2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Method of Improving WiFi SLAM based on Spatial and Temporal Coherence

Shao-Wen Yang, Sharon Xue Yang, and Lei Yang
Intel Labs, Intel Corporation
Email: {shao-wen.yang, xue.yang, lei.t.yang}@intel.com

Abstract—**The paper addresses the revisiting (loop closing) problem of simultaneous localization and mapping (SLAM) by investigating spatio-temporal coherence in inertial and perceptual inputs to improve the robustness and convergence of SLAM. The basic idea is to find out coherent subsequences of confidence in trajectory to ensure against error-prone correspondences. It is achieved by leveraging fuzzy matching based on local trajectory structure and measurement similarity. Our approach does not rely on any global features or propagation modeling, which can be unreliable in the presence of gross errors and result in divergence. Apart from WiFi SLAM, our approach can also be capable of improving generic SLAM problems by leveraging spatio-temporal coherence. The experiments show that our approach can significantly reduce the ambiguity in WiFi fingerprinting, and subsequently lead to performance improvement in terms of mapping and localization.**

I. Introduction

Deriving accurate indoor location has been an increasingly important topic in the midst of mobile application blooming [2, 6, 16, 1, 4]. Despite a decade of research and development efforts, enabling indoor location capability remains a challenging topic. There is no single location system on the market now that can provide sufficiently accurate indoor location and has been widely deployed. The main limitations of existing technologies are either in their performance inadequacy or their requirements for substantial infrastructure deployment efforts.

WiFi based indoor location technologies particularly attracted a lot of attention in the past decade given the wide deployment of wireless local area network (WLAN) infrastructure. There are three primary methods for location determination using WLAN. One method is based on propagation models, using estimated degradation of signal strength over distance in space from the known location of access points and transmit power–such as SkyHook; the second method relies on storing pre-recorded calibration WiFi measurement data (WiFi fingerprinting) in order to generate an RF map of a building such as Ekahau and Qubulus; the third method utilizes the measurement of WiFi radio wave time of flight to measure distance–time-of-arrival (ToA).

Among the three approaches, the second method relying on *WiFi fingerprinting* is most promising in providing accurate indoor location–it has been shown that it can achieve meter level accuracy as long as the pre-calibrated database has sufficiently dense calibration points (i.e., meter grids). In comparison, our test results show that indoor location derived from SkyHook has median error of 11m and mean error of 25m. ToA type of approach can reach 3-5 meter accuracy constrained by indoor multipath environment. On the other hand, existing implementations of WiFi fingerprinting (such as Ekahau or Qubulus) require initial training to correlate each location with corresponding WiFi fingerprints (a set of measured RSSIs), which leads to substantial efforts to deploy such location systems, especially when the calibration points are dense. Also, the calibration often needs to be re-performed as deployment environment changes. The substantial deployment effort is one of the primary reasons why WiFi fingerprinting based location system has not been widely adopted, despite the technology itself has been proposed for almost a decade.

II. Background

A. Motivation

To reduce the deployment effort of WiFi fingerprinting while maintaining good location accuracy, our research adopts the simultaneous localization and mapping (SLAM) framework to integrate the inertial navigation system (INS) with WiFi, with two objectives: a) The WiFi fingerprint map for the indoor environment can be automatically generated at fine granularity without human intervention. Thus, the human deployment efforts for such location system can be minimized. b) Through global optimization of GraphSLAM, we aim to achieve meter level location accuracy.

Our approach improves the robustness of existing SLAM systems to correspond identical locations while still avoiding the use of signal propagation modeling. It is motivated by the fact that fingerprinting can be unreliable without a dense location database, whereas signal strength based approaches can rely on propagation modeling or signal strength interpolatability. Instead of introducing signal strength constraints that can be unreliable due to the presence of attenuation and fading, our approach confines error propagation by reliably identifying revisiting with spatio-temporal coherence. This is a key prerequisite to make location based services independent of surrounding environment, and reduce as much human intervention as possible.

B. Related Work

SLAM [12, 9, 5] is a well-researched area in mobile robotics, which solves the problem of constructing a map for unknown environment and, at the same time, determining the location of a user with respect to the map. One of the key

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

problems in SLAM is *loop closing*, i.e., the ability to detect when a user is revisiting some area that has been mapped earlier.

Conventional SLAM schemes detect loop closure mostly based on visual sensors Newman et al. [11], Milford [10], which provide the perception of the physical environment and have rich features to identify the revisited locations. In WiFi based SLAM, however, there is no such visual sensor available; loop closing detection becomes a major challenge given the unstable nature of WiFi signal strengths. Recently, there have been several WiFi based SLAM implementations, such as Ferris et al. [7] and Huang et al. [8]. They both depend on WiFi signal strength alone to infer the revisited locations, either based on similar WiFi signal strengths, or based on physical distance constraints derived from the assumed WiFi signal propagation model. To differentiate, our approach lies in the same category with Ferris et al. [7], which is fingerprinting only and does not rely on any propagation modeling techniques. On the contrary, Huang et al. [8] used a signal strength only technique, assuming interpolatability of WiFi RSSIs within a spatial neighborhood.

Instead of relying on unreliable WiFi signal strength alone for loop closure detection, in this paper, we propose a method to detect revisited locations (note that we often use loop closing and revisit interchangeably) for SLAM by analyzing spatial and temporal coherence in inertial and WiFi inputs. The rationale is that if it is a true loop closing, the local spatial and temporal structure between two identified revisited locations would be similar. Our approach improves the robustness and location accuracy of existing WiFi based SLAM systems, as we will show theoretically and empirically.

III. Problem Formulation

SLAM is the process to build up a map for unknown environment and, at the same time, determine the location of a user with respect to the map, without prior knowledge. WiFi SLAM provides a solution to crowd-sourcing, site survey or wardriving with as less prior knowledge and human intervention as possible. The input of WiFi SLAM includes inertial measurements and WiFi signal strengths. The problem is challenging because of the poor stability and reliability of consumer-grade inertial sensors and WiFi receivers. The gross errors from inertial sensors can be far beyond meters as errors are accumulated from time to time. The ambiguity in WiFi fingerprinting cannot provide a straightforward means to correct the error reliably.

Fig. 1 illustrates the dynamic Bayesian network (DBN) for WiFi SLAM. To ensure global consistency of the posterior against divergence from gross errors, correctly inferring revisiting constraints are critical, whereas erroneous ones can result in divergence. A bad constraint can easily make the outcome far from the solution. Specifically, for sampling-based approach like Rao-Blackwellized particle filtering (RBPF), it can fail due to the limited number of samples; for least squares based approach like graph-based SLAM (GraphSLAM), it can fail due to its least squares nature.

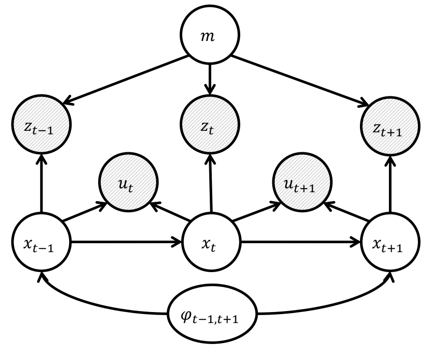


Fig. 1: DBN for SLAM. Each of the links represents a causal relationship. Solid node represents latent variables (unobserved, to be estimated) and shaded nodes are observable. x_t is the users location at time step t, m is the (signal strength) map, z_t is the WiFi signal strengths obtained at time step t, u_t is the inertial input at time step t, $\psi_{t,t'}$ denotes a revisiting constraint between locations at time steps t and t′.

Our motivation is to reliably extract revisiting constraints from finding out the coherent subsequences in the entire user trajectory $\{x_1, x_2, \cdots, x_T\}$. A fuzzy matching technique is used, instead of exact match, to take into account the various signal strength noises due to the presence of attenuation, fading and changes of the environment. It also does not rely on any global features, like absolute coordinates, which can be unreliable as the trajectory drifting can be severe. It turns out that our approach can be capable of re-calibrating the SLAM posterior even if it has diverged.

IV. Spatial and Temporal Coherence

The loop closing detection is challenging in WiFi based SLAM systems due to lack of features in determining the revisiting area. Without visual perception sensors, WiFi radio is the only source of measurements associated with the physical environment. However, WiFi signal is prone to large temporal and spatial variations, and it often fails in serving as precise *loop closing* constraints. The same WiFi measurements do not necessarily map to the same physical location and vice versa. Incorrectly identified *revisits* can lead to divergence of SLAM optimization problem, extremely harmful for SLAM performance. In this paper, in addition to WiFi signal constraints, we explore the spatial and temporal coherence of location estimates to improve the precision of loop closing detection on top of exploring WiFi measurement similarity.

Sequence alignment is a technique commonly used in bioinformatics for determining similar regions between two nucleotide or protein sequences. Many bioinformatics tasks de facto rely on successful alignments. In order to apply sequence alignment techniques to SLAM, we firstly define similarity (coherence) metrics in space and time. Once the metrics are well defined, we can perform sequence alignment on SLAM posterior to extract coherent subsequences reliably.

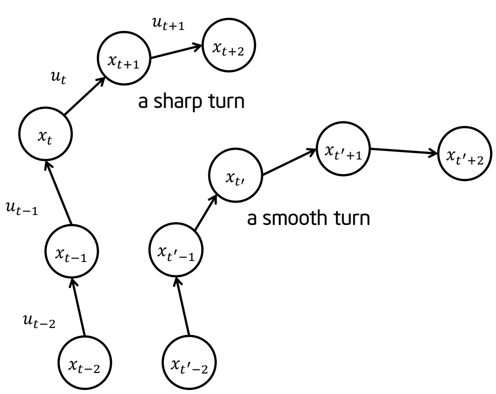


Fig. 2: Spatial coherence. u_t represents the motion from x_t to x_{t+1}, u_{t+1} represents the motion from x_{t+1} to x_{t+2}, and etc.

A. Fingerprint Similarity

WiFi fingerprinting is a coarse indicator to capture the similarity between two (almost) identical locations by quantifying the similarity in WiFi signal measurements. In our implementation, we use Euclidean norm to characterize the dissimilarity between two WiFi RSSI measurements, though other values should apply as well, which can be expressed as:

$$\rho(z_t, z_t') \quad = \quad ||z_t - z_t'||_p \tag{1}$$

The underlying assumption is that same WiFi measurements correspond to same physical locations. In reality, however, the assumption usually does not hold. Due to small scale fading (multipath propagation), two very close locations can come with significant different RSSI measurements. It turns out that locations with the same WiFi RSSI measurements are not necessarily at the same location. In our proposed solution, WiFi fingerprinting only serves as a first level coarse indicator, and does not rely on propagatioin modeling (or sensor modeling) that is hardly reliable due to unknown structure of environment and device diversity in access points. To deal with the ambiguity in WiFi RSSI measurements, we further incorporate a spatial coherence metric to capture the similarity in trajectory (sequence of locations).

B. Spatial Coherence

Spatial coherence aims at distinguishing trajectories with different local structures. A trajectory in indoor environment can be highly constrained, particularly in office-like environments primarily comprising rooms and hallways. Thus, similarity in both the spatial structure and the WiFi RSSI measurements can give us a high confidence in that a true revisiting has happened.

We define angular features, as illustrated in Fig. 2, to capture the local trajectory change that reflects the environmental structure. To deal with *disorder* and *non-atomicity* of inertial measurements, the feature is defined at a multiple time step scale manner. Disorder arises from the inaccurate detection of motions and can result in time step drift, particularly while making turns. Non-atomicity relates to the varying human motions and the smoothing nature of inertial measurements. Representing the feature at multiple time step scales makes it possible to capture changes in trajectory at a higher level of abstraction. Our current implementation uses a time step window size of 5. As illustrated in Fig. 2, the spatial coherence can be defined given the window size, and used to characterize the similarity between two subsequences. We define a scale-invariant feature, a vector consisting of the angular deviations at multiple time step scales which can be expressed as:

$$f_t \quad = \quad \begin{bmatrix} \theta(u_t) - \theta(u_{t-1}) \\ \theta(u_t) - \theta(u_{t-2}) \\ \theta(u_{t+1}) - \theta(u_{t-1}) \\ \theta(u_{t+1}) - \theta(u_{t-2}) \end{bmatrix} \tag{2}$$

where $\theta(u_t)$ denotes the angular component of u_t.

The spatial dissimilarity between pairs of locations can be captured based on the spatial coherence feature defined above using various metric. One example metric that our current implementation uses is the mean difference between spatial coherence features. The spatial coherence can serve as a strong negative indicator to disambiguate the true revisiting from the false revisiting identified by unreliable WiFi signal measurements.

$$\delta(f_t, f_{t'}) \quad = \quad \text{mean}(|f_t - f_{t'}|) \tag{3}$$

The feature defining on a time window of arbitrary size can also apply, depending on the noise level of sensor. It turns out that the more accurate the inertial input, the larger time window we can apply to capture the spatial locality. The spatial coherence can serve as a strong negative indicator to disambiguate the true revisiting from the false counterparts identified by unreliable WiFi RSSI measurements.

TABLE I: A local alignment example

Input sequences	
Sequence #1	FTFTALILLAVAV
Sequence #2	FTALLLAAV
Local alignment result #1	
Sequence #1	FTALILLAVAV
Sequence #2	FTAL-LLA-AV
Local alignment result #2	
Sequence #1	LILLAV
Sequence #2	LLLAAV
Local alignment result #3	
Sequence #1	FT
Sequence #2	FT

C. Temporal Coherence

Sequence alignment is a fuzzy matching technique working with matches that may be not perfect when finding correspondences. For example, protein or DNA matches may not be exact; sentence matches can be imperfect due to verb tenses of plurals. Spatial structure and WiFi fingerprint matches are also imperfect due to the presence of noise. Temporal coherence captures the local structure similarity in consecutive time steps

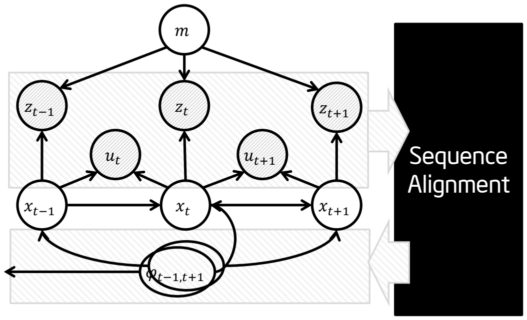


Fig. 3: DBN for stSLAM

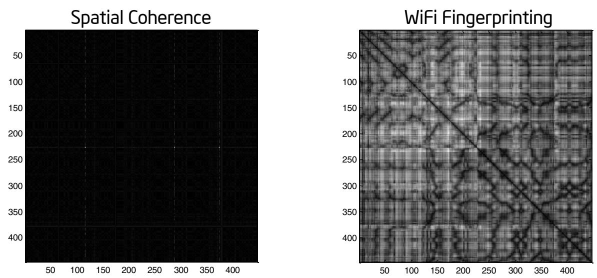


Fig. 4: Similarity matrices of WiFi fingerprinting and spatial coherence

based on WiFi fingerprinting and spatial coherence. That is, due to the highly constrained indoor physical environment, if *revisiting* has occurred at one physical location at a particular time step, it is likely that several time steps following that will be *revisited locations* as well. We apply sequence alignment techniques to extract coherent subsequences. In particular, local alignment extracts matching sequence of all possible lengths. One example of sequence alignment is illustrated in Tab. I for ease of understanding, where the similarity between pairs of entries (letters) is simply defined by equality. For our proposed SLAM system, however, each of the entries is a compound of a spatial location and a set of WiFi RSSI measurements, and the similarity between pairs is defined jointly by WiFi similarity and spatial coherence.

A threshold on matched sequence length is used as a temporal coherence measure to gauge matching confidence level. The longer the matched sequence, the more possible it is a positive match. Note that the spatial constraint is defined in a way that is independent of motion direction. We perform sequence alignment of a sequence against itself and its reverse to identify revisiting from subsequences moving in same and opposite directions. This length constraint is a tradeoff between precision and recall (i.e., detection rate), and can depend on the noise level of inertial and WiFi RSSI measurements. In our present implementation, the matched sequence length threshold is 5 time steps.

Our approach is based on Smith-Waterman local alignment algorithm [13], which takes as input the (logical) conjunction of WiFi fingerprinting and spatial coherence dissimilarities, and output the revisiting sequences of high confidence. The sequence alignment algorithm combines all of the sensory information and generates SLAM constraints using matched sequences, instead of fragmented, unreliable ones. The SLAM algorithm then incorporates all the constraints, either from sensor data or sequence alignment, to obtain a globally consistent solution.

In our algorithm, we also assume location estimates within a pre-defined time (step) window are dissimilar and search for similar sequence only outside of this window. This search window depends on moving speed of the user and the scale of the physical environment. Fig. 3 shows the DBN for spatio-temporally coherent WiFi SLAM.

V. Experimental Results

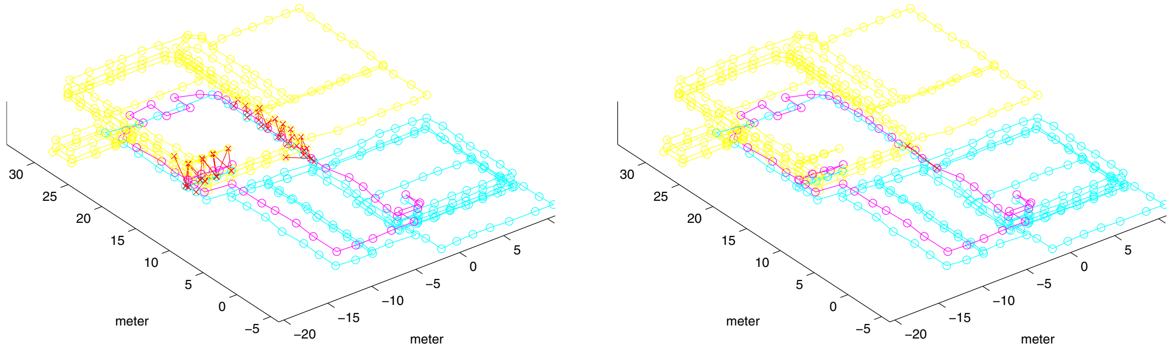
We used the GraphSLAM algorithm [15] integrating inertial sensor measurements with WiFi measurements, and compare the performance with or without the proposed spatio-temporal coherence loop closing detection algorithm. The thresholds for WiFi fingerprinting and spatial coherence are 10dbm and $60°$, respectively, which are determined empirically. For local alignment of temporal coherence, the three parameters (weights), `match`, `mismatch` and `indel`, used for determining matches are $+1$, -2 and -2, respectively. We tested our algorithm against two data sets: real WiFi data with simulated inertial data, and real inertial and WiFi data. For inertial and WiFi measurements, we used the STMicro iNEMO v2 evaluation board STEVAL-MKI062V2, and an ASUS netbook equipped with an Intel 5300 WiFi card, respectively.

A. Simulated Data Set

In this experiment, a data set of 650m travel distance containing 500 distinct measurements is used. It is collected in an office environment with ground truth locations manually annotated. In this data set, we can achieve a 96% precision (number of true positive samples out of all positive samples). In comparison, without incorporating spatio-temporal coherence, the precision is around 30%, hardly useful for realistic applications.

Fig. 4 illustrates the similarity matrices for spatial and temporal coherences using data from an office environment. The darker the entry, the higher the similarity. As can be seen from WiFi fingerprinting, we can visually identify a number of matched sequences in diagonals or reverse diagonals. On the other hand, the similarity matrix for spatial coherence comes with a quite regular grid pattern, as most of the time the travel directions are straight in an office environment. Fig. 5 shows the identified revisiting from spatio-temporal coherence. In this example, there are 36 out of 37 true revisits–a 97% accuracy. It is in particular important for GraphSLAM due to its least squares nature in which a single vulnerable revisit can result in divergence.

Fig. 6 depicts the relationship between WiFi fingerprinting threshold and the performance of our approach, including pre-



(a) True revisits (b) False revisits

Fig. 5: Revisits from spatio-temporal coherence. The graph representation of vertices in circles and edges in lines are the input trajectories of three subsets in difference colors–cyan, magenta, and yellow. Lines in red connect revisited locations from our approach. A slight deviation in z-axis is only used for better visualization.

cision, recall, and F-score. This demonstrates a nice property of our approach that WiFi fingerprinting threshold serves as a tradeoff between precision and recall. As a matter of fact that false revisiting constraints can probably lead to divergence, we chose 10dbm, a value of around 3-percentile. As can be seen, our approach can achieve a near 100% precision to ensure against divergence without the use of spatial locations in absolute coordinates. Depending on the noise level of the inertial sensors and WiFi measurements, a even looser threshold for fingerprinting will also apply.

B. Pedometric Data Set

For the data set with both real inertial and WiFi RSSI data, we manually marked 29 locations with a fixed step length of approximately 50cm in a room. A user carried the initial sensors board and the netbook, and walk around the room following the marked locations three times counterclockwisely. We implemented a pedometer algorithm based on inertial sensor measurements to estimate the motion trajectory of the user.

Fig. 7 gives the estimated trajectory from using INS dead reckoning alone, from using SLAM without spatio-temporal coherence and from using SLAM with our proposed spatio-temporal coherence, where the user has walked over the marked trajectory for three repeated loops in the experiment. As can be seen in Fig. 7a, the estimated trajectory from INS deviates from the true trajectory and three repeated loops drift away from each other in the estimated trajectory. Conventional SLAM, depicted in Fig. 7b, improves the performance to some extent, but three loops still largely deviate from each other. Our approach, shown in Fig. 7c correctly identifies most revisited locations and three loops converge well.

The location estimation error in relevance to the ground truth is shown in Tab. II. The absolute errors are calculated from aligning the estimated trajectory with the ground truth trajectory for which minimum least squares errors are obtained, so that the metric can be well defined independent of coordinate systems. Apart from the absolute metric, we also quantify the performance with the *subjective-objective* technique [3] representing the accuracy of revisiting detection. The metric can characterize the local structure of trajectories independent of coordinate systems. The threshold defining spatial neighborhood for locations is 0.5065m, the average step length. Note that all the output trajectories are scaled so that the total distances traveled are the same with the ground truth [8]. As we can see, our proposed SLAM leads to smaller location error in both RMS and mean, and much

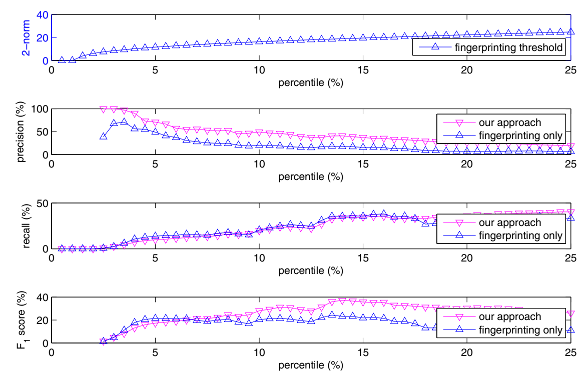
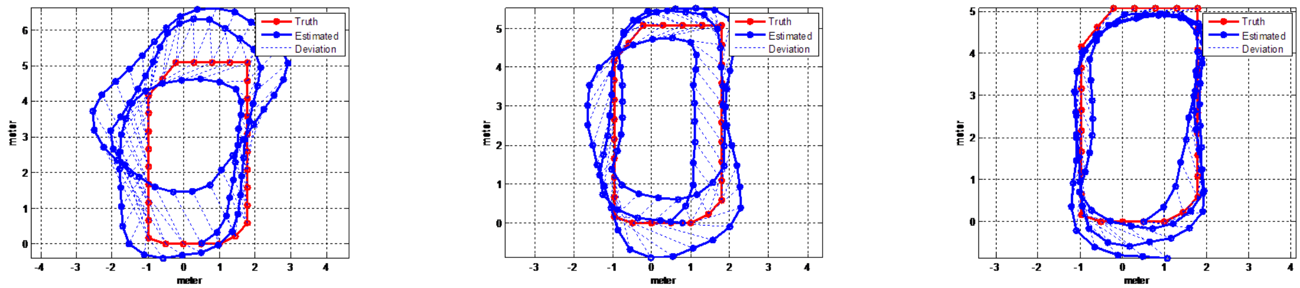


Fig. 6: Performance quantification

TABLE II: Quantitative comparison

Error (m)		INS only	SLAM	Our approach
Absolute	RMS	1.5122	1.1473	0.9181
	Mean	1.3684	1.0217	0.8695
	Std.	0.6472	0.5294	0.2962
Relative	RMS	0.9362	0.4534	0.1378
	Mean	0.7666	0.3999	0.1014
	Std.	0.5404	0.2150	0.0939



(a) INS only (b) SLAM (c) Our approach

Fig. 7: Performance comparison. The solid lines are the trajectories, and the dashed lines indicate correspondences between estimated and true locations.

smaller standard deviation.

VI. Conclusion

We have introduced the novel use of spatio-temporal coherence to improve WiFi SLAM. This is achieved by leveraging fuzzy matching techniques to associate spatio-temporally coherent subsequences of the SLAM trajectory. WiFi fingerprinting is a coarse indicator of revisiting, and the spatial coherence serves as a strong negative indicator to disambiguate between genuine and vulnerable revisits. Temporal coherence is a means for both WiFi fingerprinting and spatial coherence to work together. Our approach is particularly useful in indoor environments, where trajectories can often reflect the structure of the physical environment. While outdoor positioning can generally be achieved using GPS, indoor location awareness remains challenging, requiring significant infrastructure deployment and extensive wardriving. Robust WiFi SLAM is a key prerequisite to reduce as much the human intervention as possible to automate the creation and update of location signature database. The ample experiments show that our approach can lead to significant performance improvement over conventional SLAM techniques. The future work includes using robust SLAM algorithms Sunderhauf and Protzel [14] to improve the robustness and, as a result, precision and recall.

References

[1] M. Azizyan, I. Constandache, and R. R. Choudhury. SurroundSense: mobile phone localization via ambience fingerprinting. In *Proc. Annu. Int. Conf. Mobile Computing and Networking*, 2009.

[2] P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *Proc. Annu. IEEE Int. Conf. Computer Communications*, 2000.

[3] M. Bowling, D. Wilkinson, A. Ghodsi, and A. Milstein. Subjective localization with action respecting embedding. In *Robotics Research: Results of the 12th Int. Symp.*, pages 190–202, 2005.

[4] K. K. Chintalapudi, A. P. Iyer, and V. Padmanabhan. Indoor Localization Without the Pain. In *Proc. Annu. Int. Conf. Mobile Computing and Networking*, 2009.

[5] H. Durrant-Whyte and T. Bailey. Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. *Robot. Autom. Magazine*, 13:99–110, 2006.

[6] A. L. et al. WiFi-SLAM using Gaussian process latent variable models. In *Proc. 3rd Int. Conf. Pervasive Computing*, pages 116–133, 2005.

[7] B. D. Ferris, D. Fox, and N. Lawrence. WiFi-SLAM using Gaussian process latent variable models. In *Proc. Int. Joint Conf. Artificial Intelligence*, pages 2480–2485, 2007.

[8] J. Huang, D. Millman, M. Quigley, D. Stavens, S. Thrun, and A. Aggarwal. Efficient , Generalized Indoor WiFi GraphSLAM. In *Proc. IEEE Int. Conf. Robot. Autom.*, 2011.

[9] J. J. Leonard and H. F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Trans. Robot. Autom.*, 7(3):376–382, 1991.

[10] M. Milford. Visual route recognition with a handful of bits. In *Proc. of Robotics Science and Systems*, 2012.

[11] P. Newman, D. Cole, and K. Ho. Outdoor slam using visual appearance and laser ranging. In *Proc. IEEE Int. Conf. Robot. Autom.*, 2006.

[12] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. *Autonomous Robot Vehicles*, 1:167–193, 1990.

[13] T. F. Smith and M. S. Waterman. Identification of Common Molecular Subsequences. *J. of Molecular Biology*, 147:195–197, 1981.

[14] N. Sunderhauf and P. Protzel. Switchable constraints vs. max-mixture models vs. rrr a comparison of three approaches to robust pose graph slam. In *Proc. IEEE Int. Conf. Robot. Autom.*, 2013.

[15] S. Thrun and M. Montemerlo. The GraphSLAM Algorithm with Applications to Large-Scale Mapping of Urban Structures. *Int. J. of Robotics Research*, 25(5-6): 403–429, 2006.

[16] M. Youssef and A. Agrawala. The Horus WLAN location determination system. In *Proc. Int. Conf. Mobile Systems, Applications, and Services*, 2005.