# Surfel Cameras

**Carson Reynolds**
University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-8658 Japan
carson@k2.t.u-tokyo.ac.jp

## Abstract
Depth-camera systems provide a representation of visual scenes that combine an image with depth coordinates. These camera systems, which use principles like stereo disparity and time of flight, take samples of the surface of objects. It is convenient to model the imagery produced by depth-cameras using surfels (Pfister et al., 2000) as opposed to polygons.

Many interaction techniques exist for polygons and 2D imagery, but few have been proposed explicitly for surfels. As surfels are not explicitly connected in the same manner as polygons, interaction with them can take the form of set or point operations. Adams and Dutré (2003) previously showed how boolean operations could be implemented on solids bounded by surfels. Kim and Choi (2007) further apply surfels in creating a system to simulated application of facial makeup.

Surfel cameras fuse information normally captured separately using motion capture and texture capture systems. Unlike tomography, surfel cameras do not provide 3D imagery of the interior of objects. Interestingly, surfel cameras produce objects that can be interacted using operators more often found in 3D modeling systems than in photographic manipulation systems. Pointshop 3D (2002) provides one instance of a system for editing surfel data.

## ACM Classification Keywords
I.2.10 3D/stereo scene analysis

## Implementing a Surfel Camera
Inspired by the work of Gluckman et al. (1998), myself and others set out to develop a high-speed surfel camera system. Making use of a pair of 200 frame per second firewire cameras equipped with omnidirectional mirrors, the oversights system seeks to capture 360-degree depth images. Through the CUDA toolkit, the system uses a graphics accelerator to improve the performance of distortion removal and stereo matching processing programs.

The raw images captured from the omnidirectional mirrors are distorted. A camera model captured using Mei's omnidirectional calibration toolbox (2007) is used to provide de-warped panorama images. For each of the stereo cameras, the original image is warped into a panorama image which spans 360 degrees.

The two panoramas are in turn fed into a stereo-disparity system to compute the depth coordinates for each pixel. Kohei Yamaguchi adapted some code written by David Gallup et al. (2009) to perform stereo matching using the Graphics Processing Unit (GPU). The approach used is to search along the epipolar line between the images using sum of square differences. We found that this system is able to compute surfel images at a rate exceeding 150 Hz. By way of comparison, a system using similar algorithms but making use of a dual-core CPU operates at approximately 0.5 Hz.

The image processing is embarrassingly parallel as much of the computation takes the form of pixel operations with no data interdependencies. By taking a General Purpose GPU (Luebke et al., 2004) approach to the problem, we are able to use tens of thousands of threads to processes each omnidirectional frame. The floating point and arithmetic operation throughput offered by the comparatively larger die sizes of modern GPUs such as the Nvidia GTX 480 mean that surfel cameras can be quite fast. Moreover, surfel cameras can be made by coupling commodity GPU systems with increasingly inexpensive high-framerate cameras.

The panorama imagery and depth map are then fed into an OpenGL based rendering system developed by Elias Freider that displays the surfels. From one perspective, the viewer sees what appears to be a typical panorama. However, under rotation one can examine the surface manifold from different perspectives giving the sensation that apes viewing an object from different 3D viewpoints.

## Surfel Motes
The combination of a surfel cameras with wireless data links allows for an unusual multi-camera system. Typically, a surfel camera only views one perspective of an object. In practice this means that a surfel camera looks as if it is capturing a mold of one face of an object. A network of surfel cameras could view an object from many different but overlapping perspectives. Moreover, if the position of the surfel cameras is precisely known, then it is possible to infer the location and texture of objects in a scene in real time.

Extending of the work of Carceroni and Kutulakos (2002) we can easily imagine a world in which arbitrarily placed surfel cameras communicating with one another are able to capture the texture and motion of an entire room's worth of contents. Instead of reviewing video, a user could replay the evolution of 3D surfaces as they moved through space. Such systems would be of great use to game designers as well as those working in 3D animation.

This paper has served as a brief review of interrelated work from communities researching surfels, camera systems, and interactive graphics. The oversights system and yet to be developed surfel camera systems allow a different perspective of scenes combining motion capture with surface texture information.

## Acknowledgements
The oversights system was developed in collaboration with Kohei Yamaguchi and Elias Freider. Yoshihiro Watanabe and Hiromasa Oku also provided valuable suggestions for the mounting and camera calibration of the oversights system.

## References


Adams, B., & Dutré, P. (2003). Interactive boolean operations on surfel-bounded solids. ACM Transactions on Graphics (TOG), 22(3), 651. Retrieved from http://portal.acm.org/citation.cfm?id=882262.882320.

Carceroni, R. L., & Kutulakos, K. N. (2002). Multi-View Scene Capture by Surfel Sampling: From Video Streams to Non-Rigid 3D Motion, Shape and Reflectance. International Journal of Computer Vision, 49(2), 175. Retrieved from http://portal.acm.org/citation.cfm?id=598521.

Gallup, D., Frahm, J., & Stam, J. (2009). CUDA Stereo. Retrieved from http://www.cs.unc.edu/~gallup/cuda-stereo/.

Gluckman, J., Nayar, S. K., & Thoresz, K. J. (1998). Real-time omnidirectional and panoramic stereo. In Proceedings of the 1998 DARPA Image Understanding Workshop (Vol. 1, pp. 299-303). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.29.6913.

Kim, J.-S., & Choi, S. (2007). A virtual environment for 3D facial makeup. Lecture Notes In Computer Science, 488-496. Retrieved from http://portal.acm.org/citation.cfm?id=1770146.

Luebke, D., Harris, M., Krüger, J., Purcell, T., Govindaraju, N., Buck, I., et al. (2004). GPGPU: general purpose computation on graphics hardware. SIGGRAPH 2004: ACM SIGGRAPH 2004 Course Notes (p. 33+). New York, NY, USA: ACM. doi: 10.1145/1103900.1103933.

Mei, C. (2007). Omnidirectional Calibration Toolbox. Retrieved from http://www.robots.ox.ac.uk/~cmei/Toolbox.html.

Pfister, H., Zwicker, M., Baar, J. van, & Gross, M. (2000). Surfels: surface elements as rendering primitives. International Conference on Computer Graphics and Interactive Techniques, 335. Retrieved from http://portal.acm.org/citation.cfm?id=344779.344936.

Zwicker, M., Pauly, M., Knoll, O., & Gross, M. (2002). Pointshop 3D: an interactive system for point-based surface editing. International Conference on Computer Graphics and Interactive Techniques, 21(3), 322. Retrieved from http://portal.acm.org/citation.cfm?id=566584.