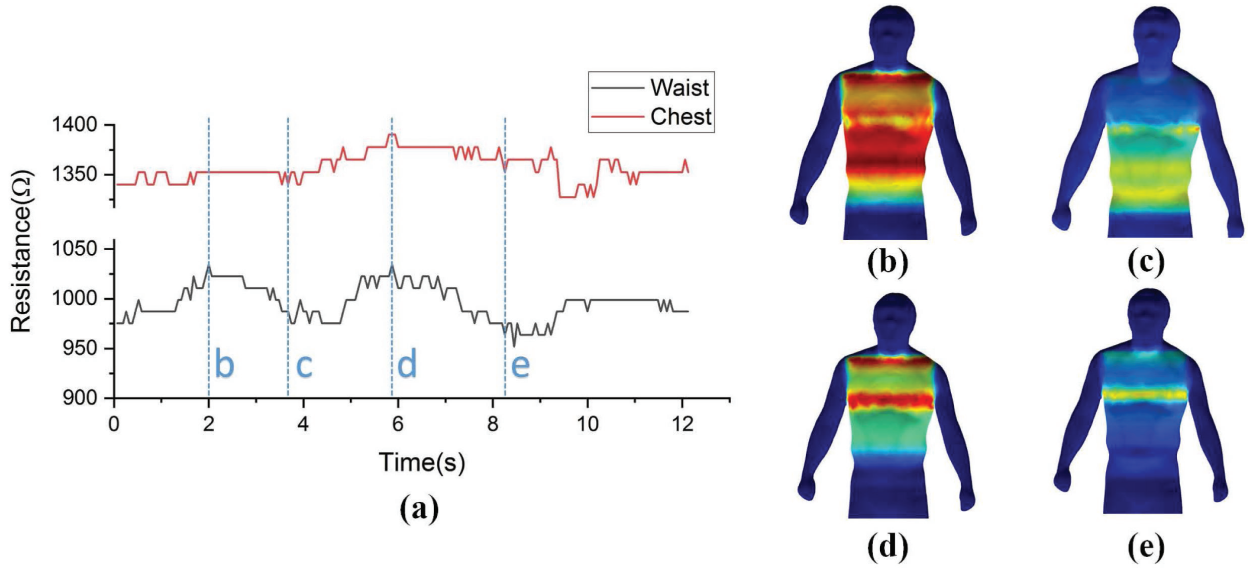


Figure 12: Breathing movement captured by our method in two breathing cycles. (a) Changes of the resistance during breathing. (b) Belly inhaling. (c) Belly exhaling. (d) Chest inhaling. (e) Chest exhaling.

3D human body mesh during breathing. Two cycles demonstrate two modes of breathing: using either chest or belly. Figure 12 showed the curves of the sensor resistance and the reconstructed mesh at multiple points in time. The color highlighted the vertex displacement at the current time point from the initial state. The color in red indicates a larger displacement, while the color in blue indicates a smaller one. We chose an initial state in which the resistance value of Sensor 1 reaches its minimal value, which indicates that the user finishes exhaling and starts to inhale in the new cycle of breathing. As shown in Figure 12(b), when inhaling, the value of Sensor 1 on waist increased dramatically, and the reconstructed human body changed most in its corresponding position. In contrast, when exhaling, each part of the reconstructed human body restored to its original shape, and the difference between the reconstructed human body and that at the initial state was very small, as shown in Figure 12(c). This can be observed from the resistance diagram (Figure 12 (a)), marked at Time-point c. In the second breathing cycle, the user mostly use the mode of chest breathing; at this time, the vertexes at the chest region demonstrate the largest displacement (Figure 12(d)). When the user exhaled again, the body returns to its original state, but the chest is not fully deflated, as shown in Figure 12(e).

6.7 Production Cost

The scalability of our work is critically related with the production cost of the complete system. The total price of the capture system costs 20 USD. The detailed breakdown is listed as follows. The average cost of the sensor is 1 USD/meter. The total length of the sensors on our prototype garment is 4 meters. The fabrication of circuit board costs 8 USD per each. The average cost of the wire is 0.01 USD/meter. The total length of the wire is 5 meters. A standard sport suit costs 10 USD. The lithium battery is 600mAh at a price of 2 USD. This low-cost device is suitable for consumer-level production. The sensors and connecting wires are all yarn-like materials, which can be seamlessly integrated into existing pipeline of textile manufacture, which is automatic and efficient compared with the lab set-up. Large scale production could further reduce the production cost. This further guarantees the scalability of our work.

6.8 Limitations

Currently, the system prototype only supports a small range of body sizes. The experimental results show that the accuracy is most accurate at the height of 175 cm, and the weight of 66 kg. Users whose body shapes are significantly different from this are expected to produce a larger reconstruction error. Constructing a customized system, with a number of size variations (S/M/L/XL, etc), may offer an improved solution for this limitation.

Our method relies on the use of a large 3D human model library. The current library which is publicly accessible is CAESAR,[33] but it contains mostly the shape models captured from American and European subjects. The differences between different ethnic groups may potentially affect the reconstruction accuracy. One of the future works is to deliver this prototype to a large number of subjects and build a human body shape covering a wide range of ethnics groups/ages/physiological states etc.

7 Conclusions

Our work explores the use of stretchable sensors for the purpose of dynamically monitoring 3D body shape. With a sparse set of stretchable sensors, our method is capable of reconstructing an upper body shape, with an average error rate of 2.79%±2.55% at the characteristic body girths. The stretchable sensors are soft and offer users with significant advantages in comfortable experiences. We only require a small number (N=5) of sensors, creating a low-cost and scalable system for consumer-level products. We conducted the across-session experiments to verify the consistency of our method. We conducted the pilot study to prove that users can wear the clothes for a relatively long period without interfering their daily tasks. Our work receives favorable user preferences over vision-based methods since there is no need of image capture with minimal on-body clothing. In contrast, with our approach, users are allowed to wear additional clothes on top of our garment prototype.

This work leads to a few directions for our future efforts. First, our current prototype focuses on the task of shape reconstruction of an upper body. It can be easily to capture the whole body by making pants integrated with the sensors. Another foreseeable application is to capture human movement, by placing the sensors at the specific joint locations. Combining both the shape and motion information of a captured subject allows us to create an identical virtual avatar. This opens up interesting applications for social networking in virtual environments. Second, another direction is to explore the capability to sense secondary deformation caused by muscle contraction. To achieve this goal, a dense sensor array is required to detect the small skin deformation. This can be integrated with other advanced sensors (e.g., to collect the electromyography signal) which collectively may offer an ideal solution to analyze the muscle activity and ultimately be applied to the scenarios of muscle rehabilitation.

8 Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities and National Natural Science Foundation of China (61702433, 61661146002).

References

[1] Aaron S Jackson, Chris Manafas, and Georgios Tzimiropoulos. 3d human body reconstruction from a single image via volumetric regression. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 64–77, 2018.

[2] Matthew Loper, Naureen Mahmood, Javier Romero, et al. Smpl: A skinned multi-person linear model. *ACM transactions on graphics (TOG)*, 34(6):248, 2015.

[3] Gul Varol, Duygu Ceylan, Bryan Russell, et al. Bodynet: Volumetric inference of 3d human body shapes. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 20–36, 2018.

[4] Longyu Zhang, Bote Han, Haiwei Dong, et al. Development of an automatic 3d human head scanning-printing system. *Multimedia Tools and Applications*, 76(3):4381–4403, 2017.

[5] Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, et al. Project jacquard: interactive digital textiles at scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4216–4227. ACM, 2016.

[6] Jussi Mikkonen and Riikka Townsend. Frequency-based design of smart textiles. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 294. ACM, 2019.

[7] Mikhaila Friske, Shanel Wu, and Laura Devendorf. Adacad: Crafting software for smart textiles design. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 345. ACM, 2019.

[8] Lea Albaugh, Scott Hudson, and Lining Yao. Digital fabrication of soft actuated objects by machine knitting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 184. ACM, 2019.

[9] Jifei Ou, Daniel Oran, Don Derek Haddad, et al. Sensorknit: Architecting textile sensors with machine knitting. *3D Printing and Additive Manufacturing*, 6(1):1–11, 2019.

[10] Morteza Amjadi, Ki-Uk Kyung, Inkyu Park, et al. Stretchable, skin-mountable, and wearable strain sensors and their potential applications: a review. *Advanced Functional Materials*, 26(11):1678–1698, 2016.

[11] Ronghui Wu, Liyun Ma, Chen Hou, et al. Silk composite electronic textile sensor for high space precision 2d combo temperature–pressure sensing. *Small*, pages 1901558/1–11, 2019.

[12] Thomas George Thuruthel, Benjamin Shih, Cecilia Laschi, and Michael Thomas Tolley. Soft robot perception using embedded soft sensors and recurrent neural networks. *Science Robotics*, 4(26):eaav1488, 2019.

[13] Mohsen Gholami, Ahmad Rezaei, Tyler J Cuthbert, Christopher Napier, and Carlo Menon. Lower body kinematics monitoring in running using fabric-based wearable sensors and deep convolutional neural networks. *Sensors*, 19(23):5325, 2019.

[14] Dooyoung Kim, Min Kim, Junghan Kwon, Yong-Lae Park, and Sungho Jo. Semi-supervised gait generation with two microfluidic soft sensors. *IEEE Robotics and Automation Letters*, 4(3):2501–2507, 2019.

[15] Dooyoung Kim, Junghan Kwon, Seunghyun Han, Yong-Lae Park, and Sungho Jo. Deep full-body motion network for a soft wearable motion sensing suit. *IEEE/ASME Transactions on Mechatronics*, 24(1):56–66, 2018.

[16] Joanne Leong, Patrick Parzer, Florian Perteneder, et al. procover: sensory augmentation of prosthetic limbs using smart textile covers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 335–346. ACM, 2016.

[17] Ali Kiaghadi, Seyedeh Zohreh Homayounfar, Jeremy Gummeson, Trisha Andrew, and Deepak Ganesan. Phyjama: Physiological sensing via fiber-enhanced pyjamas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3):1–29, 2019.

[18] Patrick Parzer, Adwait Sharma, Anita Vogl, et al. Smartsleeve: real-time sensing of surface and deformation gestures on flexible, interactive textiles, using a hybrid gesture detection pipeline. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pages 565–577. ACM, 2017.

[19] Sophie Skach, Rebecca Stewart, and Patrick GT Healey. Smart arse: Posture classification with textile sensors in trousers. In *Proceedings of the 2018 on International Conference on Multimodal Interaction*, pages 116–124. ACM, 2018.

[20] Oliver Glauser, Shihao Wu, Daniele Panozzo, et al. Interactive hand pose estimation using a stretch-sensing soft glove. *ACM Transactions on Graphics (TOG)*, 38(4):41, 2019.

[21] Hyunkyu Park, Junhwi Cho, Junghoon Park, Youngjin Na, and Jung Kim. Sim-to-real transfer learning approach for tracking multi-dof ankle motions using soft strain sensors. *IEEE Robotics and Automation Letters*, 5(2):3525–3532, 2020.

[22] Gerard Pons-Moll, Javier Romero, Naureen Mahmood, et al. Dyna: A model of dynamic human shape in motion. *ACM Transactions on Graphics*, 34(4):1–14, 2015.

[23] Thiemo Alldieck, Marcus Magnor, Weipeng Xu, et al. Video based reconstruction of 3d people models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8387–8397, 2018.

[24] Chao Zhang, Sergi Pujades, Michael J Black, et al. Detailed, accurate, human shape estimation from clothed 3d scan sequences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4191–4200, 2017.

[25] Tao Yu, Zerong Zheng, Kaiwen Guo, et al. Doublefusion: Real-time capture of human performances with inner body shapes from a single depth sensor. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7287–7296, 2018.

[26] Marc Habermann, Weipeng Xu, Michael Zollhoefer, et al. Reticam: Real-time human performance capture from monocular video. *arXiv preprint arXiv:1810.02648*, 2018.

[27] Georgios Pavlakos, Luyang Zhu, Xiaowei Zhou, et al. Learning to estimate 3d human pose and shape from a single color image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 459–468, 2018.

[28] Zifeng Wang, Yan Huang, Jinfeng Sun, Yang Huang, Hong Hu, Ruijuan Jiang, Weiming Gai, Guangming Li, and Chunyi Zhi. Polyurethane/cotton/carbon nanotubes core-spun yarn as high reliability stretchable strain sensor for human motion detection. *ACS Applied Materials & Interfaces*, 8(37):24837–24843, 2016.

[29] Naibo Lin, Liwei Cao, Qiaoling Huang, et al. Functionalization of silk fibroin materials at mesoscale. *Advanced Functional Materials*, 26(48):8885–8902, 2016.

[30] International Organisation for Standardization. *3-D Scanning Methodologies for Internationally Compatible Anthropometric Databases (ISO 20685-1:2018)*. 2010.

[31] Standardization Administration in China. *General requirements for 3-D scanning anthropometric methodologies (GB/T 23698)*. 2009.

[32] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[33] Leonid Pishchulin, Stefanie Wuhrer, Thomas Helten, et al. Building statistical shape spaces for 3d human modeling. *Pattern Recognition*, 67(C):276–286, 2017.

[34] R. W. Sumner and J. Popovió. Deformation transfer for triangle meshes. *ACM Transactions on Graphics (TOG)*, 23(3):399–405, 2004.

[35] Robert Maier, Kihwan Kim, Daniel Cremers, et al. Intrinsic3d: High-quality 3d reconstruction by joint appearance and geometry optimization with spatially-varying lighting. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3114–3122, 2017.