

COMPLEX HUMAN ACTIVITY RECOGNITION USING INDOOR POSITION, AMBIENT SOUND AND ACCELEROMETER DATA

Mridul Khan¹, Sheikh Iqbal Ahamed¹, Roger O. Smith²

¹*Ubicomp Lab, Marquette University, Milwaukee, USA*

²*R₂D₂ Center, University of Wisconsin-Milwaukee, Milwaukee, USA*

INTRODUCTION

Human activity recognition has numerous applications in context aware computing and rehabilitation engineering. Systems that can accurately identify human activities from sensor data might provide activity specific assistance or implicitly create activity logs for wellness monitoring. With the massive influx of sensor rich portable devices (Lane, Miluzzo, Lu, Peebles, Choudhury, & Campbell, 2010), this kind of applications can now be built on established mobile platforms.

Our previous work on activity recognition was about (Khan, Ahamed, Rahman, & Smith, 2011) detecting simple human activities like walking, running and jumping using a cell-phone application. We extracted features with the help of algorithms developed for EEG signal analysis and then performed classification using Support Vector Machine (SVM). The method we devised was exceptionally efficient but could only recognize activities that generated consistent motion traces. Our current work presents our preliminary work on detecting complex activities such as using the computer and watching television.

The rough information about motion provided by an accelerometer hardly provides enough data to detect basic activities. For complex human activity recognition, it is necessary to use other sensors since many activities generate similar types of accelerometer traces and can only be differentiated by context specific information such as position and ambient sound. In this work, we explain how indoor position estimation using wireless network sensing and the analysis of ambient sound can be used for complex activity detection.

INDOOR POSITIONING

Localization using GPS is only possible in outdoor environments. A robust solution for indoor positioning is still not available. A popular approach to this problem is wireless network sensing (Liu, Darabi, Banerjee, & Liu, 2007). In this approach, position is inferred from reachable wireless networks. Advances in this area have produced accuracies of up to 1.5 meters (Martin, Vinyals, Friedland, & Bajcsy, 2010). An unconventional approach to indoor positioning is to use the Acoustic Background Spectrum (Tarzia, Dinda, Dick, & Memik, 2011). This method uses ambient noise to detect position without any additional infrastructure. We hypothesized that better indoor localization can be achieved by using both wireless beacons and ambient sound. However in our test environment, the ambient sound based method performed too poorly to be useful. With enough beacons, the wireless network based approach is sufficient.

ACTIVITY RECOGNITION USING SOUND

Background sound has also been used for human activity recognition. A method based on Mel Frequency Cepstral Coefficients (MFCC) and Dynamic Time Warping (DTW) has produced accuracies of up to 92.5% for fourteen typical daily activities (Human Activity Recognition from Environmental Background Sounds for Wireless Sensor Networks, 2007). For activities like watching television, position and sound is a much better predictor than physical motion.

ACCELEROMETER BASED ANALYSIS

Using accelerometers for human activity recognition has become a common practice (Companjen, 2009). Accelerometers are generally used to detect only simple activities since the sensors available in most portable

devices provide extremely noisy data. Combining position and ambient sound data with accelerometer data can make complex activity detection possible.

METHODOLOGY

Training for Indoor Positioning

A wireless network fingerprint based indoor localization method has been described by (Martin, Vinyals, Friedland, & Bajcsy, 2010). Using this method they were able to perform room level localization with just two wireless networks with good signal strengths (RSSI above -75 dBm). We were able to programmatically detect the RSSI of available networks on an iPod using a private application programming interface (API). In the training phase, the portable device had to be moved throughout the establishment to collect the network names (SSID) and signal strengths (RSSI) at various points. Once enough data for every room is available, a nearest neighbor search using the average RSSI at an unknown position can give an estimate of position.

Activity Training

Training for activity recognition can be started after the positioning training is complete. We propose a system that uses position, ambient sound and accelerometer data to create a complete profile of an activity. Wireless network signal strengths can be used to infer position. Any device with a microphone can be used to monitor sound. The iOS platform has built in support for monitoring the average and peak power of the sound it is receiving. The accelerometer data can be collected to sense the amount of motion the device is going through. All this data can be used to build a training dataset of complex activities. Figure 1 shows how the sensed data can be fit into a feature vector that concisely describes a complex activity.

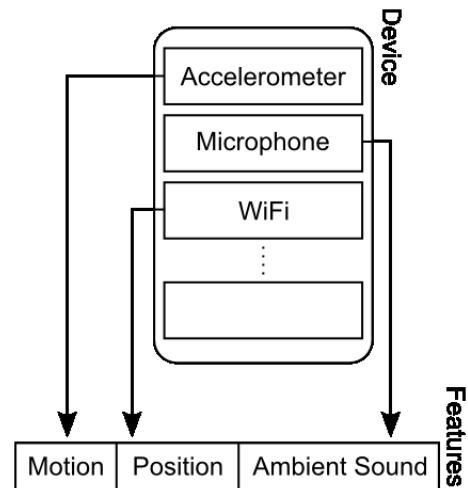


Figure 1: Compiling a feature vector from sensed data.

PRELIMINARY RESULTS

For our preliminary tests we used a fourth generation iPod touch for data collection. We used two beacons placed in two adjacent rooms. One beacon was a DLink N300 router (beacon 1) in infrastructure mode while the other was a SparkFun WiFly shield (beacon 2) in ad hoc mode. To detect whether a user is watching TV or working on a computer, we can use the portable devices proximity to the static locations of these devices. To test the practicality of this idea, we placed one beacon near the computer and another one near the television. In general propagation based models cannot provide a good distance measurement (Martin, Vinyals, Friedland, & Bajcsy, 2010). However, in this case we can simply check which beacon's RSSI value is the highest. Figure 2 makes it clear that checking the RSSI value can give us a good idea of the portable devices proximity to one of the two beacons.

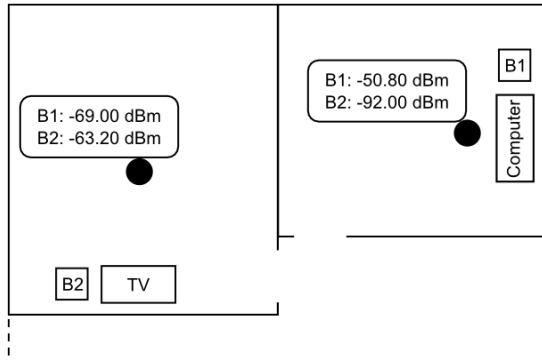


Figure 2: WiFi beacons B1 and B2 placed close to points of interest.

Table 1: Descriptive features extracted from random 5 second windows of a user performing various activities.

Activity	Features		
	Average Sound Power (db)	Standard Deviation of Peak Sound Power (db)	Standard Deviation of Acceleration Magnitude (m/s^2)
Walking	-17.49	2.23	2.93
Climbing Stairs	-15.90	2.83	3.41
Watching Television	-37.56	4.83	0.02
Using the Computer	-33.25	6.34	0.09

A sample from our preliminary findings listed in Table 1 show that activities like "Watching Television" and "Using the Computer" incur too little variation in accelerometer data to be of significance. For these activities looking for patterns in ambient sound data seems to be more sensible. Activities like "Walking" and "Climbing Stairs" on the other hand cause ambiguous ambient sound but distinctive acceleration traces. When we perform multi-level analysis that combines accelerometer, ambient sound and position data complex activity detection becomes easier.

CONSLUSIONS

Our preliminary results have led us to believe that a fundamental set of features that can describe any possible activity can eventually be derived and sensed using many portable devices available to consumers. These can be used like a DNA for complex activity recognition. Future work should be carried using a larger dataset to verify the applicability of this method.

REFERENCES

- Human Activity Recognition from Environmental Background Sounds for Wireless Sensor Networks. (2007). *Proceedings of the 2007 IEEE International Conference on Networking, Sensing and Control* (pp. 307-312). London, UK: IEEE.
- Companjen, B. (2009). Classification Methods for Activity Recognition., (p. Proceedings of Twente Student Conference on IT).
- Khan, M., Ahamed, S. I., Rahman, M., & Smith, R. O. (2011). A Feature Extraction Method for Realtime Human Activity Recognition on Cell Phones. *Proceedings of 3rd International Symposium on Quality of Life Technology (isQoLT 2011)*. Toronto, Canada.
- Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A Survey of Mobile Phone Sensing. *IEEE Communications Magazine*, 140-150.
- Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. 37, NO. 6*, 1067-1080.
- Martin, E., Vinyals, O., Friedland, G., & Bajcsy, R. (2010). Precise Indoor Localization Using Smart Phones. *Proceedings of the ACM International Conference on Multimedia (ACM Multimedia 2010)* (pp. 787-790). Florence, Italy: ACM.
- Tarzia, S. P., Dinda, P. A., Dick, R. P., & Memik, G. (2011). Indoor Localization without Infrastructure using the Acoustic Background Spectrum. *Proc. 9th Intl. Conf. on Mobile Systems, Applications, and Services (MobiSys'11)* (pp. 155-168). Washington DC, USA: ACM.