DL based Depth estimation of landmarks/salient points and Scene Reconstruction, from arbitrary pair of stereo views

Computer Vision (CS6350)

TPA-15

1. Problem Statement

Given a pair of arbitrary stereo views of a scene, use DL based methods to identify corresponding landmarks/salient points and estimate depth. Also compare results with shallow methods of rectification or Metric/Stratified reconstructions. Attempt to extend this concept for scene reconstruction (3D point cloud, say)

[Assumptions – Certain restrictions of camera poses may be considered.]

The objective is to leverage the potential of Deep Learning (DL) techniques to estimate depth and reconstruct scenes from arbitrary stereo views. A comparative analysis should be conducted, contrasting the outcomes of this DL-driven approach with the results obtained from traditional "shallow" methods, specifically those involving rectification or Metric/Stratified reconstructions.

2. Expected Input Output

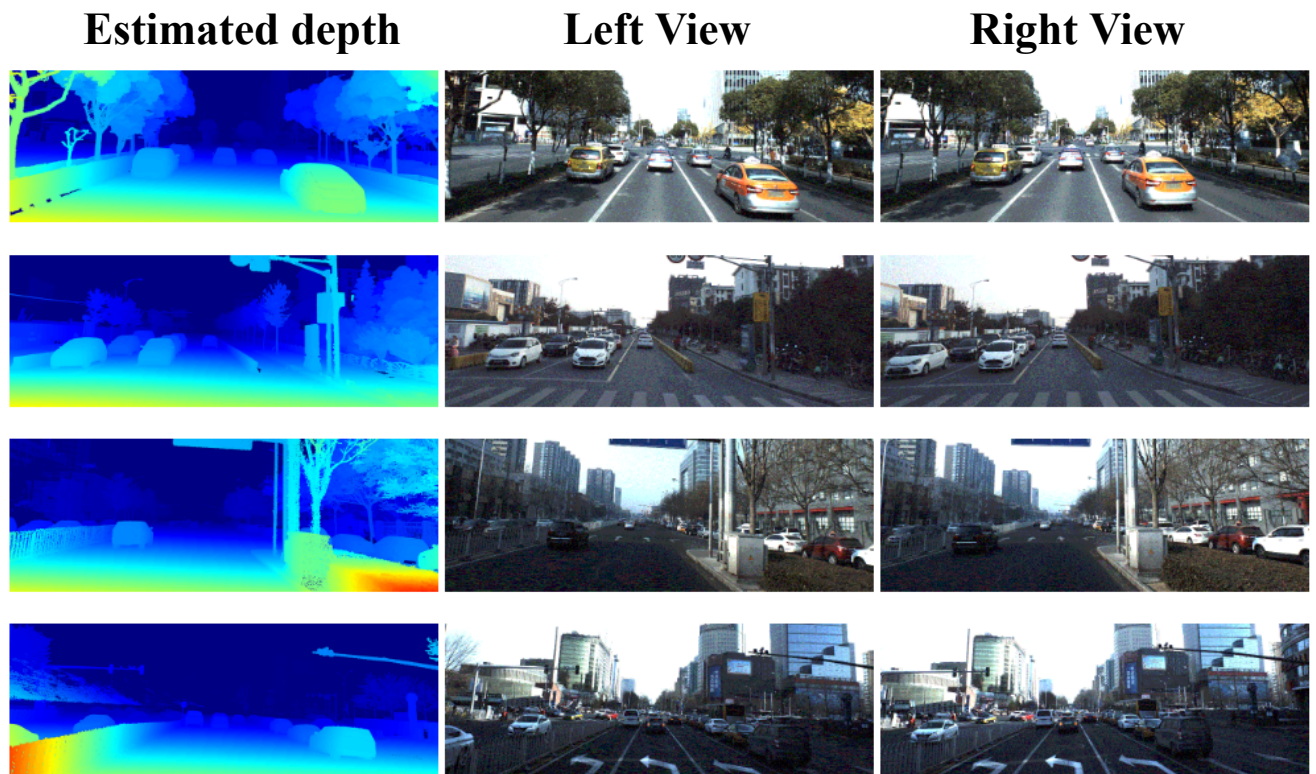**Input:** Stereo views
and **Output:** Estimated depth map

Estimated depth	Left View	Right View

Figure: Examples of stereo pairs and its predicted depth map from the ApolloScape dataset (link)**.**

3. Datasets

- Cityscapes & Foggy Cityscapes (link)
- KITTI datasets (link)
- NYU Depth Dataset V2 (link)
- DrivingStereo (link)
- Apolloscapes dataset (link)

4. Relevant code links

- https://github.com/autonomousvision/unimatch
- https://github.com/savnani5/Depth-Estimation-using-Stereovision
- https://github.com/ibaiGorordo/HITNET-Stereo-Depth-estimation
- https://github.com/ibaiGorordo/TFLite-MobileStereoNet

5. References

- Hamid, M. S., Abd Manap, N., Hamzah, R. A., & Kadmin, A. F. (2022). Stereo matching algorithm based on deep learning: A survey. *Journal of King Saud University-Computer and Information Sciences*, *34*(5), 1663-1673.
- Poggi, M., Tosi, F., Batsos, K., Mordohai, P., & Mattoccia, S. (2021). On the synergies between machine learning and binocular stereo for depth estimation from images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(9), 5314-5334.
- Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2020). Deep learning for 3d point clouds: A survey. *IEEE transactions on Pattern Analysis and Machine Intelligence*, *43*(12), 4338-4364.
- Laga, H., Jospin, L. V., Boussaid, F., & Bennamoun, M. (2020). A survey on deep learning techniques for stereo-based depth estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(4), 1738-1764.
- Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006, June). A comparison and evaluation of multi-view stereo reconstruction algorithms. In *2006 IEEE computer society conference on Computer Vision and Pattern Recognition (CVPR'06)* (Vol. 1, pp. 519-528). IEEE.
- Monasse, P., Morel, J. M., & Tang, Z. (2010, August). Three-step image rectification. In BMVC 2010-British Machine Vision Conference (pp. 89-1). BMVA Press.
- Szeliski, Richard. Computer vision: algorithms and applications. Springer Nature, 2022.
- Hartley, Richard, and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge university press, 2003.
- Ma, Y., Soatto, S., Košecká, J., & Sastry, S. (2004). An invitation to 3-d vision: from images to geometric models (Vol. 26). New York: springer.
- Fusiello, Andrea, Emanuele Trucco, and Alessandro Verri. "A compact algorithm for rectification of stereo pairs." Machine vision and applications 12 (2000): 16-22.
- Loop, Charles, and Zhengyou Zhang. "Computing rectifying homographies for stereo vision." Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149). Vol. 1. IEEE, 1999.

[Updated on 8th August 2023]