ME8135 - Preliminary Literature Review

Austin Vuong

I. Introduction

THIS review is a combination of a preliminary literature review for the State Estimation course, and for the current direction of my research. The eventual end goal of the latter requires several components working together which thus far are identified as: (1) 3D reconstruction, (2) dynamically, (3) in real-time, (4) with uncertainty modeling, (5) for monitoring and control feedback, (6) in additive manufacturing. While crucial to the end goal, aspects 2-6 can be variously factored out and integrated later as needed. The core focus is then the choice of 3D reconstruction, in particular the rapidly developing area of neural network based representations.To the best of my present knowledge, I have found no applications of these methods to additive manufacturing. The closest known work is [1] which uses a neural field to represent visibility and accessibility of a multi-axis machine.

This is a work-in-progress and its current state is more a collection of works than a coherent narrative or review. For now, wholly classical methods are excluded. Several sections are placeholders with bullet points to note a few works. Some works may be in the references but not yet incorporated into this version. Writing notes, such as todo's, are left in so that known gaps are acknowledged.

A. Neural Fields

The application of machine learning techniques to visual computing tasks has seen significant attention in recent years. The subject has been called implicit neural representations, neural implicits, or coordinate-based neural networks, while a formalization has termed them neural fields [2]. In general, these are continuous and defined at all coordinates, and can be represented as a mapping, $\mathbf{F}$, between coordinates, $\mathbf{x}$ and some quantities of interest, $\mathbf{q}$, parameterized fully or in part by a neural network.

$$\mathbf{F}(\mathbf{x}) = \mathbf{q} \tag{1}$$

As universal approximators, a sufficiently large neural network can encode continuous signals over arbitrary dimensions at arbitrary resolution [2]. While the memory of discrete parameterizations scale poorly with resolution, neural fields scale instead based on the network complexity which allows for improved adaptability.

Two popular types of fields in terms of 3D reconstructions are radiance fields and signed distance fields (SDF). The former is typically used in volumetric rendering and can be formulated as a position and camera direction input, with a color and density output.

$$\mathbf{F}(x, y, z, \theta, \phi) = (R, G, B, \sigma) \tag{2}$$

Meanwhile, SDFs are commonly used for rendering, collision detection and motion planning, and are formulated as position inputs and a scalar output representing the distance to the surface.

$$\mathbf{F}(x, y, z) = d \tag{3}$$

Neural implementations of these fields can be seen in examples such as [3], [4]. These fields can also be used as primitives for downstream tasks, such as generating meshes [5]. They can also be used in conjunction with classical methods to produce hybrid representations that can bring a range of benefits and trade-offs. (#todo expand on said benefits and trade-offs, likely in the relevant sections)

(#todo a note on identity mapping [2] (section 4.3) , where the goal is to deliberately overfit a network to "memorize" data. This addresses the general negative view of overfitting neural networks.)

B. Real-time Generation

- NGLOD [6] accelerates neural SDFs using a sparse voxel grid where smaller networks are used within the grid, rather than a monolithic network
- instant-ngp [7] accelerates neural fields (in this context called neural graphics primitives) similarly to NGLOD by using smaller networks and a spatial hash. Application to both neural SDF and NeRF is demonstrated
- [8] shares the goals of NeRF reconstructions while using a different core structure, "splatted" 3D gaussians. The result is the state-of-the-art quality of Mip-NeRF360[9] with the training times of instant-NGP[7]
- (I give a nod to NeRF-based SLAM in general which implement the aspects of real-time processing, dynamic environments, and uncertainty modeling. However, these are not as immediately relevant so I've not yet looked at them deeply. I will make note of a few works:)
  - iNeRF [10] inverts an existing NeRF to produce pose estimates
  - iMAP [11] reconstructs a monolithic NeRF for SLAM
  - NICE-SLAM [12] uses a hierarchical NeRF for faster and more stable representations than iMAP
  - NICER-SLAM [13] improves on the prior's reconstruction quality at the loss of real-time reconstruction
  - iSDF [14] constructs a map using neural (and non-truncated) SDF

C. Geometry Changes

(These are works related to changing neural fields. In general there are two broad categories: dynamic scenes, such as HyperNeRF[15], volumetric video[16], and some NeRF-based SLAM and editing geometry, such as NeuralEditor[17].

Editing geometry is in concept closer to the process of additive manufacturing. I have not looked at these in detail so I omit saying much about them at this time.)

D. Uncertainty Modeling

Modeling uncertainty of motion and sensor inputs is crucial for robots to reliably navigate environments. Neural fields, especially NeRFs, do not inherently encode uncertainties. NeRF in the wild (NeRF-W) [18] models uncertainty in dynamic environments by extending the output to include a variance value. This captures the changes of dynamic objects in a scene. However, this model is used to filter moving entities such as people and vehicles, rather than of measurement confidence. Stochastic-NeRF (S-NeRF) [19] is the earliest known work to address measurement confidence by modifying the output to be stochastic variables that follow the learned distributions. As the approach is a generalization of the original NeRF it can be combined with other NeRF frameworks to bring in Bayesian uncertainty estimations.

[20] explores vision-only navigation in an offline NeRF. Here, density is used as a strong proxy for occupancy of a position. In addition to a photometric loss based on iNeRF [10], a process loss is included in the cost function to produce state estimates.

E. Unincorporated Works

These are a set of notable works that do not yet have a home here.

- EG3D [21] is a different approach to 3D reconstruction that combines explicit and implicit (NeRF) representations as part of a generative adversarial network
- Mip-NeRF [22] borrows the mipmapping graphics technique, using multiple resolutions to improve quality and speed

References

[1] G. Harabin, A. M. Mirzendehdel, and M. Behandish, "Deep Neural Implicit Representation of Accessibility for Multi-Axis Manufacturing," *Computer-Aided Design*, vol. 163, p. 103556, Oct. 2023.

[2] Y. Xie, T. Takikawa, S. Saito, O. Litany, S. Yan, N. Khan, F. Tombari, J. Tompkin, V. Sitzmann, and S. Sridhar, "Neural Fields in Visual Computing and Beyond," *Computer Graphics Forum*, vol. 41, no. 2, pp. 641–676, May 2022.

[3] 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," in *ECCV*, 2020.

[4] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "DeepSDF: Learning continuous signed distance functions for shape representation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019.

[5] J. Tang, H. Zhou, X. Chen, T. Hu, E. Ding, J. Wang, and G. Zeng, "Delicate textured mesh recovery from NeRF via adaptive surface refinement," *arXiv preprint arXiv:2303.02091*, 2022.

[6] T. Takikawa, J. Litalien, K. Yin, K. Kreis, C. Loop, D. Nowrouzezahrai, A. Jacobson, M. McGuire, and S. Fidler, "Neural geometric level of detail: Real-time rendering with implicit 3D shapes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[7] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Trans. Graph.*, vol. 41, no. 4, pp. 102:1–102:15, Jul. 2022.

[8] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D gaussian splatting for real-time radiance field rendering," *ACM Transactions on Graphics*, vol. 42, no. 4, Jul. 2023.

[9] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "Mip-NeRF 360: Unbounded anti-aliased neural radiance fields," *CVPR*, 2022.

[10] L. Yen-Chen, P. Florence, J. T. Barron, A. Rodriguez, P. Isola, and T.-Y. Lin, "iNeRF: Inverting Neural Radiance Fields for Pose Estimation," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Prague, Czech Republic: IEEE, Sep. 2021, pp. 1323–1330.

[11] E. Sucar, S. Liu, J. Ortiz, and A. Davison, "iMAP: Implicit mapping and positioning in real-time," in *Proceedings of the International Conference on Computer Vision (ICCV)*, 2021.

[12] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, "NICE-SLAM: Neural implicit scalable encoding for SLAM," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2022.

[13] Z. Zhu, S. Peng, V. Larsson, Z. Cui, M. R. Oswald, A. Geiger, and M. Pollefeys, "NICER-SLAM: Neural Implicit Scene Encoding for RGB SLAM," 2023.

[14] J. Ortiz, A. Clegg, J. Dong, E. Sucar, D. Novotny, M. Zollhoefer, and M. Mukadam, "iSDF: Real-Time Neural Signed Distance Fields for Robot Perception," May 2022.

[15] K. Park, U. Sinha, P. Hedman, J. T. Barron, S. Bouaziz, D. B. Goldman, R. Martin-Brualla, and S. M. Seitz, "HyperNeRF: A higher-dimensional representation for topologically varying neural radiance fields," *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 1–12, Dec. 2021.

[16] A. Cao and J. Johnson, "HexPlane: A fast representation for dynamic scenes," *CVPR*, 2023.

[17] J.-K. Chen, J. Lyu, and Y.-X. Wang, "NeuralEditor: Editing neural radiance fields via manipulating point clouds," in *CVPR*, 2023.

[18] R. Martin-Brualla, N. Radwan, M. S. M. Sajjadi, J. T. Barron, A. Dosovitskiy, and D. Duckworth, "NeRF in the wild: Neural radiance fields for unconstrained photo collections," in *CVPR*, 2021.

[19] J. Shen, A. Ruiz, A. Agudo, and F. Moreno-Noguer, "Stochastic Neural Radiance Fields: Quantifying Uncertainty in Implicit 3D Representations," in *2021 International Conference on 3D Vision (3DV)*, Dec. 2021, pp. 972–981.

[20] M. Adamkiewicz, T. Chen, A. Caccavale, R. Gardner, P. Culbertson, J. Bohg, and M. Schwager, "Vision-Only Robot Navigation in a Neural Radiance World," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4606–4613, Apr. 2022.

[21] E. R. Chan, C. Z. Lin, M. A. Chan, K. Nagano, B. Pan, S. D. Mello, O. Gallo, L. Guibas, J. Tremblay, S. Khamis, T. Karras, and G. Wetzstein, "Efficient geometry-aware 3D generative adversarial networks," in *arXiv*, 2021.

[22] J. T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, and P. P. Srinivasan, "Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields," 2021.