Matthias Küver

Semi-Global Matching in close-range applications supported by artificial texture from the Microsoft Kinect

Dauer der Arbeit: **6 Monate**

Abschluss: **Mai 2012**

Betreuer: **Dipl.-Ing. Konrad Wenzel**

Introduction

Terrestrial surface reconstruction in high resolution is a task that can be accomplished by laserscanners as well as by image-based methods. Especially in close-range applications, photogrammetric techniques can deliver very accurate results even in case of using low-cost sensors. Image matching algorithms like *Semi-Global Matching* thereby allow to produce very dense pointclouds.
The quality of the surface reconstruction is depending on many factors like the geometric configuration during the aquisition, the radiometric quality of the images and also on the surface characteristics of the scene itself. One of the main disadvantages with respect to laserscanning is the need of texture on the object's surfaces. In order to assist the image matching in plain areas, it is possible to project an artificial texture into the scene.

The assigned task is to investigate the achievable accuarcy of image-based surface reconstruction in very close range applications dependent on different conditions of aquisition. For that purpose, images are being aquired with multiple off-the-shelf industry cameras while a *Microsoft Kinect* is used to project additional texture. The ambition is to investigate the influence of parameters on the quality of the surface reconstruction with such kind of a system in order to be able to pick out the optimal configuration with respect to accuracy and efficiency.

Figure 1: Sensor setup

Investigation

For the accuracy investigation, a set of planes and spheres with well-known shapes are set up into the scene at a distance of about 60 cm to the sensors. As all of the objects are approximately plain, texture is provided by the Kinect almost exclusively. After the image matching and triangulation, a best-fitting surface can be estimated for each of those objects within the resulting pointcloud. The residuals of the 3D-points with respect to the estimated surface are the delivering the basis for the subsequent investigations on reconstruction quality. For better comparison, the residuals are visualized as in a color-coded surface plot of the object. An example is given in figure 2.

A large series of different images have been aquired to investigate the error behavior of the surface reconstruction under certain external conditions, including the exposure time, external illumination and the bitdepth of the images for example. Also the influence of a change of the image scale and the baseline length has been tested. The accuracy is not only depending on the acquisition itself, but also on the way in which matches between the images are being established and handled. Therefore, several different settings for the image processing have been tested as well.

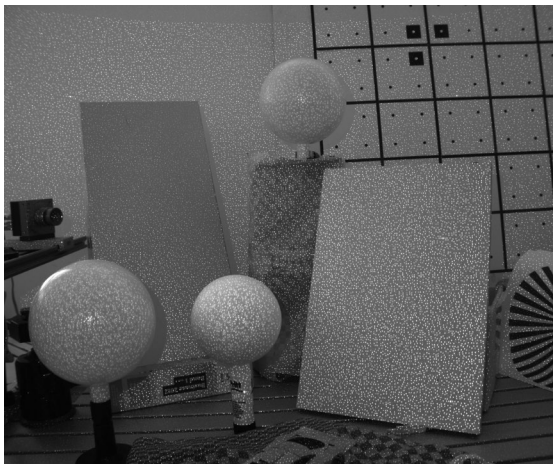

Figure 3: Scenery Assembling. The projected Kinect-pattern is visible as white dots.

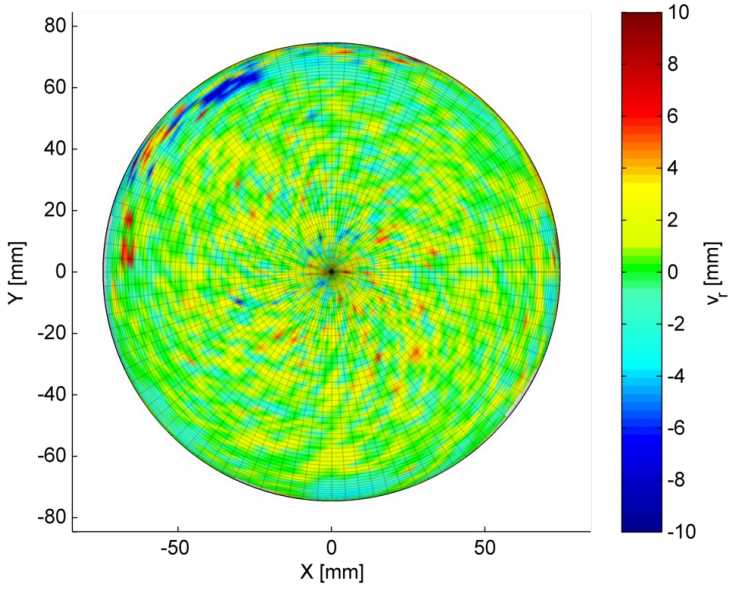

Figure 4: Visualization of the radial residuals of a reconstructed sphere

Results

To summarize the overall results of the accuracy investigation, the projection of artificial texture is highly beneficial. The reconstruction quality is becoming less sensitive against low textured surfaces, changes in external illumination or exposure time and also against individual reflection properties of the surface.
Thus, image-based surface reconstruction is getting more flexible with the assistance of the Kinect.
Although the objects have only very few texture on their own, it is possible to reconstruct the surfaces with an accuracy of less than 1.5 mm.
Using 12 Bit imagery, one can achieve reasonable results even for very short exposure times. An illustration of the improvement in using a higher bitdepth is given in figure 5. Therefore, the applied sensor setup becomes useful for hand-held point cloud recording solutions for example. For an aquisition distance of about 60 cm, the best compromise between avoiding motion-blur on the one hand and preserving sufficient signal strength on the other is found for an exposure time of 10 milliseconds.
According to the experimental results, the most important additional tasks to achieve most accurate results with such a system is the use of a wide base of about 15 cm and a redundancy restriction that a 3D-point has to be visible in at least 3 images.

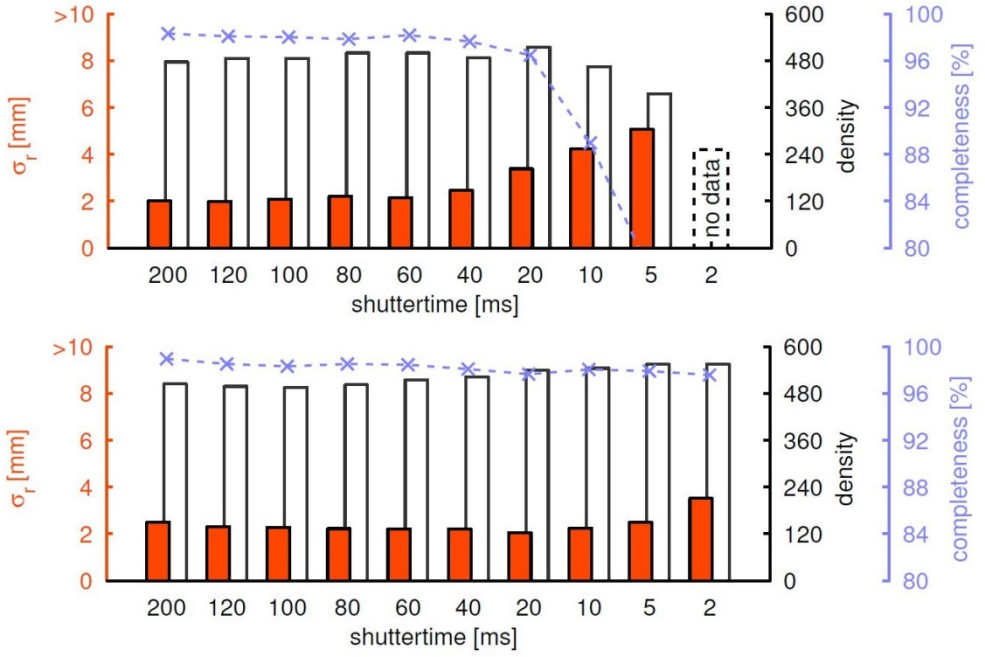

Figure 4: Summary for the reconstruction results of a sphere without natural texture using different exposure times. In the upper graph is showing the evaluation for 8 Bit, the lower one for 12 Bit images. The left y-axis depicts the standarddeviation of the radial residuals.

Conclusion

The investigation has pointed out the high quality of surface reconstruction by Semi-Global Matching in close-range applications and has presented the optimal configuration with which best and most efficient results can be achieved. With the assistance of the Kinect, a very robust and accurate system is found which can still be denoted as a low-cost solution.