Robust 3D Acquisition Using Motion Contrast 3D Scanning

O. Cossairt[1], N. Matsuda[1], M. Gupta[2]

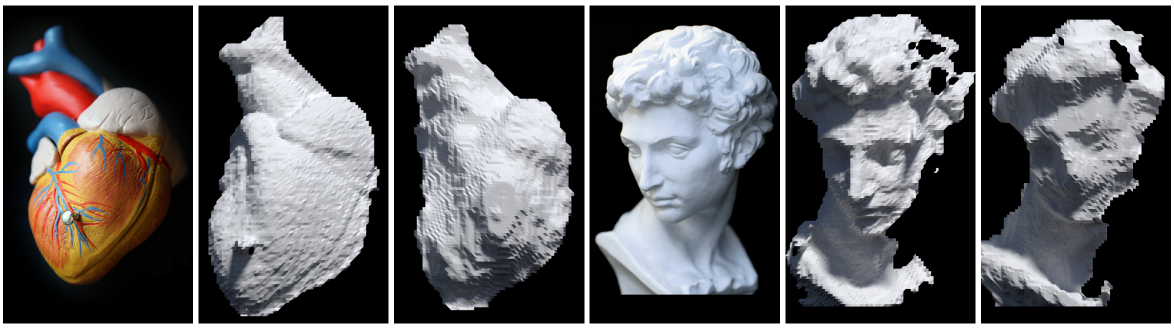
1. Northwestern University, 2133 Sheridan Road, Evanston, USA
2. Columbia University, 100 W. 120th St., New York, USA

***Abstract**: Structured light 3D scanning systems are fundamentally constrained by limited sensor bandwidth and light source power, hindering their performance in real-world applications where depth information is essential, such as industrial automation, autonomous transportation, robotic surgery and entertainment. We present a novel structured light technique called Motion Contrast 3D scanning (MC3D) that maximizes bandwidth and light source power to avoid performance trade-offs. The technique utilizes motion contrast cameras that sense temporal gradients asynchronously, i.e., independently for each pixel, a property that minimizes redundant sampling. This allows laser scanning resolution with single-shot speed, even in the presence of strong ambient illumination, significant inter-reflections, and highly reflective surfaces. The proposed approach will allow 3D vision systems to be deployed in challenging and hitherto inaccessible real-world scenarios requiring high performance using limited power and bandwidth.*

1) Introduction

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. Kinect and MC3D methods captured with 1 second exposure at 128x128 resolution (Kinect output cropped to match) and median filtered. Object placed 1m from sensor under 150 lux ambient illuminance measured at object.

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.

2) Ambient Lighting Performance

Figure 2 shows the performance of our system under bright ambient lighting conditions as compared to Kinect. We floodlit the scene with a broadband halogen lamp whose emission extends well into the infrared region used by the Kinect sensor. The ambient intensity was controlled by adjusting the lamp distance from the scene. Errors in the Kinect disparity map become signif- icant even for small amounts of ambient illumination as has been shown previously [6]. In contrast, MC3D achieves high quality results for a significantly wider range of ambient illumination. The illuminance of the laser pico-projector used in this experiment is around 150 lux, measured at the object. MC3D performs well under ambient flux an order of magnitude above that of the projector. The SHOWWX projector was used as the source in these experiments, which has a listed laser power of 1mW. The Kinect, according to the hardware teardown [7], has a 60mW laser source. The Kinect is targeted at indoor, eye-safe usage, but our experimental setup nonetheless outperforms the Kinect ambient light rejection at even lower power levels due to the light concentration advantage of laser scanning.

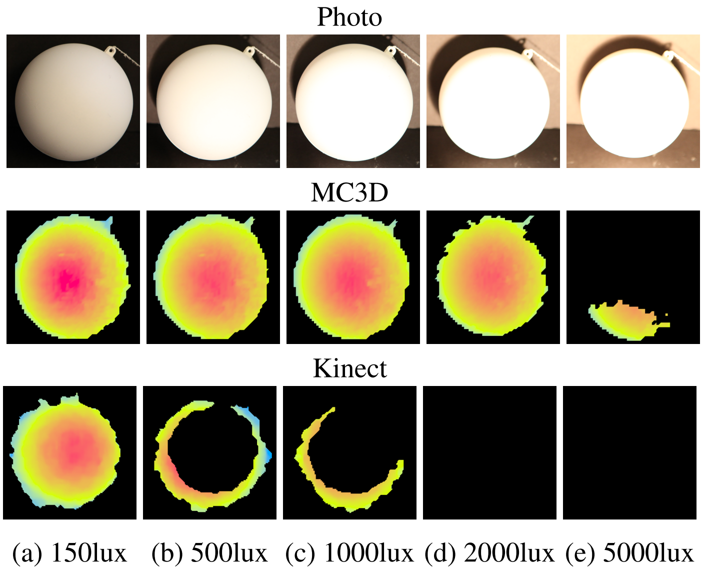


Fig. 2: Output Under Ambient Illumination. Disparity output for both methods captured with 1 second exposure at 128x128 resolution (Kinect output cropped to match) under increasing illumination from 150 lux to 5000 lux measured at middle of the sphere surface. The illuminance from our projector pattern was measured at 150lux. Note that in addition to outperforming the Kinect, MC3D returns usable data at ambient illuminance levels an order of magnitude higher than the projector power.

3) Performance with Strong Scene Interreflections

Figure 3(a) shows the performance of MC3D for a scene with significant inter- reflections. The test scene consists of two pieces of white foam board meeting at a 30 degree angle. The scene produces significant interreflections when illuminated by a SL source. As shown in the cross-section plot on the right, MC3D faithfully recovers the V-groove of the two boards while Gray coding SL produces significant errors that grossly misrepresent the shape. A Galvo laser line scanner was used as the source in these experiments.

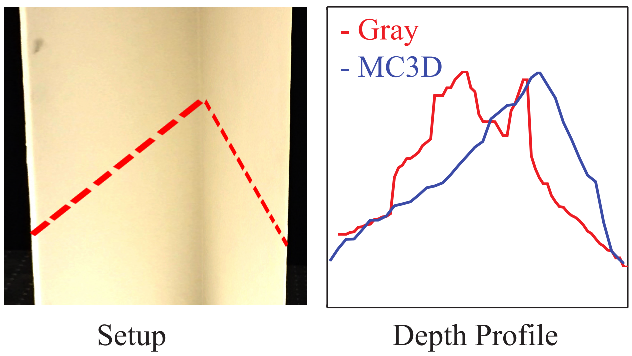


Fig. 3 : Performance with Interreflections. The image on the left depicts a test scene consisting of two pieces of white foam board meeting at a 30 degree angle. The middle row of the depth output from Gray coding and MC3D are shown in the plot on the right. Both scans were captured with an exposure time of 1/30th second. Gray coding used 22 consecutive coded frames, while MC3D results were averaged over 22 frames. MC3D faithfully recovers the V-groove shape while the Gray code output contains gross errors.

3) Performance for Specular Materials

Figure 4 shows the performance of MC3D for a highly specular steel sphere using a Galvo line scanner. The reflective appearance produces a wide dynamic range that is particularly challenging for conventional SL techniques. Because MC3D senses differential motion contrast, it is more robust for scenes with a wide dynamic range. As shown in the cross-section plot on the right, MC3D faithfully recovers the spherical surface while Gray coding SL produces significant errors at the boundary and center of the sphere.

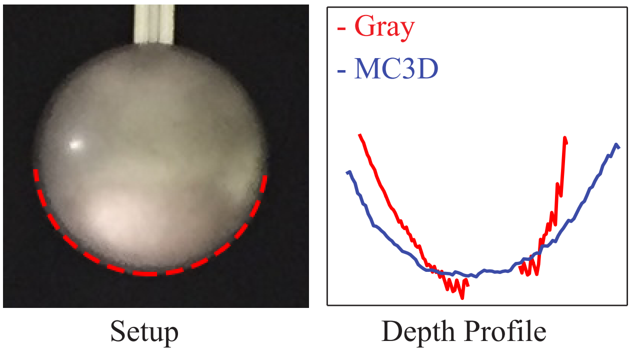


Fig. 4 : Performance with Reflective Surfaces. The image on the left depicts a reflective test scene consisting of a shiny steel sphere. The plot on the right shows the depth output from Gray coding and MC3D. Both scans were captured with an exposure time of 1/30th second. The Gray coding method used 22 consecutive coded frames, while MC3D results were averaged over 22 frames. The Gray code output produces significant artifacts not present in MC3D output.

4) Discussion

We have introduced MC3D, a new approach to SL that eliminates redundant sampling of irrelevant pixels and maximizes laser scanning speed. This arrangement retains the light efficiency and resolution advantages of laser scanning while attaining the real-time performance of single- shot methods.

There are several noise sources in our prototype system such as uncertainty in event timing due to internal electrical characteristics of the sensor, multiple event firings during one brightness change event, or downsampling in the sensors digital interface. These can be mitigated through updated sensor designs, further system engineering, and more sophisticated point cloud processing. We plan to provide a thorough noise analysis in a future publication.

Despite limitations, our hardware prototype shows that this method can be implemented using off-the-shelf components with minimal system integration. The results from this prototype show promise in outperforming existing commercial single-shot SL systems, especially in terms of both speed and performance. Improvements are necessary to develop single-shot laser scanning into a commercially vi- able product, but nonetheless our simple prototype demonstrates that the MC3D concept has clear benefits over exist- ing methods for dynamic scenes, highly specular materials, and strong ambient or global illumination.

References

[1] P. Lichtsteiner, C. Posch, and T. Delbruck. A 128× 128 120 db 15 μs latency asynchronous temporal contrast vision sen- sor. *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 ̈pflinger, R. Berner, R. Siegwart, and T. Delbruck. Adaptive pulsed laser line extraction for terrain reconstruction using a dy- namic vision sensor. *Frontiers in neuroscience*, 7, 2013.

[6] D. Castro and Mathur. Kinect outdoors. www.youtube. com/watch?v=rI6CU9aRDIo.

[7] Microsoft kinect. http://openkinect.org/wiki/Hardware_info.