Three Depth-Camera Technologies Compared

Fabrizio Pece[1] and Jan Kautz[1] and Tim Weyrich[1]

[1]Department of Computer Science, University College London, UK

1 Introduction

We present a comparison of three different depth cameras, in terms of their resolutions and precision and their suitability in VR/AR systems. With the term "depth camera" we refer to specific hardware capable of retrieving depth information about a scene either using a particular sensor or by running a stereo algorithm on the colour frames. The cameras taken into account are a *PMD[vision] CamCube*, a *PointGrey Bumblebee XB3* and a *Microsoft Kinect*. We will present for each of them a list of technical specifications, a visual representation of the depth information available and an analysis of the depth value precision. Moreover, positive and negative aspects of each camera will be analysed together with a final discussion on what is the best solution for BEAMING at the moment.

Cameras that can acquire a continuous stream of depth images are now commonly available. However choosing the best "depth camera" for a particular application is not always a trivial tasks due to the large differences between depth acquisition algorithms. Therefore, to guide and improve future development of applications that can take advantage of depth acquisition, a thorough comparison of the three most common depth acquisition technologies is now essential.

2 Results

Here we show the results of our comparison. At the end of the section, besides qualitative results, we also analyse the precision of each camera depth measurements when compared to real scenes.

The Bumblebee camera has a large frame size (1280×960 pixels) and a maximum framerate of 15 fps, supported by a very large depth (16-bits depth pixel) and working range (0.5–4.5m). Moreover, the sensors come pre-calibration and pre-registered, as the underlying technology is based on stereo algorithms. However, the camera requires parameter tuning for each new environment and usually has problems in retrieving depth information in non-textured regions. The entire depth calculation is performed on the host machine, requiring a very fast PC. The unit is also rather expensive. The Kinect camera has many positive aspects. It has a very low price, a good working range (1.2–3.5 m) supported by a very large depth range (16-bits depth pixel), a reasonable resolution (640×480 pixels) and a maximum framerate of 30 fps. Most importantly, it provides reliable depth measurements under a large variety of conditions. The depth values are available directly from the hardware, leaving the host machine free to do other computations. The depth frames acquired from the Kinect need to be registered with the RGB frames (the two sensors are located apart from each other), which is supported by the API though. The camera is based on structured light technology (SL), which makes the usage of multiple units very difficult due to interference. The PMD[Vision] CamCube, a time-of-flight (ToF) camera, acquires depth information directly on the camera hardware. Virtually no calibration is required, but can be used to improve the depth estimates. This camera has several limitations: besides being very expensive, the PMD unit can only acquire gray-scale images with a very limited frame size (200×200 pixels) and a maximum framerate of 25 fps. Moreover, due to the technology employed, the depth measurements can be very noisy and affected by the ambient light, limiting its use to indoor scenarios with controlled lighting.

We have conducted a series of experiments to establish the best performing camera in terms of depth measurements. We have compared the depth values obtained from the three cameras against the real values of several scenes. To thoroughly test the camera, we set up our scenes, such that the depth values range from a minimum of 0.5 m to a maximum of 1.5 m. The PMDs results proved to be the most precise hardware, followed by the Bumblebee and the Kinect (Table 1). However, the lack of precision of Kinect is minimal when compared to the other cameras (a maximum of 0.6 cm error against 0.3 cm for the PMD and 0.5 cm for the Bumblebee), while the depth maps available from it are more heterogeneous than the ones available from the Bumblebee. In fact, the Kinect can densely cover the full scene with its IR structured light pattern, and thus can obtain depth measurements for essentially every pixel. On the contrary, the Bumblebee discards many pixels and returns empty depth values due to the limitation of the stereo algorithm. The PMD offers a heterogeneous depth map, albeit at low resolution. Furthermore, the lack of associated colour frames makes it a limited solution for our aims.

	World (mm)	PMD (mm)	Bumblebee (mm)	Kinect (mm)
Scene 1	610	593	585	562
Scene 2	500 - 610	492 - 601	490 - 600	490 - 589
Scene 3	1200 - 1500	1195 - 1477	1189 - 1455	1175 - 1440

Table 1: Measurements comparison results

3 Conclusion

In the last few years depth acquisition has become a popular topic of research, and this has reflected in an larger availability of depth cameras that allow direct acquisition of scenes' depth information. However, the suitability of a camera to different working scenarios is strongly dependent to its technical specifications. Therefore, a technical and qualitative comparison of the three, most commonly used depth acquisition technologies can greatly help researchers when choosing "depth cameras" for particular applications (i.e., VR/AR systems).

We believe that in the near future ToF cameras will not only be extended to support colours and higher frame sizes, but also rapidly drop in price. Moreover, our experiments, as well as recent work [1], proved that ToF camera generates more accurate depth estimation than any stereo vision solution, especially in highly dynamic environments, such as a typical BEAMING session. For the time being, however, inexpensive structured-light cameras are an attractive off-the-shelf solution to perform depth estimation with limited noise and good accuracy. Due to its unique properties, stereo cameras will remain a valuable addition to any camera network, being able to augment the scene reconstruction with highly detailed depth and colour data where structured light and ToF cameras may not provide sufficient details. Overall, we find it vital to support the full range of available depth cameras within BEAMING, to allow for optimised setups, and in order to be prepared for any future developments in the emerging market of depth cameras.

References

[1] Christian Beder, Bogumil Bartczak, and Reinhard Koch. A comparison of pmd-cameras and stereo-vision for the task of surface reconstruction using patchlets. In *In Proceedings of the Second International ISPRS Workshop BenCOS*, 2007.