Where am I Touching? Solved by Matching Tactile Features with Visual Object Map

Shan Luo, Kaspar Althoefer and Hongbin Liu

Centre for Robotics Research, Department of Informatics, King's College London, WC2R 2LS, UK

shan.luo, k.althoefer, hongbin.liu@kcl.ac.uk

Abstract—**This paper proposes a novel approach to localize the contact location of a tactile sensor on an object. By taking advantage of some correspondences between visual and tactile sensing, it is proposed to localize tactile readings in visual images by sharing same sets of feature descriptors through two sensing modalities. Thus the recursive Bayesian filtering is employed and feature-based measurement model and Gaussian based motion model are built. In our tests, a tactile array sensor is utilized to generate tactile images during interaction with objects and test results have proved its feasibility.**

I. INTRODUCTION

Humans tend to utilize both vision and the tactile sensations of hands during grasping and/or manipulating objects. Key features, e.g., corners and edges, are first captured by our eyes and direct hands to an appropriate pose to handle objects. However, these visual features become unobservable during the manipulations in hand as vision is occluded. In this case, touch sensation makes up the information loss caused by the occlusion. The corresponding features are sensed in the tactile modality and poses of objects can thus be inferred. In this process prominent features act as a bridge between vision and tactile sensing, making our hand movements consequent. To apply this biological mechanism in robotics, we propose to use same feature descriptors in both vision and tactile sensing to localize the tactile features in visual images.

Earlier researchers have attempted to fuse vision and touch since decades ago but tactile sensors were only utilized to verify contacts or infer force direction due to the low resolution [1] [2]. Thanks to the increasing spatial resolution and spatiotemporal response, recently developed off-the-shelf tactile sensors have shown the ability to serve as "imaging" devices. Schneider *et al*. [3] took tactile images as features directly to recognize objects in a framework of Bag-of-Words (BoW) originated from computer vision. By use of the same streamline, Pezzementi *et al*. [4] took one step further: multiple kinds of features were extracted from tactile readings and their performances were compared. Liu *et al*. [5] recognized objects by classifying local features through the covariance analysis of pressure values in tactile images. All of these works show the potential of taking both visual and tactile readings as images and extracting features of the same type through two modalities for object contact estimation. Thus we propose to localize tactile readings in visual images using recursive Bayesian filtering, in which feature descriptors of the same type are extracted from both visual and tactile images. This novel approach is promising to be used to facilitate manipulations in hand.

Figure 1 illustrates our experimental system: a webcam (not shown here) and a Weiss tactile sensor of 14×6 sensing elements are utilized to obtain visual and tactile images of objects respectively; a Phantom Omni haptic device to which the tactile sensor is attached is employed to acquire the positions of the tactile sensor. A 3D printed gecko model is utilized to test our algorithm, which has a 2D shape that protrudes from its base 4mm. The sensor is controlled to press at multiple steps to follow its surface and the sensor plane is kept normal to the surface of the gecko.

Tactile Sensor

Tactile Patch
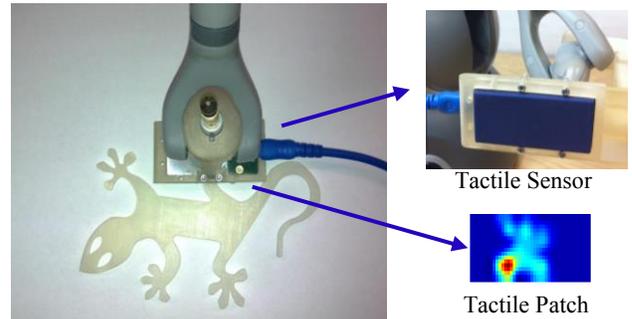
Fig. 1. Experimental setup, which consists of a webcam (not shown here) and a tactile sensor attached to a manipulator. A 3D print gecko model is selected as the test object.

II. METHODOLOGY AND EXPERIMENTS

The recursive Bayesian filtering is employed to estimate the location (state) x_t of the tactile sensor in the visual map m at each time step t, as illustrated in Fig. 2. The tactile readings z_t and movements of the tactile sensor u_t in the 3D space are collected to update the belief distributions over possible locations (states) in the visual map. Based on the input streams, it can be divided into two steps, i.e., control update and measurement update.

Tactile/ position readings

Recursive filtering

Tactile sensor

Estimation

Belief Prediction
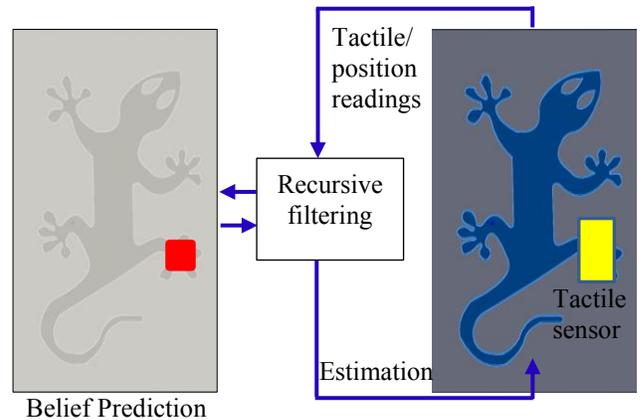
Fig. 2. Robot object interaction for localization and the data flow. Left: prediction of the contact location in the visual map, with the red region as the most probable contact location. Right: data flow of the control and measurement update.

In control update, the belief $\hat{p}(x_t|u_t)$ at $x_t=(x_1,x_2)$ is obtained by summing each product of the probability at the deterministic state (x_1-a,x_2-b) and its nearest k states ($k=8$ in our case) at time t-1, measured in Euclidean distance, and the Gaussian probability density $f(i,j)$. Here, (i,j) is the shift from the deterministic state. In the measurement update, the belief $p(x_t)$ is obtained as the product of $\hat{p}(x_t|u_t)$ and the probability $p(z_t|x_t)$ that the measurement z_t may have been observed for each hypothetical posterior state x_t. $p(z_t|x_t)$ is inverse to the Euclidean distance $d(p_{i,j},z_t)$ between descriptors extracted from segmented sub-image $p_{i,j}$ with index (i,j) and the tactile image z_t.

These features are adapted Scale Invariant Feature Transform (SIFT) descriptors [6], taking the invariance to translation and rotation into consideration. Compared to the classic SIFT algorithm, scale-space pyramids and key point localization are removed as in [7]. As shown in Fig. 3, each tactile image is segmented into three overlapped sub-images of the same size and one SIFT descriptor is extracted from each sub-image, taking sub-image centers as "key points". More details can be found in our previous work [8].
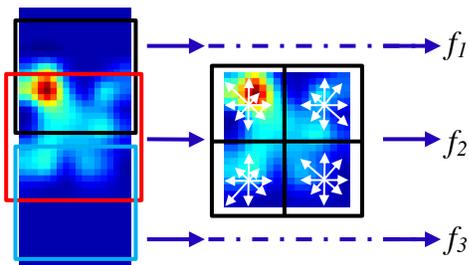
Fig. 3. Each tactile image is segmented into three overlapped sub-images and one 32-dimensional feature (f_1, f_2, f_3) is extracted from each sub-image.

The same approach is also applied to the visual images, extracting three SIFT descriptors from each sub-image. As shown in Fig. 4, a sliding window with the size of tactile image is carried out to get matching probability for each pixel.
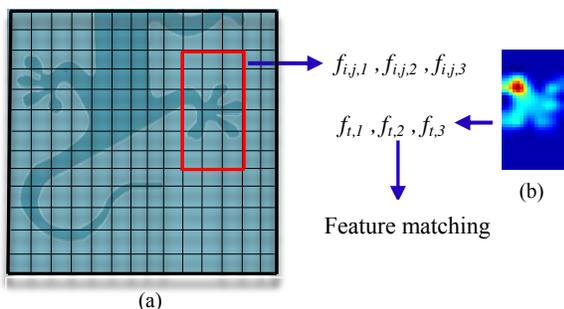
Fig. 4. Feature matching between visual and tactile images. (a) A sliding window (marked in red) in the visual image. (b) Tactile image z_t at time t.

In total, 22 experiments were carried out and in each of the experiment the tactile sensor explored the gecko model in different exploration paths. At the end of the exploration process, the robot can locate the tactile sensor in the gecko map with a large certainty. As shown in Fig. 5, it can be noticed that the localization errors tend to decrease through the localization process. To conclude, the experimental results prove the feasibility of our proposed framework in localizing the tactile sensor in a visual object map.
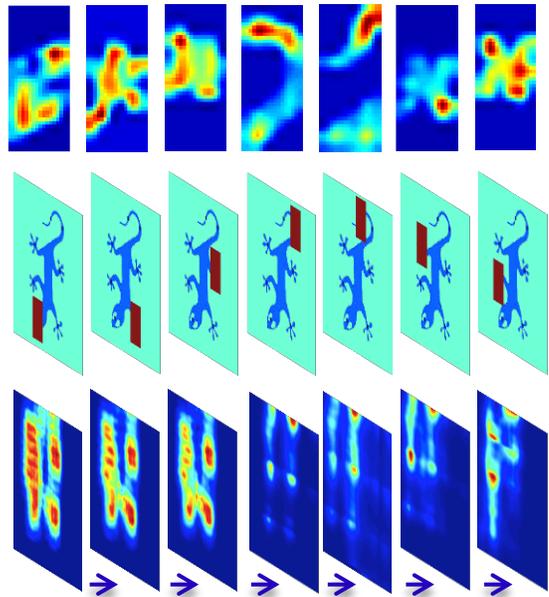
Fig. 5. Sample localization process and results. First row: obtained tactile images at each step; Second row: corresponding ground truth locations of the tactile sensor (marked as red) in the visual map. Last row: Corresponding probability distribution of locating the tactile sensor at different states.

III. CONCLUSION

This paper proposes a novel approach to integrate the vision and tactile sensing applied in localizing the tactile readings in the visual object image, which is solved by the recursive Bayesian filtering. The Gaussian noises are used to model the motion. In measurement update, revised SIFT descriptors are extracted from both the tactile and visual images and the belief distributions are updated by feature matching. Test results have proven its feasibility and it is promising to be used to facilitate robotic hand manipulations. More details can be found in [9].

REFERENCES

[1]   P. K. Allen, "Integrating vision and touch for object recognition tasks," *Int. J. Rob. Res.*, vol. 7, no. 6, pp. 15–33, 1988.
[2]   H. Wu, H. Liu, and D. Liu. "Two-Dimensional Direction Recognition Using Uniaxial Tactile Arrays," *IEEE Sens. J.*, vol. 13, no. 12, pp. 4897-4903, 2013.
[3]   A. Schneider, J. Sturm, C. Stachniss, M. Reisert, H. Burkhardt, and W. Burgard, "Object identification with tactile sensors using bag-of-features," in *IROS*, 2009, pp. 243–248.
[4]   Z. Pezzementi, E. Plaku, C. Reyda, and G. D. Hager, "Tactile-Object Recognition From Appearance Information," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 473–487, 2011.
[5]   H. Liu, X. Song, T. Nanayakkara, L. D. Seneviratne, and K. Althoefer, "A Computationally Fast Algorithm for Local Contact Shape and Pose Classification using a Tactile Array Sensor," in *ICRA*, 2012, pp. 1410–1415.
[6]   D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[7]   A. Bosch, A. Zisserman, and X. Mu, "Scene Classification via pLSA," in *ECCV*, 2006, pp. 517–530.
[8]   S. Luo, W. Mou, M. Li, K. Althoefer, and H. Liu, "Rotation and Translation Invariant Object Recognition with a Tactile Sensor," in *IEEE Sensors Conference*, 2014, pp. 1030–1033.
[9]   S. Luo, W. Mou, K. Althoefer and H. Liu, "Localizing the Object Contact through Matching Tactile Features with Visual Map", in *ICRA*, 2015, to appear.