Estimating High-Resolution Deformable Object Tactile Models using Neural Stiffness Fields

Jiaheng Han* Shaoxiong Yao* Kris Hauser

Abstract— **High-resolution visuotactile sensors provide detailed contact information that can infer the physical properties of objects in contact. This paper introduces a novel technique for high-resolution stiffness estimation of heterogeneous deformable objects using the Punyo bubble sensor. We developed an observation model for dense contact forces to estimate an object's stiffness using a visuotactile sensor and a dense force estimator. Additionally, we propose a neural Volumetric Stiffness Field (VSF) formulation that represents stiffness as a continuous function, which allows dynamic point sampling at visuotactile sensor observation resolution. The neural VSF effectively reduces artifacts commonly found in traditional point-based methods, especially in heterogeneous stiffness estimation. We incorporate a triangle mesh of the object to guide stiffness field estimation and utilize rigid-body tracking to integrate multiple visuotactile sequences from different touches on a movable object. These techniques significantly improve the quality and completeness of VSF estimation.**

I. Introduction

Tactile perception of physical properties, such as stiffness, is crucial for tasks including manipulation in clutter [1] and deformable object packaging [2]. Recent advances in vision-based tactile sensors capture high-resolution tactile images by observing the deformation of media like gels [3] or bubbles [4] using embedded cameras. These high-resolution tactile observations provide detailed contact information and are promising for fine-grained texture recognition and material parameter estimation. However, the vast majority of prior work in tactile sensing considers contact with rigid objects [5], [6]. Deformable objects undergo deformation in a manner that is coupled to the sensor medium, and material parameter estimation requires simultaneous understanding of both object and sensor geometry, as well as their respective deformations [7]. Moreover, many objects of interest for high-resolution tactile sensing are heterogeneous, with a mix of hard and soft parts, such as bagged objects or localization of hard tissues underneath soft tissues in palpation. Existing identification techniques based on analytical models, e.g., finite element models (FEM), do not scale easily to large numbers of material parameters and handling deformable-deformable contact can be challenging [8].

In this paper, we propose a system that captures both tactile models for objects composed of complex, heterogeneous materials. Our approach is based on the Volumetric Stiffness Field (VSF), a high-capacity model designed to represent heterogeneous material distributions [9], [10]. Specifically, we introduce neuralVSF, a neural-based tactile simulation acquisition system that learns from real-world tactile data to generate high-fidelity simulation assets. In summary, this paper has the following contributions: 1) We formulate an observation model for dense force estimation using visuotactile sensors combined with the VSF tactile model; 2) We propose a novel neural VSF tactile model capable of capturing high-resolution stiffness variations using a continuous function.

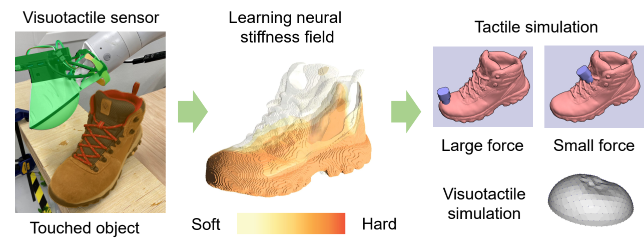

Fig. 1. Our method estimates a high-resolution stiffness field of a deformable object using a visuotactile sensor. The estimated stiffness field can simulate contact forces for novel touches and the deformation of the visuotactile sensor upon contact with the object.

II. Method

The input to our systems includes a volume containing the object, a set of trajectories from a high-resolution tactile sensor, and the corresponding sequences of sensor readings. We assume that the visuotactile sensor readings are displacements of the contact surface, and we have access to a calibrated FEM model of the tactile sensor that relates contact forces to these displacements. The object is assume to have elastic deformation. Our high-resolution VSF tactile estimation system, built on a visuotactile sensor, consists of four key components: a dense contact force estimator using [11], a VSF representation, a dense contact force observation model for the VSF, and a VSF estimation algorithm. To enhance material estimation quality, we incorporate object geometry information, significantly improving the results. To integrate tactile data of a movable object, we employ 6DoF tracking to transform touch data into the object's local frame, enabling a more complete estimation of neuralVSF.

Continuous VSF model. Our key innovation is to treat the stiffness as a continuous field $K(\cdot) : \Omega \to \mathbb{R}_+$, where we can dynamically sample points at visuotactile sensor force estimation resolution. For a tiny volume dV centered around $p \in \Omega$, there exists a Hookean spring with stiffness $K(p)dV$. When the robot pushes through this volume, the object deforms according to a time-dependent continuous deformation field $u(p,t) : \Omega \times \mathbb{R} \to \mathbb{R}^3$.

*These authors contributed equally to this work.

J. Han, S. Yao, and K. Hauser are with the Siebel School of Computing and Data Science, University of Illinois at Urbana-Champaign, IL, USA. {jh127, syao16, kkhauser}@illinois.edu. This project was partially supported by USDA/NIFA Grant #2020-67021-32799 and an equipment loan from Toyota Research Institute.

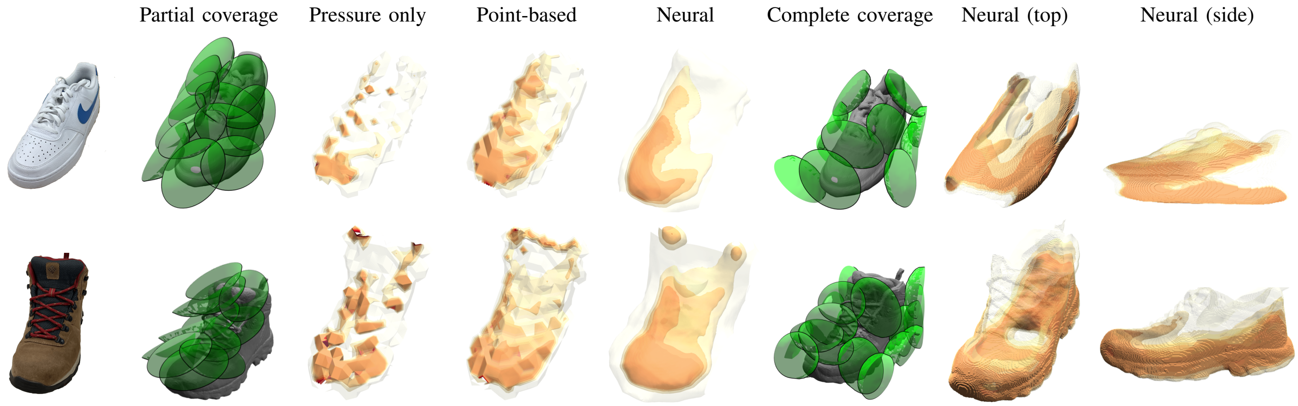

Fig. 2. Qualitative stiffness estimation comparison of Pressure-only, Point-based VSF, and Neural VSF using dense force estimation. Following partial coverage, the left columns show the estimated stiffness using top-down touches on the object. The right two columns demonstrate improved Neural VSF estimation with an SDF mask and multiple touches from different directions for complete coverage.

Dense force observation model for continuous VSF. The continuous stiffness field enables direct evaluation of dense contact forces using volumetric integration. We assumes dense contact forces are estimated on the surface of the visuotactile sensor as a triangle mesh. For each vertex k, moving along its trajectory $v_k(t) : \mathbb{R} \to \mathbb{R}^3$, we expect the VSF contact force $\hat{f}_{v,k}(t)$ is the integration of all Hookean springs within its swept volume Ω_k.

$$\begin{aligned}\hat{f}_{v,k}(t) &= \int_{\Omega_k} -K(p)u(p,t)dV \\ &= \int_{v_k(0)}^{v_k(t)} -K(v_k(\tau))u(v_k(\tau),t)s_k(\tau)\cdot dv_k(\tau) \\ &\approx \sum_{i=1}^{N} -K(v_k(\tau_i))u(v_k(\tau_i),t)s_k(\tau_i)\cdot \delta v_k(\tau_i).\end{aligned} \tag{1}$$

where $K(v_k(\tau))$ represents the local stiffness at the vertex, and $u(v_k(\tau), t)$ denotes the deformation at the vertex position $v_k(\tau)$ over time. Here the continuous line integration for $\hat{f}_{v,k}(t)$ is approximated by sampling points along vertex trajectories from 0 to time t, where $\delta v_k(\tau_i) = v_k(\tau_i) - v_k(\tau_{i-1})$ is the deformation.

Continuous VSF instantiation and training. We instantiate the continuous VSF using a NeRF-like neural network as $K(p) = g_\psi(p) : \mathbb{R}^3 \to \mathbb{R}_+$. The network g_ψ has an 8-layer architecture with sinusoidal position encoding functions, similar to NeRF [12], to capture high-frequency details of the objects. We optimize ψ by minimizing the following loss function on dense contact forces:

$$\ell(\psi) = \sum_{i=1}^{B} ||f_{v,k_i}^{t_i} - \hat{f}_{v,k_i}(t_i)||_2 + \lambda \frac{1}{n_{\text{reg}}} \sum_{i=1}^{n_{\text{reg}}} K(p_i) \tag{2}$$

Neural VSF is trained over multiple touch sequences that touch the object at different locations. We used the Adam optimizer [13] and batched force observations. Here, the second term is a free space regularization term that encourages the neural network zero stiffness outputs in untouched regions. This regularization is implemented by randomly sampling n_{reg} points within the bounding box and applying a small penalty to their stiffness values.

Improve NeuralVSF quality and completeness. Noisy dense contact forces estimation causes the VSF to show non-zero stiffness near the object boundary, while neural networks tend to smooth object edges. To solve these problems, we utilize the object's SDF as a geometric mask for the VSF by enforcing zero stiffness for points outside the object, formulated as $K(p) = \mathbb{I}_{sdf(p)<0} \cdot g_\psi(p)$. This ensures that the VSF produces sharp edges at the object boundary.

For objects placed on a table, we can only touch them from the top down, leading to incomplete estimations. The objects are freely placed on the table, and we use FoundationPose [14] to track their translation and rotation during data collection. The observed contact forces are then transformed into the object's local frame to reconstruct the VSF model.

III. Experiments and Results

Setup and baselines. We evaluate VSF variants on heterogeneous stiffness estimation with different shoes, each with higher stiffness in the toe and bottom region and lower stiffness in the tongue region. We compare three different VSF formulations, including the pressure-only method mentioned above, point-based VSF and neural VSF. For both the pressure-only model and the point-based model, we generate 20,000 VSF points from RGBD image. For neural VSF, the bounding box of the object is used to define object space.

Results. The stiffness estimation results are illustrated in Fig. 2. When comparing the point-based VSF to the neural VSF, it is evident that the point-based method exhibits more artifacts. With an SDF mask and multiple touch poses, neural VSF produces a more complete and higher-quality stiffness field that better respects object geometry.

IV. Conclusion and Future Work

In summary, this paper proposes a system designed to capture tactile models for deformable objects using visuotactile sensors. Our approach builds upon neural VSF method, which captures the heterogeneous material properties of deformable objects. We validated our system on objects made of heterogeneous materials. The estimation quality and completeness improved further by incorporating geometry information and tactile inputs from multiple directions. Moving forward, we plan to open-source high-quality tactile models to support the research community.

REFERENCES

[1] A. Jain, M. D. Killpack, A. Edsinger, and C. C. Kemp, "Reaching in clutter with whole-arm tactile sensing," *The International Journal of Robotics Research*, vol. 32, no. 4, pp. 458–482, 2013.

[2] B. Ai, S. Tian, H. Shi, Y. Wang, C. Tan, Y. Li, and J. Wu, "RoboPack: Learning Tactile-Informed Dynamics Models for Dense Packing," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024.

[3] W. Yuan, S. Dong, and E. Adelson, "GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force," *Sensors*, vol. 17, no. 12, p. 2762, Nov. 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/12/2762

[4] A. Alspach, K. Hashimoto, N. S. Kuppuswamy, and R. Tedrake, "Soft-bubble: A highly compliant dense geometry tactile sensor for robot manipulation," *2019 2nd IEEE International Conference on Soft Robotics (RoboSoft)*, pp. 597–604, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:102336499

[5] R. Calandra, A. Owens, D. Jayaraman, J. Lin, W. Yuan, J. Malik, E. H. Adelson, and S. Levine, "More than a feeling: Learning to grasp and regrasp using vision and touch," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3300–3307, 2018.

[6] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *Mechatronics*, vol. 48, pp. 54–67, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957415817301575

[7] W. Yuan, S. Wang, S. Dong, and E. Adelson, "Connecting look and feel: Associating the visual and tactile properties of physical materials," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[8] B. Frank, C. Stachniss, R. Schmedding, M. Teschner, and W. Burgard, "Learning object deformation models for robot motion planning," *Robotics and Autonomous Systems*, vol. 62, no. 8, pp. 1153–1174, 2014.

[9] S. Yao and K. Hauser, "Estimating tactile models of heterogeneous deformable objects in real time," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 12 583–12 589.

[10] S. Yao, Y. Zhu, and K. Hauser, "Structured bayesian meta-learning for data-efficient visual-tactile model estimation," in *8th Annual Conference on Robot Learning*, 2024.

[11] J.-C. Peng, S. Yao, and K. Hauser, "3d force and contact estimation for a soft-bubble visuotactile sensor using fem," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 5666–5672.

[12] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2015. [Online]. Available: http://arxiv.org/abs/1412.6980

[14] B. Wen, W. Yang, J. Kautz, and S. Birchfield, "FoundationPose: Unified 6d pose estimation and tracking of novel objects," in *CVPR*, 2024.