Tracking People using Multiple Kinects: Description, Objectives, Ethics, Resources

Chi-Jui Wu
cjw21@st-andrews.ac.uk
September 23, 2014

1 DESCRIPTION

Real time detecting and tracking multiple people have been achieved in systems that use RGB-D data(Munaro, Basso, & Menegatti, 2012), shape analysis, stereo and intensity images(Haritaoglu, Harwood, & Davis, 1998), and appearance-based tracking(Zhou & Hoang, 2005), just to name a few. Kinect is an inexpensive sensor that produces depth, RGB, infrared, and audio streams at high frame rates. The current work will build on top of Kinect's existing skeleton detection and tracking system.

The goal of the project is to design and evaluate an application that uses multiple Kinects to track moving people in real-time. Using multiple Kinects has the advantage of a larger field of view, potentially from different angles, used to capture more information about the people in sight. On the other hand, necessary Kinect calibration and data synchronization will introduce additional complexity. The tracking algorithm should be accurate and robust. It should also resolve the common problem of occlusion in detection and tracking.

Some applications of tracking multiple people include surveillance(Fuentes & Velastin, 2006), tracking with mobile robots(Choi, Pantofaru, & Savarese, 2011), facial animation of digital avatars(Weise, Bouaziz, Li, & Pauly, 2011), and the possibility of monitoring their blood oxygenation levels.

2 OBJECTIVES

The objectives are ordered by the difficulty of the tasks. They should be completed using two Kinects, unless explicitly stated otherwise.

1. Literature review
   - Identify the state of the art techniques in people tracking using RGB, depth, and infrared data
   - Identify the limitations of the current approaches
   - Identify important algorithms that can be incorporated to the current work.
2. Calibrate the Kinects using a test target
3. Synchronize the Kinects input from two different machines (May not be able to use two Kinect v2 on the same machine using the current SDK)

- RGB, infrared and depth

4. Recognize a single user (Match the skeletons from different Kinects)
5. Track a single user (Extend the Kinect's field of view)
6. Track two users (Specify the positions of the Kinects)
7. Evaluate the current system (Also applies to further objectives)
8. Automate the calibration process without using a test target
9. Track six users reliably in constraint test environment
10. Track six users who have various poses and interactions
11. Track six users in different indoor environments
    - brightness
    - various occlusion objects (Partial and full occlusion)
    - simulated physical office and home space
12. Track six users in outdoor environments

3 Ethics

Kinect is a commercial off-the-shelf hardware that has been approved for use worldwide. To help the researcher develop and verify the software, the project will require many recordings of one or more moving people. There is a small risk that people will exhibit inappropriate behaviour towards one another during the recording sessions. In such cases, the session will be immediately terminated. Participants will be required to sign a consent form, and they may leave the recording session at anytime without giving any reason. The recordings will be stored in an anonymous format and only be accessible to the researcher. The researcher will also incentivize the participants with sweets.

4 Resources

- Two Kinect for Windows v2
  - Note: May require additional Kinects and computers with 3.0 USB ports
- A large space to accommodate at least six people
  - To make recordings of moving people, for development, testing, and demonstration purposes

The supervisor will provide all of the items above.

REFERENCES

Choi, W., Pantofaru, C., & Savarese, S. (2011). Detecting and tracking people using an rgb-d camera via multiple detector fusion. *In Computer Vision Workshops (ICCV Workshops), IEEE International Conference on*, 1076-1083.

Fuentes, L. M., & Velastin, S. A. (2006). People tracking in surveillance applications. In *Image and vision computing* (Vol. 24, p. 1165-1171). Springer Berlin Heidelberg.

Haritaoglu, I., Harwood, D., & Davis, L. (1998). W4s: A real-time system for detecting and tracking people in 2 1/2 d. In *Computer vision âĂŤ eccv 1998* (p. 877-892). Springer Berlin Heidelberg.

Munaro, M., Basso, F., & Menegatti, E. (2012). Tracking people within groups with RGB-D data. *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, 2101-2107.

Weise, T., Bouaziz, S., Li, H., & Pauly, M. (2011). Realtime performance-based facial animation. *ACM Transactions on Graphics (TOG)*, *30*(77).

Zhou, J., & Hoang, J. (2005). Real time robust human detection and tracking system. *In Computer Vision and Pattern Recognition-Workshops*, 149-149.