9781941546277/2015

IN-HOME BEHAVIORAL MONITORING USING SIMULTANEOUS LOCALIZATION AND ACTIVITY DETECTION

Mateusz M. Mittek, **Jay D. Carlson**, **Francisco Mora-Becerra**, **Eric T. Psota**, and **Lance C. Pérez**

Department of Electrical Engineering
University of Nebraska-Lincoln, Lincoln, NE 68588

ABSTRACT

Shifting demographics in the U.S. has created an urgent need to reform the policies, practices, and technology associated with delivering healthcare to geriatric populations. Automated monitoring systems can improve the quality of life – while reducing healthcare costs – for individuals aging in place. For these systems to be successful, both activity detection and localization are important, but most existing research focuses on only one of these technologies – and systems that *do* collect both data treat these data sources separately. Here, we present SLAD – Simultaneous Localization and Activity Detection – a novel framework for simultaneously processing data collected from localization and activity classification systems. Using a hidden Markov model and machine learning techniques, SLAD fuses these two sources of data in real time using a probabilistic likelihood framework, which allows activity data to refine localization, and vice-versa. To evaluate the system, a wireless sensor network was deployed to collect RSSI data and IMU data concurrently from a wrist-worn watch; the RSSI data was processed using a radial basis function neural network localization algorithm, and the resulting position likelihoods were combined with the likelihoods from an IMU activity classification algorithm. In an experiment conducted in an indoor office environment, the proposed method produces 97% localization accuracy and 85% activity classification.

Keywords: RSSI localization, activity classification, indoor tracking, behavioral monitoring

INTRODUCTION

The baby boomer generation reaching retirement has dramatically increased healthcare spending both in terms of the number of patients needing healthcare, and the types of healthcare services these patients use. That the U.S. population is living longer than ever furthers the need to reduce the cost of existing long-term structured care systems, while also shifting healthcare delivery from skilled nursing care (SNC) to in-home (aging in place) care for patients who are more independent.

Behavioral monitoring systems that can determine where a patient is and which activities he or she is performing offer potential benefits to both SNC and in-home care. For SNC facilities, behavioral monitoring systems can reduce operational overhead by automating the process of performing and recording physical therapy. For in-home care, behavioral monitoring systems can supplement a home care specialist by observing activities of daily living (ADLs) and instrumental activities of daily living (IADLs), offering reminders to the patient, and alerting care providers to abnormal behaviors or decline in activity.

When focusing on in-home care, in the most useful scenario a behavioral monitoring system could detect complex, nuanced activities like cleaning, cooking, dressing, or grooming. While it is true that highly-accurate localization data might provide data at a fine enough granularity to distinguish between activities that occur in different parts of a room, some activities of interest simply cannot be distinguished based on location alone. To make matters worse, most low-cost localization systems proposed for in-home deployment lack the resolution necessary to distinguish between activities occurring in the same room.

Instead of classifying activities using a localization system, many researchers supply the patient with an accelerometer or inertial measurement unit (IMU) worn on their person to collect inertial data. By analyzing this data using statistical, time-domain, and frequency-domain methods, many activities can be identified without localization. However, many activities have similar inertial data profiles, making them difficult to distinguish.

To this end, this work proposes to solve these ambiguities by simultaneously refining localization and activity data by taking into account the dependence between these two data sources and building a hidden Markov model that is evaluated using a Viterbi decoder.

BACKGROUND

In-home behavioral monitoring research has focused on two key areas – indoor localization, and activity classification. For activity classification, much work has focused on wearable IMUs. Previous work demonstrates that even the simplest wearable IMUs equipped with only 3-axis accelerometers can provide satisfactory activity classification [1]. Different processing methods such as decision trees or machine learning techniques can be used to construct high accuracy classifiers for the IMU data [2]. Teixeira et al. [3] demonstrated that many ADLs such as cooking, eating, and cleaning can be detected accurately by modeling human arm movement.

There are many technologies which can be used for indoor localization. Many indoor localization systems use passive infrared (PIR) sensors to detect when people enter and leave different areas of the house. While being low-cost and widely-available, PIR sensors cannot differentiate between different people (or pets!), which creates difficulty when deploying in unconstrained environments. Nonetheless, the platform has been used extensively for the study of gait speed in geriatric populations (e.g. [4]).

Many attempts have been made to use IMUs for localization [5–7]. The CAR system [8], winner of the 2012 EvAAL competition [9], uses Kalman Filter to fuse the range estimated by the analysis of the RFID signal strength with the foot-worn IMU data. Their average localization error decreases as the number of RFID tags deployed in the environment increases.

RSSI localization provides similarly low cost and complexity when compared with PIR, while offering higher accuracy and the ability to distinguish between targets. The Active Badge System [10] was one of the first widely-studied systems for RSSI localization.

A Bayesian framework using the Hidden Markov Model was used previously in [11] to resolve the indoor trajectories. Attempts were made to combine the RSSI data from a wireless sensor network and the movement information from the wearable IMUs [12] but not in a context where both the activity and location were resolved simultaneously.

PROPOSED METHOD

The data for this work was collected using the Angelos Indoor Monitoring System [13]. The Angelos system is a wireless sensor network consisting of three types of nodes. The first is a TI Chronos smart watch worn by the user that collects measurements from an onboard accelerometer. These measurements are periodically broadcasted to the second type of node in the network consisting of a distribution of fixed-location sensors that are designed to perform two basic operations: 1) forward the accelerometer information, and 2) record and forward the received signal strength indicator (RSSI) information. The third type of node receives the forwarded accelerometer readings and RSSI data from each of the fixed-location nodes and forwards this data to a centralized server for data storage and processing.

The novelty of the proposed SLAD method is the use of the Viterbi decoder to resolve the maximum likelihood path between joint activity and locations. The allowable paths and their associated probabilities are derived from a combination of activity and location likelihoods and the transition probability matrix. At each instant in time, probability likelihood vectors are calculated as a product of location likelihoods and activity classification likelihoods. The transition probability matrix makes it possible to fuse activity and location data, as it captures known properties of the environment and human activity patterns.

ACTIVITY CLASSIFICATION

For the particular implementation presented here, samples of the 3-axis accelerometer on the smart watch are used. Due to the time-varying nature of human activities and the sampling rate of the smart watch, accelerometer samples are grouped together in blocks of data from which feature vectors are calculated for each instance of the time. The following 12 features are calculated for each block of data:

- Mean value of x, y, z acceleration
- Standard deviation of x, y, z acceleration
- Energy of the x, y, z acceleration
- Correlation coefficients $\rho_{x,y}$, $\rho_{x,z}$, $\rho_{y,z}$

These features have previously been shown to achieve accurate activity classifications and exhibit invariance to activity time scales [1]. A Radial Basis Function Neural Network (RBFNN) classifier is used to produce the activity data likelihood vectors. Given a possibility of m activities at each time t, an $m \times 1$ likelihood vector A_t is produced by the RBFNN containing likelihood estimates of each activity.

LOCALIZATION

The proposed method assumes that location likelihoods are in a discrete-time, discrete-location format. Thus, instead of being represented by Cartesian coordinates, the location must be a discrete value, such as a grid number or a room label. An RBFNN location classifier is trained using RSSI fingerprinting. Given a number of locations n, at each time t, an $n \times 1$ likelihood vector L_t is produced that contains the likelihood estimate for each considered location.

PATH RESOLUTION

To estimate either the current activity or the current location, one could simply select the index of the most likely elements of their respective likelihood vectors, A_t and L_t. This approach, however, does not take advantage of the the physical geometry of the environment, or that certain activities are constrained to certain locations. For example, consider a scenario where room X is connected to room Y and room Y is connected to room Z, while room X is not connected to room Z. In this scenario, selecting locations independently from each likelihood vector would allow for the possibility of moving from room X directly to room Z without passing through room Y. However, it is known that the user must pass through room Y to get from X to Z.

The Viterbi decoder is one such method that is designed to find a path through a given set of that states that has the highest probability of occurring given the observed data. The Viterbi decoder operates on a set of state likelihoods and transition probabilities. For each time instance t, the activity and location classifiers produce the likelihood vectors A_t and L_t. By assuming independence, these likelihoods can be combined into a single $[(m \times n) \times 1]$ likelihood vector that considers all combinations of activities and locations. This is done by calculating the matrix $S_t = A_t L_t^T$ and vectorizing the matrix by concatenating the columns. Hence, each element of the state likelihood vector S_t can be interpreted as the probability of performing a particular activity at a particular location at time t.

The state transition probability matrix takes the form

$$M_T = \begin{bmatrix} C_1 & p_{1,2} & p_{1,3} & \cdots & p_{1,m\times n} \\ p_{2,1} & C_2 & p_{2,3} & \cdots & p_{2,m\times n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{m\times n,1} & p_{m\times n,2} & p_{m\times n,3} & \cdots & C_{m\times n} \end{bmatrix}. \tag{1}$$

Where $C_1 ... C_{m\times n}$ are the probability coefficients of staying in the same state (when the individual keeps performing the same activity at the same location), and $p_{i,j}$ are the transition probabilities from state i to state j where each state represents the combination of activity and location. It is assumed that an activity at a particular location is likely to be followed by that same activity and location. Thus, the diagonal of the matrix M_T consists of constants C_i that are larger than the off-diagonal elements. The off-diagonal elements $p_{i,j}$ are non-zero only for allowable state transitions. The state transition matrix can be either predefined manually using known heuristics regarding the environment or trained by observing allowable activity/location pairs in a system that collects ground truth data.

Using the vector of likelihoods S_t and the transition probability matrix M_T, the Viterbi decoder produces the most likely path through the activity/location states. The Viterbi algorithm essentially operates by evaluating the probability of all allowed state transitions and only keeping track of the incoming path with the highest probability at each state. Thus the memory storage requirements and computational complexity of the Viterbi algorithm grows linearly with time. In the end, the Viterbi algorithm contains a set of $m \times n$ possible paths and selects the one that is the most likely.

Figure 1 illustrates the operation of the proposed method. Both IMU and RSSI data are recorded by the system and likelihoods are obtained using RBFNN classifiers. Then, the joint activity/location likelihoods are calculated by combining individual likelihood vectors and assuming independence. Finally, the joint likelihoods and the state transition matrix are used by the Viterbi decoder to produce the most likely path through the states.

9781941546277/2015

Figure 1: System data flow showing how IMU and localization data are combined into joint likelihoods which are evaluated by a Viterbi decoder using a joint state transition matrix.

EVALUATION

An experiment was performed to validate the proposed method using the Angelos Indoor Monitoring System. Three rooms were considered in an indoor office space and used as the possible locations. In the experiment, five activities were considered (not all of which were possible in each location).

The participant wore a smart watch that recorded 3-axis accelerometer data at a frequency of 10Hz. The watch was worn on the right hand of the user, so that activities could more easily be distinguished. The activity classifier was trained using five separately collected sets of accelerometer data. Each activity feature vector was calculated from the time window of 50 samples (5 seconds). The window was progressed by a half of the window size (25 samples / 2.5 seconds). The energy was calculated using 32-point FFT.

The RSSI measurements between fixed-location nodes and the smart watch were collected every second. The fixed-location nodes were deployed as shown in Figure 2. The total of 1078 seconds worth of data was used for training the RBFNN for location classification.

The transition probability matrix was build in such a way that the walking activity joins the locations together and allows only a particular subset of activities to be done at each location, as shown in Figure 2. The coefficients C_i were all set to a value that is $10\times$ larger than any of the off-diagonal elements $p_{i,j}$ in order to stabilize the resulting paths. This is easily justified since, at small time intervals, people are likely to maintain their current activities and locations.

A script of sequential activities and locations was defined to mimic the behavior of a typical person in the office environment. The accuracy is the percentage of properly classified feature vectors. The overall system performance (top accuracy) was: 97% for the location and 85% for the activity.

Figure 3 shows the interval graphs with raw output of the activity and location classifiers as a function of time along with the scripted ground-truth data. It is clearly visible that certain activities were not easily classified (e.g. writing was often mistaken with idle or computer use). Similarly, the localization classifier could not accurately distinguish between locations in certain instances.

Table 1: Comparative results demonstrating a significant improvement in performance when the system uses a monolithic state transition matrix combining both data sources during refinement.

Method	Activity accuracy	Localization accuracy
Separate, Raw	80%	78%
Separate, Viterbi	80%	80%
Joint, Raw	80%	78%
Joint, Viterbi	85%	97%

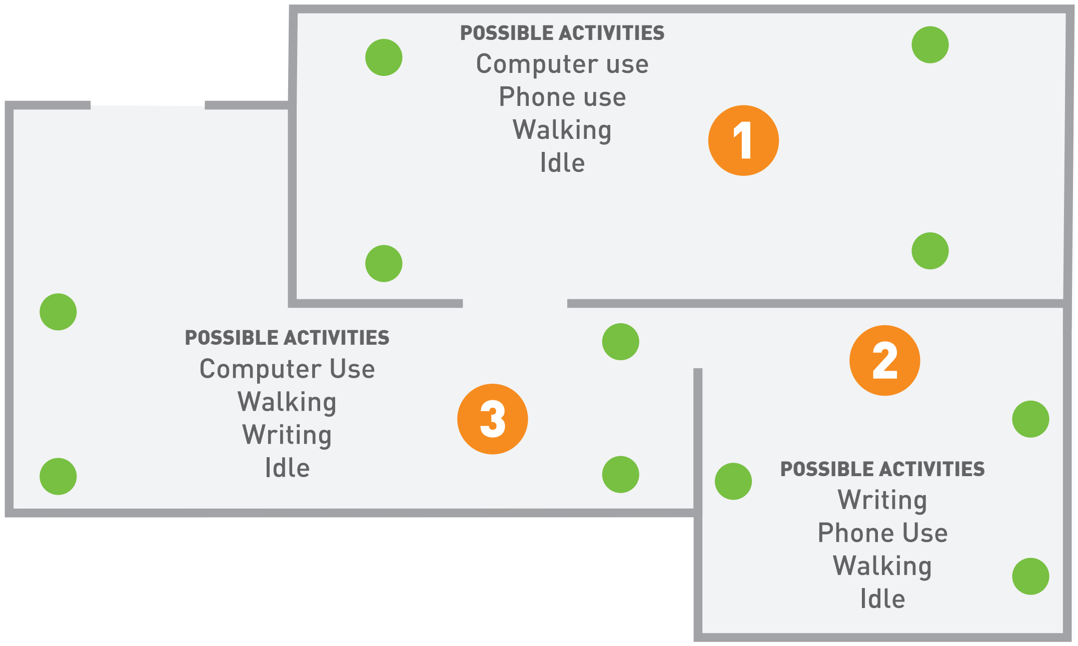

Figure 2: Map of the environment with the activities allowed by the transition matrix. The green dots represent the location of the fixed-location motes.

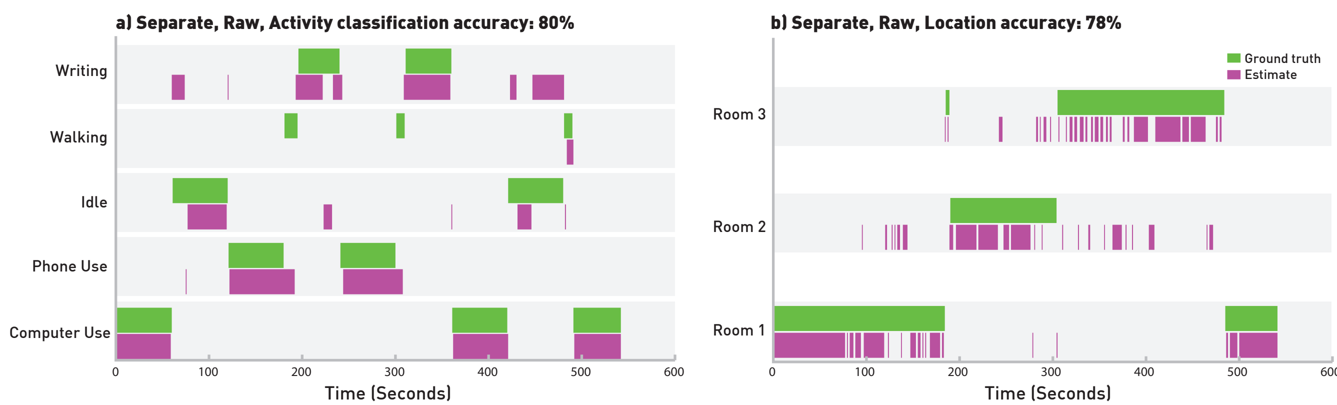

Figure 3: The interval graph showing the ground truth (green) and estimated (red) activity (left) and location (right) with the accuracy of 80% and 78% respectively. Raw separate classification data.

Figure 4 shows better path stability and increased accuracy when compared to the raw classifier output. The localization accuracy was significantly improved to 97% when the joint likelihoods were used along with the state transition matrix when compared to the raw localization classifier output (78%). The activity classification also improves from 80% to 85%.

Table 1 compares the performance improvements when using joint likelihoods and the transition state matrix. When joint likelihoods were used directly (without state transition refinement), there was no improvement to localization or activity classification accuracy. When the individual localization and activity classification inputs were treated separately and processed through separate state transition matrices, localization accuracy increased only marginally, while activity

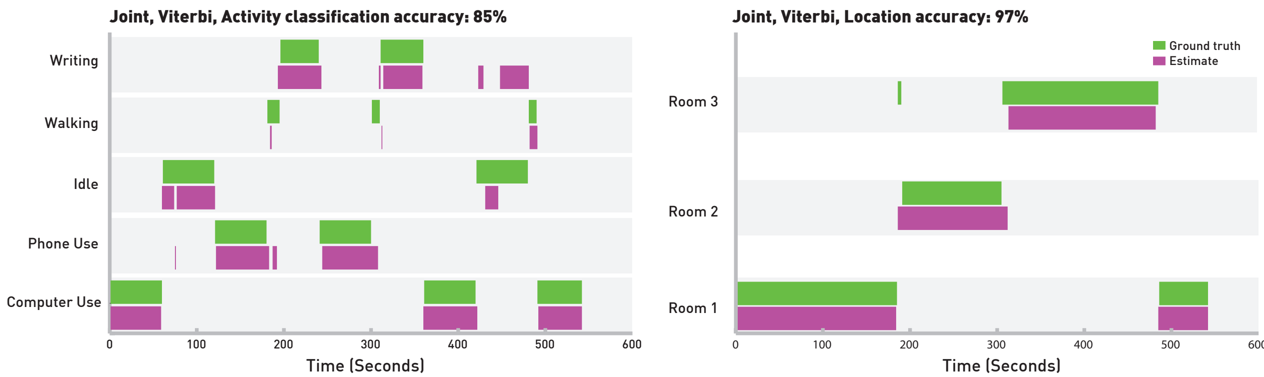

Figure 4: The interval graph showing the ground truth (green) and estimated (red) activity (left) and location (right) with the accuracy of 85% and 97% respectively. Joint likelihoods are estimated through the Viterbi decoder.

classification accuracy saw no improvements.

However, when the separate data sources were treated as joint data and processed using a monolithic state transition matrix, localization accuracy improved significantly to 97%, while activity classification accuracy improved modestly to 85%.

CONCLUSION

This work proposes simultaneous localization and activity detection (SLAD) as a method for accurate activity classification and localization. By treating localization data and activity classification data as joint likelihoods, the proposed method resolves noisy location and activity data using a joint state transition matrix, which offers a compact means of representing geometric and functional constraints of the environment. Experimental results show that this method allows a low-cost, unobtrusive system to gain significant improvement in identifying activities and localization.

REFERENCES

[1] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *AAAI*, vol. 5, pp. 1541–1546, 2005.

[2] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, pp. 119–128, Jan. 2006.

[3] T. Teixeira, D. Jung, G. Dublon, and A. Savvides, "Recognizing activities from context and arm pose using finite state machines," in *Third ACM/IEEE International Conference on Distributed Smart Cameras, 2009. ICDSC 2009*, pp. 1–8, Aug. 2009.

[4] J. A. Kaye, S. A. Maxwell, N. Mattek, T. L. Hayes, H. Dodge, M. Pavel, H. B. Jimison, K. Wild, L. Boise, and T. A. Zitzelberger, "Intelligent systems for assessing aging changes: Home-based, unobtrusive, and continuous assessment of aging," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 66B, pp. i180–i190, July 2011. PMID: 21743050.

[5] Y. Nam, "Map-based indoor people localization using an inertial measurement unit.," *J. Inf. Sci. Eng.*, vol. 27, pp. 1233–1248, 2011.

[6] F. De Cillis, F. De Simio, L. Faramondi, F. Inderst, F. Pascucci, and R. Setola, "Indoor positioning system using walking pattern classification," in *2014 22nd Mediterranean Conference of Control and Automation (MED)*, pp. 511–516, June 2014.

[7] S. Grzonka, A. Karwath, F. Dijoux, and W. Burgard, "Activity-based estimation of human trajectories," *IEEE Transactions on Robotics*, vol. 28, pp. 234–245, Feb. 2012.

[8] A. R. J. Ruiz, F. S. Granja, J. C. Prieto Honorato, and J. I. G. Rosas, "Accurate pedestrian indoor navigation by tightly coupling foot-mounted imu and rfid measurements," *Instrumentation and Measurement, IEEE Transactions on*, vol. 61, no. 1, pp. 178–189, 2012.

[9] J. A. lvarez Garca, P. Barsocchi, S. Chessa, and D. Salvi, "Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 EvAAL competition," *Journal of Ambient Intelligence and Smart Environments*, vol. 5, pp. 119–132, Jan. 2013.

[10] R. Want, A. Hopper, V. Falco, and J. Gibbons, "The active badge location system," *ACM Trans. Inf. Syst.*, vol. 10, p. 91102, Jan. 1992.

[11] S. Yi, P. Mirowski, T. K. Ho, and V. Pavlovic, "Pose invariant activity classification for multi-floor indoor localization," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 3505–3510, Aug. 2014.

[12] A. Fink, H. Beikirch, M. Voss, and C. Schroder, "RSSI-based indoor positioning using diversity and inertial navigation," in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–7, Sept. 2010.

[13] J. D. Carlson, M. Mittek, S. A. Parkison, P. Sathler, D. Bayne, E. T. Psota, L. C. Perez, and S. J. Bonasera, "Smart watch rssi localization and refinement for behavioral classification using laser-slam for mapping and fingerprinting," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 2173–2176, IEEE, 2014.