Integrating Multiple Depth Sensors into the Virtual Video Camera

Kai Ruhl, Kai Berger, Christian Lipski, Felix Klose, Yannic Schroeder, Alexander Scholz, Marcus Magnor *
Computer Graphics Lab, TU Braunschweig

Abstract

In this ongoing work, we present our efforts to incorporate depth sensors [Microsoft Corp 2010] into a multi-camera system for free view-point video [Lipski et al. 2010]. Both the video cameras and the depth sensors are consumer grade.

Our free-viewpoint system, the Virtual Video Camera, uses image-based rendering to create novel views between widely spaced (up to 15 degrees) cameras, using dense image correspondences. The introduction of multiple depth sensors into the system allows us to obtain approximate depth information for many pixels, thereby providing a valuable hint for estimating pixel correspondences between cameras.

Keywords: image-based rendering, data acquisition

1 Motivation

Free-viewpoint video systems render new viewpoints from a sparse set of given input views, e.g. from simultaneously recorded video streams. While most of these systems rely on an approximate scene geometry to guide the view interpolation, the system [Lipski et al. 2010] considered in this poster is purely image-based. It allows for interpolation both in space and time by treating the temporal and spatial domain equally.

While this versatile system already proved to manage challenging scene recordings, e.g. a firebreather or an outdoor skateboarding scene, it still suffers from small rendering artifacts (e.g. smearing) which are mainly caused by inaccurate pixel correspondences due to ambiguities in the optical flow between two input images. In this poster, we propose to enhance the optical flow by adding depth sensors to the setup. With this additional information we do not reconstruct the scene geometry fully, but instead we reduce ambiguities in low-textured regions and provide occlusion information to guide the image correspondence estimation.

2 Our approach

We calibrate the depth sensors together with the existing multi-view setup, then use world space points reprojected into all cameras as spatial flow information.

Calibrating multiple depth sensors with color cameras In order to simultaneously align a depth sensor with a consumer-grade RGB camera, we introduce a new calibration pattern, consisting of a large reflective patch (e.g. aluminum foil) and small white paper patches attached to it in alternating order, forming a checkerboard. The reflective patches deflect the emitted IR pattern, while the diffuse regions reflect it back to the sensor. This results in a distinguishable pattern, which can be used for robust alignment, Fig. 1 (left).

Image correspondence estimation with depth hints In image correspondence estimation (and optical flow), two major challenges are low-textured regions and occlusions. Depth information, when accurate, helps to address these issues. However, the resolution of contemporary consumer depth sensors is considerably lower than camera resolutions and thus less accurate. We use the potentially inaccurate depth information merely as a hint to the correspondence estimation. The depth points are backprojected into world space, and then projected into all cameras’ image spaces, Fig. 1 (right), upsampling the depth points via splatting. In effect, foreground objects are correctly covered with depth information, while border regions may be covered wrongly. Then, image correspondences are obtained by calculating the displacement between two splatted projections of a depth point. However, some of this information is incorrect due to the generous foreground coverage; more inaccuracies are added because our cameras are only loosely synchronized. Depth information can strictly only be considered as a hint. Therefore, our current work focuses on methods for the required filtering. One of the main ideas is to use the depth hints only when it is consistent with the data term of the image correspondence algorithm.

Using multiple depth sensors Due to the comparatively wide horizontal angles of our setup, a single depth sensor does not cover the scene appropriately. The inclusion of additional sensors improves the depth point coverage. However, due to the active light of the depth sensors, interference has to be expected. In practice, we found that a maximum of two depth sensors can be operated concurrently, giving us the desired wide angle coverage.

References


Microsoft Corp. 2010. Kinect for Xbox 360, November. Redmond WA.

Figure 1: The integration of low-resolution depth sensors into the Virtual Video Camera. Our new calibration pattern consists of a reflective-diffuse-checkerboard, which is clearly visible both in the depth and image sensors (two left). The additional depth information can be used as a soft hint to guide the image correspondence estimation needed for spatial image interpolation (two right)