ECCV #3555

TensoRF: Tensorial Radiance Fields

We thank the reviewers for their encouraging comments. We are glad to see that the reviewers generally appreciate our good result quality and high (memory and time) efficiency, and think the paper is "inspiring to the researchers in this field and can potentially make a big impact" (R1) and the idea "seems general and can be extended to the related tasks" (R2). We now respond to reviewers' comments.

Good rendering quality. (R1, R2, R3) We utilize a multi-channel feature grid to represent a radiance field; other concurrent works (like DVGO, Plenoxels) also found similar grids can lead to fast reconstruction. Our central idea is to consider the grid as a 4D tensor and adopt low-rank tensor factorization for efficient modeling. This naturally leads to high compactness in addition to fast reconstruction, and we believe this also benefits the reconstruction quality.

Note that, although we leverage tensor decomposition, we are not addressing a decomposition/compression problem, but a reconstruction problem based on gradient decent, since the feature grid / tensor is unknown. In essence, our CP/VM decomposition offers low-rank regularization in the optimization, leading to better quality. In fact, with a dense feature grid, this reconstruction problem is relatively over-parameterized/under-determined; e.g., a 300^3 grid with 27 channels has >700M parameters, while one hundred 800×800 images provide only 64M pixels for supervision. Therefore, many design choices – including pruning empty voxels, coarse-to-fine reconstruction, and adding additional losses, which have been similarly used in TensoRF and concurrent works (DVGO, Plenoxels) – are all essentially trying to reduce/constrain the parameter space and avoid over-fitting. In general, low-rank regularization is crucial in addressing many reconstruction problems, like matrix completion [1], compressive sensing [2], denoising [4]; tensor decomposition has also been widely used in tensor completion [5, 3], which is similar to our task. Tensor decomposition naturally provides low-rank constraints and reduces parameters; this similarly benefits the radiance field reconstruction as demonstrated by our work.

Moreover, TensoRF represents a 5D radiance field function that expresses both scene geometry and appearance; hence, we believe our 4D tensor is generally low-rank, because a 3D scene typically contains a lot of similar geometry structures and material properties across different locations. Note that, in various appearance acquisition tasks, similar low-rank constraints have been successfully applied for reconstructing other functions, including the 4D light transport function in relighting [7] and the 6D SVBRDF function in material reconstruction [8, 6] (where a common idea is to model a sparse set of basis BRDFs; this is similar to our modeling of vector components in the feature dimension in the matrix $\mathbf{B}$). We combine low-rank constraints and neural networks from a novel perspective, in tensor-based radiance field reconstruction. TensoRF essentially models the scene with global basis components, discovering the scene geometry and appearance commonalities across the spatial and feature dimensions. As pointed by R1 and R2, we hope our findings in tensorized low-rank feature modeling can inspire other modeling and reconstruction tasks.

R2. Theoretical analysis on gradient-decent-based nonlinear optimization with neural modules is always a challenge. We hope the discussion about our work and other low-rank optimization tasks is able to address your concern. We will also address the writing and formatting issues as suggested.

R3. We thank the reviewer for pointing out those valuable references; we will add and discuss all of them in the paper. Besides, our model is not very sensitive to different numbers of feature channels and we thus chose 27 to be consistent with the SH coefficients. Specifically, TensoRF-VM-192 achieves PSNRs of 33.07/33.14/33.27 with 13/27/54 channels respectively on the NeRF Synthetic dataset.

One current limitation is that we do not handle unbounded scenes, since we consider a regular bounding box. We believe this can be addressed by applying spherical coordinates (like NeRF++) and leave such applications in future work. This has been discussed in the supplementary material. Besides, we didn't find any clear failure cases in the four common datasets we evaluate, but similar to NeRF, our quality can be relatively lower with blurriness/noises, if a scene contains highly specular materials or very detailed structures; we can add this discussion. Nonetheless, as shown by the per-scene results in the supplementary material, TensoRF achieves reasonable reconstruction on every scene and leads to the best quality on most scenes.

References

[1] E. J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010. 1

[2] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang. Compressive sensing via nonlocal low-rank regularization. *IEEE transactions on image processing*, 23(8):3618–3632, 2014. 1

[3] S. Gandy, B. Recht, and I. Yamada. Tensor completion and low-n-rank tensor recovery via convex optimization. *Inverse problems*, 27(2):025010, 2011. 1

[4] H. Ji, C. Liu, Z. Shen, and Y. Xu. Robust video denoising using low rank matrix completion. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1791–1798. IEEE, 2010. 1

[5] J. Liu, P. Musialski, P. Wonka, and J. Ye. Tensor completion for estimating missing values in visual data. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):208–220, 2012. 1

[6] G. Nam, J. H. Lee, D. Gutierrez, and M. H. Kim. Practical svbrdf acquisition of 3d objects with unstructured flash photography. *ACM Transactions on Graphics (TOG)*, 37(6):1–12, 2018. 1

[7] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo. Kernel nyström method for light transport. In *ACM SIGGRAPH 2009 papers*, pages 1–10. 2009. 1

[8] Z. Zhou, G. Chen, Y. Dong, D. Wipf, Y. Yu, J. Snyder, and X. Tong. Sparse-as-possible svbrdf acquisition. *ACM Transactions on Graphics (TOG)*, 35(6):1–12, 2016. 1