South by South-East or sitting at the desk. Can orientation be a place?

Ulf Blanke
MPI Informatics, Saarbrücken
blanke@mpii.de

Robert Rehner
TU Darmstadt
rehner@dvs.tu-darmstadt.de

Bernt Schiele
MPI Informatics, Saarbrücken
schiele@mpii.de

Abstract

Location is a key information for context-aware systems. While coarse-grained indoor location estimates may be obtained quite easily (e.g. based on WiFi or GSM), finer-grained estimates typically require additional infrastructure (e.g. ultrasound). This work explores an approach to estimate significant places, e.g., at the fridge, with no additional setup or infrastructure. We use a pocket-based inertial measurement sensor, which can be found in many recent phones. We analyze how the spatial layout such as geographic orientation of buildings, arrangement and type of furniture can serve as the basis to estimate typical places in a daily scenario. Initial experiments reveal that our approach can detect fine-grained locations without relying on any infrastructure or additional devices.

1. Introduction

Location has been identified as a key component for context-aware systems [3] and may enable a variety of applications [7, 13]. WiFi fingerprinting might be used to obtain a coarse-grained location estimate (10's of meters). Finer-grained estimates (10's of cm) can be obtained, e.g. by ultrasound [10], but require additional and expensive sensors. On the other hand *dead reckoning* does not require any infrastructure. Observing walking direction and speed through inertial measurement, the relative displacement from a starting point can be estimated. With special sensor placement at the foot, the error can be reduced to less than 1% of the distance travelled [4].

Using more convenient sensor placements such as the pocket or the belt others estimate the walking direction [5] for dead reckoning [11] and detect re-visited location coordinates [6]. However, the inherent limitation of these approaches is the accumulation of errors over time. This typically results in large estimation errors in scenarios without absolute location updates as it is often the case when the user stays within the same environment (such as his apartment) for several hours. Rather than to estimate the user's location incrementally we, therefore, explore the use of absolute orientation as well as orientation traces to determine the location of the user within a certain environment.

In contrast to previous work aiming for geometric or *topographic* localization of a user [6, 11], we take a *topological* view on location. We are not interested in precise estimation of users' coordinates, but rather in detecting meaningful places that are routinely visited.

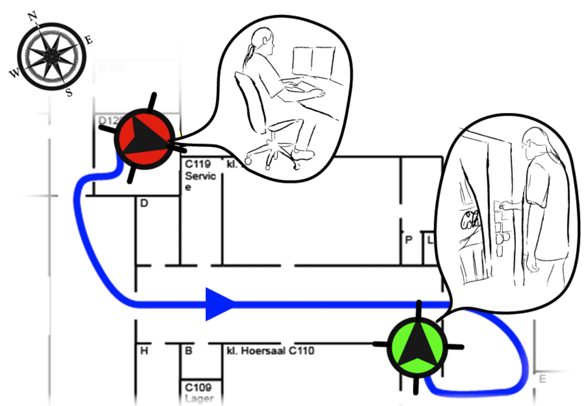

Figure 1. Example of two connected places.

The environmental layout predefines potential places that we can reside in. The layout of buildings, the arrangement and type of furniture or significant landmarks (such as billboards) structure the space in which we are active. The facing direction of a person as characteristic for location has been mentioned [1], but not analyzed before. For example, when a person is sitting at her working desk she typically faces a certain *orientation* (red mark in Fig. 1). Likewise transiting between different places is characterized by a specific *orientation trace* [2] as our movements are constrained by walls, doors, stairs, etc. The blue arrow in Fig. 1 illustrates a place transition characterized by a sequence of orientation changes. The user first turns to the right at the door, passes the door and turns left, followed by several other turns until he reaches the vending machine. Using human orientation seems intuitively promising for place detection and thus location estimation.

In this work we present a method to detect places from a continuous stream of data with no additional requirements for current-generation phones. Placing an inertial sensor into the pocket, data is collected, segmented and classified. We collected data from two typical scenarios: *at home* and *at work*. These include typical everyday places and tran-

1550-4816/11 $26.00 © 2011 IEEE
DOI 10.1109/ISWC.2011.18

sitions between such places. We show that orientation can indeed be used to discriminate between places. Combining orientation and orientation traces when changing places improves place detection.

2. Detecting significant places

The space which we navigate through is naturally constrained by our environment. We propose a place recognition system that is based on two main characteristics: *(I) static orientation* while at a place and *(II) orientation traces* while transiting between places. Based on these characteristics we formulate the following research question:

- Can environmental constraints be used for detecting significant *places* of the user?

Our approach consists of four steps. First, we segment the continuous data stream into *static segments*, assuming the user resides (sitting or standing) at a place P_i and into *in-transit segments*, assuming the user changes his place P_i (by walking or using stairs) to another place P_j. In a second step we train a classifier based on the orientation of the user in order to classify static segments into a known set of places. In third step, we classify in-transit segments based on the orientation trace from known transitions. In a last step, we finally combine the detectors for static places and place transitions to estimate significant places jointly.

Segmentation. According to Mathie et al. [8] we assume that the variance of the acceleration over a window is a suited cue for estimating walking motion. We evaluated different thresholds and set the threshold in favor of gaining an over-segmentation. Note, that an over-segmentation can be handled easier than merged segments. Such interruptions can also occur naturally, for example by having a short chat with a colleague on the corridor. To handle the over-segmentation, we do a pairwise combination of all segments similar to Zinnen et al. [14] and obtain overlapping potential in-transit segments (see Fig 2). In the later transition classification task, we perform non-maximum-supression by choosing the segment with highest score.

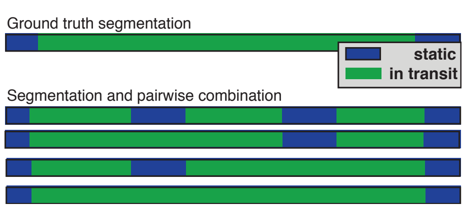

Figure 2. Top: ground truth segments. Bottom: obtained segmentation by thresholding the acceleration variance and pairwise combined segments.

Place detection using static orientation. Given *static* segments and global 3D-orientation we train a naïve Bayes classifier. Given place labels, mean and variance are calculated for each orientation entry to fit a Gaussian model $P(x) = \mathcal{N}(x, \mu, \sigma)$. The Gaussians are then combined in naïve Bayes fashion. Samples of unknown static segments can then be classified into the place with maximum likelihood.

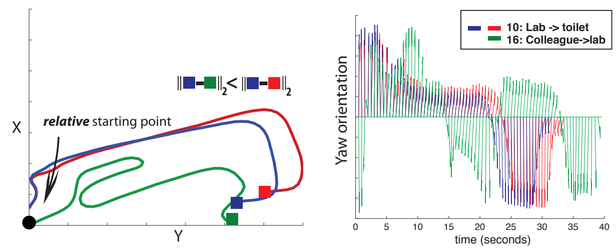

Figure 3. Example of two instances of equal transitions (blue and red) and a third different transition (green). Left: accumulated displacement estimate. Right: orientation trace as feather plot.

Detecting place transitions. Given *in-transit* segments, a sequence of orientations (or turns) is given. Previous work accumulates this sequence and estimates a transition's displacement to a relative starting point [6]. Given a set of known transitions, a nearest neighbor classifier on the displacement's coordinates is used to assign the closest known transition. However, we observed that using displacement coordinates only can be error prone as illustrated in Fig. 3. Given a relative starting point, the transition *(colleague→lab)* results in a lower distance to the transition *(lab→toilet)* than a second instance of *(lab→toilet)*.

In contrast to [6] we therefore use the complete orientation trace. Then similarity between instances of *(lab→toilet)* and dissimilarity *(colleague→lab)* becomes more evident (Fig. 3, right) than using displacement coordinates only.

Changes in walking speeds or path variation, e.g., by collision avoidance with other persons or opening doors do not allow place transitions to be exactly reproduced. Dynamic Time Warping (DTW) [9] offers an excellent similarity measurement that accounts for such variability in the signal. For the warping function we use a $p = 0$ symmetric slope constraint. Given the similarity measurement of the DTW we use kNN-classification to assign an unknown place transition to known transition classes.

Combined place detection. Given topological information in terms of connected places, we can use information of originating places, transitions and destination places jointly. We implement a simple voting-algorithm. First we replace the kNN-assignment of the place transitions with a soft assignment to obtain scores. We normalize the DTW distances using a sigmoid function $y = \frac{1}{1+e^{a\cdot(x+b)}}$, where a and b are estimated from training data. We use a linear weighting $P(place) = \alpha \cdot score(transition) + (1 - \alpha) \cdot score(place)$ as voting technique. That means, e.g., if $\alpha = 1$ the place detection would be ignored completely and only the transition detection would be used.

3. Datasets

In order to analyze whether environmental constraints provide relevant characteristics for detecting significant places, we recorded two datasets in typical environments: at work and at home. The dataset was recorded on three non-consecutive days, where each place was visited at least once. In total 20 places with 37 place transitions were recorded, which corresponds to 4.5-5km of travelled path.

In the office. During a day an office worker visits typical places such as *colleagues*, *the lab* or *the cafeteria*. In total we used 10 such places. Theoretically 90 transitions between these places are possible, but we limited our dataset to 24 typical transitions. Fig. 4 depicts the places and transitions.

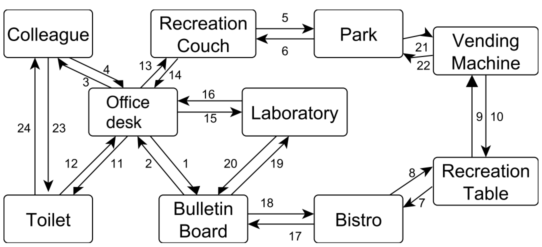

Figure 4. Places and place transitions at the office.

At home. As a second dataset, we recorded a typical evening situation which constitutes 13 transitions between 10 places. Examples of places include being *at the fridge*, *on the couch* or *at the dining table*. In contrast to the office dataset (minimum transition time of 17s), place transitions can be much shorter up to just 2s of total duration, corresponding to about 1 to 1.5m of travelled distance.

Both datasets where recorded continuously and annotated manually. Additionally we recorded a third dataset containing irrelevant places outside of the home and the office, for instance, when walking around in other buildings or downtown in a shopping area.

3.1 Hardware

For the data collection, we use the Xsens MTx inertial measurement unit[1]. Based on fusing 3D-accelerometer, gyroscope and magnetometer data, the global orientation of the sensor is estimated. The sampling rate is set to 120Hz. We equipped the sensor with a bluetooth module and a battery, both worn in the right pocket. A belt-attached smartphone was used to collect the data via bluetooth connection. In this work we assume equal orientation of the sensor. This marks a limitation to the system, but in combination with recalibration methods from, e.g., [2, 5] we expect to relax this assumption in future work.

4. Experimental results

We present our analysis in three different steps. First, we visualize the data for static places. Then, we report on a discriminative analysis of place transitions using the segmentation from groundtruth. Finally, we report on the performance of our approach on continuous data.

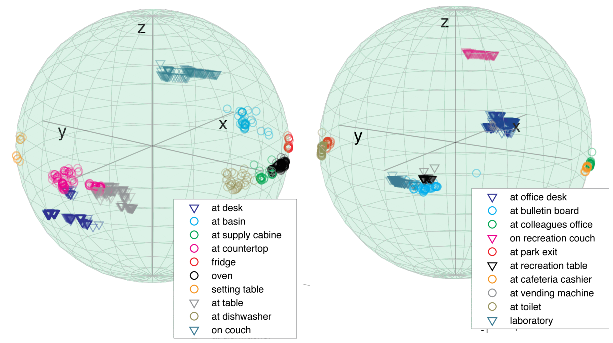

Figure 5. Examples of global facing direction in the office (right) and at home (left). Triangles symbolize sitting and circles standing posture.

Visual inspection of *places*. Fig. 5 illustrates the facing direction on a 3D sphere, i.e., the vector originating in the pocket and pointing to the front. Each color specifies a significant place. Note, that we aggregated 3D-orientation to the facing direction for sitting and standing as triangles and circles respectively in this visualization only. Since the posture is reflected in the data we make use of the full 3D-orientation for the later classification task.

For the home dataset (left), most places separate well, except *at oven* and *at supply cabinet*. These places face the same direction and cannot be disambiguated by orientation. For the office dataset we can observe stronger overlaps between places. Here *at park exit* and *at toilet* have a very similar orientation, as well as *at colleague* or *at cafeteria*. Interestingly sitting *on couch* at home or in the office has a slightly different sitting angle than sitting *at the desk*. This is a useful characteristic to disambiguate both places.

Discriminative analysis of *place transitions*. Fig. 6 shows the DTW-measured similarity between each transition. The majority can be distinguished well. We can also see intra-class similarity between class 1 (*office→bulletin board*), 16 (*lab→office*) and 20 (*lab→bulletin board*). Here the walking paths contain equal sequences of turns. Despite of different path lengths, the orientation traces are warped in length, leading to confusion for paths with similar sequences of turns. A more sophisticated warping function penalizing different trace lengths could distinguish similar sequences of different lengths. While home and office transitions can be distinguished using additional context such as WiFi localization it is worthwhile to note that transitions within the office do not resemble the transitions at home.

Continuous detection. Fig. 7 shows the results for continuous detection on both datasets. Using $120Hz$ we obtain a mean class precision of about 70% for the office place detection. Places with similar orientation are confused. *At*

[1]http://www.xsens.com

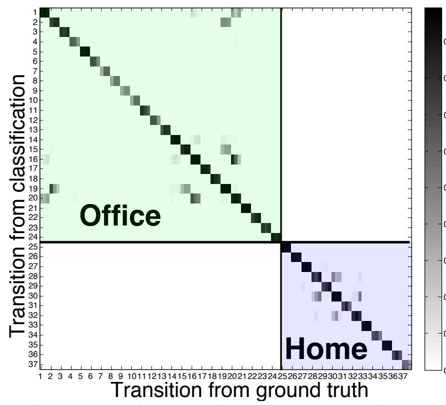

Figure 6. Normalized soft assignment of place transitions.

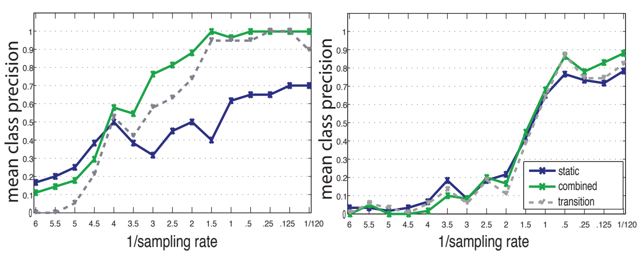

Figure 7. Mean class accuracy for place, transition, and joint place detection.

toilet is frequently confused with at *park exit*, *at cafetaria* or *at the colleague*. Adding transition information the results are improved 30% to perfect recognition. As expected transitions help significantly due to their richer information and less ambiguity compared to places alone.

At home, we observe nearly 80% for the place detection. At home, the arrangement of furniture is less uniform than in the office buildings. The improvement of the joint detection is lower with about 10%. Here, transitions are shorter than in the office, and do not contain complex enough information. This is reflected in the reduction of the sampling rate. While for the office scenario performance remains relatively stable down to $0.5Hz$, performance drops significantly with less than $1Hz$ for the home dataset.

4.1 Discussion

We saw that orientation can be an interesting cue for location. The ambiguity between places seems to be less than expected. Here, the interplay between different conditions plays a key role. While sitting in an ambiguous orientation due to uniform building layouts, the properties of furniture can help to disambiguate. For example the sitting height on a couch or an offset of a few degrees for the yaw axis, can lead to different sitting angles (and therefore 3D orientation) than in an office chair.

Detecting places in the office is slightly worse than for the home scenario. Probably the uniform building layout and furniture arrangement make orientation more ambiguous. Adding transition information to recognize places yields a significant improvement. The results for the home scenario seem almost complementary. While the place detection is better overall, transitions do not contain sufficient information to yield a significant improvement. Transitions are shorter and less complex than in the office scenario.

5. Conclusion and future work

We present a method for detecting significant places in a continuous data stream. Static orientation in places and absolute orientation traces between places can serve as characteristics for detecting places. Our method does not require any special hardware and can be applied using recent phones such as an iPhone 4 worn in the pocket.

Given the promising results we are planning a long-term multi-user dataset to allow extraction of significant places using clustering and motif discovery techniques. In another direction we will explore map matching techniques similar to work in [12]. The topology can literally be drawn on the map and reduce or even replace training of the classifier.

Acknowledgements. This work is funded by the DFG project *"Methods for Activity Spotting With On-Body Sensors"*.

References

[1] M. Azizyan, I. Constandache, and R. Roy Choudhury. Surroundsense: mobile phone localization via ambience fingerprinting. In *MobiCom09*.

[2] U. Blanke and B. Schiele. Sensing location in the pocket. In *Adjunct Proceedings of Ubicomp08*.

[3] A. Dey. Understanding and using context. *Personal and ubiquitous computing*, 5(1):4–7, 2001.

[4] E. Foxlin. Pedestrian tracking with shoe-mounted inertial sensors. *Computer Graphics and Applications 05*.

[5] K. Kunze, P. Lukowicz, K. Partridge, and B. Begole. Which way am I facing: Inferring horizontal device orientation from an accelerometer signal. In *ISWC09*.

[6] S. Lee and K. Mase. Activity and location recognition using wearable sensors. *Pervasive Computing*, 1(3):24–32, 2002.

[7] L. Liao, D. Fox, and H. Kautz. Location-based activity recognition. In *NIPS05*.

[8] M. Mathie, A. Coster, N. Lovell, and B. Celler. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, 25:R1, 2004.

[9] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43–49, 1978.

[10] P. Steggles and S. Gschwind. The Ubisense smart space platform. In *Adjunct Proceedings Pervasive05*.

[11] U. Steinhoff and B. Schiele. Dead reckoning from the pocket - an experimental study. In *PerCom10*.

[12] O. Woodman and R. Harle. Pedestrian localisation for indoor environments. In *UbiComp08*.

[13] Q. Yang. Activity recognition: linking low-level sensors to high-level intelligence. In *ICCAI09*.

[14] A. Zinnen, C. Wojek, and B. Schiele. Multi activity recognition based on bodymodel-derived primitives. In *LoCA09*.