

Motion Contrast 3D Scanning

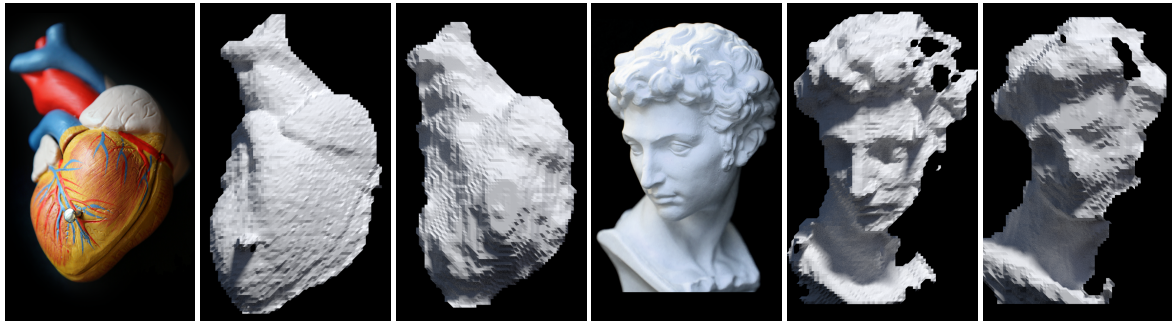
O. Cossairt¹, N. Matsuda¹, M. Gupta²

1. Northwestern University, 2133 Sheridan Road, Evanston, USA

2. Columbia University, 100 W. 120th St., New York, USA

We present a new method for structured light 3D scanning called Motion Contrast 3D scanning (MC3D). The key principle behind MC3D is the conversion of spatial projector-camera disparity to temporal events recorded by a motion contrast sensor [1]. The idea of mapping disparity to time has been explored previously in the VLSI community, where several researchers have developed highly customized CMOS sensors with on-pixel circuits that record the time of maximum intensity [2-4]. The use of a motion contrast sensor in a 3D scanning system is similar to these previous approaches with two important differences: 1) The differential logarithmic nature of motion contrast cameras improves performance in the presence of ambient illumination and arbitrary scene reflectance, and 2) motion contrast cameras are currently commercially available while previous techniques required custom VLSI fabrication, limiting access to only the small number of research labs with the requisite expertise.

MC3D consists of a laser line scanner that is swept relative to a DVS sensor. The event timing from the DVS is used to determine scan angle, establishing projector-camera correspondence for each pixel. The DVS was used previously for SL scanning by Brandli et al. [5] in a pushbroom setup that sweeps an affixed camera-projector module across the scene. This technique is useful for large area terrain mapping, but ineffective for 3D scanning of dynamic scenes. We have designed a SL system capable of 3D capture for exceptionally challenging scenes, including those containing fast dynamics, significant specularities, and strong ambient and global illumination.



(a) Reference Photo (b) MC3D (c) Kinect (d) Reference Photo (e) MC3D (f) Kinect

Fig. 1 Comparison between Motion Contrast 3D Scanning (MC3D) and Microsoft Kinect.

We captured these objects with our system and the Microsoft Kinect depth camera. The Kinect is based on a single-shot scanning method, and has a similar form factor and equivalent field of view when cropped to the same resolution as our prototype system. For our experimental results, we captured test objects with both systems at identical distances and lighting conditions. We fixed the exposure time for both systems at 1 second, averaging all input data during that time to produce a single disparity map. We applied a 3x3 median filter to the output of both systems. The resulting scans, shown in Fig. 1, clearly show increased fidelity in our system as compared to the Kinect.

References

- [1] P. Lichtsteiner, C. Posch, and T. Delbruck. A 128× 128 120 db 15 μ s latency asynchronous temporal contrast vision sensor. *Solid-State Circuits, IEEE Journal of*, 43(2), 2008.
- [2] K. Araki, Y. Sato, and S. Parthasarathy. High speed rangefinder. In *Robotics and IECON'87 Conferences*, pages 184–188. International Society for Optics and Photonics, 1988.
- [3] T. Kanade, A. Gruss, and L. R. Carley. A very fast vlsi rangefinder. In *IEEE ICRA*, pages 1322–1329. IEEE, 1991.
- [4] Y. Oike, M. Ikeda, and K. Asada. A cmos image sensor for high-speed active range finding using column-parallel time-domain adc and position encoder. *IEEE Transactions on Electron Devices*, 50(1):152–158, 2003.
- [5] C. Brandli, T. A. Mantel, M. Hutter, M. A. Ho pfinger, R. Berner, R. Siegwart, and T. Delbruck. Adaptive pulsed laser line extraction for terrain reconstruction using a dynamic vision sensor. *Frontiers in neuroscience*, 7, 2013.