2021 Fifth IEEE International Conference on Robotic Computing (IRC) | 978-1-6654-3416-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/IRC52146.2021.00015

FingerVision with Whiskers: Light Touch Detection with Vision-based Tactile Sensors

1st Akihiko Yamaguchi
Graduate School of Information Sciences
Tohoku University
Sendai, Japan
info@akihikoy.net

***Abstract*—In order to increase the sensitivity of the vision-based tactile sensor FingerVision especially to normal force, we explore an idea of introducing *whiskers* as an alternative to markers. Since whiskers deforms more largely against light touch than markers, the image change of the FingerVision camera is larger, and thus the resolution of force is increased. In this paper, we fabricate and compare some versions of FingerVision with whiskers. The empirical comparisons demonstrate that the idea to introduce whiskers work as we expect. This paper also demonstrates intuitive applications, pulling a tissue paper grasped by a robot and poking a standing pencil by a robot.**

***Index Terms*—Tactile sensor, FingerVision, Robotic manipulation**

I. Introduction

We are exploring vision-based tactile sensing [1]–[11]. Because of high resolution, high reliability, and ease of fabrication, we think this is a promising approach for robotic manipulation. Our version of such a vision-based tactile sensor is named FingerVision that has a completely transparent skin which enables multimodal sensing (force, slip, etc.) [10], [11]. We have made many applications of robotic manipulation with FingerVision such as slip-feedback grasping.

FingerVision consists of a transparent elastic skin where markers (dots) are placed around the surface, a transparent hard layer made with acrylic, and a camera; see Fig. 1(a). Some different computer vision methods are used to process captured video from the camera. In our implementation [11], we used a blob tracking method to detect the movement of the markers for estimating force distribution, and a background subtraction to detect movement of proximity objects for estimating slip distribution.

An issue of FingerVision is that although it is sensitive and accurate to shear force, it is less sensitive and less accurate to normal force. The reason is that shear force is estimated from horizontal movement of the markers which is large displacement in FingerVision images, while normal force is estimated from marker size change which is subtle in FingerVision images. As explored in [11], FingerVision is capable to detect slip sensitively which enables standard grippers to grasp unknown objects without breaking them even if they are fragile and/or lightweight, such as origami arts. Slip-feedback control is used to find a good grasp, however slip is a phenomenon caused by robot or object motion, i.e. the feedback control using slip takes time. One way to solve this issue is learning grasp force for an unknown object. For that purpose, the sensitivity and accuracy of normal force is necessary.

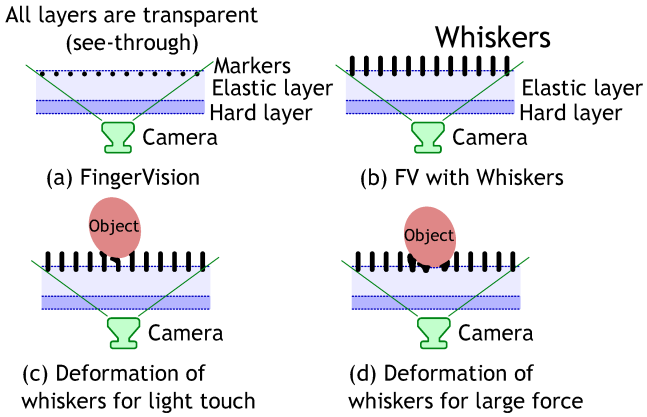

Fig. 1. (a) Conceptual design of FingerVision, (b) replacing its markers with whiskers, (c) deformation of whiskers for light touch, and (d) deformation of whiskers for large force.

Animals have *whiskers*, and much of them are used as sensors. Since they are light weight, they are sensitive to light touch or contact with very small force. Such a functionality is also attractive from engineering point of view. For example, a light weight whisker-like sensor was developed which was able to sense light touch [12].

In order to make FingerVision sensitive and accurate to normal force, we explore an idea to introduce whiskers onto the surface of FingerVision. Fig. 1(b) shows a conceptual illustration. Those whiskers are thin, soft, and placed outside of the skin. When an object contacts some whiskers, they will easily deform (Fig. 1(c)). When a stronger force is applied from the object, the whiskers deform more together with the surrounding elastic skin (Fig. 1(d)). These deformations will be shown in the FingerVision camera, and computer vision will be able to track them. With those whiskers, we can improve the sensitivity of FingerVision to normal force.

This paper explores the concept of whiskers by prototyping some versions, and empirically comparing them with the other types. In the experiments, we also introduce an approach to remove the hard layer (acrylic plate) from FingerVision, which

978-1-6654-3416-4/21/$31.00 ©2021 IEEE
DOI 10.1109/IRC52146.2021.00015

increases the sensitivity since the skin deforms more.

We fabricated three versions of FingerVision with whiskers, FingerVision without hard layer, and a standard FingerVision, and conducted experiments to compare the sensitivity. The experimental results demonstrate that the idea to introduce whiskers work. We also show two examples, pulling a tissue paper grasped by a robot and poking a standing pencil by a robot. They intuitively demonstrate the advantages of FingerVision with whiskers.

The rest of this paper is organized as follows: Section II introduces related work. Section III shows preliminary prototypes. Section IV proposes FingerVision with whiskers. Section V reports the experiments. Section VI concludes this paper.

II. Related Work

The idea of using imaging sensors for tactile sensing is decades old. A recent survey [13] introduces many of them. An initial attempt was measuring the frustration of total internal reflection within a waveguide on a sensor surface caused by contact [14]–[17]. This idea eventually leads to a product [18].

In these years, using cameras becomes popular since the price becomes cheaper, the size becomes smaller, computer vision becomes more handy, and high-resolution data can be obtained. In general, the components are camera(s), a transparent hard layer, and a transparent elastic layer (skin). An approach is obtaining object information contacting with the sensor by directly seeing the object through the skin [6], [10], [11], [19]. An idea to cover the skin with a reflective membrane was proposed in [4], [20], which was effective to precisely obtain the surface texture.

Placing markers on the skin is a widely-used approach [1]–[3], [5], [7]–[11], [21]–[24]. These markers make it easy to detect the deformation of the elastic skin caused by force. Marker displacements are proportional to the external force as the displacements are directly caused by the external force. They use computer vision methods to detect and track the markers.

The forms of markers vary: a lattice pattern is used in [1], an array of pins is used in [1], arrays of two-colored dots are used in [3], a single dot is used in [21], and an array of single-colored dots are used in [2], [5], [7]–[11], [22]–[24].

Some of them look similar to our approach in the sense of using whiskers. An array of pins was introduced in [1], [8], and an array of nodule markers was introduced in [7], [22]. The largest difference of our whiskers from those work is that the whiskers proposed in this paper are created *outward*, while those pins and nodule markers are created *inward*. Inward pins and nodule markers do not work like whiskers since each pin or marker does not react to external force independently as they are supported by the surrounding elastic skin. Outward whiskers can be easily deformed by external force as the most part of the whiskers are not contacting with the skin and they move independently with each other. Thus, the proposed whiskers in this paper dramatically increase the force sensitivity. Additionally, the FingerVision principle, transparent skin, makes it easy to detect and track the deformation of outside whiskers as they can be observed directly by the camera. This reduces the complexity of the computer vision method to process FingerVision image sequence.

III. Preliminary Experiments

In order to establish a design strategy of FingerVision with whiskers, we made preliminary prototypes. We fabricated FingerVision with a mold whose bottom had an array of small holes. After casting silicone with this mold, whiskers were added to the surface of FingerVision. The material of whiskers is the same as that of the elastic skin (silicone). Fig. 2(a)(b) shows the prototypes of FingerVision with whiskers. In order to track the deformation of the elastic skin and the whiskers, we placed markers in two ways as shown in the third picture of Fig. 2(a).

We applied the computer vision methods designed for the original FingerVision [11]. The proximity vision program worked to detect a nearby object and its movement (slippage). This was because the skin has a good transparency. However, the marker tracking program did not work well. We expected that the markers placed both at the tips of whiskers and on the skin are detected and tracked. Actually, the markers at the tips of whiskers sometimes *disappeared* on the video from the FingerVision camera when they move largely because of the refraction caused by transparent whiskers (see Fig. 2(c)).

IV. FingerVision with Whiskers

A. Design

Based on the preliminary experiments, we design whiskers that work as the markers of FingerVision to make it sensitive to normal force as well as shear force. In order to avoid the refraction issue of the whiskers, we design each whisker to be colored entirely. The slight touch on a whisker tip deforms the whisker, and that deformation will be captured by the FingerVision camera through the transparent skin. With this

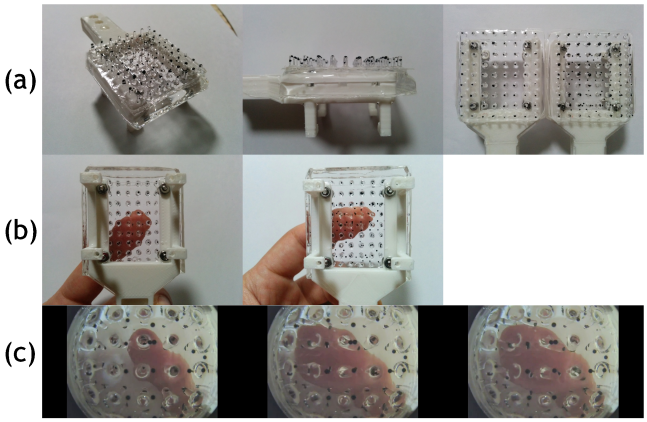

Fig. 2. (a) Prototypes of FingerVision with whiskers, (b) views of whiskers from camera side, (c) views from FingerVision cameras.

design, detecting deformation of whiskers is much easier than that of the preliminary whiskers.

The whiskers are made with the same material with the elastic skin, i.e. silicone. When casting silicone to form whiskers, we add black dye into the silicone resin. The amount of dye affects the optical and physical properties of whiskers. Adding more dye forms darker black, slightly softer, and more sticky whiskers. Adding less dye forms lighter black (gray), slightly harder, and less sticky whiskers.

The design parameters of FingerVision with whiskers include the hardness of whiskers, the whisker diameter, the whisker pattern (e.g. rectangular array of a certain interval), and the whisker shape and length. The design parameters of the original FingerVision are also inherited, such as the thicknesses of the elastic skin and the acrylic layer. As exploring all of these parameters is not practical, we focus on the most critical parameters: the whisker length and the content of the black dye which changes the hardness of the whisker. We use XP-565 silicone from Silicones Inc. for making both the elastic skin and the whiskers. The hardness of XP-565 after curing is A-16, and its transparency is very high. The thicknesses of the elastic skin and the acrylic layers are 4 mm and 2mm respectively, which are the same as the original FingerVision. We design the diameter of whiskers to be same as the marker diameter of the original FingerVision, i.e. 1 mm. We also use the same rectangular pattern of whiskers as the marker pattern of original FingerVision, i.e. 4 mm x 4 mm array. Since it is difficult to precisely cast the whisker shapes, we design the shape of each whisker to be capsule where the length of a whisker is an adjustable parameter.

B. Force Estimation

This section describes a computer vision method to track whisker deformation. The tracking result is used to estimate the force applied to the whisker. Since the objective of this work is proving the concept of FingerVision with whiskers, we consider a simple and minimum approach for the whisker tracking. In our original work of FingerVision [11], we used a blob detection method to detect and track the markers. This idea still works with whiskers since in the FingerVision view, the whiskers are also blobs. The difference from the original FingerVision is the shapes of blobs; in the original FingerVision, they are circles in images, while whiskers have different shapes. This difference can be handled by adjusting the parameters of blob detection method. Concretely, we need to configure the method so that non-circle blobs are detected.

1) Whisker Detection and Tracking: We introduce the blob detection and tracking method used in [11]. The basic idea is applying a blob detection locally for each blob. We consider a small region around the previous blob position, and apply the blob detection to obtain the current blob position. For the blob detection, we use a function implemented in OpenCV: the `cv::SimpleBlobDetector` class. The entire procedure is illustrated in Fig. 3.

The actual procedure consists of two phases: calibration and tracking. In the calibration phase, we apply blob detection to an entire image to get initial blob positions and sizes. In this process, we use some (e.g. 10) frames to check the stability of the blobs. If some blobs move, they are considered as noise and removed from the tracking targets.

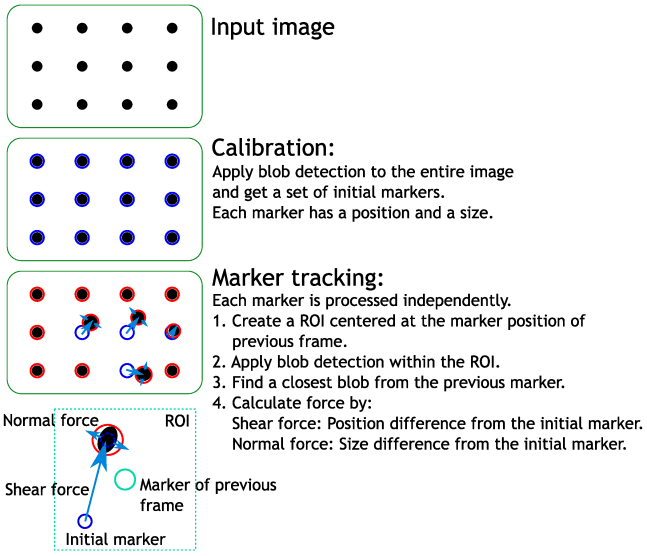

Fig. 3. Overview of marker tracking method to estimate force.

We track each blob frame by frame starting from the initial blob positions. We consider a small (e.g. 30x30) region of interest (ROI) around the previous blob position. First we count the non-zero pixels in the ROI and compare it with the non-zero points of the initial blob. If there is a large difference, we do not perform blob tracking (i.e. a detection failure). Otherwise we apply the blob detection method to the ROI. Only one blob is expected; otherwise it is considered a failure. We compare the previous and current blob positions and sizes, and if their differences are large, it is considered a failure. Otherwise the blob location and size are considered as the new blob location and size. In case of a detection failure, we keep the previous blob position and size, and if the failure continues for some frames, we reset the blob position and size to the initial ones. The reset also makes the force estimate to be zero.

This algorithm relies on the blob detection method. Although it is robust in typical use cases, the detection fails in situations where the surrounding is too dark, and the target object color is too close to the marker color. These issues can be engineeringly solved, for example we introduce LED for dark situations, and we combine multiple colors to form the markers.

2) Example: Fig. 4 shows the examples of the marker tracking for whiskers. In the left images, an object is not touching the FingerVision. In the middle images, the object is touching the whiskers where small marker movements are detected. In the right images, the object is strongly pushing the FingerVision surface where the markers move largely and the movements are detected.

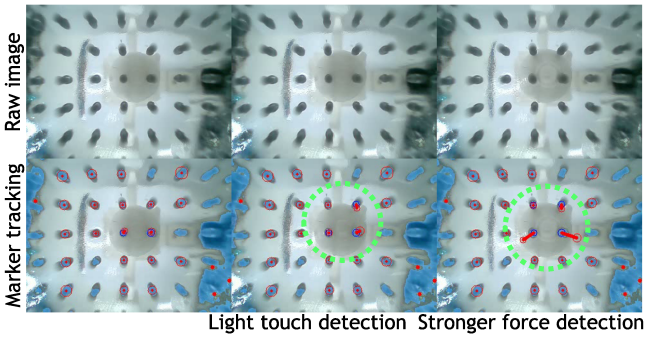

Fig. 4. Examples of marker tracking for whiskers. Each top image corresponds with the bottom image; they are an input image and the marker tracking result, respectively.

3) Force Estimation: From the blob movement, we estimate an array of forces. The blob detection and tracking provides a position and a size of each blob. The position change is caused by a horizontal (shear) force, while the size change is caused by a normal force. In this work, we define horizontal force to be proportional to the position change of each blob, and normal force to be proportional to the size change of each blob.

Let d_x and d_y denote the horizontal blob movement from the initial position, and let d_s denote the size change of blob from the initial size. The force estimate at each blob is given by:

$$[f_x, f_y, f_z] = \big[c(d_x, \mathbf{c}_x), c(d_s, \mathbf{c}_y), c(d_y, \mathbf{c}_z)\big] \tag{1}$$

where c is a conversion model, and $\mathbf{c}_x, \mathbf{c}_y, \mathbf{c}_z$ are the model parameters. Note that f_y is the normal force in our setup.

The model c is used to convert raw FingerVision reading to an engineering unit (gram force, in this paper) where the model parameters are calibrated with a reference force sensor. As the conversion model c, we use a linear regression or a two-mode linear model. We empirically select a model for each FingerVision. The reason of introducing the two-mode linear model is that FingerVision with whiskers is considered to have two modes (Fig. 1(c) and (d)). A linear model is defined as:

$$c(x, \mathbf{c}) = f_0 + f_1 x \tag{2}$$

where the model parameter is $\mathbf{c} = [f_0, f_1]$. A two-mode linear model is defined as:

$$c(x, \mathbf{c}) = f_{10} + (f_{11} - f_{21})\Big(\frac{\ln \sigma(x, \beta, x_t)}{\beta} + x_t\Big) + f_{21} x \tag{3}$$

where the model parameter is $\mathbf{c} = [x_t, \beta, f_{10}, f_{11}, f_{21}]$, and σ is sigmoid function: $\sigma(x, \beta, x_t) = (1 + \exp(-\beta(x - x_t)))^{-1}$. Two linear models $f_{10} + f_{11}x$ and $C + f_{21}x$ are smoothly connected at x_t by the sigmoid function where β controls the smoothness. The above function is derived so that its derivative becomes $(1 - \sigma(x))f_{11} + \sigma(x)f_{21}$.

For calibrating the model parameters $\mathbf{c}_x, \mathbf{c}_y, \mathbf{c}_z$, we use the iterative closest point (ICP) algorithm. For given two trajectories, one is raw FingerVision readings and the other is force sensor values, we iteratively optimize the model parameter so that the distance between the trajectories is minimized. The calibration is done independently for each axis.

V. Experiments

We implement some versions of FingerVision with whiskers in different design parameters, as well as FingerVision without acrylic layer, and an original design of FingerVision. In order to reduce the artifact, we fabricate all sensors in a limited condition; making them in the same day with the same silicone and the other materials, and the same curing time of silicone. FingerVision sensors compared here are following five versions:

W08-Black: With whiskers of 0.8 mm length, containing sufficient amount of black dye.

W12-Black: With whiskers of 1.2 mm length, containing sufficient amount of black dye.

W08-Gray: With whiskers of 0.8 mm length, containing less amount of black dye.

No-Acrylic: Removing the hard layer (acrylic plate) from the original design of FingerVision.

Flat: Original FingerVision design.

No-Acrylic is introduced as a simple way to improve the sensitivity and accuracy. Since the elastic layer is not supported by the hard layer (supported only by the side frame), the elastic layer deforms more against the external force, i.e. the markers move more largely in the image. This is an easy solution to improve the FingerVision sensitivity, however it is weak against large force; the sensor may break. Additionally, the sensitivity changes according to the location on the sensor surface. It is most sensitive around the center of the sensor since the skin deforms most around the center, while it is less sensitive near the side frame. Thus, the practical use cases of this approach will be more limited than the whisker approach.

Fig. 5 shows the structural difference of each design. The fabricated FingerVision sensors and the zoomed views of the surfaces are also displayed. From these zoomed views, we can see that there are differences of the quality of whisker fabrication. W08-Black and W12-Black look to have some fabrication errors of whiskers. There are irregular shapes and chipped whiskers. They might be due to the content of the dye; they are fabricated with sufficient amount of dye, which made the whiskers softer and more sticky. They made the fabrication of whiskers unstable. On the other hand, W08-Gray looks to have the best quality. This would be because of the less amount of dye, which made the whiskers harder and less sticky.

A. Force-Response Comparison

We compare the force response of the versions of FingerVision. Each FingerVision is attached on a finger of a parallel gripper Robotis RH-P12-RN mounted on a robotic arm Universal Robots UR3e. We put a fixture on a digital weight scale A&D EJ-6100B of 0.01 gf resolution. This weight scale can be connected to a PC for recording the vertical

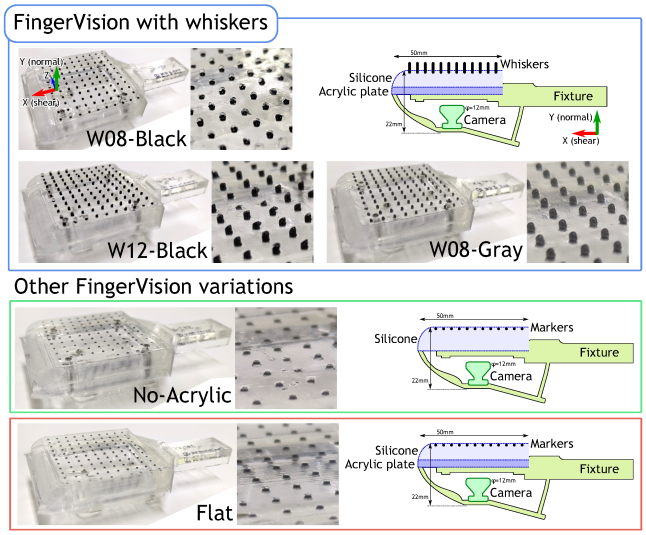

Fig. 5. Variations of FingerVision compared in this experiment. For each FingerVision, the entire sensor view and a zoomed view of the surface are shown as well as its structure design.

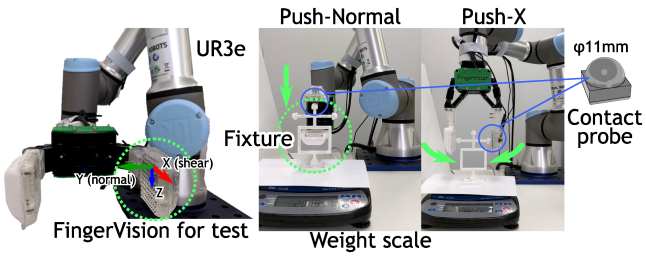

Fig. 6. Setup of force evaluation. Each FingerVision to be tested is attached on a finger of the gripper. The other finger has a dummy FingerVision that has the similar shape. The right images show the tests of Push-Normal and Push-X with the weight scale and the fixture.

force at 10 Hz. This data is also used as a reference force sensor to calibrate the conversion models. We let the robot to push the fixture in two ways. **Push-Normal**: Pushing the fixture downward by moving the FingerVision to its normal direction. **Push-X**: Holding the fixture vertically, and closing the gripper. Since the fingers move roundly because of the gripper structure, this gripper motion creates horizontal push on FingerVision surface, as well as push on the normal direction. We do not move the arm. Fig. 6 illustrates the setup of force evaluation. The contact probe of the fixture has a hemisphere shape as shown in the figure. Although the probe shape may affect the force readings, using a different shape would not change the comparisons among the FingerVision versions since the same marker resolution is used among them. Thus, we use a single shape as the probe.

1) Motion Details: The Push-Normal motion starts where the FingerVision is not touching the fixture (i.e. the weight scale reading shows the weight of the fixture). The motion is kept until the weight scale reaches 300 gf with the speed 0.2 mm/s. The robot waits a few seconds, and then moves back to the initial pose with the same speed.

The Push-X motion starts where the FingerVision is not touching the fixture (i.e. the weight scale reading shows the weight of the fixture). The gripper closing motion is kept until the weight scale shows 150 gf. The robot waits a few seconds, and then opens the gripper to the initial width with the same speed.

In order to make the experiments easy, we implemented a simple keyboard interface to command each motion: one key to go forward and another to go backward. Although the force profile applied to the FingerVision sensors will slightly differ among the trials, we can still fairly compare them since the motions are slow and we consider only the relation between the FingerVision readings and the reference weight scale values.

2) Data Analysis: After executing each motion, we obtain sequences of FingerVision and weight scale readings. As well as the calibration of the conversion models, they are used to analyze the sensing capability. We calculate two values for each pair of sensor and pushing direction: noise and minimum force detection. The noise is computed as follows: We apply a moving average filter (filter size 101) to get a smooth average curve. Then we compute the noise as an average of absolute difference between the filtered curve and the original data. The minimum force detection is calculated in increasing the force applied to the weight scale. It is defined as a weight scale value when the filtered FingerVision reading exceeds the zero-force noise level. As the filter, we use a moving average filter of filter size 10. The zero-force noise level is a peak value of FingerVision reading when there is no force. Note that the minimum force detection reflects the sensitivity of the sensor (smaller minimum force detection means more sensitive).

3) Results and Discussion: Fig. 7 and Fig. 8 show the results of Push-Normal and Push-X where the weight scale reading and the force estimates of FingerVision are plotted. The results of calculating the minimum force detection are also plotted. Fig. 9 and Fig. 10 are the plots of FingerVision reading vs. force estimate. Both the conversion function and the converted result are plotted. Fig. 11 is a comparison of minimum force detection (sensitivity) and noise of each sensor and each pushing direction.

From these results, we found:

(1) Overall, the FingerVision estimates and the weight scale readings correspond (Fig. 7, 8).

(2) In most cases of FingerVision with whiskers, we used the two-mode linear models. Overall, they provide good conversions, but there are exceptions in low force range of W12-Black and W08-Gray.

(3) Fig. 11 (b): Overall, the noise in normal direction is larger than that of shear (X) direction. An exception is the case of No-Acrylic; the reason would be that the skin deforms largely since there is no hard-layer support under the elastic skin. Longer whiskers seem to increase the noise (X noise of W12-Black is larger than those of W08-Black and W08-Gray).

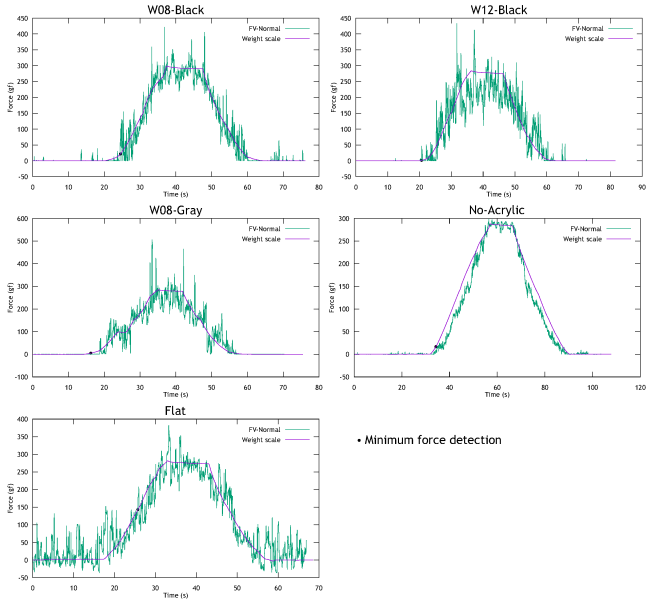

Fig. 7. Results of Push-Normal where the weight scale reading and the normal force estimate of FingerVision over time are plotted.

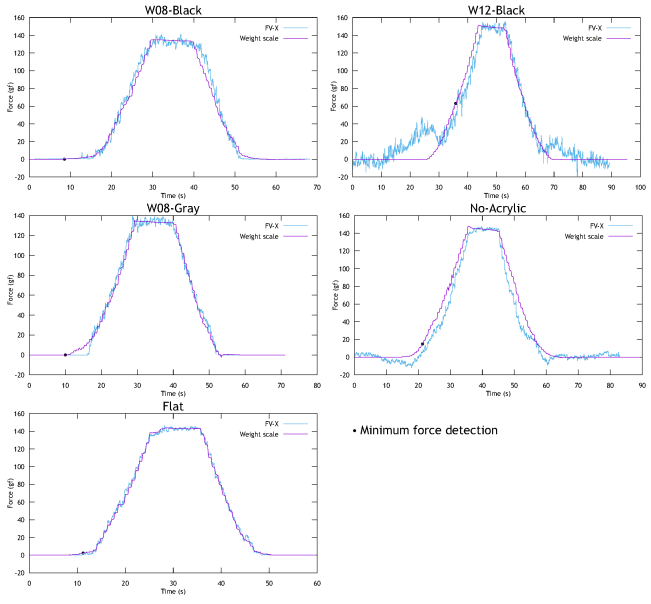

Fig. 8. Results of Push-X where the weight scale reading, and the force estimate (X-direction) of FingerVision over time are plotted.

(4) Fig. 11 (a): Minimum force detection of Flat (the original design of FingerVision) in normal direction is much larger than that in X-direction. In the other designs, minimum force detection in normal direction is improved. W12-Black is the most sensitive in the normal direction, which would be due to the longer whiskers. On the other side, the minimum force detection values W12-Black and No-Acrylic in X-direction are unexpectedly larger. From Fig. 8, we think this is due to calculation failure of minimum force detection as we can see the significant change of the FingerVision estimates in the beginning of pushing. In other words, the sensitivities of W12-Black and No-Acrylic in X-direction should be much higher.

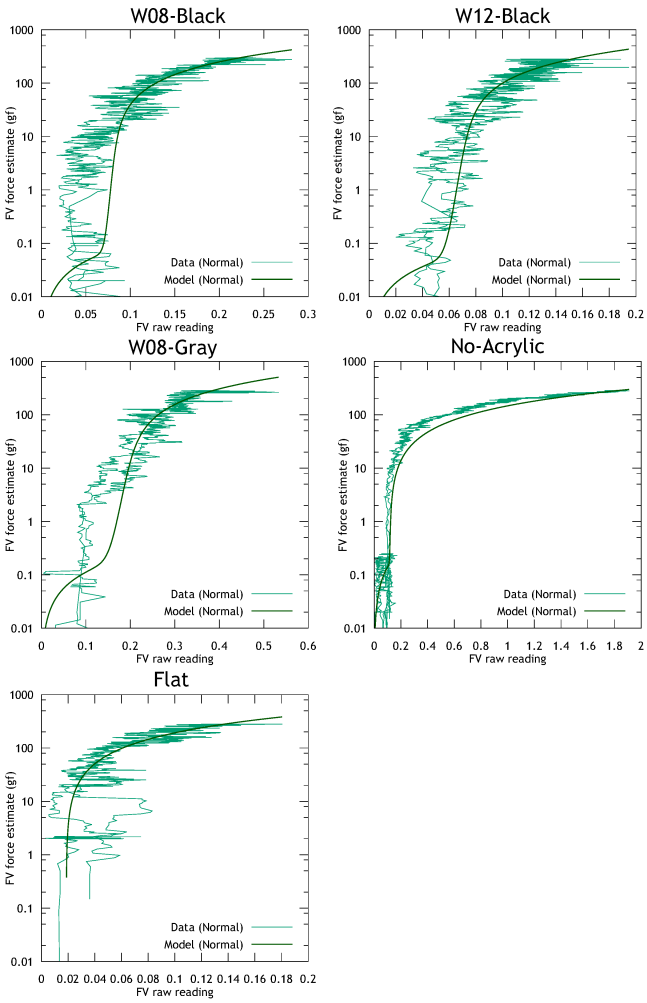

Fig. 9. Push-Normal: FingerVision reading vs. force estimate. The vertical axis is in log scale. Both the conversion function (Model) and the converted result (Data) are plotted.

(5) From Fig. 9 and 10, there is not much hysteresis. W08-Gray has slight hysteresis in X direction, and W12-Black has slight hysteresis in X direction.

The finding (2) was due to the whisker structure. Those two modes are when whiskers have room to bend, and when whiskers completely squished. In the first mode, FingerVision is very sensitive to small force. The existence of two modes means that the increase of FingerVision estimate may not be monotonic although the weight scale increases monotonically. This issue cannot be solved by using two-mode linear models. Thus, it caused the calculation failure of minimum force detection in (4). For solving this issue, we would need to

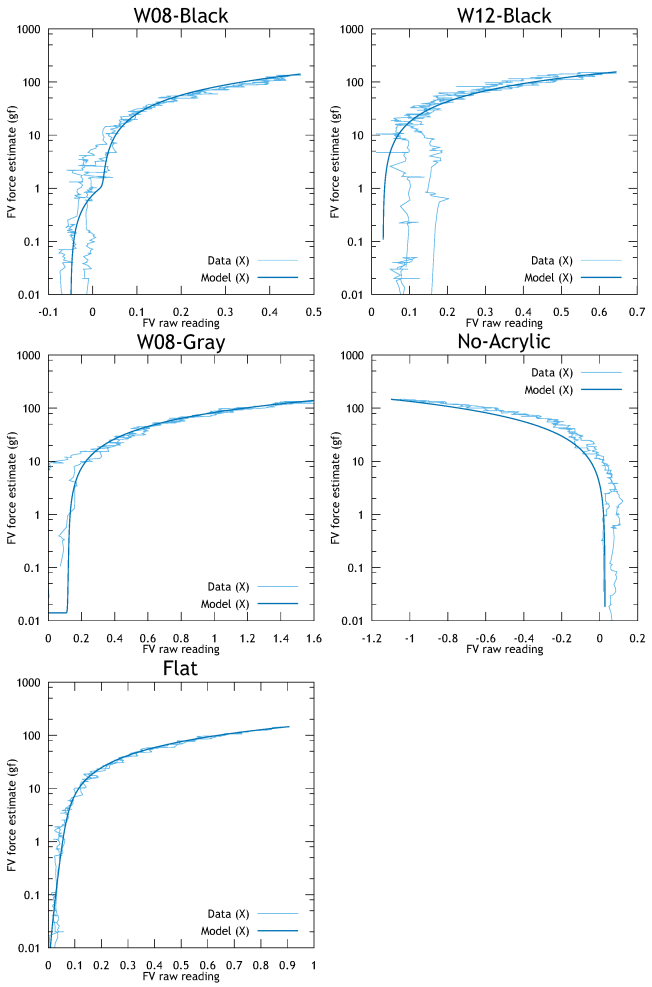

Fig. 10. Push-X: FingerVision reading vs. force estimate. The vertical axis is in log scale. Both the conversion function (Model) and the converted result (Data) are plotted.

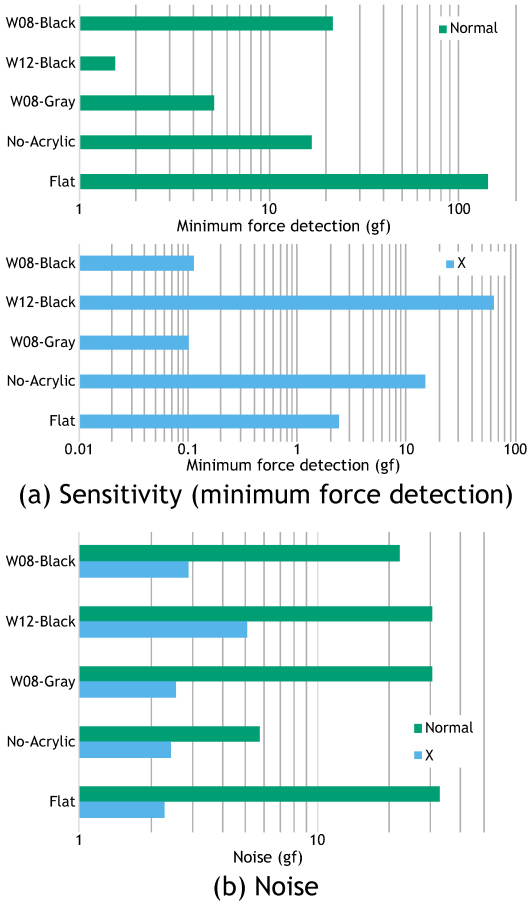

Fig. 11. Comparison of minimum force detection (sensitivity) and noise. The horizontal axis is in log scale.

develop an improved computer vision method for whisker tracking especially in longer whisker cases (such as W12-Black). In shorter whisker cases (W08-Black and W08-Gray), we do not see much necessity of that. W08-Gray seems to have the most practical performance at the current implementation.

From these results, we can conclude that introducing whiskers improves the force sensitivity of FingerVision. No-Acrylic structure also improves, but using whiskers improves more.

B. Demonstrations with W08-Gray

This section presents intuitive demonstrations of FingerVision with whiskers. We use W08-Gray since it is most stable in the previous experiments.

1) Pulling a Tissue Paper: The first demonstration is pulling a tissue paper. Fig. 12 shows the setup of pulling a tissue paper. The robot UR3e grasps a tissue, and a human operator pulls the tissue from some directions. The order of

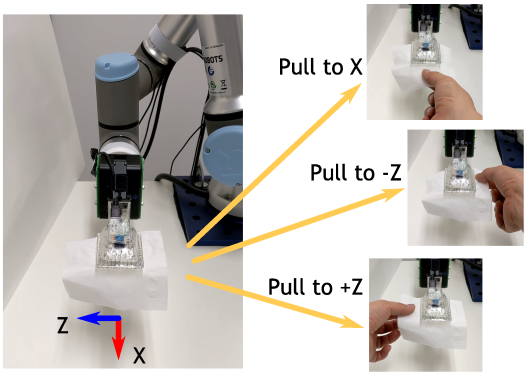

Fig. 12. Setup of pulling tissue paper.

pulling directions are as follows: twice +X, twice −Z, and twice +Z. Fig. 13 shows the obtained force profile. Although the human operator needed to pull the tissue paper weakly since it is fragile, the graph shows that the FingerVision captured the forces in correct directions.

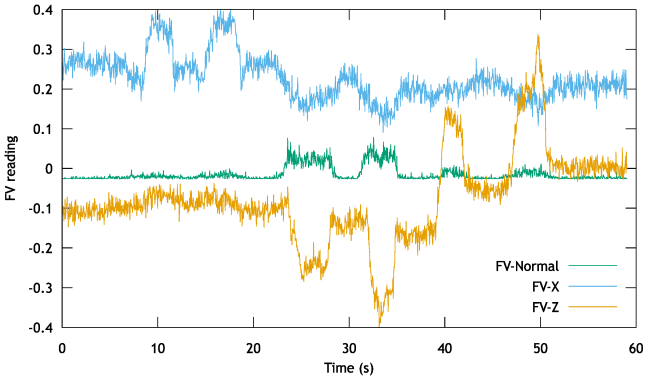

Fig. 13. Force profile in pulling tissue paper.

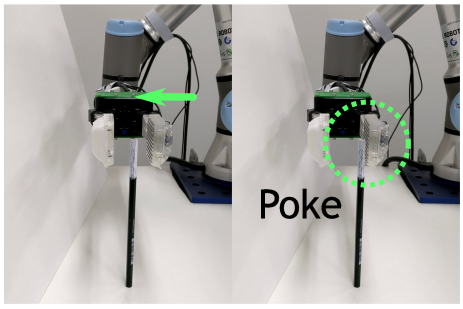

Fig. 14. Scene of the pencil poking. We move the robot leftward to poke the standing pencil by FingerVision.

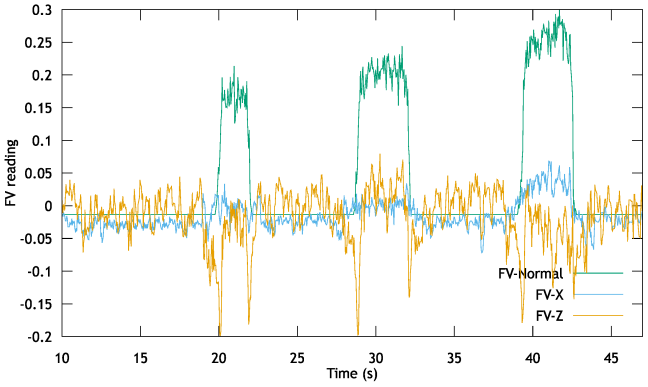

Fig. 15. Force profile in pencil poking.

2) Poking a Standing Pencil: The next demonstration is poking a standing pencil whose weight is 5 gf. Fig. 14 shows the scene of pencil poking. The robot moves to the pencil, and stops when the FingerVision touches the pencil. This contact event is detected by the operator. Fig. 15 shows the force profile of three poking trials. Although the pencil is very light weight, the FingerVision with whisker was able to detect the normal forces of pokes.

VI. Conclusion

This paper explored an idea of introducing whiskers into a vision-based tactile sensor FingerVision. Since whiskers deform easily against a light touch of objects, they improve the sensitivity of FingerVision for normal force. We fabricated some versions of FingerVision with whiskers and empirically compared it with the other approaches. The experimental results showed that the idea to introduce whiskers worked as we expected. We also demonstrated intuitive applications, pulling a tissue paper grasped by a robot and poking a standing pencil by a robot.

The idea of using whiskers as markers explored in this paper would work with the other vision-based tactile sensing. However if we use non-transparent skin, we need to handle the occlusion of whiskers. Since the skin of FingerVision is transparent, we was not suffered from this problem.

In this paper, We did not focus on proximity vision, but it is still available since whiskers on FingerVision skin do not ruin the transparency of the skin. Implementing tactile behaviors with FingerVision with whiskers is a future work. The future work also involves developing a new algorithm to track whisker deformation.

Acknowledgment

Yamaguchi was supported in part by NSK Foundation for the Advancement of Mechatronics, and the Canon Foundation 10th Research Grant Program (K18-0105).

References

[1] K. Nagata, M. Ooki, and M. Kakikura, "Feature detection with an image based compliant tactile sensor," in *1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1999, pp. 838–843.

[2] D. Hristu, N. Ferrier, and R. W. Brockett, "The performance of a deformable-membrane tactile sensor: basic results on geometrically-defined tasks," in *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*, vol. 1, 2000, pp. 508–513.

[3] K. Kamiyama, H. Kajimoto, N. Kawakami, and S. Tachi, "Evaluation of a vision-based tactile sensor," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 2, 2004, pp. 1542–1547.

[4] 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*, 2009, pp. 1070–1077.

[5] Y. Ito, Y. Kim, and G. Obinata, "Robust slippage degree estimation based on reference update of vision-based tactile sensor," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 2037–2047, 2011.

[6] K. Shimonomura and H. Nakashima, "A combined tactile and proximity sensing employing a compound-eye camera," in *SENSORS, 2013 IEEE*, 2013.

[7] T. Assaf, C. Roke, J. Rossiter, T. Pipe, and C. Melhuish, "Seeing by touch: Evaluation of a soft biologically-inspired artificial fingertip in real-time active touch," *Sensors*, vol. 14, no. 2, p. 2561, 2014.

[8] N. F. Lepora and B. Ward-Cherrier, "Superresolution with an optical tactile sensor," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, 2015, pp. 2686–2691.

[9] 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)*, 2015, pp. 304–311.

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

[11] ——, "Tactile behaviors with the vision-based tactile sensor fingervision," *International Journal of Humanoid Robotics*, vol. 16, no. 03, p. 1940002, 2019.

[12] W. Deer and P. E. I. Pounds, "Lightweight whiskers for contact, pre-contact, and fluid velocity sensing," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1978–1984, 2019.
[13] A. Yamaguchi and C. G. Atkeson, "Recent progress in tactile sensing and sensors for robotic manipulation: can we turn tactile sensing into vision?" *Advanced Robotics*, vol. 33, no. 14, pp. 661–673, 2019.
[14] S. Begej, "Planar and finger-shaped optical tactile sensors for robotic applications," *IEEE Journal on Robotics and Automation*, vol. 4, no. 5, pp. 472–484, 1988.
[15] H. Maekawa, K. Tanie, K. Komoriya, M. Kaneko, C. Horiguchi, and T. Sugawara, "Development of a finger-shaped tactile sensor and its evaluation by active touch," in *Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on*, vol. 2, 1992, pp. 1327–1334.
[16] H. Yussof, J. Wada, and M. Ohka, "Sensorization of robotic hand using optical three-axis tactile sensor: Evaluation with grasping and twisting motions," *Journal of Computer Science*, vol. 6, no. 8, pp. 955–962, 2010.
[17] T. Ikai, S. Kamiya, and M. Ohka, "Robot control using natural instructions via visual and tactile sensations," *Journal of Computer Science*, vol. 12, no. 5, pp. 246–254, 2016.
[18] OptoForce Co., "White paper: Optical force sensors — introduction to the technology," Tech. Rep., January 2015. [Online]. Available: http://optoforce.com/
[19] K. Shimonomura, H. Nakashima, and K. Nozu, "Robotic grasp control with high-resolution combined tactile and proximity sensing," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016.
[20] 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 *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 3988–3993.
[21] J. Ueda, Y. Ishida, M. Kondo, and T. Ogasawara, "Development of the NAIST-Hand with vision-based tactile fingertip sensor," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005, pp. 2332–2337.
[22] C. Chorley, C. Melhuish, T. Pipe, and J. Rossiter, "Development of a tactile sensor based on biologically inspired edge encoding," in *Advanced Robotics, 2009. ICAR 2009. International Conference on*, 2009, pp. 1–6.
[23] T. Sakuma, F. Von Drigalski, M. Ding, J. Takamatsu, and T. Ogasawara, "A universal gripper using optical sensing to acquire tactile information and membrane deformation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.
[24] 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)*, 2018, pp. 6437–6444.