Exploring a Multi-Sensor Picking Process in the Future Warehouse

Alexander Diete
September 9, 2016

University of Mannheim
About the project
Problem

Figure 1: Picking process in warehouses
Use sensors and video data to enhance the process
Hardware

- Data glass (Vuzix M100)
- Wristband (Custom 3D Print)
- Depth Sensor (Project Tango Tablet)
Data gathering
Data collected

- Data glass
  - IMU data
  - Video stream
- Wristband
  - IMU data
  - RFID read
- Tango
  - Point cloud data
Figure 2: Different parts being recorded
Figure 3: 3rd person depth view
Figure 4: Sensor Data Collector App
Activities to be recognized

- Navigation (walking to shelf)
- Locating shelf
- Grabbing into shelf
Problems

- Time synchronization
- Consistent recording rate for the sensors
- Start and endpoint of labels
Solutions

- Zero lining for time synchronization
  - Align datasets in post-processing
- Manual sensor rate adjustment for glasses
- Use observation video to pinpoint start and end of activities
Solutions - Alignment tool
Solutions - Labeling tool
Dataset
• First recording session resulted in 2.7 GB
• Different processes recorded
  • Picking from one shelf
  • Picking from multiple shelves
  • Picking with different hands
Figure 5: Accelerometer data from wristband
Future Work
Recording optimization

- Switch to full client server architecture
  - Synchronized start of all devices recording
  - Health status of sensors
  - Reduce the overall setup time
- Better live preview of data
  - Video stream and plot of data
  - Includes health status of sensors
Machine Learning

- Video stream
  - Object recognition (boxes, shelves)
  - Motion detection

- Sensor data
  - Activity recognition (walking, standing, arm movement)
  - Combination of both data streams
Depth information

- 3rd person perspective vs. 1st person perspective
- 3rd person perspective feasible for recognition but hard to deploy.
- 1st person perspective: minimum distance of depth sensor is 30cm
  - Means that detection of objects is not feasible
  - **But:** Can recognize if background is blocked by some object
  - Thus grabbing detection should be possible
Conclusion
Summary

- Created a framework for collecting multiple data sources
- Built tools to align and label data
- Proposed multiple approaches for activity recognition
Open Questions

- Is the selection of sensors sufficient for task?
- Can machine learning be applied to the combination of data?
- Semi supervised learning applicable for different warehouse locations?
Thank you for your attention