2020 16th IEEE International Conference on Automation Science and Engineering (CASE)
August 20-21, 2020, Online Zoom Meeting

VTacArm: A Vision-based Tactile Sensing Augmented Robotic Arm with Application to Human-robot Interaction

Yazhan Zhang*[1], *Student Member, IEEE,* Guanlan Zhang*[1], Yipai Du[1] and Michael Yu Wang[2], *Fellow, IEEE*

***Abstract*— Endowing artificial sense of touch comparable to human's has been challenging, yet significant to enabling adaptive and collaborative interaction in contact-rich tasks. This work is dedicated to proposing a novel vision-based tactile sensor augmented robot arm (VTacArm) design with full surface coverage and developing algorithms to retrieve contact information which is essential for down-stream feedback control. We first introduce the robot arm design and its accompanying fabrication process. Then, to convert the contact signals in the image space to the arm coordinate system, a calibration procedure and method are proposed. Finally, the tactile robot arm and the contact information extraction algorithm are integrated into a control system for collaborative interaction tasks. Bumping detection/reaction and contact motion following experiments are presented to justify that the designed tactile robot arm and proposed contact sensing method are beneficial and give robot capabilities to adapt to human contacts, which is vital for workers' safety. Our work can be informative for developing novel full-body vision-based tactile sensing on robots as a new concept with significantly lower cost and manageable fabrication complexity.**

I. INTRODUCTION

Tactile sensing has been investigated and proven to play critical roles in human interaction with the environment. For robotic systems, artificial skin that enables robots to feel and react to physical contact is attracting more and more attention and becomes an inevitable demand as robots come increasingly close contact with people. Manufacturers worldwide use a growing ratio of industrial robots to human workers to perform collaborations in a contact-rich fashion in recent years. In consequence, safe interaction emerges as a major concern. By detecting and regulating the contact force, robots with artificial skin can achieve tasks without dealing damage to objects and human workers.

However, for most current robots, the sense of touch is absent and underdeveloped in two levels. First, the conventional tactile sensor is too expensive to be fully deployed onto robot systems due to the complex fabrication processes and clumsy signal acquisition systems developed for the tactile sensors [1]. Although a full-body tactile sensing system [2] is very appealing due to its perception ability to external environment surrounding the robot, this scale of skin can be much more expensive than the robot arm itself and probably leads to tedious fabrication and being break-down prone. Besides, signal processing and inference based on electronic readings to physical contact information can be very computational demanding for large amount of tactile sensor units on a robotic system. Second, most tactile sensors fabricated in the past lacked multidimensional deformation sensing capability. For example, sensors with capacitive, piezoelectric or piezoresistive transducing interfaces usually can only respond to normal pressing, but not to tangential dragging, which is, in contrast, common for human skin. Therefore, a low-cost, multidimensional appreciable, full-body deployable, computational efficient artificial skin is crucial for bringing robots closer to human working space.

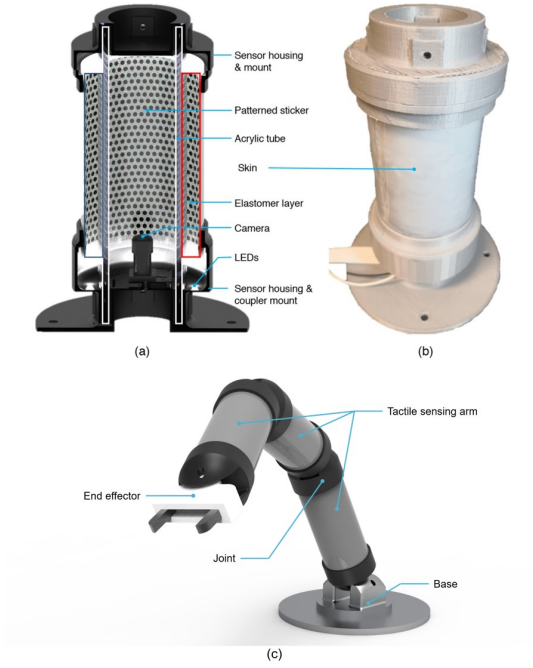

Fig. 1. VTacArm robot arm design skematic. (a). Rendered 3D model (in a cutaway view). (b). VTacArm sensor prototype. (c). A robot arm consists of three VTacArm sensors (rendered by software)

Recently, as an advance on tactile sensor technologies, vision-based tactile sensors have been emerging in various robotic systems with advantages of easy fabrication, high resolution, and multi-axial deformation sensing capability, e.g. Gelforce [3], FingerVision [4], Gelsight [5] and a compact version called Gelslim in [6]. Algorithms for inferencing contact information are developed in multiple levels such as contact detection, force sensing [7] and contact events detec-

This work was supported by the Hong Kong Innovation and Technology Fund (ITF) ITS-018-17FP.

[1]Y. Zhang, G. Zhang and Y. Du are with the Department of Mechanical and Aerospace Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: yzhangfr@connect.ust.hk; guanlan.zhang@connect.ust.hk; yipai@ust.hk).

[2]M. Y. Wang (corresponding author) is with the Department of Mechanical and Aerospace Engineering and the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (tel.: +852-34692544; e-mail: mywang@ust.hk).

* denotes authors with equal contribution.

tion and prediction [8]. These applications of vision-based sensors demonstrated the superiority as sensor units on end-effectors mainly for object manipulation. Attaching robot arm with vision-based sensors is unexplored previously. In this work, we expand the application of the vision-based tactile sensor to a full-body skin robot system by proposing novel sensor design called **VTacArm** and accompanying contact information acquisition method to reflect that the vision-based sensor is of potential to cover the range of human sense of touch on body surface beyond fingertips.

As shown in Fig. 1, the VTacArm design is composed of two major components: torus-shape elastomer layer as skin and camera installed at one end as the transduction interface. Dot grid pattern is attached on the outer surface of the clear gel layer as the tracking target. Once external force is applied onto the skin, the elastomer deforms accordingly following continuum mechanics and the displacements of dots are captured by the monitoring camera to obtain a deformation field. Afterwards, the deformation field in the image domain is further back projected to the sensor coordinate frame via the calibrated projection matrix and sensor model in order to facilitate correct regulation to the force stimuli.

The proposed tactile augmented robotic arm features multidimensional force sensing, which provides necessary input to various tasks in human-robot interaction. There are two important tasks in contact-rich human-robot interaction: 1) bumping detection/reaction; 2) contact motion following. Bumping detection/reaction involves contact detection and sending command afterwards to the robot to stop motion or retract. In the task of contact motion following, human co-worker grab and hold a certain part of the robot and exert force aiming to transform the robot between configurations, mimicking an adult human teaching a child handwriting by holding the child's hand. In our paper, with the touch sensation acquired from the VTacArm, our robot adapts to the contact made by human worker tasked with specific human-robot interaction goals.

This work contributes in mainly two perspectives:

- Presenting a novel robotic arm design fully covered with vision-based tactile sensing skin and demonstrating the feasibility of the full-body contact information acquisition and extraction.
- Developing calibration process and devising necessary calibration instruments. Integrating the design sensor and contact information extraction algorithm into human-robot interaction tasks.

The paper is structured as follows. Section II introduces previous works related to to robotic full-body skin, vision-based tactile sensors and human-robot interaction. In section III, we extensively describe the design and fabrication of the sensor proposed and method to transform image signal to contact information. In section IV, experiments including calibration evaluation and two human-robot interaction tasks are elaborated and analyzed. In section V, conclusion are drawn finally.

II. Related Works

A. Whole-body Artificial Skin

Human body has tactile sensation distributed with varying density, achieving trade-off between functional effectiveness and energy/attention efficiency [9]. So, for robotic system, artificial skin is not just a patch of sensing matrix with high resolution, but a highly distributed sensor system [1][2]. There are few cases of the realization of whole-body artificial skin to application level. M. Toshiharu et al. in [10] implemented normal force sensing units on a nursing robot's body and arms for tasks like hugging and lifting of humans reactively. I. Kumagai et al. in [11] presented a full-body multi-axis tactile sensor suit for a humanoid robot to detect contact states in an object receiving and releasing tasks collaborated with human. Beyond the normal force sensing, the sensor developed in this work provides tangential force sensation. These two works installed standalone sensor units onto robots' surfaces with different spacing at different positions. And Each sensor unit was powered and signal was acquired separately, which would be challenging if the density of sensors were to be increased.

There are previous works emphasizing on conformation of sensors onto robots' body surface, as human skin does. Y. Ohamura and Y. Kuniyoshi in [12] introduced a humanoid with whole body tactile sensing capability to manipulate heavy objects or human-shape objects. The sensors used in their work are printable on compliant sheet and could be "cut and paste" to robot body surface conveniently. Besides, the sensor patches were connected serially, which alleviates the burden of power wiring and signal multiplexing. Advancing on systematic approach to realize whole-body tactile sensing on robots in both sensor capabilities and algorithms to process numerous units efficiently, P. Mittendorfer et al. [2] developed a humanoid robotic system covered with multi-modal tactile cells connected with revolute joints to retain compliance. And event-based control framework used reduced the frequency required for reading processing and response.

The previous attempts to construct whole-body artificial skin generally suffered from one of these drawbacks if not all: lower density, high data processing bandwidth requirements and complex power wiring and signal multiplexing. In our work, we design a robot arm skin using only one camera provide robot with sensation of multidimensional contact force. Our design is superior in both fabrication and signaling.

B. Vision-based Tactile Sensors

Vision-based tactile sensors are gaining more and more attention for its superior tactile resolution, easy fabrication, robust electronics, compact form-factor and simple multiplexing peripherals. In the perspective of the versatility in feature extraction operation in different levels, vision-based tactile sensor is capable of acquiring multi-modal contact information including deformation [4], texture [13]–[15], contact area localization [6][16], geometry reconstruction [13][17] and force estimation [3][4][17][18][7]. Beyond

these low-level contact information, vision-based tactile sensors have been performing effectively in high-level tasks like object recognition [15], localization of dynamic object [19], simultaneous localization and mapping of the sensor on objects [20], slip detection [5][21][14][22] and a fine-grained contact events classification [8].

Most of the previous implemented vision-based tactile sensors are targeting at resolving contact signals at fingertips [3][8][15][23]. However, there exists other forms of vision-based tactile sensors. L. Cramphorn et al. in [24] fabricated a dome-shape tactile sensor called TacTip that was installed as an end effector on a manipulator. Similarly, B. W. McInroe et al. [18] presented an arm tip with integrated tactile sensing and pneumatic actuation. Tactile sensors on fingertips or on arm tips are unable to cover sufficiently large body area that is preferred by agile adaptation to external contacts in human-robot interaction tasks. In this work, we show that a skin covering a whole arm can also be realized using vision-based tactile sensor with a novel design and proper signal processing.

C. Contact-rich Human-robot Interaction

To ensure safety during human-robot interaction, apart from artificial skin, adaptation policies of robot actions are essential. In [2], the robot armored with whole-body tactile sensors executed configuration transformation following the trajectory generator's commands and adjusted simultaneously by tactile events including approach, contact, load and pull states. G. Cannata et al. in [25] devised a control strategy to maintain contact made by human so as to following the trajectories of contact points. J. M. Romano et al. [26] were inspired by human control scheme of tactile sensation and proposed method to process artificial tactile information to generate action primitives during a manipulation action. In our work, we design simple human-robot interaction tasks including bumping detection and motion following to demonstrate the agility advantage in terms of due to the dense and multi-axial tactile information provided by the VTacArm sensor.

III. System and Methods

In this section, a detailed description of the sensor design, fabrication process and processing algorithms to retrieve useful contact information are elaborated. As illustrated in Fig. 1, the VTacArm sensor is composed of soft skin embedded with marker array, webcam module for monitoring, and peripheral illumination components and housing parts for convenient mounting. Each VTacArm sensor section can measure contact information individually. And if they are connected like (c), a robot arm with full-body vision-based tactile sensing is achieved.

The principle of the sensor is straightforward. Once the skin makes contact with external objects and deforms, the camera captures the changed image of the marker array. The image of the marker array is processed through marker localization and the positions of markers are associated by a tracking algorithm to obtain the spatiotemporal deformation field of the skin surface. More specifications will be detailed in the following subsections.

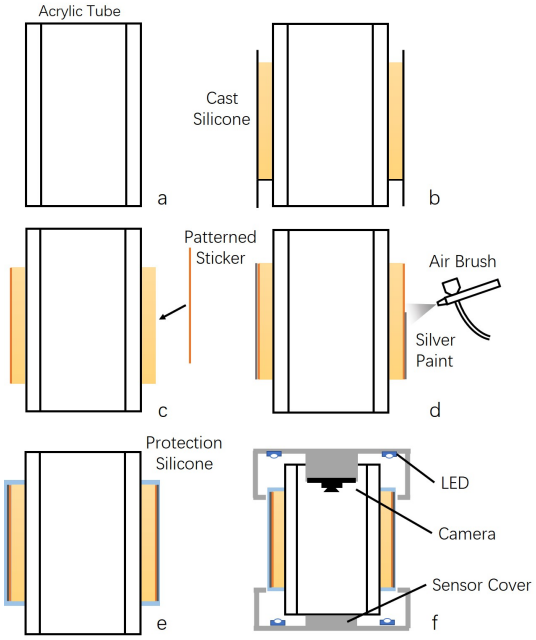

Fig. 2. Sensor fabrication process. (a). An acrylic tube is cut into desired length. (b). Proper amount of silicone rubber solvent is poured into a mold to form a elastomer layer with uniform thickness around the acrylic tube. (c). A patterned sticker is adhered onto the rubber surface (d). Mixture of silver pigment is sprayed onto the sticker layer and baked till dry. (e). Another silicone rubber is coated on pigment layer and cured. (f). Camera, LED, tube with skin are assembled.

A. Sensor Fabrication

Fabrication of the tactile robot arm takes account of various aspects. First, the camera's field of view (FOV) is required to cover the large area of an 100mm-long tubular elastomer with radius of 25mm from the interior of the acrylic tube. We chose Raspberry Pi Camera Module v2 with a fisheye lens, which has a FOV of 160 degree and an adjustable focal length. The camera is installed at one end of a transparent acrylic tube, with length of 175mm, outer diameter of 50 mm and thickness of 4 mm, so that the strength of the sensor's main body to withstand external impact is guaranteed.

The medium of deformation in this sensor is a transparent, hyperelastic and durable silicone rubber (shore hardness of 20A close to the hardness of human skin). The solvents of two-part silicone rubber are mixed in an $1:1$ ratio and cured in a two-piece mold for casting. After formed, an elastomer layer is adhered firmly onto the outer surface of the acrylic tube and serves as the deformation interfacing substrate. In order to acquire the deformation of the skin, a thin patterned sticker is adhered onto the surface of elastomer layer. This sticker with marker array is printable on a thin flexible adhesive film by laser jet printers. Being both trade-off problems, the density of the markers are chosen as high as possible while constrained by tracking error rate, and the marker size is the smaller the better while constrained by the

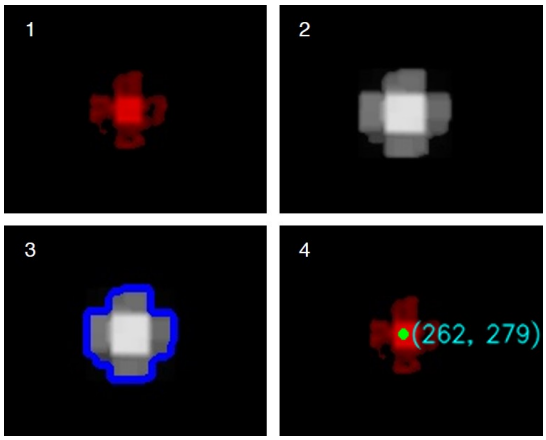

Fig. 3. Results generated from the image processing pipeline. (1). Optical flow obtain from DIS on image sequence (image cropped to zoom in). (2). Image obtained from maximum filter. (3). Patch contour identification. (4). Contact Patch center localization.

marker localization inaccuracy. Therefore, the marker density and size need to be settled empirically in experiments. A thin frosted paint layer made of limonene, silicone and silver pigment is sprayed on the patterned sticker. It isolates potential disturbing lights, blocks background scenes and also disperse lights from the LEDs to form a relatively more uniform illumination inside the internal of the sensor. A more durable silicone rubber is coated as the outermost layer to reinforce the sensor.

Since the sensor is fully covered with opaque material, proper illumination is necessary. 8 LEDs are mounted on both side of the tube. A semi-transparent frosted plastic plate is placed in front of each LED to generate a more diffused light inside the tube. Two 3D printed parts are used to fix the tube, LEDs and camera and provide coupling mount to the robot arm. The whole schematic procedures of the fabrication are shown in Fig. 2.

B. Extraction of Contact Information

The main functionality of the VTacArm is to retrieve deformation on the cylindrical elastomer surface when making contact with objects, from which useful features, i.e. contours and centers of contact patches can be extracted for downstream applications. Fig. 3 shows the pipeline of processing raw images to acquire contact information. Images sequence is captured by the camera at a rate of 40 fps with an initial resolution of 1640×1232. A mask is superimposed to the image, trimming out the region of interests (ROI). Then the masked image is resized to 512×288 to increases the computation speed while still preserving a decent accuracy of contact patch estimation.

Fast Optical Flow with Dense Inverse Search (DIS) We utilize dense inverse search (DIS) optical flow algorithm [27] featuring more efficient computation to obtain a deformation vector field from the image sequence at a higher frequency. The algorithm yields optical flow $\mathbf{U}_s$ by finding a warping vector $\mathbf{u} = (u, v)$ for each template patch T in reference image I_t, which minimize the difference between patches in

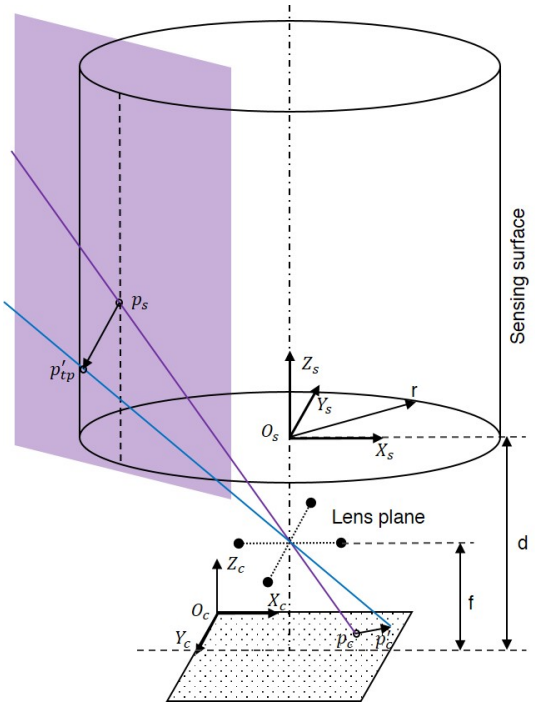

Fig. 4. The perspective projection schematic diagram of the sensor imaging and back projection of point and vector from image domain.

reference image and query image.

$$\mathbf{u} = \text{argmin}_{\mathbf{u}'} \sum_{x} [I_{t+1}(\mathbf{x} + \mathbf{u}') - T(\mathbf{x})]^2 \tag{1}$$

where $\mathbf{x} = (x, y)^T$ is the center of patch T in image I_t, which in our case is the initial static frame of pattern, and $I_{t+1}(\mathbf{x}+\mathbf{u}')$ is the best matched sub-window of $T(\mathbf{x})$ in query image I_{t+1}, which in our case, is the n^{th} frame.

$\mathbf{U}_s$ is iteratively generated from coarsest level (with largest patch size) to finest level (with smallest patch size). And in each iteration, the quality of optical flow is improved by Variational Refinement. Given the vector field, vector magnitudes larger than zero represent the displacement of pattern, which are used to infer the deformation of elastomer.

Contact Identification and Center Localization This part involves multiple image processing methods. To estimate the contact patches, the deformation field is first thresholded with a hand-picked magnitude value to filter out noises that can stem from lighting reflection, imaging process factors and etc. Then we apply a maximum filter to dilate the value of local maxima of vector magnitude into its neighbors to compensate loss of contact area due to the previous thresholding operation and then binarize the deformation field image with regions of ones denoting active contact patches and regions of zeros denoting non-contact area. Next, a contour finder marks the boundary of contact areas. After the contact patches are identified, the geometric center is obtained by averaging the locations of all the points inside an active contact patch.

Calibration of Contact Points It is needed to project center of contact patches $p_c = (x_c, y_c)^T$ on image domain

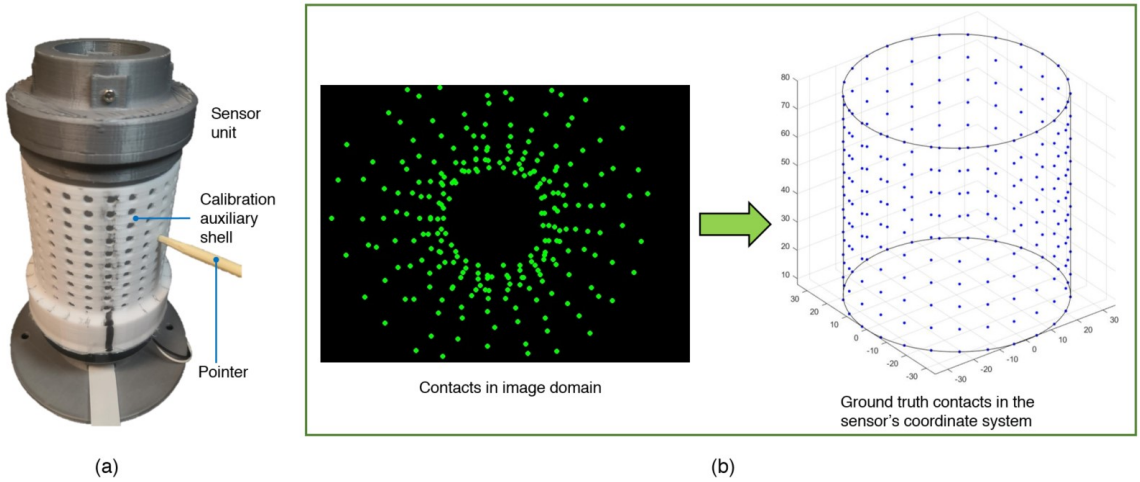

Fig. 5. Experimental setup of the data collection for calibration. (a). Instruments for positioning the contact points. (b). Acquired contact data in image domain (2D) and predetermined Cartesian positions on sensor surface (3D).

back to the points on the surface of the cylindrical elastomer $p_s = (x_s, y_s, z_s)^T$ in the sensor's coordinate system, which is defined as 20 mm above the center of the camera lens surface and Y-axis flipped, to reconstruct contact in physical space. To fulfill this goal, we propose calibration data collection procedure as follows:

1) A 3D printed shell is fixed w.r.t the sensor and covers the active sensing area tightly without making interfering contacts. The shell has 250 holes of 4 mm diameter with predetermined position on a grid w.r.t C_s as $p_h = (x_h, y_h, z_h)^T$ on its wall. The instruments are given in Fig. 5(a).
2) A stick with 3.8mm diameter probes the sensor surface through the holes one by one in order. Each time, the contact point p_s on the sensing surface is calculated based on p_h (there is a small gap between the shell inner surface and the sensor's outer surface):

$$p_s = \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot p_h, \tag{2}$$

where $k = 0.97$ is the ratio between the inner diameter of the shell and the outer diameter of the elastomer layer.

Using this dataset collected, we build up the correspondence between points in image domain and points in 3D sensor surface, shown in Fig. 5(b). Fig. 4 illustrates the perspective projection with additional cylindrical constraint in the camera system. Given a set of 2D-3D point correspondences (e.g. $p_c - p_s$), the projection matrix that maps points in 3D space onto 2D image domain is formulated as the equation below [28].

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R\mathbf{p}_s - \tilde{C} \tag{3}$$

$$\begin{aligned} x' &= x/z \\ y' &= y/z \end{aligned} \tag{4}$$

$$\begin{aligned} x'' &= x' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2p_1 x' y' + p_2(r^2 + 2x'^2) \\ y'' &= y' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1(r^2 + 2y'^2) + 2p_2 x' y' \end{aligned} \tag{5}$$

where

$$r^2 = x'^2 + y'^2.$$

Finally

$$\begin{bmatrix} \mathbf{p}_c \\ 1 \end{bmatrix} = K \begin{bmatrix} x'' \\ y'' \\ 1 \end{bmatrix} \tag{6}$$

where K, R and $\tilde{C}$ represent 3×3 camera calibration matrix, 3×3 rotation matrix of the camera coordinate O_c and the coordinate of camera center (3×1) w.r.t. the sensor coordinate frame, respectively. Here we model the lens distortion using rational model provided in [29], which includes 8 distortion parameters. With redundant point pairs provided, K, R, $\tilde{C}$ and 8 distortion parameters can be solved using Direct Linear Transformation (DLT) [28]. As shown in Fig. 4, when the sensor is used in experiments, points and vectors in image domain are supposed to be back projected to be points and vectors on the cylindrical surface. The function of the cylindrical surface is formulated as a constraint of x and y coordinate values in coordinate frame O_s, which is determined by matrix R and $\tilde{C}$.

To map a vector that represents the surface stress or total force in a contact patch, we first construct a vertical tangential plane (in purple in Fig. 4) at the back projected point p_s of the vector's start point p_c in the image domain, formulated as $p_{sx}x + p_{sy}y = p_{sx}^2 + p_{sy}^2$. Considering vectors on the image plane all correspond to tangential vectors on the sensor surface, the end point such as p_c' is back projected

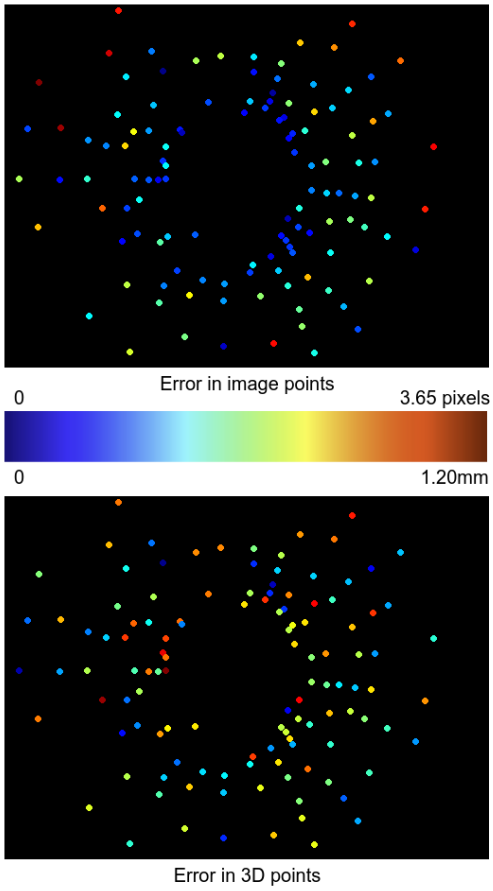

Fig. 6. The forward and backward projection error of individual data points that we used for calibration. They are both normalized and plotted using the same color map.

as a ray and the ray intersects with the tangential plane to obtain the corresponding vector's end point p'_{tp} in 3d space, as sketched in Fig. 4.

IV. Experiments and Results

In this section, we demonstrate the principle of VTacArm by using one sensing section and present the utilization of the dataset obtained from the calibration experiments in the computation of the projection matrix of the imaging system in a quantitative way. Based on the back projection of points and vectors in image domain, the contact patches and contact stress vectors can be mapped onto the sensor surface for both visualization and down-stream adaptive control during the interaction between human and robot.

A. Camera Calibration and Evaluation

In the process we described in section III-B, there are a total of 250 candidate point-pairs that can be utilized for camera calibration. Compared to the 11 unknown parameters in the calibration matrix (5 for intrinsic parameters, 3 for rotation and 3 for translation), this is an over-determined system. Moreover, in the dataset there might exist outliers that cannot reflect the true projection relationship accurately. Instead of using sampling based methods like RANSAC, we consider a simpler procedure for outlier removal. First, we perform the aforementioned calibration process over the full dataset containing 250 data pairs. Then we use the obtained calibration result to back project the 2D image points into the 3D space. Compared with ground truth 3D locations, we remove all point pairs exceeding the $2mm$ error threshold. 131 points are removed and 119 points remain. Finally, we reiterate the calibration process with the remaining data points. With the obtained calibration matrix, we achieved an average error of **3.65** pixels from 3D to 2D, and an average of **1.20**mm from 2D to 3D. The point locations and their corresponding errors are shown in Fig. 6. Note that these points are filtered by the outlier removal procedure. For the case of 3D to 2D, when the point is further away, its projection on the image is more accurate because it tends to accumulate around the center of the image. For 2D to 3D, it's the opposite. Inaccuracy in the measurement of pixel coordinate will cause a drastic change in the 3D location for points closer to the center of the image. For our application our main interest is in projecting image points to 3D world, and this error plot indicates our working range should be focused on the lower part of the arm (closer to the camera). How to enlarge the working range to cover higher parts of the arm might be an interesting future work.

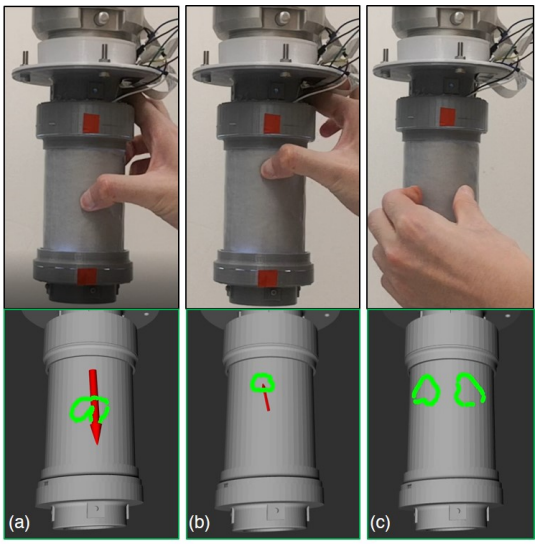

Fig. 7. Visualization of contact contours and contact displacement vector during a contact session. Real contact at the top row and simulated visualizations at the bottom row. (a). Contact contour at the start. The red vector is the contour displacement vector of last contact session. (b). Contour at the end and contact displacement vector. (c). Multiple contact patches.

B. Contact Information Visualization

Before integrating the sensor and contact information extraction unit into actual feedback control loop, we first visualize the contact information to show the benefit by using the sensor as one section of a robot arm during interaction. As captured in Fig. 7, the tactile sensor is mounted onto a UR robot arm's end flange. An human operator place a finger on the sensor sensing surface and then move the finger relative to the sensor surface without losing contact, shown in Fig. 7(ab). The visualized contours (in green color) are

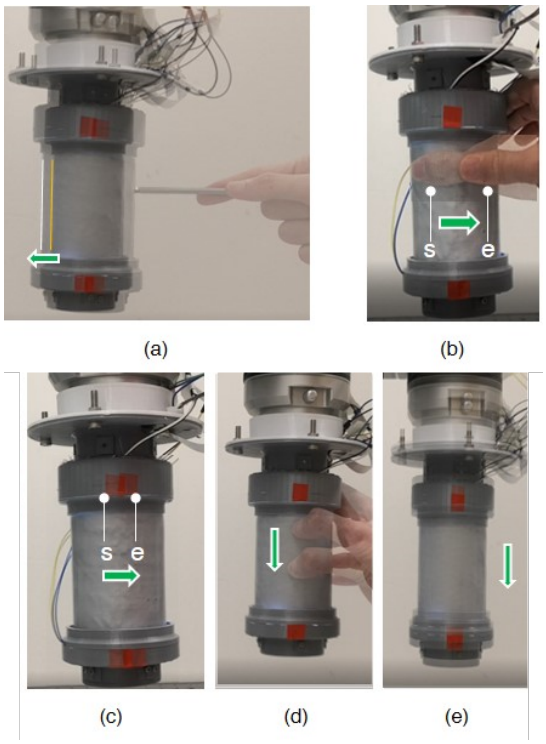

Fig. 8. Bumping detection/reaction and motion following experiments. Motions of the robot arm are drawn with arrows. (a). Bumping detected and the robot arm retrieves to avoid being damaged. (b). Fingertip contacts the sensor surface and drive around a horizontal curve. s: start, e: end. (c). The robot arm responses with rotation. (d). Fingertip contacts the sensor surface and drives vertically across a track. (e). The robot arm responses with a vertical translation.

illustrated in a real-time updated interface (via RViz and ROS) in the bottom row of Fig. 7(ab). When a contact and moving period is over, a vector (in red color) denoting the contour center displacement is rendered on the sensor surface in the simulated interface Fig. 7(b). The sensor is also capable of detecting multiple contacts and projecting contacts in image domain to the sensor surface. As given in Fig. 7(c), the human operator makes contact with the sensor surface at two contact sites, simultaneously, the contours are shown on the sensor surface. Via the experimental results, it is clear that the visualization resolution obtained with our sensor is much higher compared with flexible sleeves embedded with electric-transduction based tactile sensors.

C. Adaptation in Interaction Experiments

In this experiment, two sub-experiments showing human-like behaviors that are desired in human-robot interaction (HRI) tasks will be implemented: bumping detection/reaction and contact motion following.

For bumping detection and reaction experiment, a human operator pokes the sensor surface with a screwdriver at a contact point as shown in Fig. 8(a). Then the contact position is detected by contact patch identification module. The contact point localized is back projected to the sensor surface for calculation of the retrieving vector that points to the direction of the robot arm motion. For demonstration purpose, the retrieving vector is simply the inverse normal vector at the contact point passing through the axis of the cylinder. As illustrated in Fig. 8(a), after the poking, the robot instantly move in the direction of the poking to avoid further poking which might cause damage to the sensor, mimicking a natural action human might take after receiving contact capable of causing harm.

In the experiment of motion following at contact point, a human operator makes contact with the sensor's surface and drives certain motion for the robot arm to follow. In Fig. 8(bc), when a surface traction with a moment w.r.t. the cylinder axis is applied, the robot arm rotate along the cylinder axis by an angle determined by the contact patch displacement from the start site to the end site. In Fig.8(de), the human operator move above the sensor surface along the cylinder axis without exerting moment, the robot arm response with a pure translation along the cylinder axis. The covered distance equals to the displacement of contact patch from the start to the end. From these two experiments, we successfully code the robot arm to behave like a human given different interaction tasks including damage avoiding and drag following. Furthermore, these interactions can be triggered at any location on the sensor surface, which is superior compared tactile array on robot arm with much lower density.

V. Conclusion

In this work, we develop a full body vision-based tactile sensing augmented robot arm section called VTacArm with a novel imaging system. The sensor features easy fabrication, simple multiplexing and multi-axial deformation sensing. To extract contact information, we develop a pipeline of image processing to not only identify contact patches but also localize the contact sites and degrees of deformation. We devise a system with unique instruments and method to calibrate the sensor, from which sensor model is estimated and further utilized to back project contact features in the image domain to the sensor coordinate frame. After integrating the sensor and contact information extraction modules, experiments that showcasing the robot's capability to mimic human actions when interacting with environment are implemented. Specifically, bumping detection and reaction, motion following at contact sites are presented. Experimental results reflects the merits of the sensor designed including better resolution, multi-axial sensing ability, larger sensible area, pushing full-body tactile sensing one step forward.

References

[1] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing-from humans to humanoids." *IEEE Trans. Robotics*, vol. 26, no. 1, pp. 1–20, 2010.

[2] P. Mittendorfer, E. Yoshida, and G. Cheng, "Realizing whole-body tactile interactions with a self-organizing, multi-modal artificial skin on a humanoid robot," *Advanced Robotics*, vol. 29, no. 1, pp. 51–67, 2015.

[3] K. Sato, K. Kamiyama, N. Kawakami, and S. Tachi, "Finger-shaped gelforce: sensor for measuring surface traction fields for robotic hand," *IEEE Transactions on Haptics*, vol. 3, no. 1, pp. 37–47, 2010.

[4] A. Yamaguchi and C. G. Atkeson, "Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 1045–1051.

[5] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, "Measurement of shear and slip with a gelsight tactile sensor," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 304–311.

[6] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1927–1934.

[7] Y. Zhang, Z. Kan, Y. Yang, Y. A. Tse, and M. Y. Wang, "Effective estimation of contact force and torque for vision-based tactile sensors with helmholtz–hodge decomposition," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4094–4101, 2019.

[8] Y. Zhang, W. Yuan, Z. Kan, and M. Y. Wang, "Towards learning to detect and predict contact events on vision-based tactile sensors," *arXiv preprint arXiv:1910.03973*, 2019.

[9] R. S. Johansson and A. Vallbo, "Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin." *The Journal of physiology*, vol. 286, no. 1, pp. 283–300, 1979.

[10] T. Mukai, M. Onishi, T. Odashima, S. Hirano, and Z. Luo, "Development of the tactile sensor system of a human-interactive robot "riman"," *IEEE Transactions on robotics*, vol. 24, no. 2, pp. 505–512, 2008.

[11] I. Kumagai, K. Kobayashi, S. Nozawa, Y. Kakiuchi, T. Yoshikai, K. Okada, and M. Inaba, "Development of a full body multi-axis soft tactile sensor suit for life sized humanoid robot and an algorithm to detect contact states," in *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*. IEEE, 2012, pp. 526–531.

[12] Y. Ohmura and Y. Kuniyoshi, "Humanoid robot which can lift a 30kg box by whole body contact and tactile feedback," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 1136–1141.

[13] M. K. Johnson and E. H. Adelson, "Retrographic sensing for the measurement of surface texture and shape," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009, pp. 1070–1077.

[14] S. Dong, W. Yuan, and E. H. Adelson, "Improved gelsight tactile sensor for measuring geometry and slip," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 137–144.

[15] W. Yuan, S. Wang, S. Dong, and E. Adelson, "Connecting look and feel: Associating the visual and tactile properties of physical materials," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5580–5588.

[16] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. Adelson, "Cable manipulation with a tactile-reactive gripper," 2019.

[17] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[18] B. W. McInroe, C. L. Chen, K. Y. Goldberg, R. Bajcsy, and R. S. Fearing, "Towards a soft fingertip with integrated sensing and actuation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 6437–6444.

[19] R. Li, R. Platt, W. Yuan, A. ten Pas, N. Roscup, M. A. Srinivasan, and E. Adelson, "Localization and manipulation of small parts using gelsight tactile sensing," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 3988–3993.

[20] M. Bauza, O. Canal, and A. Rodriguez, "Tactile mapping and localization from high-resolution tactile imprints," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3811–3817.

[21] Y. Zhang, Z. Kan, Y. A. Tse, Y. Yang, and M. Y. Wang, "Fingervision tactile sensor design and slip detection using convolutional lstm network," *arXiv preprint arXiv:1810.02653*, 2018.

[22] K. Van Wyk and J. Falco, "Slip detection: Analysis and calibration of univariate tactile signals," *arXiv preprint arXiv:1806.10451*, 2018.

[23] A. Yamaguchi and C. G. Atkeson, "Implementing tactile behaviors using fingervision," in *Humanoid Robotics (Humanoids), 2017 IEEE-RAS 17th International Conference on*. IEEE, 2017, pp. 241–248.

[24] L. Cramphorn, J. Lloyd, and N. F. Lepora, "Voronoi features for tactile sensing: Direct inference of pressure, shear, and contact locations," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2752–2757.

[25] G. Cannata, S. Denei, and F. Mastrogiovanni, "Contact based robot control through tactile maps," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 3578–3583.

[26] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1067–1079, 2011.

[27] T. Kroeger, R. Timofte, D. Dai, and L. Van Gool, "Fast optical flow using dense inverse search," in *European Conference on Computer Vision*. Springer, 2016, pp. 471–488.

[28] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[29] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.