Literature Review 2
COMP.5460 Computer Graphics I, Spring 2018

Yang Meng
Yang_Meng@student.uml.edu

Primary Paper:
"3D Motion Sensing of Any Object Without Prior Knowledge" by L. Miyashita, R. Yonezawa, Y. Watanabe, and M. Ishikawa.

Secondary Paper:
"Automatic registration of overlapping 3D point clouds using closest points" by Y. Liu.

Paper Review:
Motion sensing is designed to detect motion by using sensors. Researchers have tried to propose different systems to achieve motion sensing. Some of those researchers have attempted to attach sensors (e.g., magnetic sensor) to a human to detect their poses. The sensors are called contact sensors and the human is called target. The drawback is that, not all targets could have sensors attached to detect their motion such as water. Other researchers proposed non-contact sensors, such as 2D camera and Microsoft Kinect stereo camera, to detect motion. However, this is usually computational intensive and the resulted delay is fatal in real-time situations. In those two approaches, the developer also has to know the physical information of the target for effective detection: unknown objects are out of capability.

To tackle the problem, authors in the primary paper proposed a new system to detect motion of unknown objects by using laser-based sensors [2]. The system consists of a laser range finder, a laser Doppler velocimeter, a galvanometer scanner and a dichroic mirror. The laser sensor and the velocimeter both emit beams and are used together to measure the velocity of a target in different beam directions. The galvanometer scanner and a dichroic mirror are to synchronize the beams emitted from the laser sensor and the velocimeter. A computationally inexpensive matrix formularization is proposed to get the rotation and translation matrices of the moving target. The authors also employed a regularization technique to remove noise and achieve numeral stability.

For the strengths of the paper, the system is fast because the authors turned motion detection a linear matrix problem and it becomes feasible to detect the motion of unknown targets. This brings possibilities for applications. For example, user interface designers can use this technique to turn every physical object to control software or games: use a spoon to fast forward a YouTube video, use a paper gun to play a shooting game. However, one weakness lies in the system setup. Compared to only one sensor in most researches, the system requires four sensors. It also requires careful placement to ensure beam synchronization and the transformation matrix accuracy.

In the bibliography, I picked a paper about point cloud registration [1]. Point cloud is simply an image with depth information and is the output from stereo cameras, such as Microsoft Kinect. To detect motion using point clouds, the physical information, such as height, width, and depth, of the object is needed to find points that represent the same object between continuous point cloud capturing. The finding process is called registration. As aforementioned, this motion detection technique needs prior knowledge of the target and is computational intensive. Previously, state-of-the-art registration algorithms (e.g., SoftAssign, EMICP) for automatic registration of overlapping free form objects require $O(n^2)$ time, which makes those algorithms expensive to compute on continuous point clouds with thousands of points in each.

To solve the problem, Liu proposed to combine either SoftAssign or EMICP with iterative close point algorithm. Originally, SoftAssign or EMICP examine all combination of points of two point clouds. A K-dimensional tree is used to find closest points, and these correspondences of the established points are represented by probabilities estimated by the entropy maximization principle and optimized by deterministic annealing scheme. In addition, slack variables are introduced to model point outliers. Experiments are conducted to show the accuracy of the extended algorithm.

Still, motion sensing using this technique need registration pre-processing step, which is more computational intensive than solving a linear algebra equation. This is the biggest weakness for 3D camera approach. However, in terms of point cloud registration, the proposed algorithm increases registration accuracy to find overlapping objects. For the objects found in most tested images, the standard deviation data are mostly within 0.3 mm. Nonetheless, it is unfortunate to see that the time complexity is not discussed in the paper, leaving the impression that the time complexity is the same as or higher than SoftAssign or EMICP's.

References

[1] Y. Liu, "Automatic registration of overlapping 3D point clouds using closest points," Image and Vision Computing, vol. 24, no. 7, pp. 762–781, 2006.
[2] L. Miyashita, R. Yonezawa, Y. Watanabe, and M. Ishikawa, "3D Motion Sensing of Any Object Without Prior Knowledge," in Proceedings of ACM SIGGRAPH Asia 2016, pp. 218:1–218:11.