Supplementary material for Multi-Kinect Avatar Capture with Robust Calibration and Alignment

No Author Given

No Institute Given

1 Camera synchronization

Due to the bandwidth limitation of the USB controller on a commodity desktop PC, simply connecting all 8 cameras through the USB interface is not enough to capture and collect data simultaneously. The capture delay between cameras can easily be several seconds due to the data transfer conflict, making the capture process inconvenient. To this end, several recent works on Kinect-based capture such as [4] and [7] connect each Kinect to one PC and require a server-client structure to synchronize between cameras. However, this largely increases the system complexity and violates our goal of affordability and portability. We solve the synchronization problem by adding an extra USB-PCI adaptor card to the host PC, and connecting the cameras to the PCI card, where the increased bandwidth is sufficient to capture raw data using our current camera setting. This way, the capture process lasts for around 3-4 seconds.

A common measuring error in time-of-flight (ToF) cameras is the multi-path interference [3], where multiple sources of light are detected on a pixel, other than the originial signal. This is a problem, especially if multiple ToF sensors are used, as signals from one could be received by another. We take a number of steps for minimizing this problem. First, by using just one frame of data we reduce the risk that the ToF scanning process overlaps between cameras (as Kinects can not be frame-synchronized). Second, we remove incorrect measurements which usually occur at the corners in the outlier removal step - the number of cameras ensures that the body is fully scanned even after removing some data. Third, we make sure that the room is constantly lit, without external sources such as bright lights or windows.

2 Data preprocessing

To eliminate the influence of surrounding objects and only keep the meaningful human body shape, we first filter the raw data using a distance threshold from each camera (1.3m in our experiments). However, the remaining human body data usually suffer from outliers, especially in the areas where the surface normal is perpendicular to the camera viewing direction (see Figure 1 left). We adopt the DBSCAN clustering algorithm [2] to segment the point cloud into clusters, and take the largest cluster as the cleaned data (see Figure 1 right). The cleaned up data are further smoothed to reduce random noise. We use the Moving Least

Fig. 1: Outlier removal and denoising.

Squares [1] algorithm to estimate a continuous surface in a local neighborhood and then reproject the data points on that surface. We also estimate the normal information for each point using PCA on their local neighborhood within a fixed radius, which will be used in the later reconstruction step of our pipeline. The neighborhood size is an important factor that determines the quality of estimated normal, and we empirically found that radius of 2.5cm works well in our experiments.

3 Additional results

	ICP	Non-rigid ICP	our method
Male 3 (Figure 6)	0.00475686	0.00438185	0.00419615
Woman 2 (Figure 6)	0.00437787	0.00375167	0.00385607
Woman 1 (Figure 6)	0.00490372	0.00418299	0.0041061
Male 2 (Figure 6)	0.00428622	0.00433294	0.00420506
Male (Figure 4)	0.00487387	0.00475633	0.0043115
Male 1 (Figure 6)	0.00417856	0.00403136	0.00393687
Male (Figure 2)	0.00422222	0.00411214	0.00390228

Table 1: A quantitative evaluation of our system

In order to quantitatively evaluate our registration results, we sample the resulting patches and measure the average distance to the closest point on any other patch. We compare our results on 7 patches with those obtained using the correspondence based non-rigid registration algorithm used by Li et al [5], and to rigid ICP [6].

References

1. M. Alexa, J. Behr, D. Cohen-Or, S. Fleishman, D. Levin, and C. T. Silva. Computing and rendering point set surfaces. *IEEE TVCG*, 9(1):3–15, 2003.

2. Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. International Conference on Knowledge Discovery and Data Mining, 1996.
3. David Jiménez, Daniel Pizarro, Manuel Mazo, and Sira Palazuelos. Modeling and correction of multipath interference in time of flight cameras. *Image and Vision Computing*, 32(1):1–13, 2014.
4. Marek Kowalski, Jacek Naruniec, and Michal Daniluk. Live scan3d: A fast and inexpensive 3d data acquisition system for multiple kinect v2 sensors. In *3DV*, pages 318–325. IEEE, 2015.
5. Hao Li, Etienne Vouga, Anton Gudym, Linjie Luo, Jonathan T. Barron, and Gleb Gusev. 3d self-portraits. ACM SIGGRAPH Asia, 2013.
6. Kok-Lim Low. Linear least-squares optimization for point-to-plane icp surface registration. *Chapel Hill, University of North Carolina*, 4(10), 2004.
7. Vahid Soleimani, Majid Mirmehdi, Dima Damen, Sion Hannuna, and Massimo Camplani. 3d data acquisition and registration using two opposing kinects. In *3DV*, pages 128–137. IEEE, 2016.