Comparison of the performance of few SOTA stereo depth estimation techniques (both shallow and deep) for a foreground object, to achieve high accuracy

Computer Vision (CS6350)

TPA-2

1. Introduction

Depth estimation from stereo images is a critical problem in computer vision with widespread applications, including autonomous driving, robotics, and augmented reality. Accurate depth estimation is vital for the functionality and safety of these applications. The methods for stereo depth estimation can be broadly divided into shallow learning-based methods, which rely on traditional computer vision techniques, and deep learning-based methods, which utilize neural networks to learn depth information from large datasets.

2. Problem Statement

The objective of this project is to explore and compare state-of-the-art (SOTA) stereo-depth estimation methods, including both shallow learning-based and deep learning-based approaches. The comparison will focus on their accuracy and computational efficiency.

3. Methods

Shallow Methods

- Stereo Vision (taught in class) [2]
- Block Matching [3]
- Semi-Global Block Matching [4,5]

Deep Learning Method

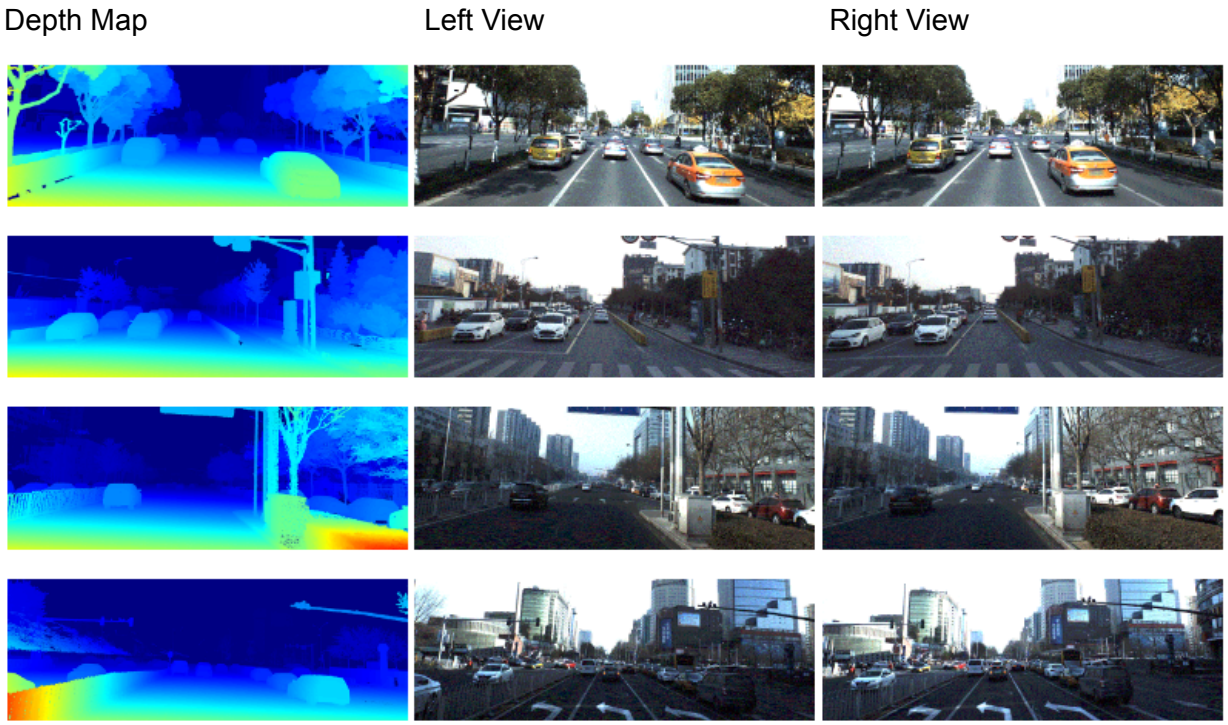
- Collection of DL methods: Stereo Demo [6]
- Other methods [1] [7]
- Accurate and Efficient Stereo Matching via Attention Concatenation Volume [1]
- Dusk Till Dawn: Self-supervised Nighttime Stereo Depth Estimation using Visual Foundation Models [7]

Note:

- You are free to explore more methods.
- And also special attention may be given to determine examples (cases), where DL (or both) methods fail to estimate.

4. Expected Input and Output

Input: Stereo View
Output: Estimated Depth Map



5. Dataset

- KITTI includes KITTI 2012 and KITTI 2015 Link
- NYU Depth Dataset V2 Link
- ETH3D Stereo Dataset Link
- Apolloscapes Link
- DrivingStereo Dataset Link

Explore the dataset and choose any 2 to compare the methods and demonstrate the results.

6. References

1. Madhu Vankadari, , Samuel Hodgson, Sangyun Shin, Kaichen Zhou Andrew Markham, Niki Trigoni. "Dusk Till Dawn: Self-supervised Nighttime Stereo Depth Estimation using Visual Foundation Models." (2024).
2. https://docs.opencv.org/4.x/dd/d53/tutorial_py_depthmap.html
3. https://learnopencv.com/depth-perception-using-stereo-camera-python-c/
4. https://www.digitalnuage.com/disparity-and-depth-estimation-from-stereo-camera
5. https://www.mathworks.com/help/visionhdl/ug/stereoscopic-disparity.html
6. https://github.com/nburrus/stereodemo

7. G. Xu, Y. Wang, J. Cheng, J. Tang and X. Yang, "Accurate and Efficient Stereo Matching via Attention Concatenation Volume" in IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 46, no. 04, pp. 2461-2474, 2024.
8. 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.
9. 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.
10. 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
11. 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.
12. 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.