Multi-person Identification and Localization for Ambient Assistive Living

Georgios Galatas[1 2], Shahina Ferdous[1], and Fillia Makedon[1]

[1] Heracleia Human Centered Computing Lab, Computer Science and Engineering Dept.
University of Texas at Arlington, USA
[2] Institute of Informatics and Telecommunications, NCSR "Demokritos", Athens, Greece
{georgios.galatas,shahina.ferdous}@mavs.uta.edu,
makedon@uta.edu

Abstract. In this paper, we present a novel, non-intrusive system that uses RFID technology and the Kinect sensor in order to identify and track multiple people in an assistive apartment. RFID is used for both identification and location estimation while information from the Kinect sensor is used for accurate localization. Data from the various modalities is fused using two techniques. During the experiments conducted, our system exhibited high accuracy, thus proving the effectiveness of the proposed design.

Keywords: Person localization, context-awareness, multi-sensory fusion, depth information, Microsoft Kinect, RFID.

1 Introduction

Successful multi-person identification and localization is a fundamental step towards activity monitoring, emergency detection and ultimately context-awareness. The proliferation of ambient-intelligent environments has triggered research related to applications, such as monitoring Assistive Daily Living (ADL), fall detection, risk prevention and surveillance [1, 2]. Accurate person localization plays an essential role in all these applications and has been dealt with using many different approaches. Video cameras are the most commonly used devices since they are affordable and provide abundant information about people's activities and their surroundings. Nevertheless, when used domestically, they can be considered invasive, and the segmentation and tracking problems using planar video in a multi-person setting are very challenging [3, 4]. On the other hand, RFID systems are consistently used to keep track of medicine and patients in hospitals [5]. However, radio frequency signal propagation suffers from various issues such as multi-path attenuation, diffraction and reflection in an indoor environment [6] and therefore, RFID cannot be considered sufficiently accurate for localization purposes.

Therefore, in our approach, we have utilized the identification capabilities of RFID and combined that with precise 3D tracking from the Kinect to create an accurate identification and localization solution. RFID is used for both discerning between

N. Streitz and C. Stephanidis (Eds.): DAPI/HCII 2013, LNCS 8028, pp. 109–114, 2013.

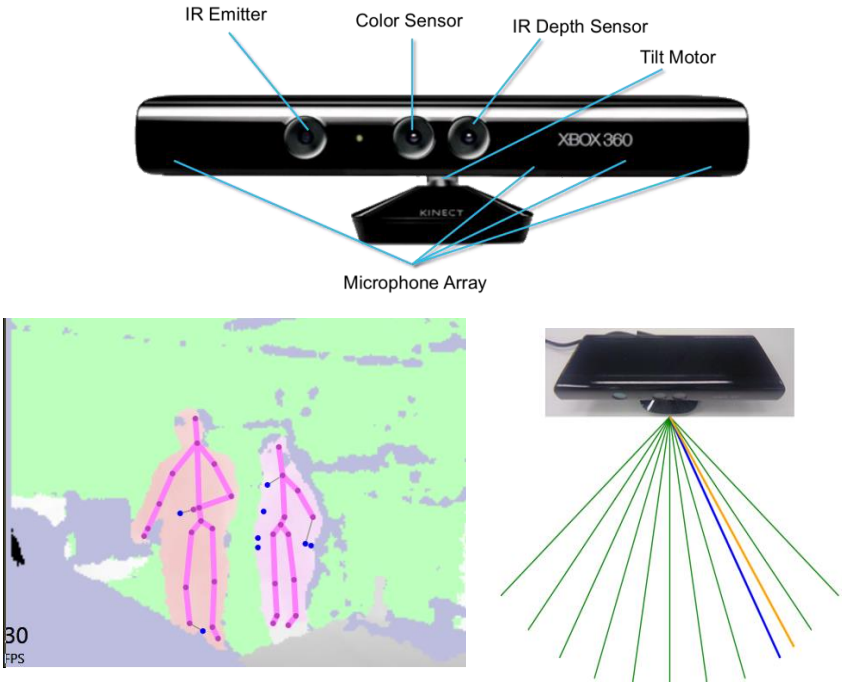

Fig. 1. The Kinect sensor (top) and examples for skeletal tracking (bottom-left) and audio localization (bottom-right)

users and providing a rough estimate of their location. Skeletal tracking is carried out using the Kinect sensor's 3D depth images and sound source localization is conducted utilizing its microphone array to deduce accurate location information. The two approaches we have used to fuse the data from all sources are classification-based and proximity-based, with the later exhibiting the highest accuracy.

In section 2 of this paper, we present the methodology used to build our system. Section 3 describes the experimental setup and results and section 4 concludes our work.

2 Methodology

The Kinect is a new device released by Microsoft that incorporates a color camera, a structured light 3D depth sensor and a microphone array (fig. 1). In our system, we utilized information captured by both the depth sensor and the microphone array. More specifically, we used skeletal tracking implemented using the MS Kinect SDK, in order to locate and track people in the field of view (FOV) of the sensor. Each detected skeleton has a unique identifier for a specific session, which is defined by the 3D space coordinates of its joints. In addition to the depth sensor, the Kinect has a microphone array comprised of 4 microphones in order to localize sounds. The information acquired using the array is the direction of the incoming sound, as well as a

confidence indicator for the estimated direction. In our system, this information is utilized as a rough estimate, only if the other two sources fail to determine the exact location of the person.

The RFID system that we used was comprised of two antennas and a tag reader. Its main role was to identify the person in its field of sense (FOS), but also to provide a rough estimate of her/his location using the received signal strength indicator (RSSI) from each antenna. The mapping between the RSSI values and the actual position of the tag is accomplished through a calibration process that accounts for both the directionality of the antennas and the specific layout of the room. Multiple people are identified using their unique RFID tag and tracked as long as they remain in the FOS of the system. Skeletal tracking alone may not be able to discern between different people since a new tracking id is issued each time a person is lost from the FOV of the Kinect and then re-enters. Therefore, we improved our system's accuracy by matching the new RFID tag with the new tracking id as soon as an individual enters the room. This technique allows identification of each individual detected by the skeletal tracker. Finally, when no skeleton is detected in the FOV, but a tag is still being detected, audio localization is utilized in order to increase accuracy (e.g. triangulation when only one antenna reads the tag).

In order to combine the location information from all sources, we implemented 2 techniques: 1) proximity-based and 2) classification-based as shown in fig. 2. The proximity-based approach uses a proximity database that contains the position signature properties of: 1) the RSSI of a tag from an RFID antenna within a region, 2) the antennas that detect the particular tag and 3) the (x, y, z) coordinates range of the region, for a mesh of discrete regions in the deployment space. In the test phase, each

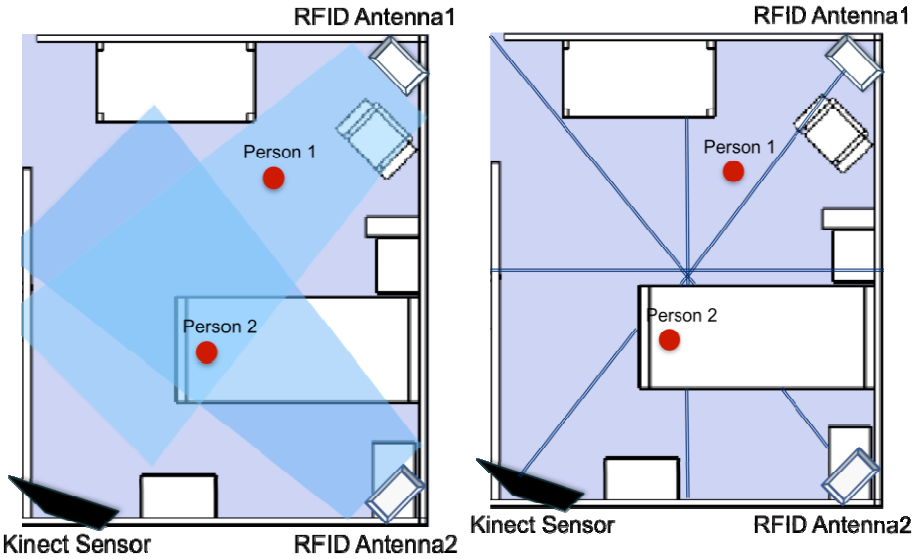

Fig. 2. Proximity-Based approach (left), Classification-Based approach (right)

detected signature vector is compared to this database and the closest match found is returned as the current position. After the region has been defined, the precise coordinates of the person within this region are estimated using the Kinect skeletal tracking. The classification approach is based on a training phase during which statistical regression is applied on pre-specified position signatures (RSSI in our case) in order to build a classifier. This classifier is then used to classify the current signature to a particular sector of the deployment space. This sector is then mapped to the location coordinates detected from the Kinect sensor. As afore-mentioned, in both approaches we use the sound from the microphone array as another modality besides skeletal tracking to resolve ambiguities in mapping.

Fig. 3. RFID system Devices

3 Discussion of Results

The equipment used for our experiments are an MS Kinect sensor, an Alien ALR-9900+ RFID tag reader and 2 Alien ALR-9611 circular polarization antennas (fig. 3). The range of the

Table 1. Experimental results

Approach	Accuracy
Classification-based (statistical regression) (4-person)	60%
Classification-based with tag-id matching (4-person)	65%
Classification-based with tag-id matching (2-person)	67%
Proximity-based (4-person)	68%
Proximity-based with tag-id matching (4-person)	76%
Proximity-based with tag-id matching (2-person)	86%

System is sufficient for a domestic environment and circular polarization ensures that tag detection is orientation-invariant. The experiments were carried out at the Heracleia assistive apartment. The antennas were placed at the 2 corners of the bedroom and the Kinect on a 3rd corner in order to maximize the FOS and FOV. Real-time tracking data is properly converted and displayed on a visualization tool (fig. 4). We conducted extensive experiments in our realistic domestic setup for 2 and 4 individuals, and the results are shown in Table 1. Accuracy denotes the percentage of correctly estimated locations for all individuals present in the room and also accounts for misidentifications and mismatches between the detected tag sector and skeletal id location. Accuracy was higher for the 2 person scenario with both methods. In

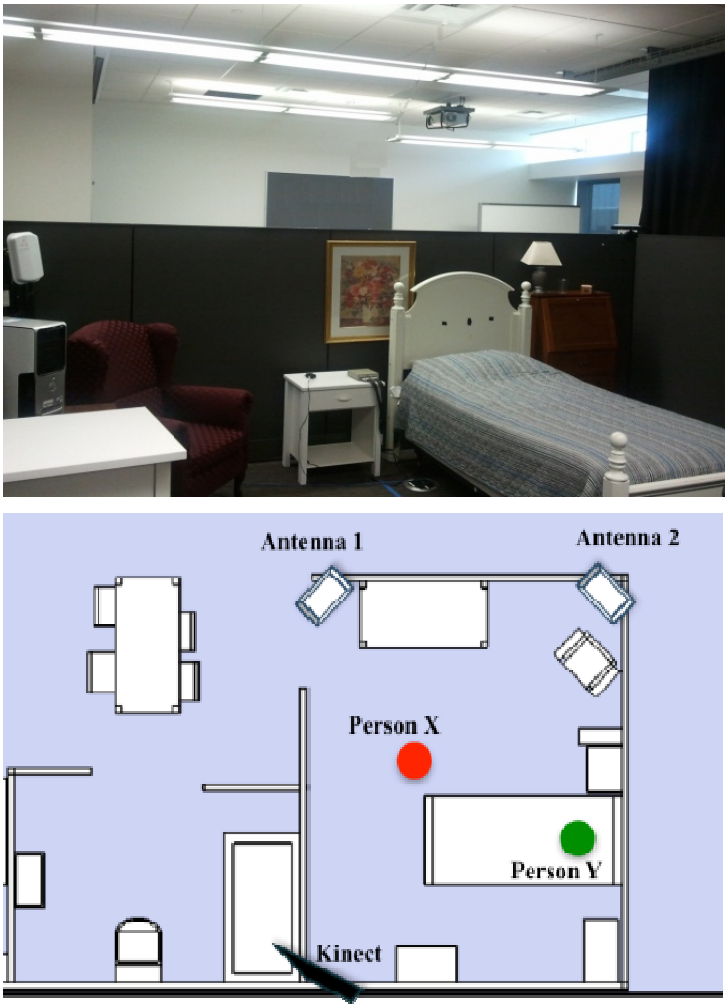

Fig. 4. The Heracleia assistive apartment (top) and the localization visualization tool (bottom)

addition, matching each RFID tag to the corresponding skeletal id resulted in a slight increase in accuracy, by minimizing false identifications. The highest accuracy was 86%, achieved using the proximity-based approach. The reason that the proximity-based approach performed better overall is that it utilizes both the RSSI and the Kinect information for sector mapping. On the other hand, the classification-based approach only uses the RSSI for mapping.

4 Conclusions

In conclusion, we combined the identification capabilities of RFID with accurate tracking from the Kinect in order to create an accurate multi-person identification and localization system for assistive environments. We used 3 types of data, RSSI, 3D depth and audio to solve the localization problem using 2 methods. The experiments conducted, proved the effectiveness of our system for this scenario. Further experiments will include additional Kinect sensors for more robust tracking.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grants No. NSF-CNS 1035913, NSF-CNS 0923494.

References

1. Teixeira, T., Jung, D., Savvides, A.: Tasking networked cctv cameras and mobile phones to identify and localize multiple people. In: Proc. Ubicomp, pp. 213–222 (2010)
2. Cucchiara, R., Fornaciari, M., Prati, A., Santinelli, P.: Mutual calibration of camera motes and RFIDs for people localization and identification. In: Proc. ICDSC (2010)
3. Mitzel, D., Horbert, E., Ess, A., Leibe, B.: Multi-person tracking with sparse detection and continuous segmentation. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 397–410. Springer, Heidelberg (2010)
4. Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M., Shafer, S.: Multi-camera multi-person tracking for easy living. In: Proc. IEEE IWVS (1997)
5. Ashar, B.S., Ferriter, A.: Radiofrequency identification technology in health care: benefits and potential risks. JAMA: The Journal of the American Medical Association 298, 2305–2307 (2007)
6. Hahnel, D., Burgard, W., Fox, D., Fishikin, K., Philipose, M.: Mapping and Localization with RFID Technology. In: Proc. IEEE ICRA, pp. 1015–1020 (2004)