Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

Master's Thesis

Multi-view Stereo by Temporal Non parametric Fusion

Manoj Kolpe Lingappa

Submitted to Hochschule Bonn-Rhein-Sieg,
Department of Computer Science
in partial fullfilment of the requirements for the degree
of Master of Science in Autonomous Systems

Supervised by

Prof. Dr Nico Hochgeschwender
Second Supervisor
MSc. Deebul Sivarajan Nair

Month 20XX

I, the undersigned below, declare that this work has not previously been submitted to this or any other university and that it is, unless otherwise stated, entirely my own work.

Date

Manoj Kolpe Lingappa

Abstract

Stereo vision is one of the field of computer vision that targets to construct the 3D model of a object using images. Over the past years a large number of algorithms and architectures have been proposed to find the 3D geometry of the object. However, a lack of dataset taken at varying environmental conditions made it difficult to compare the performance of the algorithms. It takes a lot of time to process large images and bad reconstruction with the low textured images. Most of the state of the art depth estimation algorithm are computationally heavy and cannot be deployed on the edge device. A light weight architecture with reasonable performance needs to be developed to deploy in a low computational power devices. Conventional approaches uses two view stereo rigs for reconstruction. However, estimation of depth from unconstrained monocular camera images is a challenging task. There are advantage of using the moving camera. Firstly with larger baseline the accuracy of the distant object can be improved. Secondly with multiple varying point images are able to fuse all the information for robust and stable depth estimation. This work concentrate on the depth estimation from unconstrained monocular camera images and extension of the disparity map estimation architecture to the segmentation.

Acknowledgements

Thanks to ....

Contents

List of Figures

List of Tables

1

Introduction

Computer vision aims to understand the surrounding environment using various mathematical modelling techniques. First generation of depth estimation was based on pixel matching between multiple images taken from a calibrated cameras [Laga et al.(2020)Laga, Jospin, Boussaid, and Bennamoun]. With the development of 3D reconstruction, depth sensors are becoming increasingly popular in areas such as self-driving cars. These sensors are used to obtain the information of the surrounding environment. However, the acquired depth maps from these sensors are sparse in nature due to low computational power resulting in information loss of the captured depth map. Another approach to reconstruct the 3D scene of an object is with the help of high quality images captured from the camera where the texture and lighting information are captured [Zhu et al.(2021)Zhu, Min, Wei, Chen, and Wang]. Reconstruction of three dimensional view from images is a classic problem in the computer vision domain. Multi view stereo algorithms can reconstruct the disparity maps or three dimensional view of an object from the images [Chen et al.(2021)Chen, Xu, Zhao, Zhang, Xiang, Yu, and Su]. It is the process of reproducing the 3D scenes from the multiple images given the camera poses and internal camera matrix. Number of areas take advantage of the reconstruction such as 3D mapping, 3D printing, video games, online shopping in the consumer domain, visual effect industry, digital mapping [Furukawa and Hernández(2015)], vehicle tracking, aircraft estimation and positioning [Manuel et al.(2018)Manuel, Manuel, Edith, Ivone, and Ramirez], depth estimation [Yao et al.(2018)Yao, Luo, Li, Fang, and Quan]. Depth estimation is the process of extracting the depth of objects present in the images by capturing and processing multiple images of the object taken from different locations. Images can be obtained from a stereo camera or a monocular camera. This work is based on the monocular camera images. Estimation of depth from the unconstrained monocular camera images is a challenging task. Most of the state of the art depth estimation algorithms are based on deep learning and compute cost volume according to the hypothesized depths. 3D convolution is applied to this cost volume to regress and predict the depth map [Gu et al.(2020)Gu, Fan, Zhu, Dai, Tan, and Tan]. This work aims to reproduce the result and deploy depth disparity estimation algorithms in a mobile device Finally end the research work with feasibility study of depth estimation architecture to segmentation.

1.1 Motivation

1.1.1 ...

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

1.1.2 ...

1.2 Challenges and Difficulties

1.2.1 ...

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

Etiam ac leo a risus tristique nonummy. Donec dignissim tincidunt nulla. Vestibulum rhoncus molestie odio. Sed lobortis, justo et pretium lobortis, mauris turpis condimentum augue, nec ultricies nibh arcu pretium enim. Nunc purus neque, placerat id, imperdiet sed, pellentesque nec, nisl. Vestibulum imperdiet neque non sem accumsan laoreet. In hac habitasse platea dictumst. Etiam condimentum facilisis libero. Suspendisse in elit quis nisl aliquam dapibus. Pellentesque auctor sapien. Sed egestas sapien nec lectus. Pellentesque vel dui vel neque bibendum viverra. Aliquam porttitor nisl nec pede. Proin mattis libero vel turpis. Donec rutrum mauris et libero. Proin euismod porta felis. Nam lobortis, metus quis elementum commodo, nunc lectus elementum mauris, eget vulputate ligula tellus eu neque. Vivamus eu dolor.

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

1.2.2 ...

1.2.3 ...

1.3 Problem Statement

Multi view stereo is one of the field of computer vision that targets to construct the most likely 3D model of a object using images. Reconstruction of the true 3D geometry is a ill posed problem. Over the past years a large number of algorithms and architectures have been proposed to find the 3D geometry of the object. However, a lack of dataset taken at varying environmental conditions made it difficult to compare the performance of the algorithms [Seitz et al.(2006)Seitz, Curless, Diebel, Scharstein, and Szeliski]. It takes a lot of time to process large images and with the low textured images a bad reconstruction is observed [Jancosek and Pajdla(2009)], [Seitz et al.(2006)Seitz, Curless, Diebel, Scharstein, and Szeliski], [Strecha et al.(2008)Strecha, Von Hansen, Van Gool, Fua, and Thoennessen]. Most of the state of the art depth estimation algorithm are computationally heavy and cannot be deployed on the edge device. A light weight architecture with reasonable performance needs to be developed to deploy in a low computational power devices. Conventional approaches uses two view stereo rigs for reconstruction. However, estimation of depth from unconstrained monocular camera images is a challenging task. There are advantage of using the moving camera. Firstly with larger baseline the accuracy of the distant object can be improved. Secondly with multiple varying point images are able to fuse all the information for robust and stable depth estimation [Hou et al.(2019)Hou, Kannala, and Solin]. This work concentrate on the depth estimation from unconstrained monocular camera images, deployment on the edge device, and extension of the disparity map estimation architecture to the segmentation.

1.3.1 ...

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetuer quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a, augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

Vivamus eu tellus sed tellus consequat suscipit. Nam orci orci, malesuada id, gravida nec, ultricies vitae, erat. Donec risus turpis, luctus sit amet, interdum quis, porta sed, ipsum. Suspendisse condimentum, tortor at egestas posuere, neque metus tempor orci, et tincidunt urna nunc a purus. Sed facilisis blandit tellus. Nunc risus sem, suscipit nec, eleifend quis, cursus quis, libero. Curabitur et dolor. Sed vitae sem. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Maecenas ante. Duis ullamcorper enim. Donec tristique enim eu leo. Nullam molestie elit eu dolor. Nullam bibendum, turpis vitae tristique gravida, quam sapien tempor lectus, quis pretium tellus purus ac quam. Nulla facilisi.

Duis aliquet dui in est. Donec eget est. Nunc lectus odio, varius at, fermentum in, accumsan non, enim. Aliquam erat volutpat. Proin sit amet nulla ut eros consectetuer cursus. Phasellus dapibus aliquam justo. Nunc laoreet. Donec consequat placerat magna. Duis pretium tincidunt justo. Sed sollicitudin vestibulum quam. Nam quis ligula. Vivamus at metus. Etiam imperdiet imperdiet pede. Aenean turpis. Fusce augue velit, scelerisque sollicitudin, dictum vitae, tempor et, pede. Donec wisi sapien, feugiat in, fermentum ut, sollicitudin adipiscing, metus.

Donec vel nibh ut felis consectetuer laoreet. Donec pede. Sed id quam id wisi laoreet suscipit. Nulla lectus dolor, aliquam ac, fringilla eget, mollis ut, orci. In pellentesque justo in ligula. Maecenas turpis. Donec eleifend leo at felis tincidunt consequat. Aenean turpis metus, malesuada sed, condimentum sit amet, auctor a, wisi. Pellentesque sapien elit, bibendum ac, posuere et, congue eu, felis. Vestibulum mattis libero quis metus scelerisque ultrices. Sed purus.

Donec molestie, magna ut luctus ultrices, tellus arcu nonummy velit, sit amet pulvinar elit justo et mauris. In pede. Maecenas euismod elit eu erat. Aliquam augue wisi, facilisis congue, suscipit in, adipiscing et, ante. In justo. Cras lobortis neque ac ipsum. Nunc fermentum massa at ante. Donec orci tortor, egestas sit amet, ultrices eget, venenatis eget, mi. Maecenas vehicula leo semper est. Mauris vel metus. Aliquam erat volutpat. In rhoncus sapien ac tellus. Pellentesque ligula.

Cras dapibus, augue quis scelerisque ultricies, felis dolor placerat sem, id porta velit odio eu elit. Aenean interdum nibh sed wisi. Praesent sollicitudin vulputate dui. Praesent iaculis viverra augue. Quisque in libero. Aenean gravida lorem vitae sem ullamcorper cursus. Nunc adipiscing rutrum ante. Nunc ipsum massa, faucibus sit amet, viverra vel, elementum semper, orci. Cras eros sem, vulputate et, tincidunt id, ultrices eget, magna. Nulla varius ornare odio. Donec accumsan mauris sit amet augue. Sed ligula lacus, laoreet non, aliquam sit amet, iaculis tempor, lorem. Suspendisse eros. Nam porta, leo sed congue tempor, felis est ultrices eros, id mattis velit felis non metus. Curabitur vitae elit non mauris varius pretium. Aenean lacus sem, tincidunt ut, consequat quis, porta vitae, turpis. Nullam laoreet fermentum urna. Proin iaculis lectus.

Sed mattis, erat sit amet gravida malesuada, elit augue egestas diam, tempus scelerisque nunc nisl vitae libero. Sed consequat feugiat massa. Nunc porta, eros in eleifend varius, erat leo rutrum dui, non convallis lectus orci ut nibh. Sed lorem massa, nonummy quis, egestas id, condimentum at, nisl. Maecenas at nibh. Aliquam et augue at nunc pellentesque ullamcorper. Duis nisl nibh, laoreet suscipit, convallis ut, rutrum id, enim. Phasellus odio. Nulla nulla elit, molestie non, scelerisque at, vestibulum eu, nulla. Ut odio nisl, facilisis id, mollis et, scelerisque nec, enim. Aenean sem leo, pellentesque sit amet, scelerisque sit amet, vehicula pellentesque, sapien.

1.3.2 ...

1.3.3 ...

2

State of the Art

Multi view stereo tries to reconstruct the 3D geometry of a target area using set of captured images and poses from different viewpoints. Multi view stereo originated from solving the stereoscopic matching problem as a computation problem [Marr and Poggio(1979)]. Two view stereo is a active research area till date and the research is evolved into multi view stereo problem. Instead of capturing the two images from two viewpoints, multi view stereo captures the images in between to increase the robustness of the reconstruction algorithm [Tsai(1983)], [Okutomi and Kanade(1993)].

2.1 Multi view stereo

Multi view stereo used in the multiple domains such as visual perception in autonomous driving, virtual reality and augmented reality [**?**], [**?**]. Multi view stereo (MVS) works with same principle as the stereo matching but with a enhancement of dealing with very large number of images [**?**].

2.1.1 3D reconstruction methods

- Active Depth sensors
- Passive Image based

2.1.2 Output representation

- Volumetric reconstruction
- Point cloud reconstruction
- Depth map based

2.2 Estimation of 3d geometry using traditional method

- COLMAP

2.3 Estimation of 3d geometry using deep learning method

- Multi-View stereo by temporal nonparametric Fusion
- MVSNet
- DeepMVS
- MVDepthNet
- DeepTAM
- DPSNet

2.4 Deployment on edge device

Deployment of depth estimation algorithm on a android device.

2.5 Limitations of previous work

3 Methodology

How you are planning to test/compare/evaluate your research. Criteria used.

3.1 Setup

3.2 Experimental Design

4

Solution

Your main contributions go here

4.1 Proposed algorithm

4.2 Implementation details

5 Evaluation

Implementation and measurements.

6

Results

6.1 Use case 1

Describe results and analyse them

6.2 Use case 2

6.3 Use case 3

7
Conclusions

7.1 Contributions

7.2 Lessons learned

7.3 Future work

A

Design Details

Your first appendix

B
Parameters

Your second chapter appendix

References

[42()] Google maps.

[43()] Microsoft bing maps.

[44()] Apple maps.

[Chen et al.(2021)Chen, Xu, Zhao, Zhang, Xiang, Yu, and Su] Anpei Chen, Zexiang Xu, Fuqiang Zhao, Xiaoshuai Zhang, Fanbo Xiang, Jingyi Yu, and Hao Su. Mvsnerf: Fast generalizable radiance field reconstruction from multi-view stereo. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14124–14133, 2021.

[Furukawa and Hernández(2015)] Yasutaka Furukawa and Carlos Hernández. Multi-view stereo: A tutorial. *Foundations and Trends® in Computer Graphics and Vision*, 9(1-2):1–148, 2015.

[Gu et al.(2020)Gu, Fan, Zhu, Dai, Tan, and Tan] Xiaodong Gu, Zhiwen Fan, Siyu Zhu, Zuozhuo Dai, Feitong Tan, and Ping Tan. Cascade cost volume for high-resolution multi-view stereo and stereo matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2495–2504, 2020.

[Hou et al.(2019)Hou, Kannala, and Solin] Yuxin Hou, Juho Kannala, and Arno Solin. Multi-view stereo by temporal nonparametric fusion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2651–2660, 2019.

[Jancosek and Pajdla(2009)] Michal Jancosek and Tomás Pajdla. *Segmentation based multi-view stereo*. Citeseer, 2009.

[Laga et al.(2020)Laga, Jospin, Boussaid, and Bennamoun] Hamid Laga, Laurent Valentin Jospin, Farid Boussaid, and Mohammed Bennamoun. A survey on deep learning techniques for stereo-based depth estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[Manuel et al.(2018)Manuel, Manuel, Edith, Ivone, and Ramirez] Mendoza Guzmán Vıctor Manuel, Mejıa Munoz José Manuel, Moreno Márquez Nayeli Edith, Rodrıguez Azar Paula Ivone, and SRE Ramirez. Disparity map estimation with deep learning in stereo vision. *Proceedings of the Regional Consortium for Foundations, Research and Spread of Emerging Technologies in Computing Sciences (RCCS+ SPIDTEC2)*, pages 27–40, 2018.

[Marr and Poggio(1979)] David Marr and Tomaso Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156):301–328, 1979.

[Okutomi and Kanade(1993)] Masatoshi Okutomi and Takeo Kanade. A multiple-baseline stereo. *IEEE Transactions on pattern analysis and machine intelligence*, 15(4):353–363, 1993.

[Seitz et al.(2006)Seitz, Curless, Diebel, Scharstein, and Szeliski] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, 1:519–528, 2006. ISSN 0302-2345.

[Strecha et al.(2008)Strecha, Von Hansen, Van Gool, Fua, and Thoennessen] Christoph Strecha, Wolfgang Von Hansen, Luc Van Gool, Pascal Fua, and Ulrich Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *2008 IEEE conference on computer vision and pattern recognition*, pages 1–8. Ieee, 2008.

[Tsai(1983)] Roger Y Tsai. Multiframe image point matching and 3-d surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2):159–174, 1983.

[Yao et al.(2018)Yao, Luo, Li, Fang, and Quan] Yao Yao, Zixin Luo, Shiwei Li, Tian Fang, and Long Quan. Mvsnet: Depth inference for unstructured multi-view stereo. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 767–783, 2018.

[Zhu et al.(2021)Zhu, Min, Wei, Chen, and Wang] Qingtian Zhu, Chen Min, Zizhuang Wei, Yisong Chen, and Guoping Wang. Deep learning for multi-view stereo via plane sweep: A survey. *arXiv preprint arXiv:2106.15328*, 2021.