**Review paper**

“Depth estimation using stereo vision and structured light”

Name: -Adhiraj Mahesh Alatage

USN: -2GI20EC009

1. This paper addresses the limitations of RGB-D sensors like the Microsoft Kinect or Asus Xtion, which are affordable 3D sensors used for depth sensing. While these sensors are useful, they have restrictions. They have a limited measurement range and face difficulties when dealing with transparent, shiny, or matte objects that cause reflection problems. Additionally, when multiple RGB-D cameras are used together, the infrared (IR) patterns emitted by each camera can interfere with each other, resulting in a significant loss of depth information.
2. This paper presents a new technique for reconstructing 3D shapes using a single image and a fringe pattern projected onto the target object. The method utilizes three deep convolutional neural network (CNN) models: FCN, AEN, and UNet, to quickly reconstruct the 3D shapes. The CNN models are trained and validated using data obtained from a high-accuracy multi-shot FPP technique.
3. This paper provides an overview of the current state-of-the-art in high-speed 3D shape measurement techniques based on structured light methods. These techniques have gained significant popularity in recent years, thanks to advancements in computing speed and hardware affordability
4. This paper investigates how the Microsoft Kinect, a type of structured light depth sensor, can be used to improve indoor scene segmentation. The authors propose a model based on Conditional Random Fields (CRF) and examine various representations of depth information. They also introduce a new dataset specifically designed for indoor scene analysis, which includes accurate depth maps and comprehensive labels
5. This paper focuses on monocular depth estimation, which is crucial for understanding scenes and enabling various applications. The paper introduces different deep learning models and provides an overview of monocular depth estimation algorithms based on deep learning. It covers aspects such as training methods and types of tasks.
6. This manuscript discusses the importance of acquiring high-resolution, real-time 3D surface data of moving objects in various fields. It focuses on structured light profilometry methods, which offer non-invasive and non-contact measurements.
7. This paper introduces a new approach to depth sensing by combining the advantages of two commonly used structured light techniques: time multiplexing (TM) and spatial neighborhood. The authors propose a set of hybrid structured light patterns that incorporate phase-shifted fringe and pseudo-random speckle.
8. This paper addresses the limitations of conventional binocular stereo vision in depth estimation, particularly in areas without clear features. To overcome this, the authors propose a method that combines encoded structured light and binocular stereo vision. However, calibrating the projector for structured light can be challenging.
9. Nighttime stereo depth estimation is a challenging task due to various factors like low light, noise, glare, and non-uniform light distribution. Obtaining accurate disparity ground-truths for nighttime images is difficult. To tackle this problem, this paper introduces a network that combines day/night image translation and stereo depth estimation.
10. Depth information is crucial for accurate image measurement, 3D reconstruction, and image recognition. Different methods, such as 3D laser scanners, structured light, and depth cameras, can be used to obtain depth information. Traditional binocular camera-based methods rely on disparity between left and right views but suffer from occlusion and mismatched points.

Reference: -

[1] Carlo Dal Mutto, Pietro Zanuttigh, and Guido Maria Cortelazzo. Probabilistic tof and stereo data fusion based on mixed pixels measurement models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(11):2260–2272, 2015. 5

[2] James Davis, Ravi Ramamoorthi, and Szymon Rusinkiewicz. Spacetime stereo: A unifying framework for depth from triangulation. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., volume 2, pages II–359. IEEE, 2003. 5, 7

[3] Shivam Duggal, Shenlong Wang, Wei-Chiu Ma, Rui Hu, and Raquel Urtasun. Deeppruner: Learning efficient stereo matching via differentiable patchmatch. In Proceedings of the IEEE International Conference on Computer Vision, pages 4384–4393, 2019. 2

[4] Barak Freedman. Depth mapping using projected patterns. US Application Publication, US 2010/0118123 A1, 2010. 3

[5] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pages 3354–3361. IEEE, 2012.

[6] Xinjing Cheng, Peng Wang, and Ruigang Yang. Depth estimation via affinity learned with convolutional spatial propagation network. In Proceedings of the European Conference on Computer Vision (ECCV), pages 103–119, 2018. 2

[7] Wei Bao, Wei Wang, Yuhua Xu, Yulan Guo, Siyu Hong, and Xiaohu Zhang. Instereo2k: a large real dataset for stereo matching in indoor scenes. Science China Information Sciences, 63(11):1–11, 2020. 1

[8] Michael Bleyer, Christoph Rhemann, and Carsten Rother. Patchmatch stereo-stereo matching with slanted support windows. In British Machine Vision Conference, volume 11, pages 1–11, 2011. 2

[9] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo matching network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5410– 5418, 2018. 2, 4, 5, 6

[10] Shuo Cheng, Zexiang Xu, Shilin Zhu, Zhuwen Li, Li Erran Li, Ravi Ramamoorthi, and Hao Su. Deep stereo using adaptive thin volume representation with uncertainty awareness. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 2524–2534, 2020.