Calibration of distributed multimodal sensor networks using cross-correlation of arrival processes

Parisa Supitayakul (Kasetsart University, Thailand), *Zeynep Yucel (Okayama University), Francesco Zanlungo (ATR), Akito Monden (Okayama University), Pattara Leelaprute (Kasetsart University, Thailand)

1. Introduction

Ambient intelligence is a rapidly growing field, particularly together with the recent developments in sensing technologies, decreasing costs and increasing interest in assisted living, service, surveillance etc applications [1]. Although, in the early days, it was confined to sparsely populated indoor areas such as domestic environments, recently highly dynamic public spaces, both in and outdoor are addressed [2].

These intelligent environments are often equipped with multi-modal sensors, where most common modalities include vision and lidar, but depending on the application domain also haptic, acoustic, radar or wifi-signals may be used among other technologies [4]. In some cases, sensors may be part of the infrastructure (or the environment), while in many other cases there are also on-board sensors located on autonomously moving platforms (such as robots, electric wheelchairs or self-driving vehicles).

The purposes of these sensors are usually tracking of mobile obstacles and self-localization of moving platforms. For continuous tracking and/or target association, the spatial relation of the sensors, i.e. positional calibration, needs to ascertained. However, the calibration of such a network over a large dynamical space is a challenge due to the following issues:

- Privacy concerns may not permit collection of target-specific information.
- Field of views (FoV) do not necessarily overlap.
- Sensor readings from various modalities may not permit detection of features of same nature.

To account for these issues, we propose a simple yet efficient method for topological calibration of distributed sensors with various modalities. For that purpose, we use the arrival times of targets into each FoV and compute the correlation between different sensing ranges based on a set of assumptions. Subsequently, we test the resiliency of the method in real world settings. Our results demonstrate the efficacy and functional domain of the approach under certain limiting factors.

2. Related work

The proposed event correlation based approach presents similarities to the binary sensor networks, which have the advantages of being simple, easy to generalize, fast to process and thus are used in various studies. For instance, Aslam et al. define a binary data structure, where sensor readings indicate whether the target is moving towards the sensor or away from it [3]. Given the sensor locations, they apply a particle filter based approach on these readings to obtain target trajectories. Other important works in this area involve [6] and [7], which assume a binary reading as in/out of FoV but require sensor density to be high enough for the FoVs to overlap. In comparison to these studies, we treat a more fundamental problem. Namely, given only sensor readings without their relative positioning, we try to derive the topological relation [1].

Although from a sensor localization point of view, Kwon et al. treat a similar problem [8], they consider a particular case, where the events registered by individual sensors are generated from a global event (as in the case of seismic events) and propagate through the FoVs of various sensors, where in our case, the events (arrivals) are independent. Besides, our study distinguishes itself from the wellknown time difference of arrival (TDOA) methods, since TDOA assumes the signal to travel with a known velocity, where our method does not require any such condition [9]. On the other hand, cooperative localization methods require a subset of nodes with known locations[10], in contrast to our method, which treats all nodes in the same manner (with unknown locations).

In addition to these distinctions, we can list the advantages of our approach as follows:

- It does not need object identification, which means anonymous data can be used.
- It offers the possibility of fast online estimation, since it uses a plain array of numbers instead of high level object features such as pattern, shape, color etc.
- It can easily be incorporated with different modalities.
- It enables simultaneous calibration and probabilistic target association [2].

However, there are several requirements to have satisfactory performance using this approach. Namely, this method would suffer degraded performance rates under congestion and high variation on target velocity. Nevertheless, in what follows, we will

[1] Of course, once we achieve the calibration, the proposed method can be extended to probabilistically associate targets and obtain piece-wise continuous tracking.

[2] Although the latter is out of the scope of this paper.

demonstrate that it achieves satisfactory performance under uncontrolled real-world settings based on an extensive dataset recorded in a dynamic public space.

The outline of the paper is as follows. We will first formulate the problem in a formal manner and describe the solution strategy. Then, we will introduce our dataset and apply the method on it to prove the successful performance. Finally, we will list our conclusions and describe the future work.

3. Problem formulation and solution strategy

Consider that an interconnected system (such as a web of streets) is scanned using a multi-modal sensor network at various locations. In order to express this sensing environment analytically, we may use graph terminology [5]. Assume that there are N_n edges on this graph, which represent the sensor locations.

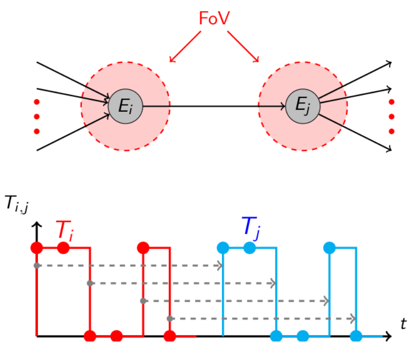

Fig.1 Example of a simple subgraph and relation of sensor readings.

Assume that we are particularly interested in the relative position of two edges like E_i and E_j on the (sub)graph given in Figure 1. Let these edges collect information from a limited space around them, i.e. their FoV. Provided that E_i and E_j are equipped with position tracking sensors, we can record the arrival and departure times (in and out of the corresponding FoVs) of each target (or equivalently, agent).

Assume that the vertices are directed in such a way that any agent that arrives in E_i goes to E_j and no agent arrives from the opposite direction. In this sense, E_i acts as a source and E_j acts as a sink in the subgraph. Suppose that a (mobile) agent α appears at E_i subject to an arrival distribution of P_i and travels with a velocity of v_α towards E_j. We express the arrival times observed at E_i with a function $T_i[n]$,

$$T_i[n] = \begin{cases} 1, & t_\alpha = nt_s \\ 0, & \text{otherwise} \end{cases}$$

where t_α is the arrival time of agent α and t_s is the sampling interval of the sensor. Clearly, for the double edge $\{E_i, E_j\}$ and single vertex directed subgraph given in Figure 1, where the agents have a constant velocity, T_i is free from any sort of disturbance and thus its exact replica will be observed at E_j after a time delay of t_d [3],

$$T_j[n] = T_i[n-k],$$

where $t_d = kt_s$. The delay term depends on v_α and the vertex length L_{ij} with the following relation,

$$t_d = \frac{L_{ij}}{v_\alpha}.$$

Assuming that the topological relation of the edges (i.e. L_{ij}) are not known but only the observations of T_i and T_j are available, we can solve for the value of t_d and discover the connectivity relation and distance between the edges taking a correlation stand-point.

Remember that the cross-correlation of two signals $f[n]$ and $g[n]$ is defined in discrete time domain as follows,

$$(f * g)[l] = \sum_{n=-\infty}^{\infty} \left(f[n] - \bar{f}\right) \left(g[n-l] - \bar{g}\right),$$

as $\bar{f}$ and $\bar{g}$ stand for the expected values of f and g, respectively. Without loss of generality, we can replace g by f and obtain the auto-correlation coefficients for f,

$$(f * f)[l] = \sum_{n=-\infty}^{\infty} \left(f[n] - \bar{f}\right) \left(f[n-l] - \bar{f}\right).$$

This function is integrable provided that f is of finite energy, i.e. finite support. In addition, it obviously achieves the maximum at $l = 0$, which is equal to the energy of the normalized signal $\tilde{f}[n]$,

$$E_{\tilde{f}} = \langle \tilde{f}[n], \tilde{f}[n] \rangle = \sum_{l=-\infty}^{\infty} |\tilde{f}[n]|^2,$$

where

$$\tilde{f}[n] = f[n] - \bar{f}.$$

Similarly, it is trivial to prove that correlation of a signal $f[n]$ with a delayed version of itself by any n_0, $g[n] = f[n-n_0]$, peaks exactly at n_0,

$$n_0 = \arg\max_l (f * g)[l].$$

Here instead of f and g, we can use the time series of arrivals T_i and T_j. For such arrival time series, it is plausible to assume that the observations come from a limited time interval of $[t_0, t_f]$, due to the practical limitations of real world applications. It directly follows that the signal is of finite support. Therefore, the estimation of t_d is guaranteed, provided that the

[3]Note that T_i and T_j are functions defined in discrete time domain with binary values and thus do not have a derivative.

observation duration is sufficiently long $t_d < t_f - t_0$. Then, the delay term can be solved as,

$$k = \arg\max_l (T_i * T_j)[l], \quad (1)$$

and the relative position of the edges is obtained through,

$$L_{ij} \propto k * t_s.$$

If v_α is available, the length of the vertex can also be calculated within an accuracy of $v_\alpha t_s$.

However, these findings are valid under a set of assumptions, which introduce several limitations. Below, we list some causes of such limitations:

- Variation of velocity: The cross-correlation term $(T_i * T_j)$ might be subject to clutter if v_α has a large variance.
- Directionality of traffic: The vertices may be bidirectional such that the traffic on L_{ij} is interlaced.
- Bifurcations: Edges E_i and E_j may not be connected directly and some agents leaving E_i may be bifurcated towards edges other than E_j.

We may express the impact of such factors on T_i with a disturbance term such that,

$$T_j[n] = T_i[n-k] + d[n].$$

However, it is not trivial to express $d[n]$ analytically taking in consideration all the dynamic factors. Therefore, in what follows, we directly apply the cross-correlation approach on a real world dataset and prove resiliency of the proposed approach even under various violations of the assumptions listed in the solution strategy.

4. Implementation on real world data

We consider the tracking dataset introduced by [11, 12] to demonstrate the application of our method in real world cases. This dataset is recorded in a network of underground pedestrian streets in Umeda, Osaka. The experiments are carried out at an intersection of two particular streets of this underground network (see Figure 2) and a total of 12791 people are tracked over a 6-hour time window.

The agents (pedestrians) arrive in and depart from the environment at four zones (the ends of the two streets). We denote these four zones with letters A∼D. In addition, although we scanned the entire environment, in order to create the effect of non-overlapping FoV, we assume that sensory information is available only from four sensing zones, i.e. A∼D.

We call the amount of time required to go from one sensing zone (or equivalently an edge in our graph theoretic approach) to another one as "travel time". This corresponds to the delay term t_d as described in Section 3. In order to understand the distribution pattern of t_d, we compute its histogram with bin size

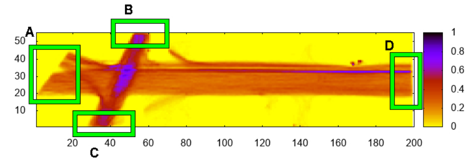

Fig.2 Normalized density map of the environment.

of 2 sec for each route (i.e. source/sink pair) as in Figure 3. We consider peak values of these curves to constitute the ground truth of t_d for respective routes. Note that the nonzero standard deviation already indicates that the variation of agent velocity has a more prominent impact on longer routes.

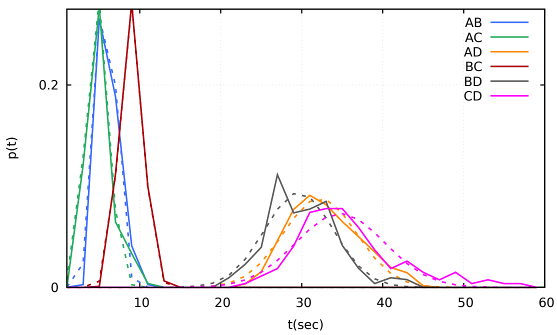

Fig.3 The pdf of travel times for all possible source-sink pairs

Next, we apply our method on the trajectory segments from sensing ranges A∼D and compute the correlation values. We illustrate the correlation terms over a given time window in Figure 4. In this case, the clutter is not only due to deviation of target velocity but also directionality of the traffic and bifurcations. Therefore, the patterns of the curves are more distorted and significantly different than the ones in Figure 3.

However, despite the clutter, we still observe significant peaks on four curves, i.e. the ones relating AB, AC, AD, and BC. In addition, the locations of these peaks are similar to those presented in Figure 3, which indicates that the method is capable of estimating the travel times with satisfactory accuracy. For the remaining two cases of BC and CD, there are no prominent peaks and the pattern of the signal is dominated by clutter.

The reason for this is the unbalance of traffic between various routes. Namely, there are few agents traveling between these routes as presented in Table 1 in comparison to the other routes with accurate estimations.

5. Conclusion and future work

In this study, we describe a simple and efficient positional calibration method for a distributed sensor

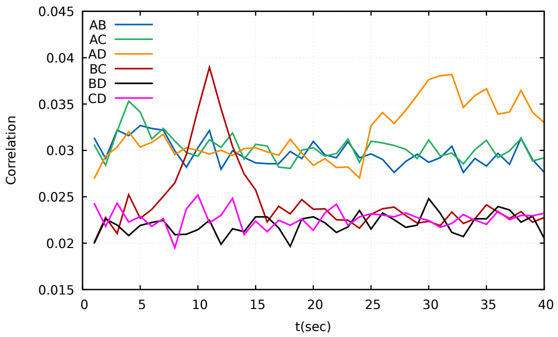

Fig.4 Correlation terms for all possible routes.

Table 1 Number of agents traveling on all routes.

Route	AB	AC	AD	BC	BD	CD
N_{agents}	563	782	2732	2199	265	135

network with various modalities. The proposed approach does not require target-specific information, respects privacy concerns, offers the possibility of fast and online calibration, can easily be incorporated with various sensor modalities (provided that presence/absence of the targets in FoV can be registered), and has the potential of simultaneous calibration and target association. Although, it is developed based on a list of assumptions, it is demonstrated using a real world dataset that even though there are reasonable violations, it can still cope with the dynamic factors. Of course, in order to assess the exact limitations of the method, we cannot carry out experiments in the real world addressing each possible case. Therefore, in the future we will design a simulation framework for evaluating the effect of common restrictions at various degrees illustrating the functional domain of the approach and the extend of limitations.

References

[1] Sadri, F.: "Ambient intelligence: A survey," ACM Computing Surveys (CSUR), 43(4), p.36.

[2] Ducatel, K., Bogdanowicz, M., Scapolo, F., Leijten, J., and Burgelman, J. C.: "Scenarios for ambient intelligence in 2010," Luxembourg: Office for official publications of the European Communities, 2001.

[3] Aslam, J., Butler, Z., Constantin, F., Crespi, V., Cybenko, G., and Rus, D. : "Tracking a moving object with a binary sensor network," International conference on Embedded networked sensor systems, pp. 150-161, ACM, 2003.

[4] Buratti, C., Conti, A., Dardari, D., and Verdone, R.: "An overview on wireless sensor networks technology and evolution," Sensors 9, no. 9 (2009): 6869-6896.

[5] Diestel, R.: "Graduate texts in mathematics: Graph theory," 2000.

[6] Mechitov, K., Sundresh, S., Kwon, Y., and Agha, G.: "Cooperative tracking with binary-detection sensor networks," International conference on Embedded networked sensor systems, pp. 332-333, ACM, 2003.

[7] Wang, Z., Bulut, E., and Szymanski, B. K.: "Distributed energy-efficient target tracking with binary sensor networks," ACM Transactions on Sensor Networks v. 6, no. 4 (2010): 32.

[8] Kwon, Y., and Agha, G. : "Passive localization: Large size sensor network localization based on environmental events," International Conference on Information Processing in Sensor Networks, pp. 3-14. IEEE, 2008.

[9] Gustafsson, F., and Gunnarsson, F. : "Positioning using time-difference of arrival measurements," IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 6, pp. VI-553. IEEE, 2003.

[10] Patwari, N., Ash, J. N., Kyperountas, S., Hero, A. O., Moses, R. L., and Correal, N. S.: "Locating the nodes: cooperative localization in wireless sensor networks," IEEE Signal processing magazine 22, no. 4 (2005): 54-69.

[11] Zanlungo, F., Ikeda, T., and Kanda, T.: "Potential for the dynamics of pedestrians in a socially interacting group," Physical Review E 89, no. 1 (2014): 012811.

[12] Yucel, Z., Zanlungo, F., Ikeda, T., Miyashita, T., and Hagita, N.: "Deciphering the crowd: Modeling and identification of pedestrian group motion," Sensors 13, no. 1 (2013): 875-897.