Shape Reconstruction Task for Transfer of Haptic Information between Robotic Setups

Shubhan P. Patni and Matej Hoffmann

***Abstract*— Robot morphology, which includes the physical dimension and shape but also the placement and type of actuators and sensors, is highly variable. This also applies to different robot hand and grippers, equipped with force or tactile sensors. Unlike in computer vision, where information from cameras is robot and largely camera-independent, haptic information is morphology-dependent, which makes it difficult to transfer object recognition and other pipelines between setups. In this work, we introduce a shape reconstruction and grasping task to evaluate the success of haptic information transfer between robotic setups, and propose feature descriptors that can help in standardizing the haptic representation of shapes across different robotic setups.**

I. INTRODUCTION

There are numerous different grippers and end-effectors available for robotic arms both commercially and for research purposes. Further, each gripper may be fitted with sensors for tactile and proprioceptive feedback. There are also a variety of different tactile sensors available which can be integrated with grippers. We can classify the most commonly used tactile sensors into two classes based on their working principle—first, optical sensors like Gelsight [1], DIGIT [2] and TacTip [3] which use vision processing from a camera embedded behind the sensor's surface membrane to extract information from object interations. Second, transductive sensors that convert mechanical changes to electric signals for feedback, like the SynTouch BioTAC [4], SINGLEX [5], Contactile [6] and uSkin [7]. The transductive sensors can further be split into sub-types based on how the mechanical interaction is converted to electrical signals. Some grippers like the BarrettHand [8] and Shadow Hand [9] may also come readily integrated with their own in-house tactile sensors.

For example, in Figure 1, we can see different setups aiming to haptically explore the same object. The GelSight sensor will give feedback in the form of an image of its membrane, which is not interpretable by other sensors. The BioTac sensor will provide feedback as a time series of measured reactive force at each fingertip, while the RG6 gripper will report back the force at its actuator. The feedback from the iCub hand will comprise of a series of values for each of its 104 tactile sensing units—12 on each finger and 44 in the palm. With the large variety of gripper and tactile sensing options available, data collection to build larger datasets of tactile robot-object interaction and manipulation faces some challenges. It is highly dependent on the robotic setup, since each morphological combination of {robotic arm, gripper, tactile sensor} will have different data structures for how interaction feedback is reported and stored. Compiling data from different setups for morphology-independent learning and exploration requires the development of processing methods for each setup individually. These factors also hinder the development of multimodal sensing datasets [10]–[13], even though such datasets exist from independent works, since it is difficult to compile and compare them because of the varying nature of tactile feedback data.

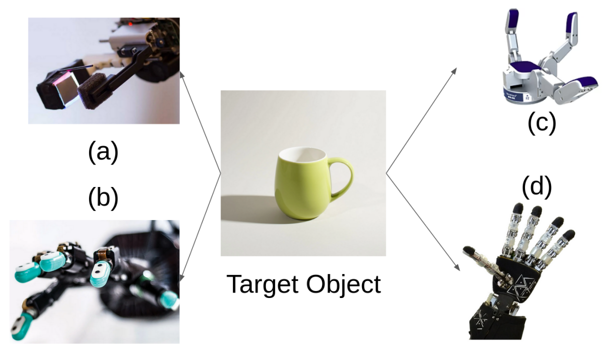


Fig. 1: Different robotic setups explore the same object but provide different feedback. (a) The GelSight sensor mounted on a two finger gripper provides image feedback. (b) The BioTac sensor mounted on a Shadow Hand provides a single time series force variation per finger. (c) The Barrett Hand and (d) the iCub Hand provide force feedback from 96 and 104 tactile units respectively.

This work suggests a shape reconstruction and grasping task to develop methods for transferring haptic object information among robotic setups. The objects used are from the YCB object dataset [14], although the set of object models will be unknown to the robot. We use one robotic setup to haptically explore the target object and aim to reconstruct its shape. Then, we transfer the learned shape to other setups to carry out grasping tasks, and evaluate the success of the grasping task. To transfer haptic information, a concept of morphology-independent haptic shape representation is introduced. The representation is based on how three-dimensional

S.P. and M.H. were supported by OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16_019/0000765 "Research Center for Informatics", and the Czech Technical University in Prague, grant no. SGS22/111/OHK3/2T/13.

Shubhan P. Patni is a PhD student with the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague patnishu@fel.cvut.cz

Matej Hoffmann is an Associate Professor with the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague matej.hoffmann@fel.cvut.cz

shapes are constructed in computer graphics [15]—simple features that may be extracted using any gripper and tactile sensor combination. Since it is possible to extract the same features for many different robotic setups, they should also act as a common interpretation of object shape information between the different setups.

II. LITERATURE REVIEW

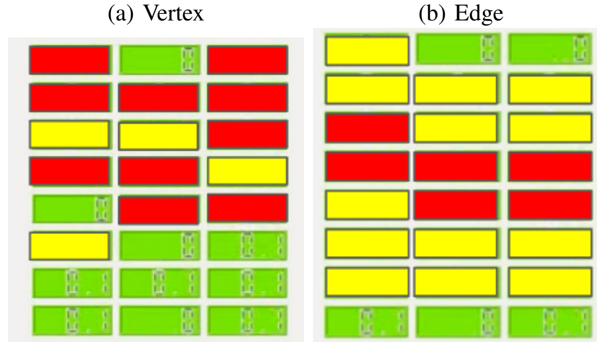
Most of the research including transfer learning and haptics focuses on transferring a network trained on visual data to infer haptic data, and vice versa. Recent works [16]–[19] investigate and implement different networks that can find correlations between visual and haptic input for object recognition. Navarro-Guerrero et al. [20] have done a comprehensive literature review of this connection between visual and haptic perception across different fields, from cognition and neuroscience to robotics. Our interpretation of the "transfer" of knowledge in "transfer learning" focuses less on the transfer between different sensory modalities, and more on the transfer between different physical embodiments of the same tactile modality.

Towards object reconstruction, Luo et al. [21] introduced iCLAP, an iterative touch exploration method to recreate objects accurately that decided exploration based on possible feature completion. Pezzementi et al. [22] used feature recognition from tactile arrays mounted on a gripper to formulate an exploration plan to reconstruct the shapes of objects. Rustler et al. [23] accomplish shape reconstruction by using a tactile "poking" action to detect contact with objects and subsequently filling the Cartesian space. The tactile exploration is paired with vision to generate object models with an Implicit Geometric Regularization Network (IGR) [24]. Very good accuracy was achieved for shape reconstruction within five to ten touches when visual data was also provided. Other works include edge following with the iCub robot [25] and a simulator for shape reconstruction [26].

To standardize tactile sensing across different robotic setups, there have been attempts in recent years to correlate it to visual data. Le et al. [27] and Takahashi et al. [28] attempted to correlate time series of vision and tactile data by using Decoder-Encoder agents that found common latent spaces between the two modalities and found it possible to predict next-frame touch from video, and images of object surfaces from the tactile data. Zambelli et al. [29] created a multimodal datset with the iCub robot to correlate all visual, haptic, and audio sense to each other. Each of these works perform very well, but are their learning is limited to their own setups. An interesting attempt towards aggregating and interpreting raw data from large tactile arrays on grippers was presented in [30], which used convolutional neural networks (CNNs) fine tuned for their morphology and data stream coming from their setup to achieve good physical property estimation results. The network was named the "Morphology Specific CNN", which showcases the difficulty in generalizing learning networks or data collection across different robotic setups.

III. TOUCH PRIMITIVES

(a) Vertex (b) Edge



(c) Flat Surface (d) Curved Surface

Fig. 2: Detection of different shape primitives on the fingers and palm of the Barrett Hand.

A fundamental way to describe the shapes of objects is to break them down into vertices, edges and surfaces [15]. Although rudimentary, these four primitives are the foundation to defining three-dimensional shapes. This simplistic method does not take into account detailed features like surface texture or the roughness of materials, but by recognizing these features humans are able to achieve successful object recognition rapidly. We propose the identification of these same features by grippers as "touch primitives". Each exploration of the object can give multiple touch primitives depending on the number of fingers in the gripper and number of regions fitted with tactile arrays, and with sufficient sampling we are able to collect enough primitives to create a volumetric boundary for the object. The possible types of primitives are: vertex, edge, flat surface, and curved surface. As long as a gripper/sensor combination is able to detect these features, they should be able to record data in the suggested morphology-independent format below. An object exploration O is described as a set of primitives P, where each primitive is described by:

$$P = \{\vec{x}, i, \vec{d}_i\} \tag{1}$$

where $\vec{x}$ is the position and orientation of the finger or tactile array that detected the primitive, i is the type of primitive, holding values 1 to 4 for "vertex", "edge", "flat_surface" and "curved_surface" respectively. The primitive descriptor $\vec{d}_i$ helps us orient the primitive in 3D space with respect to $\vec{x}$. For the "vertex" primitive, it is the same as $\vec{x}$. For "edge"

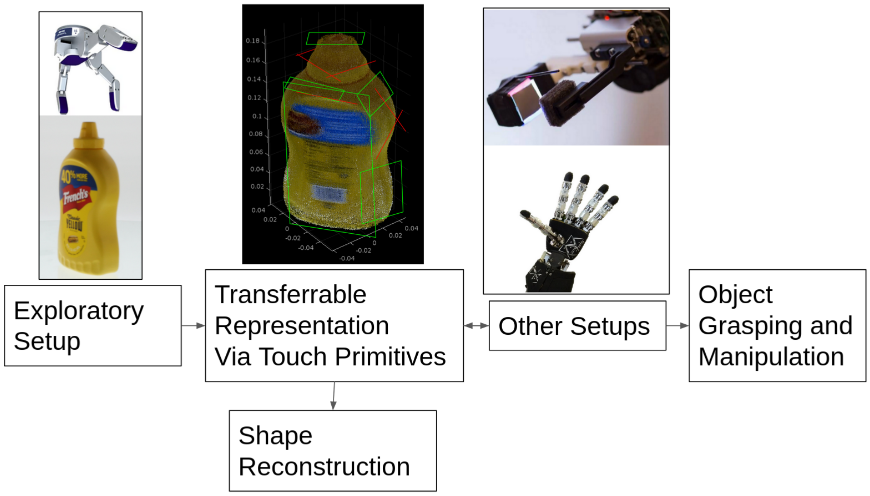


Fig. 3: Complete task for the transfer of object information between setups.

primitives, it is the slope of the detected edge on the surface of the tactile array. For the surface primitives, it is the normal to the tactile array at the point of maximum force. As a first implementation for the extraction of touch primitives, we choose the Barrett Hand, a three-fingered anthropomorphic hand with torque and position feedback, as well as arrays of tactile sensors on each finger and the palm. The alignment of the tactile sensors on the hand allows for the recognition of the simple touch primitives, as shown in Figure 2.

This concept of using touch primitives for object representation opens up many avenues for future research as well. First, how these primitives can be used to recreate and recognize object shapes with accuracy. This is discussed as our target task in the next section as well. Further, how different grippers can extract the same primitives. Integrating grippers with various tactile sensors provides them with such capabilities, however, we can also investigate whether the same primitives can be recognized by the grippers as a result of sequential exploration without rich tactile feedback. Maye et al. [31] describe how sensorimotor contingency theory can be used to infer context and information from a robot based on its state and action exploration sequence. This theory may allow grippers without tactile sensing capabilities to extract touch primitives from target objects as well. Next, once the extraction of geometric touch primitives is accomplished, there will also be the possibility to integrate physical properties into the same description vector with stiffness and friction maps as seen in [32]. Finally, if the proposed method of representing objects is accurate and useful in transferring object information between robotic setups, generating grasp proposals using touch primitives as inputs can be explored.

IV. SHAPE RECONSTRUCTION AND EVALUATION TASK

Our proposed task to develop methods for haptic object information passing between different robotic setups is the task of shape reconstruction and grasping for unknown target objects. This will first be carried out on the primary setup—a BarrettHand gripper mounted on a KUKA LBR IIWA arm, extracting touch primitives from target objects in a pre-defined sequential exploration. A pre-defined sequence is easier for exploration when vision is not being used. Then the collected information may be communicated to other setups, (a) the Robotiq 2F-85 gripper, equipped with the DIGIT tactile sensor, mounted on a KUKA LBR IIWA robotic arm, and (b) the iCub robot integrated with its own tactile array. Successful task completion will be evaluated via the following metrics:

- The shape reconstruction on the primary setup is measured by evaluating the Intersection over Union metric, comparing the reconstructed shape to ground truth point cloud models of the target objects. These models will not be known beforehand to the robot, and will only be used during evaluation.
- Successful information transfer via touch primitives is observed via successful grasping of the target object by the other robot setups. A practical metric will be used to test grasp success, as suggested in Le et al. [33]. The test comprises of grasping the object using proposed grasps, shaking it, and measuring the time it takes for the object to lose contact with the gripper. They show that better grasping proposals improve the shake time as well.

As a smaller goal, preliminary work is restricted to recognition of the target object from a pool of already-known object models stored in memory, rather than attempting reconstruction of unknown objects, which is a more complex exploration activity and more susceptible to environment and computational noise. For object recognition from touch primitives, we are planning to compare the methods noted below. Exploring these shape completion methods with our touch primitives as input will also help with the eventual task of shape reconstruction of unknown objects.

References

[1] 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.

[2] 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.

[3] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, and N. F. Lepora, "The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies," *Soft robotics*, vol. 5, no. 2, pp. 216–227, 2018.

[4] N. Wettels, J. A. Fishel, Z. Su, C. H. Lin, G. E. Loeb, and L. SynTouch, "Multi-modal synergistic tactile sensing," in *Tactile sensing in humanoids—Tactile sensors and beyond workshop, 9th IEEE-RAS international conference on humanoid robots*, 2009.

[5] "Singlex sensorsfrom seed robotics," https://www.seedrobotics.com/fts-tactile-pressure-sensors.

[6] H. Khamis, B. Xia, and S. J. Redmond, "A novel optical 3d force and displacement sensor–towards instrumenting the papillarray tactile sensor," *Sensors and Actuators A: Physical*, vol. 291, pp. 174–187, 2019.

[7] T. P. Tomo, A. Schmitz, W. K. Wong, H. Kristanto, S. Somlor, J. Hwang, L. Jamone, and S. Sugano, "Covering a robot fingertip with uskin: A soft electronic skin with distributed 3-axis force sensitive elements for robot hands," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 124–131, 2017.

[8] W. Townsend, "The barretthand grasper–programmably flexible part handling and assembly," *Industrial Robot: an international journal*, vol. 27, no. 3, pp. 181–188, 2000.

[9] "Shadowhand sensor," https://www.shadowrobot.com/.

[10] A. Burns, X. Fan, J. Pinkenburg, D. Lee, V. Isler, and D. Lee, "Multimodal dataset for human grasping," in *The 29th International Conference on Robot and Human Interactive Communication Workshop*, vol. 3, 2020.

[11] V. Chu, I. McMahon, L. Riano, C. G. McDonald, Q. He, J. M. Perez-Tejada, M. Arrigo, T. Darrell, and K. J. Kuchenbecker, "Robotic learning of haptic adjectives through physical interaction," *Robotics and Autonomous Systems*, vol. 63, pp. 279–292, 2015.

[12] R. Calandra, A. Owens, M. Upadhyaya, W. Yuan, J. Lin, E. H. Adelson, and S. Levine, "The feeling of success: Does touch sensing help predict grasp outcomes?" *arXiv preprint arXiv:1710.05512*, 2017.

[13] O. Kroemer, C. H. Lampert, and J. Peters, "Learning dynamic tactile sensing with robust vision-based training," *IEEE transactions on robotics*, vol. 27, no. 3, pp. 545–557, 2011.

[14] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: The ycb object and model set and benchmarking protocols," *arXiv preprint arXiv:1502.03143*, 2015.

[15] V. E. Arriola-Rios, P. Guler, F. Ficuciello, D. Kragic, B. Siciliano, and J. L. Wyatt, "Modeling of deformable objects for robotic manipulation: A tutorial and review," *Frontiers in Robotics and AI*, vol. 7, p. 82, 2020.

[16] P. Falco, S. Lu, C. Natale, S. Pirozzi, and D. Lee, "A transfer learning approach to cross-modal object recognition: from visual observation to robotic haptic exploration," *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 987–998, 2019.

[17] J. M. Gandarias, A. J. Garcia-Cerezo, and J. M. Gomez-de Gabriel, "Cnn-based methods for object recognition with high-resolution tactile sensors," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6872–6882, 2019.

[18] C. Sferrazza and R. D'Andrea, "Transfer learning for vision-based tactile sensing," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 7961–7967.

[19] M. Kaboli, D. Feng, and G. Cheng, "Active tactile transfer learning for object discrimination in an unstructured environment using multimodal robotic skin," *International Journal of Humanoid Robotics*, vol. 15, no. 01, p. 1850001, 2018.

[20] N. Navarro-Guerrero, S. Toprak, J. Josifovski, and L. Jamone, "Visuo-haptic object perception for robots: an overview," *Autonomous Robots*, pp. 1–27, 2023.

[21] S. Luo, W. Mou, K. Althoefer, and H. Liu, "Iterative closest labeled point for tactile object shape recognition," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 3137–3142.

[22] Z. Pezzementi, E. Plaku, C. Reyda, and G. D. Hager, "Tactile-object recognition from appearance information," *IEEE Transactions on robotics*, vol. 27, no. 3, pp. 473–487, 2011.

[23] L. Rustler, J. Lundell, J. K. Behrens, V. Kyrki, and M. Hoffmann, "Active visuo-haptic object shape completion," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5254–5261, 2022.

[24] A. Gropp, L. Yariv, N. Haim, M. Atzmon, and Y. Lipman, "Implicit geometric regularization for learning shapes," *arXiv preprint arXiv:2002.10099*, 2020.

[25] U. Martinez-Hernandez, G. Metta, T. J. Dodd, T. J. Prescott, L. Natale, and N. F. Lepora, "Active contour following to explore object shape with robot touch," in *2013 World Haptics Conference (WHC)*. IEEE, 2013, pp. 341–346.

[26] E. Smith, D. Meger, L. Pineda, R. Calandra, J. Malik, A. Romero Soriano, and M. Drozdzal, "Active 3d shape reconstruction from vision and touch," *Advances in Neural Information Processing Systems*, vol. 34, pp. 16 064–16 078, 2021.

[27] M. A. Lee, Y. Zhu, K. Srinivasan, P. Shah, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8943–8950.

[28] K. Takahashi and J. Tan, "Deep visuo-tactile learning: Estimation of tactile properties from images," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8951–8957.

[29] M. Zambelli, A. Cully, and Y. Demiris, "Multimodal representation models for prediction and control from partial information," *Robotics and Autonomous Systems*, vol. 123, p. 103312, 2020.

[30] S. Funabashi, G. Yan, F. Hongyi, A. Schmitz, L. Jamone, T. Ogata, and S. Sugano, "Tactile transfer learning and object recognition with a multifingered hand using morphology specific convolutional neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[31] A. Maye and A. K. Engel, "A discrete computational model of sensorimotor contingencies for object perception and control of behavior," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 3810–3815.

[32] T. N. Le, F. Verdoja, F. J. Abu-Dakka, and V. Kyrki, "Probabilistic surface friction estimation based on visual and haptic measurements," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2838–2845, 2021.

[33] T. N. Le, J. Lundell, F. J. Abu-Dakka, and V. Kyrki, "A novel simulation-based quality metric for evaluating grasps on 3d deformable objects," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 3123–3129.