2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Multidirectional slip detection and avoidance using dynamic 3D tactile meshes from visuotactile sensors

Peng Song[1], Juan Antonio Corrales Ramón[2], and Youcef Mezouar[1]

***Abstract*— Visuotactile sensors have gained attention during the last years in robotics because they are able to reconstruct with high precision the 3D contact shape (or mesh) between the robotic fingers and the object. A new slip detection and avoidance algorithm is proposed based on the dynamic variation of the height of the contact mesh. Firstly, the contact mesh is reconstructed in real time by applying a neural network that estimates normal vectors from color variations along all the pixels of the images recorded by the camera inside the tactile sensor. The contact mesh corresponding to this height map is used for detecting slip with higher success rates in comparison with previous approaches based on machine learning methods directly applied to contact images or the analysis of markers integrated into the sensor's surface. The proposed algorithm is validated experimentally in multiple directions not only for different types of objects (volumetric/planar/linear, deformable/rigid) but also with different resolutions of the contact mesh.**

I. Introduction

Slippage should usually be avoided during robotic manipulations, as the occurrence of slip can easily cause operation failure. To avoid slippage, robots should be able to detect its occurrence at first. Slippage detection is a popular topic for the application of tactile sensors. Previous works explored different technical principles [1] to realize a sense of slip for robotic hands. One main methodology is to monitor the ratio between the tangential contact force and the normal contact force [2], [3]. If this ratio increases beyond the static friction coefficient between the robotic hand and the object, a slip occurs. To measure multi-axial forces, a variety of sensors is designed based on sensing technologies like piezoresistive [4] and capacitive [5]. Acquiring multiaxial forces requires a more complicated sensor structure [1]. In addition, the simplified friction model adopted is not valid for soft materials. The vibration generated by the sliding between two contacted surfaces can be an indicator of slip. A piezoelectric sensor [6] can produce a dense fluctuation signal when slipping occurs. However, temperature has a high influence on its performance. Some work exploits the velocity/acceleration signal to detect slip [7], [8], but may be sensitive to external disturbances [1]. Signal processing technologies like transform operations [9] and filters [10] are used to isolate and reveal the slip signal.

[1]First Author and Third Author are with CNRS, Clermont Auvergne INP, Institut Pascal, Université Clermont Auvergne, Clermont-Ferrand, France songpsp123@163.com, youcef.mezouar@sigma-clermont.fr

[2]Second Author is with Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain jacr56@gmail.com

For visuotactile sensors, there are also some slip detection strategies. One method is to integrate physical markers inside the deformable layer on the top of the sensor. Tracking and analyzing the movement of markers can help detect slips [11]–[14], but these markers can be noise in tactile images for microgeometry, object recognition, object classification, and shape measurement [15]. Markers may provide some directional information related to slip in the 2D surface plane, but in practical slip prevention, usually only the occurrence of slip is important to control the action of the gripper. In addition, the integration of markers complicates fabrication. For visuotactile sensors without markers, slip detection can be treated as a video classification problem based on machine learning [16], [17]. To improve the generality of the model, a large dataset with a wide variety of objects should be labeled, which is time-consuming. Another drawback for this kind of method is that it usually needs 12 frames of tactile images to output a result for DIGIT running at $30Hz$ [16]. In contrast, the proposed method only needs 2 frames.

By 3D reconstruction of the contact imprint of visuotactile sensors, the proposed slip detection approach can provide customizable high sensing resolution without physical markers. Compared to learning-based methods [16], [17], its prediction is more physically intuitive and faster using only two consecutive image frames. Its effectiveness and generality are validated by different kinds of unknown objects. An avoidance scheme is implemented to regulate the action of the gripper accordingly. Regarding the overall detection accuracy, the machine learning method in [17] can achieve up to 88% and the method in [16] can reach 90% using a sequence of 12 tactile frames as input. The methods analyzing markers can obtain 84% in [13] and 86% in [18]. The proposed method are tested around 50 times in total on different objects as partly shown in Figure 4. In all cases, the slip is successfully indicated when the slip occurs.

In the remainder of this work, Section II explains the method for implementing real-time 3D reconstruction of tactile images. Then the slip detection method and its performance are presented in Section III. In Section IV, the density of the proposed tactile mesh is discussed. Section V demonstrates the strategy of slip avoidance. Concluding remarks are presented in the last section.

II. 3D tactile imprint

The sense of touch can be digitalized with high resolution tactile images. Recovering 3D contact information from 2D tactile images enhances the perception of robots.

A. Principle of 3D reconstruction

To achieve 3D contact shape, the photometric stereo algorithm [19] is used for the series of GelSight [14], [20], [21]. As the similarity between GelSight and DIGIT [22] adopted in this research as its availability and robust mechanical design, the same theory is applied in 3D reconstruction. According to this theory [23], the normal surface of objects can be estimated by observing the object under different illuminations. In a vision-based tactile sensor, the deformation of the elastomer changes the reflectance of the membrane layer coating on the top of the sensor. The RGB 3-channel light intensities captured by the camera is influenced by this reflectance. The relationship between the light intensity and the normal surface of deformation is used to implement the 3D reconstruction.

The height map of a surface could be described by:

$$z = H(x, y) \tag{1}$$

where (x, y, z) refers to the 3D coordinates of one pixel in the tactile image. The surface normal at the location (x, y) can be expressed as [24]:

$$N(x, y) = (\frac{\partial H}{\partial x}, \frac{\partial H}{\partial y}, -1) \tag{2}$$

where $\frac{\partial H}{\partial x}(x, y)$ and $\frac{\partial H}{\partial y}(x, y)$ represent the height gradients along x and y directions, respectively. The relationship between the light intensity $I(x, y)$ received by the camera and the shape of the reflective surface can be described by:

$$I(x, y) = R(\frac{\partial H}{\partial x}(x, y), \frac{\partial H}{\partial y}(x, y)). \tag{3}$$

The inverse function of R can take the observed light intensity $I(x, y)$ as input and output the height gradients. However, this nonlinear function is hard to be explicitly established.

Once known the height gradients, they can be integrated by solving the Poisson equation to obtain the height map H. The Poisson Equation has the form as:

$$\nabla^2 H = \frac{\partial(\frac{\partial H}{\partial x})}{\partial x} + \frac{\partial(\frac{\partial H}{\partial y})}{\partial y} \tag{4}$$

where ∇^2 refers to the Laplace operator.

B. A mapping neural network

To get height gradients from color intensities, a calibrated lookup table can work if the light distribution on the sensing surface is quite uniform like GelSight. The table has the array form $(r', g', b', 2)$ where r', g', b' are three indexes related to color intensities and 2 is reserved for the two gradients. The buildup of the table needs a calibration process as demonstrated in [20]. If the light distribution on the sensing surface is quite uneven, using the same table on all pixels of tactile images cannot guarantee stable reconstruction. For DIGIT, its light distribution is not uniform and quite uncertain, which requires a more robust 3D reconstruction method.

In our use case, a multilayer perceptron (MLP) is adopted. The network has five input variables, including the coordinates (x, y) of a pixel in the tactile image and the color variation $(\Delta r, \Delta g, \Delta b)$ at a pixel between the current tactile image and the non-contact reference image. The addition of (x, y) is helpful in balancing the performance of 3D reconstruction under different lighting conditions on the sensing surface. Using the color variation instead of the absolute color value is proved useful in experiments to reduce the noise in the non-contact area. The network has three output elements as the normal vector (n_x, n_y, n_z) of a location on the sensing surface that corresponds to a pixel in the tactile image. This normal vector is normalized. Between the input layer and the output layer, three hidden layers are added to compose a neural network in the form of $(5-32-32-32-3)$.

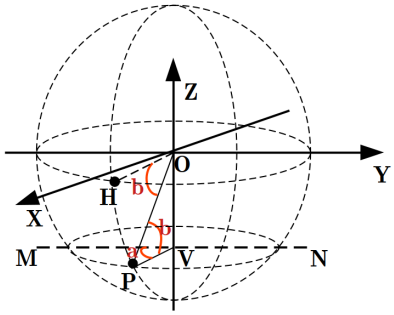

Fig. 1. Computation of the normal of one contact point.

To prepare a dataset, 30 tactile images pressed with a ball bearing are collected. In each tactile image, the contact circular imprint pressed by the ball bearing is labelled by Hough Circle Transform and manual adjustment. Inside the circle, the normal (n_x, n_y, n_z) corresponding to each pixel can be calculated using the geometrical model of the ball bearing as shown in Figure 1. A frame is placed in the center O of the ball bearing. The goal is to calculate the normal vector at an arbitrary contact point P that corresponds to a pixel in the tactile image. The surface XOY defined by the x-axis and y-axis is parallel to the sensing surface. In the section $MNPV$ of the ball bearing, the line MN is parallel to the y-axis and the line PV is parallel to the line HO in the surface XOY. In the base frame, if the point P is (x, y, z) and the center O is (x_c, y_c, z_c), then the value a of the angle PVM can be computed as:

$$a = \arctan \frac{x - x_c}{y - y_c} \tag{5}$$

The distance PV is:

$$d_{PV} = \sqrt{(x - x_c)^2 + (y - y_c)^2} \tag{6}$$

The distance PO is the radius r of the ball bearing, then the value b of the angle VPO is:

$$b = \arccos \frac{d_{PV}}{r} \tag{7}$$

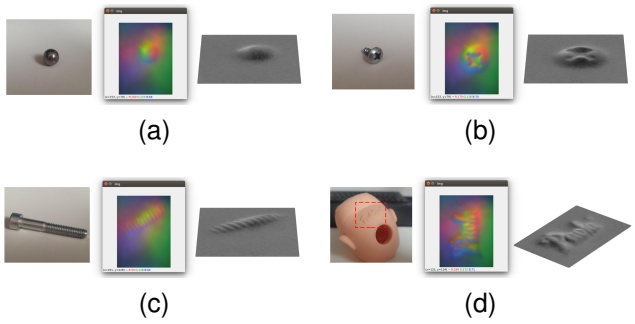

Fig. 2. 3D reconstruction (on the right) from RGB tactile images (in the middle) of several objects (on the left). (a) Ball. (b) Screw top. (c) Screw thread. (d) Signature on the toy.

Given a and b, the normal (n_x, n_y, n_z) on the point P can be determined by:

$$\begin{aligned} n_x &= \cos a \cos b \\ n_y &= \sin a \cos b \\ n_z &= \sin b \end{aligned} \tag{8}$$

Outside the circle, the normal value is set as $(0, 0, 1)$ for the flat area. The number of pixels per millimeter is calibrated by pressing a caliper on the sensor. The data $(x, y, \Delta r, \Delta g, \Delta b, n_x, n_y, n_z)$ is gathered in a file as the dataset. The proposed neural network is trained in pixel-wise way. Each image can provide 320×240 pixels. Hence, 30 tactile images are enough to provide a large dataset. The labelling and training is fast.

In real-time 3D reconstruction, the normal vector output by the neural network can be converted into the gradients of surface by the equations:

$$\begin{aligned} Grad_x &= n_x / n_z \\ Grad_y &= n_y / n_z \end{aligned} \tag{9}$$

Then we use Fast Poisson Algorithm [25] to integrate the gradients to get the entire height map of the sensing surface of DIGIT. To be noted, we use a small fixed value to replace n_z when it is close to zero. Some reconstructed 3D contact shapes are shown in Figure 2. It takes around 0.0156 seconds to generate a height map from a tactile image. To measure the precision of the proposed neural network, 36 tactile images are collected pressed by ball bearings. For each sampled image, we compute the root mean square error (RMSE) between the estimation of the neural network and the contact shape generated by the geometrical model of the ball bearing. The average RMSE is around 0.0262. In contrast, the neural network trained by the absolute color values has the average RMSE around 0.0349. This precision is already adequate to implement an effective slip detection algorithm.

III. Slip Detection

To enable slip detection and avoidance, we exploit the real-time 3D reconstruction of DIGIT. The resolution of a tactile image is 320×240. In maximum, 76800 pixels can be regarded as virtual tactile units because each pixel is associated with a reconstructed height value. As needed we can easily adjust the density of these virtual tactile units, especially when a lower resolution is enough for one use-case and some computation can be saved. For instance, by activating one pixel every ten pixels in a row and in a column, a evenly distributed tactile mesh 32×24 can be formed on the sensing surface with 768 tactile units in total. This is usually more numerous than the number of physical markers in some visuotactile sensors and much more numerous than most traditional tactile sensors such as piezoresistive or capacitive sensors. Two tactile meshes with custom densities are visualized in Figure 3.

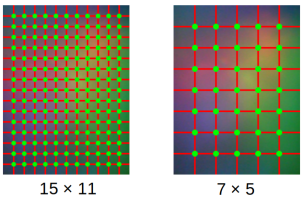

Fig. 3. Two tactile meshes with custom densities. The marked dots represent the selected tactile units.

To judge the occurrence of slip, we compare two consecutive frames. The absolute variation of the height value on each tactile unit is summed up as a slip signal S, which is formulated as:

$$S = \sum_{i=1}^{n} |h_t^i - h_{t-1}^i| \tag{10}$$

where n denotes the number of the used tactile units, h refers to the height value at one pixel and t means the current tactile frame. If this signal is larger than a threshold, a slip is found. If the signal is always lower than the threshold, grasp is deemed stable. The threshold is usually set to twice the average value of the static slip signal when no slip occurs, which is proved effective in our practice. DIGIT runs at $30Hz$ to stream tactile images, and thus the signal updates at $30Hz$ as well. In the plots, the signal is illustrated over time at this frequency. Using the overall variation of the contact shape to monitor slip is physically intuitive, as humans can feel the deformation of the tissue of the fingers.

A. Slip detection with unknown objects

A variety of common items are used to test our method, such as tissue, socks, paperboard, plastic package, USB cable, etc. Some of them are listed in Figure 4. Their dimensions, materials, and properties are quite different for testing the applicability of the proposed method.

For a firmly gripped object, its reconstructed contact shape on the tactile sensor is unique and stable. When a slip is triggered, this contact shape varies. The distributed tactile units capture this variation and the slip signal can reflect it quantitatively. For USB cables, its slippage is an evident relative motion on fingertips. Thus, its variation in contact shape is also apparent. However, for layer-like objects like fabrics, how do we capture the subtle variation of the contact shape and generate an adequate slip signal when slip occurs?

Thanks to the high resolution of the 3D reconstruction, this is achieved. We can measure the contact details, even the texture of fabrics. Although the general shape of contact is less variable, sliding of texture can stimulate the distributed tactile units and produce a sufficiently large slip signal, as shown in Figure 5. The 3D reconstruction on some tested objects is visualized in Figure 6. In each sub-figure, the raw 2D tactile image from DIGIT is on the left and its 3D reconstruction is on the right. For the USB cable, the general contact shape is clear. For other layer-like deformable objects without a solid general shape, we can still get their 3D textures. When they slide in grip, the distribution of the texture changes, leading to a detectable change in the slip signal. Identifying the change of the physical contact shape as the indicator of slip provides high intuitivity and universality.

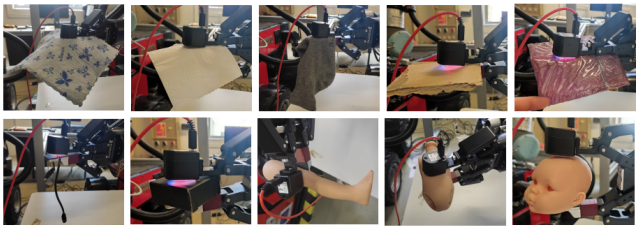

Fig. 4. Some objects used in the slip test. First row: towel, tissue, sock, paperboard, plastic package. Second row: USB cable, carton box, toy leg, toy body, toy head.

B. Slip detection in different directions

For the objects less volumetric, like tissue, cable, etc, a slippage is usually introduced by pulling it. However, for volumetric objects, contact conflicts with the environment could introduce a slip in all directions with reference to the gripper. In this subsection, we study how the height-based method performs when a slip happens in different directions. Four different objects are studied, including the carton box, toy leg, toy body and toy head as shown in Figure 4.

The carton box is a hexahedron and its faces are quite flat. In the tests, the grip is initiated with the two largest surfaces to avoid the edges. We tried to control the relative motion of the carton box in the horizontal plane which is parallel to the surface of fingertips. Firstly, the motion of the box is faster and it is moved more slowly in the second trial. In the direction vertical to the sensing surface, a motion like pressing the box toward the sensor is applied. This is also regarded as a kind of slip since the contact between the object and the other fingertip becomes less tight. The corresponding slip signals are depicted in Figure 7. In these situations, the signal marking the slip status is easy to be identified from the normal status. For the slip caused by sudden interference, a signal impulse can be observed. For a slower relative motion, a continuous signal variation presents in a longer period.

Besides the carton box in a regular shape, we also tested the toy components in complicated shapes. To define the direction of slip, a coordinate frame is placed as illustrated

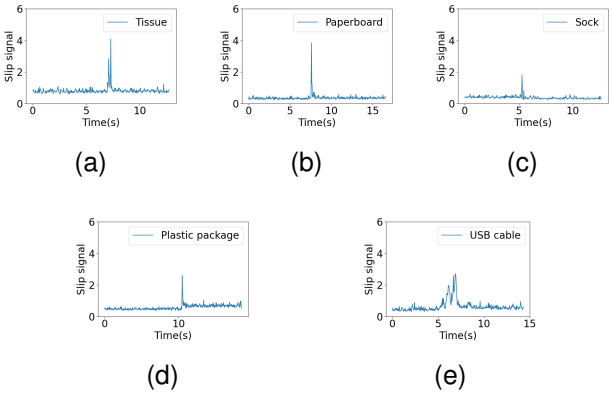

Fig. 5. Slip signal in tests with some common items. (a) Tissue. (b) Paperboard. (c) Sock. (d) Plastic package. (e) USB cable.

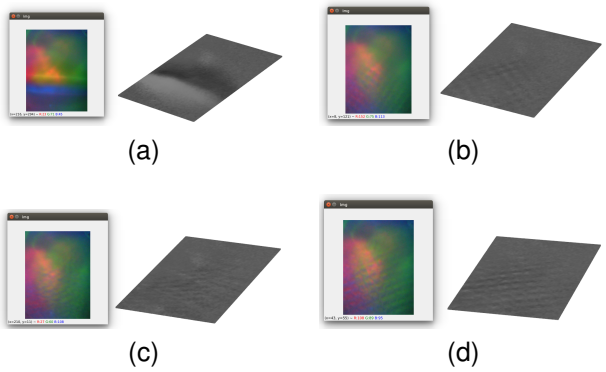

Fig. 6. 3D reconstruction of some tested objects. (a) USB cable. (b) Tissue. (c) Sock. (d) Towel.

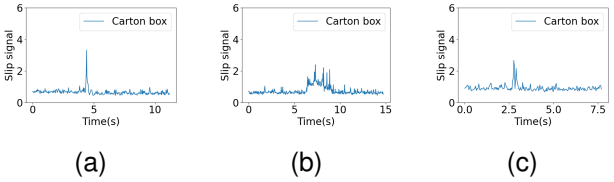

Fig. 7. Slip signal in tests on carton box. (a) Horizontal fast motion. (b) Horizontal slow motion. (c) Vertical motion.

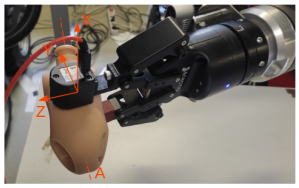

Fig. 8. Coordinate frame defined for slip tests on toy components.

in Figure 8. The Z-axis is parallel to the axis of gripper. The X-axis is perpendicular to the top side of the tactile sensor, and the Y-axis is perpendicular to the other two axes. The dashed A-axis denotes the geometrical axis of the toy component itself. After each toy component is gripped, a small translation is applied on the toy component along each

axis in the coordinate frame, respectively. A small translation along the X-axis is possible since the surface of the sensor is deformable. Besides, a small rotation is applied around its own axis. We did not test the rotation around the axes of the coordinate frame because this kind of motion is similar to the motion previously tested on a small scale. We performed the same tests on toy leg, toy body and toy head. In total, 12 tests are performed and all can produce a clear signal change to indicate the occurrence of slippage. In these tests, the minimum ratio between the signal peak value and the average signal value before slip is greater than 3 and the maximum ratio can be greater than 17. We can conclude that the height-based method is capable of detecting slips that occur in different directions.

IV. Density of the Tactile Mesh

The density of the tactile units can be easily adjusted by choosing a part of the pixels in the tactile image. The selected tactile pixels make up a sensing array, which is termed a tactile mesh. In the densest case, all the pixels are used to form a 320×240 tactile mesh. Picking up some pixels in a certain pattern can create a tactile mesh in the form 160×120, 32×24, or 4×3, etc. The height values on the unused pixels are discarded. The impact of the density of the tactile mesh is studied in two groups of experiments.

Firstly, we proved that a denser tactile mesh is necessary to detect a subtle slippage of a small gripped object. Three different sized hex keys are tested as shown in Figure 9a. Initially, a grip is formed on the small hex key, the medium and the large, respectively. Their tactile images in grip are shown in Figure 9c. Then we slightly push one end of the gripped hex key to introduce a slight slippage. The tactile meshes 32×24 and 8×6 are compared. If a tactile mesh 32×24 can work, the denser meshes are supposed to work as well. The tactile mesh is evenly distributed on the sensing surface.

The results of the tests are shown in Figure 10. For the large hex key, both tactile meshes are able to generate an identifiable signal peak to mark the occurrence of slip. For the small key and the medium key, the denser tactile mesh can still function. However, the sparse tactile mesh cannot produce a distinguishable signal at the moment of slip. Thus, the density of the tactile mesh influences its capacity to detect the slight slip of objects having a small contact with the sensing surface. In the case of a sparse tactile mesh and a small contact region, a slight slip cannot stimulate sufficient tactile units to produce a large signal peak. In contrast, denser tactile meshes are more sensitive as even a small change of the contact region can swipe on a lot of tactile units.

Secondly, we found that if the contact between objects and the sensor occupies the entire sensing surface, an evenly distributed tactile mesh can work equivalently well. The towel shown in Figure 4 is tested. The contact between the towel and the sensor takes up the whole sensing surface and pulling the towel is able to stimulate the entire sensing surface. The tactile meshes in the form of 320×240, 64×48, 32×24, 16×12 and 4×3 are compared. Without changing

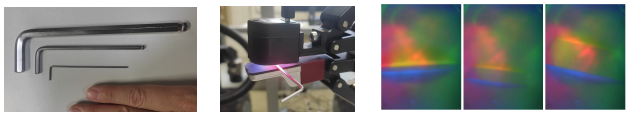

Fig. 9. Tested hex keys. (a) Hex keys in three dimensions. (b) Gripped small hex key. (c) Tactile images of hex keys.

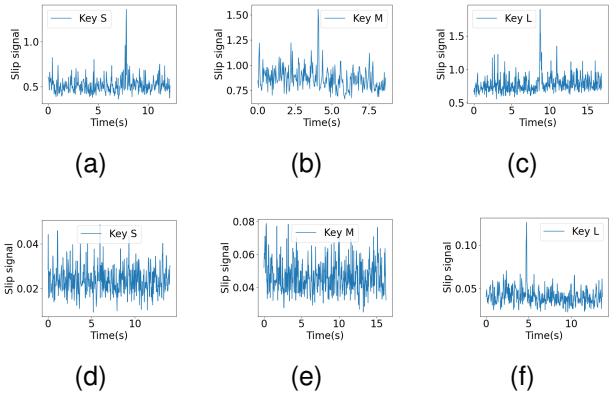

Fig. 10. Slip signals of hex keys. (a) Small hex key, 32×24. (b) Medium hex key, 32×24. (c) Large hex key, 32×24. (d) Small hex key, 8×6. (e) Medium hex key, 8×6. (f) Large hex key, 8×6.

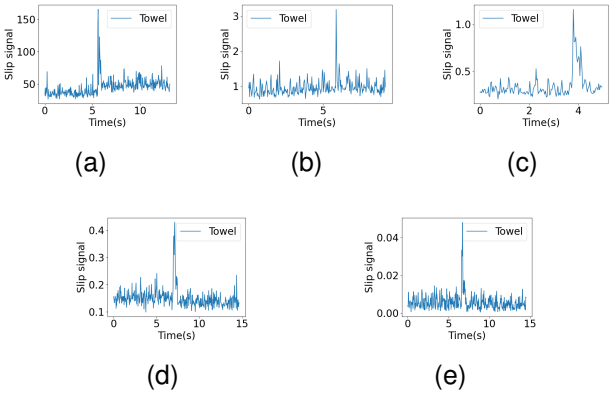

Fig. 11. Slip signals of towel. (a) 320×240. (b) 64×48. (c) 32×24. (d) 16×12. (e) 4×3.

hardware, these tactile meshes are switched to measure slip signals when a slip is introduced by pulling the gripped towel. The parameters to control the initial grip are the same for all tests. The recorded slip signals are plotted in Figure 11.

We can find that the values of these slip signals are different. The slip signal of a denser tactile mesh is larger in all stages. Before and after slip, more used tactile units can generate a larger accumulated noise value. During slip, more influenced tactile units produce a higher signal peak. However, the scale between the signal peak and the signal before or after slip is quite similar for these tactile meshes. This means that using any of them can easily make the slip distinguishable. The tactile mesh density can scale the slip

signal value but the pattern of signal keeps in this case.

The guidance is that if the manipulated objects have a large contact area with tactile sensors, the low density can be used, but if the contact is relatively small, the high density is necessary to guarantee the high sensitivity of slip detection. In future work, the tactile mesh density should automatically adapt to the sensed contact area between the object and the sensor.

V. Slip Avoidance

Upon detection of slip, the robot should react accordingly to prevent further slippage and even drop of the gripped object. A slip avoidance pipeline is developed as depicted in Figure 12.

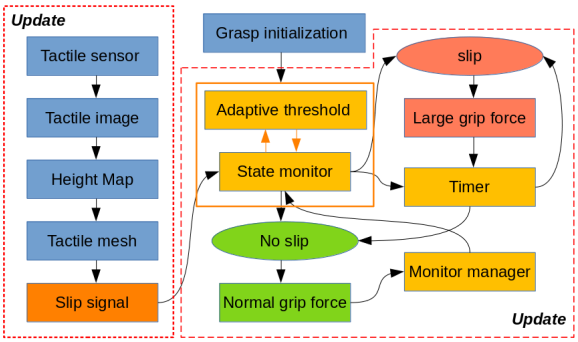

Fig. 12. Pipeline of slip avoidance.

The left column describes the workflow to achieve the slip signal from the tactile image. The real-time slip signal is fed to the state monitor, which decides the slip state. The state monitor is launched when the grasp is formed on the object. Before the state monitor is able to work, it sends a request to the module named ”Adaptive threshold” to get the proper threshold value to judge the slip signal. Usually the ”Adaptive threshold” module calculates the average of the 30 consecutive slip signals (in one second) in the normal state without slip and returns twice this average value to the state monitor.

This two-fold scale between the threshold and the average signal value in normal state is determined by trial and error. This scale is proved effective in the numerous tests on various objects. The idea is to find a suitable small value to distinguish the slip signal from the normal signal. If the scale is too small, the protective response to slippage could often be triggered by mistake. If the scale is too large, some subtle slip may not be detected. Upon a good 3D reconstruction, the variation of the signal noise in normal state is within a stable range. Thus, it is not difficult to find a proper threshold scale.

For a preset tactile mesh, the returned threshold could be a little different while grasping different kinds of objects. The impact of this small variation can be handled by this adaptive threshold. An alternative way is to set a fixed threshold for all objects. The fixed threshold should be slightly larger than the adaptive threshold. The risk is to miss some subtle slips.

Given a threshold, the state monitor compares it with the real-time slip signal. If the signal is always lower than the threshold, the program stays in the ”No slip” mode. The normal grip force is applied as the initial grip and no additional action is taken. The robot can execute its operation as normal. When a signal larger than the threshold is found, it turns on the ”slip” mode and a larger grip force is applied by the gripper. Then the timer starts to count if the state monitor does not output ”slip” state in continuous three seconds. This duration can be customized. If the ”slip” is counted, the program is still in the ”slip” mode, the large grip force is not released, and the timer restarts the counting in another three seconds. If the timer receives no more slip output from the state monitor in three seconds, the slippage is judged gone, and thus the ”No slip” mode is switched back on. The normal grip force is reapplied to avoid the possible damage or any negative effect on the gripped object.

However, changing grip force can cause the variation of the contact shape and result in a large signal that could be mistakenly regarded as ”slip” by the state monitor. To avoid this misjudgement, the ”Monitor manager” module is added. When the gripper is in motion, the state monitor is disabled temporally by the monitor manager. After the motion is finished and the gripper becomes static again, the monitor manager can relaunch the state monitor.

This pipeline is tested with different toy parts, including arms, legs, heads and bodies. We manually interfere with the gripped object to induce slips and then observe the reaction of the gripper. In the slip mode, the object is gripped more. If the interference persists, the slip mode keeps. When the timer and the state monitor decide to turn on the no-slip mode, the grip force is reduced. New interference can turn on the slip mode again. This process runs smoothly and the different modes can be told by observing the action of the gripper and the resulted grip status.

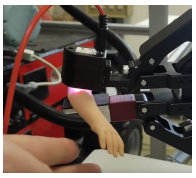

Fig. 13. Initial grip on toy arm.

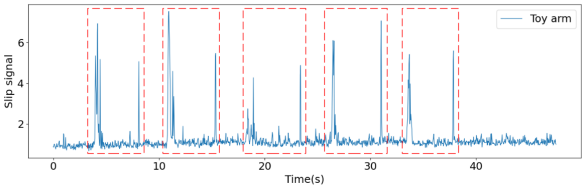

Fig. 14. Slip signal in slip avoidance.

From the slip signal, we can also validate the effectiveness of the proposed pipeline. The toy arm is initially gripped, as shown in Figure 13. Then we interfere with the grip five

times from different directions. Each interference caused the slip mode, and the no-slip mode is recovered after a period without interference. The recorded slip signal can reflect this mechanism, as depicted in Figure 14. From left to right, every two signal peaks are in a pair as highlighted by the red rectangle. We have five pairs in total corresponding to five interferences. In each pair, the first peak represents the signal variation caused by the slip and the stronger grip, and the following peak represents the signal variation caused by the reduction of grip force when there is no more slip. With the help of the module "Monitor manager", the second signal peak in each pair is not treated as slip. The proposed pipeline of slip avoidance works efficiently when slip occurs and flexibly when slip is gone.

VI. Conclusion

A novel slip detection method is proposed based on the tactile mesh derived from the reconstructed height map on the tactile sensing surface. The effectiveness of the proposed method is validated by various objects with different physical properties. The slippage occurring in different directions can be detected. As the elemental tactile unit in the tactile mesh is a pixel, we can change the density of the mesh by using certain part of the pixels in the tactile image. The impact of this density is studied. An avoidance strategy for slips is implemented to react flexibly to the varying slip states in manipulation.

This slip detection method exploits the real-time variation of the 3D contact shape, which is quite general and physically intuitive.

Acknowledgment

This work is supported by a funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.869855. JACR was funded by the Spanish Ministry of Universities through a 'Beatriz Galindo' fellowship (Ref. BG20/00143) and by the Spanish Ministry of Science and Innovation through the research project PID2020-119367RB-I00.

References

[1] R. A. Romeo and L. Zollo, "Methods and sensors for slip detection in robotics: A survey," *Ieee Access*, vol. 8, pp. 73 027–73 050, 2020.

[2] X. Song, H. Liu, J. Bimbo, K. Althoefer, and L. D. Seneviratne, "A novel dynamic slip prediction and compensation approach based on haptic surface exploration," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 4511–4516.

[3] T. Okatani, A. Nakai, T. Takahata, and I. Shimoyama, "A mems slip sensor: Estimations of triaxial force and coefficient of static friction for prediction of a slip," in *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*. IEEE, 2017, pp. 75–77.

[4] A. Mingrino, A. Bucci, R. Magni, and P. Dario, "Slippage control in hand prostheses by sensing grasping forces and sliding motion," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94)*, vol. 3. IEEE, 1994, pp. 1803–1809.

[5] J. L. Novak, "Initial design and analysis of a capacitive sensor for shear and normal force measurement," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), Tech. Rep., 1988.

[6] S. Shirafuji and K. Hosoda, "Detection and prevention of slip using sensors with different properties embedded in elastic artificial skin on the basis of previous experience," *Robotics and Autonomous Systems*, vol. 62, no. 1, pp. 46–52, 2014.

[7] G. A. AL, B. El Achab Oussallam, and U. Martinez-Hernandez, "A single-chip multimodal tactile sensor for a robotic gripper," in *2021 20th International Conference on Advanced Robotics (ICAR)*, 2021, pp. 364–369.

[8] A. Karamipour, M. Babaeihassankolou, and S. A. A. Moosavian, "A novel tactile sensor for slip detection based on relative angular velocity," *IEEE Sensors Letters*, vol. 6, no. 8, pp. 1–4, 2022.

[9] I. Agriomallos, S. Doltsinis, I. Mitsioni, and Z. Doulgeri, "Slippage detection generalizing to grasping of unknown objects using machine learning with novel features," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 942–948, 2018.

[10] L. Zollo, G. Di Pino, A. L. Ciancio, F. Ranieri, F. Cordella, C. Gentile, E. Noce, R. A. Romeo, A. Dellacasa Bellingegni, G. Vadalà *et al.*, "Restoring tactile sensations via neural interfaces for real-time force-and-slippage closed-loop control of bionic hands," *Science robotics*, vol. 4, no. 27, p. eaau9924, 2019.

[11] 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.

[12] 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.

[13] 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.

[14] I. H. Taylor, S. Dong, and A. Rodriguez, "Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 10 781–10 787.

[15] A. C. Abad and A. Ranasinghe, "Visuotactile sensors with emphasis on gelsight sensor: A review," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7628–7638, 2020.

[16] M. Lambeta, H. Xu, J. Xu, P.-W. Chou, S. Wang, T. Darrell, and R. Calandra, "Pytouch: A machine learning library for touch processing," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 208–13 214.

[17] J. Li, S. Dong, and E. Adelson, "Slip detection with combined tactile and visual information," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 7772–7777.

[18] S. Dong, D. Ma, E. Donlon, and A. Rodriguez, "Maintaining grasps within slipping bounds by monitoring incipient slip," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3818–3824.

[19] 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.

[20] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, p. 2762, 2017.

[21] S. Wang, Y. She, B. Romero, and E. Adelson, "Gelsight wedge: Measuring high-resolution 3d contact geometry with a compact robot finger," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 6468–6475.

[22] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer *et al.*, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.

[23] J.-D. Durou, M. Falcone, Y. Quéau, and S. Tozza, "A comprehensive introduction to photometric 3d-reconstruction," *Advances in Photometric 3D-Reconstruction*, pp. 1–29, 2020.

[24] E. W. Weisstein, "Normal vector," *https://mathworld. wolfram. com/*, 2002.

[25] "Fast poisson reconstruction in python," https://gist.github.com/jackdoerner/b9b5e62a4c3893c76e4c.