Thermal Radiance Fields

Xin-Yi Pan

1 Introduction

Thermal imaging is very useful for a various applications where imaging in the visible domain is limited. The ability to capture and visualize temperature variations and heat signatures from the Infrared (IR) spectrum provides additional information of a scene and the objects within it [1]. Many applications would benefit from 3D thermal field reconstruction rather than purely 2D thermal imaging [2] [3] [4]. In particular, this project focuses on 3D thermal field reconstruction in Long-Wave Infrared (LWIR) with wavelengths of 8 to 14 µm.

Applications of 3D thermal imaging include security and surveillance, preventive maintenance, building inspection, monitoring rock masses, and archaeology [1] [5] [6] [7] [8] [9].

3D radiance field reconstruction has made great strides using visible cameras, but 3D reconstruction from thermal images is challenging because thermal cameras are often pixel-limited and have lower resolution compared to RGB cameras [10]. This makes it difficult to find robust 2D features from which to recover camera poses by structure from motion (SfM) / Colmap. [xy: This will be demonstrated in an experiment that will be done.] Hence, thermal cameras are often used together with other sensors to reconstruct a scene with a higher resolution than what the thermal camera can provide [11] [12].

Directly extending radiance field models to visible and thermal images produces limited quality reconstructions even with correct thermal poses. This is because of the inherent difference in the way materials in general interact with thermal and visible spectrum. Hence, we would like to propose a strategy to combine information from the two spectra, while respecting and recovering these material-specific properties. The final goals are:

- To develop a strategy that uses both visible and thermal information for 3D reconstruction
- To demonstrate improved 3D thermal reconstruction quality across different scenes
- If time permits, we would also like to explore the idea of thermal superresolution.

2 Related Work

2.1 Thermal Imaging

Thermal cameras detect and measure the heat signature of objects, where emitted infrared energy is converted into thermal images with varying levels of IR radiation [13]. This provides insights to scenes and the objects that visible cameras do not [1]. The contactless nature of thermal imaging adds to its attractiveness in a diverse range of applications as mentioned in section 1.

In these wavelength ranges, materials for lenses and sensors differ from those used in visible light cameras, which make it more costly and difficult to produce high-resolution cameras. Thermal cameras use very expensive special germanium lenses that transmit IR spectrum and block visible light [14] [15]. The longer wavelength also implies the need for each element in the detector array to be larger than those in the visible light spectrum [16], hence, reducing the number of detector elements which contributes to the significantly lower resolution [14] [10] and higher production cost of thermal cameras [17].

Considering the general pixel-limitation and low resolution of thermal cameras, they are often used together with other sensors to reconstruct a scene with a higher resolution than what the thermal camera can provide [11] [12].

While there have been 3D approaches in the IR regime [18], past methods of reconstruction tend to be less than ideal in terms of the region within the IR spectral as well as the quality of the reconstructed images, as mentioned in section 2.3. Additionally, other thermal methods/representations such as ContactDB [19] and Non-Line of Sight Imaging [20] are inherently different in their approaches and purposes. ContactDB focuses on contact maps from functional grasping which differs from our non-contact 3D reconstruction. Non-Line of Sight Imaging detects the reflection of the object, instead of the object, which differs from our approach which involves direct imaging.

2.2 3D Reconstruction and Novel-View Synthesis

3D reconstruction has been used to aid in the visualization, survey and analysis of large and difficult or inaccessible objects or landscapes, in ways that 2D images are unable to provide [2] [3] [4].

Structure from Motion (SfM) with Multi-View Stereo is a technique that reconstructs 3D surface models from 2D images. SfM computes projection matrices and 3D points using corresponding points in each view from 2D images [21], while MVS uses calibrated image for dense 3D construction of scenes. Some limitations include the quality of the reconstructed scene being limited to that of the input images and camera parameters computed from SfM algorithms, as well as reconstruction assumptions such as scene rigidity [22] [23].

Novel-View Synthesis creates new images from arbitrary view points. Differential rendering is a process that predicts a scene's radiance from any viewpoint with 3D object gradients via differentiable ray marching [24]. One method is to model scenes as implicit representations, with the development of Neural Radiance Field (NeRF). This greatly facilitates the synthesis of views by projecting output colors and densities obtained from querying 5D coordinates comprising of 3D position coordinates and 2D viewing directions along camera rays [25]. Recent advances in NeRF-based implicit 3D reconstruction have achieved closer photo-realistic results to view-synthesis problems in aspects as such improving NeRF's performance in representing finer details [26] [27] [28].

However, most work is in the visible domain which limits the information that can be gleaned from scenes. Extending NeRF capabilities to represent other parts of the spectrum which is invisible to the human eye, would allow for details that were previously unavailable, to be available [29].

2.3 Multispectral NeRF

Capturing data beyond the visible spectrum can be helpful in identifying features that might be transparent in the visible spectrum [30] [31] [32] [33] [34]. Work has been done to incorporate other sensors with RGB sensors to incorporate information beyond the visible range with integrated sensors.

For instance, X-NeRF [18] tackles setting of multispectral images by optimizing transform between RGB/other cameras including Near IR (NIR) cameras [35]. NIR images tend to have higher resolution as compared to LWIR and MWIR images due to its shorter wavelength [36] [37]. However, the reconstructed images presented lack sufficient features. This suggests room for improved robustness to IR images. Additionally, this approach assumes knowledge of the camera intrinsics as well as the shared density assumption of classic NeRF which is not always the case.

Aside from potential improvements in the reconstruction, there are limitations to the aforementioned method. One key limitation is inherent in the workings of the camera. NIR cameras are based on reflected energy. While NIR cameras may work in the day where there is light, an external light source is required at night [37]. This is unlike thermal imaging cameras such as LWIR and MWIR cameras which detect thermal emissions from objects [38]. Although NIR based imaging allows for resolution similar to visible cameras and have been used in Night vision goggles and LIDAR [39], it does not fully capture the advantages associated to thermal imaging.

Consumer-grade thermal cameras are often lower cost and have poorer image resolution [8] [9]. This poses multiple challenges when to comes to both obtaining the camera poses and subsequent 3D reconstruction results [40] [18]. There is a need for a combination of both visible and Infrared (IR) wavelength ranges, with points from RGB augmented with thermal information, for higher accuracy reconstruction with sparse input images [3]. We propose a method for thermal 3D reconstruction of scenes using both RGB and thermal images. While have been approaches that used similar insights, such as dehazing [41], hyperspectral imaging [42] and 3D reconstruction of a person via reflections [43], we demonstrate a method for 3D thermal field reconstruction that separately models material interactions with thermal and visible spectra to improve reconstruction quality.

3 Timeline

3.1 Week 8

- To collect at least 2 sets of data (one indoor and one outdoor)
- To fix our code
- To run the sets of data through our code

3.2 Week 9

- To implement baseline comparison. In particular, I am keen to use NeRF-based method (X-NeRF [18]) as a baseline comparison as it involves work across spectra which is similar to what we are doing. For the IR wavelength, they are working with NIR while we are working with LWIR. I am curious to find out how the results with our code would compare with theirs for the above datasets.

3.3 Week 10

- To conduct ablation study on both density loss and total variation loss

References

[1] Fluke, "What is thermal imaging? thermal cameras and how they work," Jan 2024. [Online]. Available: https://www.fluke.com/en-us/learn/blog/thermal-imaging/how-infrared-cameras-work

[2] C. Collaro and M. Herkommer, "Research, application, and innovation of lidar technology in spatial archeology," in *Encyclopedia of Information Science and Technology, Sixth Edition*. IGI Global, 2025, pp. 1–33.

[3] J. M. Jurado, A. López, L. Pádua, and J. J. Sousa, "Remote sensing image fusion on 3d scenarios: A review of applications for agriculture and forestry," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102856, 2022.

[4] C. Collaro, C. Enríquez-Muñoz, A. López, C. Enríquez, and J. M. Jurado, "Detection of landscape features with visible and thermal imaging at the castle of puerta arenas," *Archaeological and Anthropological Sciences*, vol. 15, no. 10, p. 152, 2023.

[5] J. Casana, A. Wiewel, A. Cool, A. C. Hill, K. D. Fisher, and E. J. Laugier, "Archaeological aerial thermography in theory and practice," *Advances in Archaeological Practice*, vol. 5, no. 4, pp. 310–327, 2017.

[6] C. Brooke, "Thermal imaging for the archaeological investigation of historic buildings," *Remote Sensing*, vol. 10, no. 9, p. 1401, 2018.

[7] A. Adán, B. Quintana, J. Garcia Aguilar, V. Pérez, and F. J. Castilla, "Towards the use of 3d thermal models in constructions," *Sustainability*, vol. 12, no. 20, p. 8521, 2020.

[8] G. Grechi, M. Fiorucci, G. M. Marmoni, and S. Martino, "3d thermal monitoring of jointed rock masses through infrared thermography and photogrammetry," *Remote Sensing*, vol. 13, no. 5, p. 957, 2021.

[9] O. González, M. I. Lizarraga, S. Karaman, and J. Salas, "Thermal radiation dynamics of soil surfaces with unmanned aerial systems," in *Pattern Recognition: 11th Mexican Conference, MCPR 2019, Querétaro, Mexico, June 26–29, 2019, Proceedings 11*. Springer, 2019, pp. 183–192.

[10] F. L. Liu, *Single-Shot 3D Microscopy: Optics and Algorithms Co-Design*. University of California, Berkeley, 2022.

[11] "Flir one pro thermal imaging camera for smartphones — teledyne flir." [Online]. Available: https://www.flir.com/products/flir-one-pro/?vertical=condition%2Bmonitoring&segment=solutions

[12] FLIR, “Flir one® series thermal imaging cameras for ios® or android™ smartphones.” [Online]. Available: https://www.flir.com/flir-one/

[13] Fluke, “What is thermal imaging? how a thermal image is captured.” [Online]. Available: https://www.fluke.com/en-us/learn/blog/thermal-imaging/how-infrared-cameras-work

[14] F. Nilsson, *Intelligent network video: Understanding modern video surveillance systems*. crc Press, 2008.

[15] B. Mesnik, “Thermal versus optical ip cameras.” [Online]. Available: https://kintronics.com/thermal-versus-optical-ip-cameras/

[16] “Comparing sensitivity of thermal imaging camera modules.” [Online]. Available: https://www.flir.com/discover/cores-components/Comparing-Sensitivity-of-Thermal-Imaging-Cameras-Modules/

[17] R. Schmidt, “How patent-pending technology blends thermal and visible light.” [Online]. Available: https://www.fluke.com/en-us/learn/blog/thermal-imaging/how-patent-pending-technology-blends-thermal-and-visible-light

[18] M. Poggi, P. Z. Ramirez, F. Tosi, S. Salti, S. Mattoccia, and L. Di Stefano, “Cross-spectral neural radiance fields,” in *2022 International Conference on 3D Vision (3DV)*. IEEE, 2022, pp. 606–616.

[19] S. Brahmbhatt, C. Ham, C. C. Kemp, and J. Hays, “Contactdb: Analyzing and predicting grasp contact via thermal imaging,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 8709–8719.

[20] T. Maeda, Y. Wang, R. Raskar, and A. Kadambi, “Thermal non-line-of-sight imaging,” in *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2019, pp. 1–11.

[21] D. Robertson and R. Cipolla, “Practical image processing and computer vision,” in *chapter Structure from Motion*. John Wiley & Sons Australia, 2009.

[22] Y. Furukawa, C. Hernández *et al.*, “Multi-view stereo: A tutorial,” *Foundations and Trends® in Computer Graphics and Vision*, vol. 9, no. 1-2, pp. 1–148, 2015.

[23] S. Wang, H. Jiang, and L. Xiang, “Ct-mvsnet: Efficient multi-view stereo with cross-scale transformer,” in *International Conference on Multimedia Modeling*. Springer, 2024, pp. 394–408.

[24] H. Kato, D. Beker, M. Morariu, T. Ando, T. Matsuoka, W. Kehl, and A. Gaidon, “Differentiable rendering: A survey,” *arXiv preprint arXiv:2006.12057*, 2020.

[25] 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.

[26] M. Suhail, C. Esteves, L. Sigal, and A. Makadia, “Light field neural rendering,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 8269–8279.

[27] K. Gao, Y. Gao, H. He, D. Lu, L. Xu, and J. Li, “Nerf: Neural radiance field in 3d vision, a comprehensive review,” *arXiv preprint arXiv:2210.00379*, 2022.

[28] 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,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 5855–5864.

[29] H. Zhu, Y. Sun, C. Liu, L. Xia, J. Luo, N. Qiao, R. Nevatia, and C.-H. Kuo, “Multimodal neural radiance field,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 9393–9399.

[30] Y. Zhang, S. Müller, B. Stephan, H.-M. Gross, and G. Notni, “Point cloud hand–object segmentation using multimodal imaging with thermal and color data for safe robotic object handover,” *Sensors*, vol. 21, no. 16, p. 5676, 2021.

[31] J. Á. S. Carmona, E. Quirós, V. Mayoral, and C. Charro, “Assessing the potential of multispectral and thermal uav imagery from archaeological sites. a case study from the iron age hillfort of villasviejas del tamuja (cáceres, spain),” *Journal of Archaeological Science: Reports*, vol. 31, p. 102312, 2020.

[32] M. McLeester, J. Casana, M. R. Schurr, A. C. Hill, and J. H. Wheeler III, “Detecting prehistoric landscape features using thermal, multispectral, and historical imagery analysis at midewin national tallgrass prairie, illinois,” *Journal of Archaeological Science: Reports*, vol. 21, pp. 450–459, 2018.

[33] G. Patrucco, A. Gómez, A. Adineh, M. Rahrig, and J. L. Lerma, “3d data fusion for historical analyses of heritage buildings using thermal images: The palacio de colomina as a case study,” *Remote Sensing*, vol. 14, no. 22, p. 5699, 2022.

[34] N. Sutherland, S. Marsh, G. Priestnall, P. Bryan, and J. Mills, “Infrared thermography and 3d-data fusion for architectural heritage: A scoping review,” *Remote Sensing*, vol. 15, no. 9, p. 2422, 2023.

[35] Microsoft, “Azure kinect dk depth camera.” [Online]. Available: https://learn.microsoft.com/en-us/azure/kinect-dk/depth-camera

[36] J. Oncea, “Swir, mwir, and lwir: One use case for each.” [Online]. Available: https://www.photonicsonline.com/doc/swir-mwir-and-lwir-one-use-case-for-each-0001

[37] I. Electro-Optics, “Nir (near-infrared imaging (fog/haze filter).” [Online]. Available: https://www.infinitioptics.com/technology/nir-near-infrared

[38] A. Optics, “The differences between swir, mwir, and lwir cameras.” [Online]. Available: https://www.axiomoptics.com/blog/the-differences-between-swir-mwir-and-lwir-cameras/

[39] O. Optical, “Optical material: Infrared optics.” [Online]. Available: https://www.emf-corp.com/optical-materials/optical-material-infrared-optics/

[40] A. López, J. M. Jurado, C. J. Ogayar, and F. R. Feito, “An optimized approach for generating dense thermal point clouds from uav-imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 182, pp. 78–95, 2021.

[41] F. Dümbgen, M. El Helou, N. Gucevska, and S. Süsstrunk, “Near-infrared fusion for photorealistic image dehazing,” Tech. Rep., 2018.

[42] S. Hu, R. Hou, L. Ming, S. Meifang, and P. Chen, “A hyperspectral image reconstruction algorithm based on rgb image using multi-scale atrous residual convolution network,” *Frontiers in Marine Science*, vol. 9, p. 1006452, 2023.

[43] R. Liu and C. Vondrick, “Humans as light bulbs: 3d human reconstruction from thermal reflection,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 12 531–12 542.