A Method for Estimating Position and Orientation with a Topological Approach using Multiple Infrared Tags

Yoshiyuki NAKAMURA
Information Technology Research Institute, AIST, Tokyo, Japan
Email: nmura@ni.aist.go.jp

Yuko NAMIMATSU
Alpha Systems, Inc., Kanagawa, Japan
Email: namimay@alpha.co.jp, miyazan@alpha.co.jp

Nobuo MIYAZAKI
Alpha Systems, Inc., Kanagawa, Japan

Yutaka MATSUO
and Takuichi NISHIMURA
Information Technology Research Institute, AIST, Tokyo, Japan
Email: y.matsuo@aist.go.jp, taku@ni.aist.go.jp

Abstract—This paper presents a method for estimating an object’s two-dimensional (2D) position and orientation based on topological information collected using infrared tags without any special location sensors or direction sensors. Estimating a user’s and articles’ location, irrespective of circumstances, is an important issue for context-aware systems. Users are present in a location with some purpose or intention. Therefore, a user’s position and orientation clearly reflect their context. Especially, orientation information can reflect a more detailed context than that obtained merely according to location: people standing face-to-face or back-to-back would have vastly different contexts. The analyses explained in this paper particularly examine an object’s orientation and describe a new method for estimating an object’s position and orientation in an indoor, real-world environment. Using a simulation and an implemented prototype system, the experimental results demonstrate the feasibility of our topological estimation method.

I. INTRODUCTION

A salient concern in research of real-world oriented systems, such as those used for wearable computers or ubiquitous computing [1], is location estimation. Particularly for human interaction systems, three aspects of location-associated context are important information: where you are, whom you are with, and what resources are nearby [2]. Obtaining and using these contexts, we specifically tackled the issue of community support at a conference site or exhibition hall [3], [4]. Community support simply means activating communities through their human relations.

To infer human relations in the real world, we specifically examine orientation and position to discover who is facing a user or what is in front of a user: such information implies the user context. For example, if an object’s position and orientation are obtained, situations can be inferred, such as people talking to each other in a circle or people who are auditing a poster presentation together. Using supporting information based on that context, users might easily find a common topic and have a nice chat at the site.

Present location systems [5] can detect user locations using various sensors. Multi-camera, pseudo-GPS, and ultrasonic systems can determine a user’s location with high accuracy. Along with a direction sensor, such as a digital compass, it is considered that those location systems can detect both position and orientation. However, those systems require a complicated infrastructure, arranged in advance, in an environment to estimate the user position. Events such as academic conferences require that costs of time and installation be kept as low as possible because the event period is limited. Moreover, the site is often a limited temporal space. Therefore, there is no great need for application of three-dimensionally accurate navigation; moreover, the simplicity of user devices and methods is an important consideration.

We have presented an overview of our method briefly in [6] to address the temporal-spatial issue specific to the event. The present paper provides detailed definitions and methods for estimating an object’s two-dimensional (2D) position and orientation using a topological approach. The proposed method estimates them using multiple infrared tags without any special location sensors or direction sensors.

II. POSITION AND ORIENTATION ESTIMATION METHOD USING A TOPOLOGICAL APPROACH

A. Topological Estimation Problem

This paper defines the local relationship as the positional and angular relationship between two objects: humans or things. The global relationship is a set of objects’ relationships represented by the entire local relationship. The global relationship also represents a set of constraints among objects. The term topology in this paper is a set of these constraints. Eventually, it is necessary to locate the position and orientation for each object, thereby satisfying the constraints of the global relationship as much as possible. This is a problem to solve by estimating global position and orientation considering the topology; we call this problem a topological estimation problem and designate such a method as topological estimation. Note that the topological estimation problem and the topological estimation are not identical in meaning.

Fig. 1 shows an overview of the topological estimation problem. In this study, the local relationships, which constitute the global relationship, are assumed to be gathered from multiple infrared tags associated with its emitting orientation.
Each data item is displayed as a “global relationship” in Fig. 1, and is represented as its constituent four-tuple data of \(a_1/\theta_{a_1} \ a_2/\theta_{a_2}\), where \(a_1\) is the detecting object’s name, \(\theta_{a_1}\) is the detecting angle, \(a_2\) is the detected object’s name, and \(\theta_{a_2}\) is the detected angle by \(a_1\). The unit of orientation is degrees (deg); it is associated with its tag ID in advance.

The infrared tag offers stable operation in indoor environments to collect position relationships using a proximity approach, as represented by Active Badge [7]. This feature was also demonstrated in our prior activity [8]. Infrared light also has easier controlling directivity features than either radio or ultrasound. By associating the directivity with its tag ID, it can simply collect orientation relationships. Conversely, it is difficult to measure distance using a simple circuit. However, its effective range can vary easily. In our method, relative local relationships can be collected using infrared directivity and the effective range.

The objects each have a unique ID that is separate from the infrared tag ID associated with its orientation. The other objects’ positions and orientation can be determined relatively and recursively based on the unique ID and local relationships if some objects have a known position and orientation at the site. This paper proposes a method for estimating these topological positions and orientations using a dynamics model that leverages the spring force and the repulsive force.

**B. Replacement Issue on Topological Estimation Problem**

Presume that two objects have position and orientation relationships like those shown in Fig. 2. Although each object has a reference point that expresses the front, the orientation is represented by a difference of angles from the reference point. We also assume that two different types of object exist in the site, which are movable objects such as humans or robots, and static objects, such as exhibits or walls, for which the position and orientation are already known.

Here, the topological estimation problem is described as follows. Given a distance \(l\), the angle \(\theta\), and that its time changes are visible with high accuracy, estimating the position and orientation becomes a simple replacement issue. In consideration of the local relationships collected from real infrared tags, however, these parameters of \(\theta\), \(l\), and time changes should be assumed to have extremely low accuracy. Therefore, a method to allocate the objects is needed.

We address the solution strategy of the topological estimation problem in detail using an example of Fig. 3. For this example, assume that each object has four directional resolutions of front, back, left, and right. Using the infrared tag, Obj6 should detect Obj2, which is showing its back in front of the left. In addition, Obj6 might detect the display panel Obj1, but against long odds because Obj1’s infrared tag is sometimes occluded by Obj2. Similarly, Obj6 detects Obj3, which is showing its back in front, Obj5, which is standing face-to-face, and Obj4, which is facing the back. The local relationships for the other objects are detected similarly with the infrared tag from Obj6’s perspective.

Topological estimation, which is a method to solve the topological estimation problem, establishes each object’s position and orientation in a 2D plane to reduce nonconformity with local relationships. Here, again referring to the Fig. 3 example, robot Obj4’s feasible area seems to be a wide area that is indeterminate, as seen from Obj6. However, if Obj4 does not detect Obj2 and Obj5, the feasible existence area can be focused into the shadow of Obj6. Similarly, Obj3 and Obj5 can be estimated in front of Obj6, and the mutual relationship between Obj3 and Obj5 can be determined. The estimation error can be reduced using these constraint conditions.

**C. Dynamics Model of Topological Estimation**

We have designed a dynamics model that extends the spring force model to achieve topological estimation. Fig. 4 shows an overview of the dynamics model. In our dynamics model, objects having a local relationship are assumed to be linked together with a spring with certain natural length because they...
are nearby. It is assumed that repulsive forces will be activated among objects that have no relations because they are probably far away.

The detailed algorithm follows. First, for non-static objects, randomly arrange objects’ initial positions and orientations. Next, replace the local relationships among objects with springs. These local relationships that are replaced by springs are the dashed-line relations in Fig. 1. Then, calculate the objects’ positions and orientations when the total forces of \( P_i(o_i) \) and \( P_j(o_i) \) become the most minimum and have the least nonconformity. The method described in this paper approximates the calculation as a numerical analysis using repetitive operations for which the total distance of each operation step is least. Although the numerical analysis causes a local minimum problem that falls into a wrong solution, certain external stimulations that break the equilibrium state help to alleviate the problem. In our model, next sensing data input is equivalent to stimulation. Once the first solution is obtained, the next calculation uses the solution as the initial arrangement that estimates the new situation when the new sensing data arrive.

Next, we define force \( P_i \), for which each object is at a certain position, as follows.

\[
P_i(o_i) = \sum_{j=1}^{N} \text{spring}(o_i, o_j) + \sum_{j=1}^{N} \text{repulsion}(o_i, o_j)
\]

In that equation, \( \text{spring}(o_i, o_j) \) is the spring force function that pertains between objects \( o_i \) and \( o_j \) with certain factor \( k_s \) if they share a local relationship; it is proportional to their distance. In addition, \( \text{repulsion}(o_i, o_j) \) is the repulsive force function with certain factor \( k_r \) and is inversely proportional to the squared distance. Finally, \( N \) represents the number of objects.

The spring force works with direction as well as distance. This force \( P_i(o_i) \) works on the difference from local relationship orientation and is defined as

\[
P_i(o_i) = \sum_{j=1}^{N} \text{roll}(o_i, o_j)
\]

where \( \text{roll}(o_i, o_j) \) is the spring force function that works between the detected orientation and current orientation with certain spring force factor \( k_s \) if they have a local relationship between \( o_i \) and \( o_j \). It is proportional to the difference of angles between the sensed orientations and the current computational orientations. For example, if two objects are sensed as being in a face-to-face situation, i.e. 0 degrees relative to each other, and the current computational orientation is also face-to-face, and the roll force will be zero. In addition to the spring force, a stress swings the other object dealing with the roll shown in Fig. 5. This swing stress works for moving the object position, which has more than two different relations to fit the most likely situation. The result of this force is shown with Fig. 10 in Section III-C.

### III. Computation Simulation Evaluation

In this section, we examine the effectiveness of the proposed method of topological estimation for estimation of the global position and orientation through computation simulation. We first describe the mode of simulation, and then show the result for orientation estimation.

#### A. Simulation Settings

We first prepare a virtual space, which is a 500 \( \times \) 500 pixel rectangular space. Fig. 6 shows the space’s initial setting. Static objects, which are the anchor points, are arranged on four sides of the center. The symbol strings arranged in the left side of Fig. 6 indicate their coordination as four-tuple data of “\( a : (x, y, o) \)”, where \( a \) is the object name, and \( x \) and \( y \) respectively represent the X-axis and Y-axis coordinate values; \( o \) is its orientation. The coordinate origin is shown above left. The orientation increases in a clockwise direction at a unit of degrees and the origin is the direction of the X-axis. The projection shown on the objects indicates the front of each.

We next assume that the local relationships among movable objects, especially those among humans, are collected from nametag-type devices shown in Fig. 7. The nametag has four infrared tags that face in different directions. The infrared tag ID offers its associated direction. In advance, tags are
When a sensing condition is satisfied, the exchanged IDs represent the local relationship. Those four directions of front, back, left and right are simply created meaning that the front is a direction in which interesting things are usually present, the left and right directions correspond to equal positions of a thing, and the back is not applicable to any of the previous. The number of the infrared tags in this method prescribes the direction resolution.

The exchanged local relationships among the infrared tags are reported via active RFID tags and are stored in a database. These local relationships are used as input for topological estimation. Note that the simulation deals with a certain static snapshot scene because the purpose is the examination of accuracy of the proposed dynamics model. A dynamic and time-variable simulation remains as a subject for our future work.

B. Simulation Conditions and Method

The error of estimation $E$ for global position and orientation is calculated as follows:

1) Several $N$ objects are allocated randomly in a certain position and orientation in the simulation field shown in Fig. 6.

2) The local relationships between two objects are determined among all $N$ allocated objects and four static objects based on a criterion with a sensing effective range $l$ and effective angle $\phi$.

3) Global position and orientation are estimated from the local relationships determined in the above process using the dynamics model described in Section II-C.

4) The estimation errors $E$ are calculated comparing the estimated global position and orientation with the initial position and orientation that were allocated randomly in (1).

In that simulation, the local relationships are determined when a sensing condition is satisfied. Fig. 8 shows the sensing condition by which Obj2 sensed the signal of Obj1. The sectors shown in the figure offer the sensing range and the signal range. Here, $\phi_r$ is defined as the effective angle of the sensing range, $\phi_e$ is defined as the effective angle of the signal range, and $l$ is defined as the effective length of the signal. In the simulation, these parameters are set such that the angle $\theta_r$ is 120 deg ($\pm$60 deg from the center) and angle $\theta_e$ is 40 deg ($\pm$20 deg from the center). These parameters resemble those of the general infrared sensor and LED, but the range shape of the actual sensor and LED is a teardrop shape. The simulator approximates this range using the sector shape. Because of the occlusion, sensing fails when any object having radius $r$ exists on the send/receive signal line.

The estimation error of $E$ is divided into two detailed errors of $E_d$ and $E_0$, where $E_d$ represents the arithmetic average of total displacement between original positions and estimated positions. It is defined as follows.

$$E_d = \frac{1}{M N} \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{(x_{oi} - x_{si})^2 + (y_{oi} - y_{si})^2}$$

Here, position coordinate $(x_o, y_o)$ is the estimated position and $(x_s, y_s)$ is the original position in which the objects are allocated randomly. In addition, $N$ represents the number of the objects and $M$ represents the number of the trials. Moreover, $E_0$ represents the arithmetic average for the angle in a similar manner; it is defined as follows.

$$E_0 = \frac{1}{M N} \sum_{i=1}^{M} \sum_{j=1}^{N} |\theta_{oi} - \theta_{si}|$$

Therein, angle $\theta_o$ is the estimated angle, whereas $\theta_s$ is the original angle that is allocated randomly before applying the topological estimation.

The natural length of the spring is 125 pixels. It is assumed that a repulsive force acts between objects within 100 pixels of distance. For preventing oscillation, the factor of spring $k_s$ and $k_a$, and the factor of repulsive force $k_r$ decreases its value depending on distance in the topological estimation repetitive operation.

The purpose of this simulation is examination of topological estimation behavior. For that reason, we chose scenes in which all objects have at least one local relationship to the other object as the parent population of the simulation. Fig. 9 shows a typical estimation result scene that satisfies this assumption. The assumption offers the minimum condition that all objects are captured in the simulation. Conversely, an object that
has no local relationship becomes indeterminate in terms of position and orientation.

C. Simulation Results and Discussion

This subsection shows the result of the orientation estimation. We discuss the results based on the number of the objects as follows.

We first discuss the simplest instance, in which the number of the objects \( N \) is 1. The local relationship established between only one static object causes one spring. The topological estimation therefore estimates the object’s position distance from 125 pixels, which is the natural length of the spring, irrespective of the original position. Consequently, there are few advantages to knowing the position of only one object because the position error \( E_d \) depends strongly on the effective sensing distance of \( l \). The angle error \( E_\theta \) also depends strongly on the sensing condition, as shown in Fig. 8; the greatest error is determined geometrically. For \( \phi_r \) of 40 deg and \( \phi_r \) of 120 deg, the greatest error of \( E_\theta \) becomes 40 deg, as calculated geometrically.

However, the topological estimation is peculiar in the case in which a movable object has two local relationships with other static objects that are located respectively on the X-axis and Y-axis. With the model that replaced local relations with only a spring, the movable object should be estimated on a line tied to two static objects. However, the result shows it out of line, as shown in Fig. 10 because the swing stress is activated, as shown in Fig. 5.

We next discuss the fact that the number of the allocated objects \( N \) changes from 1 to 25. Fig. 11 shows the estimation error of distance \( E_d \), and Fig. 12 shows that of the angle. Each unit is [pixel] and [degree]. As those figures show, the estimation errors decrease for \( N \). In this simulation, the effective sensing range \( l \) is 250 pixels, where \( l \) is one-half of the room, and the number of trials \( M \) is 1000.

It is peculiar to the algorithm that the angle error is very small, although the distance error is large. The estimated angle is the average of 8.4 deg on \( N=23 \). In that case, the number of the local relationships is the average of 6.1 lines. This means that each object has three local relationships, on average. This result underscores the effectiveness of the proposed method because that characteristic is an important factor among actual situations to infer the human relations.

D. Divergence

On the other hand, this simulation highlighted the issue of topological estimation. As the number of the objects \( N \) increased, the divergence of computation increased. Fig. 13 shows an example of divergence when \( N=10 \). The percentage of divergence was 19% when \( N=10 \); it was 37% when \( N=20 \), which is mainly attributable to the fact that the number of static objects, which represent anchoring points, decreases relative to \( N \). Therefore, one solution for this issue is to simply increase the static objects. For example, a new and fifth static object
IV. IMPLEMENTATION AND REAL DEVICE EVALUATION

In this section, we examine the characteristic features of the real device. We first show the communication protocol and the implementation details; then we evaluate the topological estimation with the real devices and infer the human relations from the data.

A. Local Relationship Collection Device

We implemented a prototype device that collects local relationships in the real world to realize topological estimation. Fig. 14 portrays that device. We presume that local relationships are collected using a nametag type device like that shown in Fig. 7 at an actual site such as an academic conference or exhibition.

Fig. 15 shows the protocol by which the device collects local relationships. In the protocol, a local relationship between two objects can be detected intermittently at some interval. The detected local relationship is reported immediately from the active RFID tag to the gateway-like sink node. In our usage, the active RFID tags offer only a communication channel and do not mutually communicate.

Fig. 16 shows a block diagram. We adopted a small, low-power-consumption-type active RFID tag [9]. The sending ID is encoded by PIC16LF88, the infrared tag micro-controller, which directly controls the infrared LED. It also decodes the receiving signal to the ID data. In the prototype implementation, the data length of the infrared tag is 8 bits, with 4 bits for the device ID, 3 bits for the orientation ID, and 1 bit for control. The carrier frequency is 38 kHz, whereas the transmission average time is 22 ms, including the header and invert code for error detection. The transmission frequency is approximately once per second, but it is controlled by software and is customizable. Each infrared tag is connected as a daisy chain. Therefore, the object can be operated with more than four tags.

B. Sensing Property of Collection Device

Before the actual equipment experimentation, we measured the fundamental data for the local relationship collection. It is important for collection that the device show sufficient direc-
tivity for angle resolution and a high sensing rate within the directivity and the effective range. The maximum transmission range is useful as a parameter for topological estimation as the natural length of the spring.

In the experiment, we first examined directivity. It showed the half-value angle of $\pm 20$ deg of LED for transmission as a result. This value conforms to the computation simulation parameter in Section III-B. However, under physical circumstances, it can be detected weakly within this angle because it is a half-value, even though the simulation is set so as not to detect the sensor out of the range. Overlapping with effective range $\pm 60$ deg of the light-receiving element, the issue of receiving a signal from an unexpected angle was revealed. To fit the prior simulation characteristics, we therefore attached a cylindrical component to the top of the light-receiving element shown in Fig. 17 to suppress incident light physically.

We next measured the effective range of the distance. Fig. 18 shows the result. Reception drops sharply at a distance of 3.8 m. This effective range is controllable from 3.8 m to 6.5 m by tuning the variable resistor. We also measured the active RFID tag’s effective range. Results show that the coverage is approximately 15 m.

We finally evaluated power consumption. Considering operation at a conference site or exhibition site, it is desirable that the run length be at least 1 day of 8 hours. In the experiment, we observed how long the device can operate with one button-type CR2032 lithium battery under the most extreme conditions to the device. The device can run about 3.5 h continuously with four infrared tags, transmission and reception of infrared once per second, and reporting of detected IDs once per second via radio, and at the maximum effective range, which is 6.5 m. Using an extra sleep operation and reducing the effective range to 3.8 m, the time was extended to about 4.5 h of continuous operation. Therefore, the device can operate for more than 8 h using two CR2032 batteries.

C. Estimation Characteristics and Evaluations

We evaluated estimation characteristics using the real device and the proposed algorithm. For simplification, we undertook an experiment on the same height of 2D plane. In fact, although the height varies for the posterior length and layout of the static object, it poses only a negligible problem for actual operation of the infrared tag [8] because it does not offer high precision. Conversely, our method offers higher accuracy from low-precision data like that obtained from infrared tags.

In accordance with the simulation condition and effective distance of 3.8 m described in the previous Section IV-B, we arranged a rectangular space of 7 m on a side, which is about double that of 3.8 m, as the experiment target field. Fig. 17 shows the actual local relationship collection device used in the field experiment.

Fig. 19 shows the topological estimation. The left side of the figure represents the result of topological estimation from the local relationships collected at the left figure’s position and orientation. The lines among objects on the right figure are springs of the topological estimation dynamics model: local relationships. These results show that the topological estimation reconstructed the situation shown in the left figure.

The situation depicted in Fig. 20 subsumes that three per-
sons are talking and facing each other in a circle. No local relationship exists among static objects, which serve as anchor points. Therefore, the position of three persons has fallen into indeterminacy. However, this situation, in which three people are facing each other, is estimated properly\(^1\). Therefore, the proposed topological estimation should estimate the situation as “Three persons are near each other and facing each other” → “It is likely that they became acquainted because it is very possible they are talking.”

V. RELATED WORKS

The topological estimation described in this paper features the capability to estimate objects’ orientation. It also provides both local and global relative positioning of indoor objects. However, much other research on location systems has been presented [5].

In the sense of a device mechanism among the location systems, our work is related to those systems that use simple proximity information [7]. With the proximity information, a metaphor of swapping business cards can explicitly relate human-human or human-thing information [10], [11], [12] to enhance human interactions or collaborations at conference sites or institute open houses as targeted by our system. Focusing on speech sound can be an efficient trigger for automatically relating proximal objects with the help of microphone devices [13]. Some systems obtain and visualize human activity at the site by localizing users within the specific area with an active RFID tag effective range [14] or by recording activity using multiple heterogeneous sensors [15].

These precedent systems, however, address only one of the two relations of local peer-to-peer or global relative positioning. Topology estimation differs from these in that it can obtain both local and global relative positions. However, the previous systems offer little environmental device load and provide easy set-up features with the use of simple infrared or radio wireless communication technology. These features are important at the site because an event such as an academic conference or exhibition is not a regular use for such a venue. Devices on those systems have been well designed considering this aspect, as have our systems.

\(^1\)In this example, the positions of 12 and 13 are estimated properly \(\frac{1}{2}\) of the time, although no local relationship exists among them.

In networked sensing and wearable computing research fields, localization is also an important trend. Some systems estimate orientation as well as location, although these systems mostly address the sensing and tracking of the location of the target node. The approach of these systems is to use a method of angle-of-arrival (AoA) with the help of a digital compass [16], using model-based matching [17] and using multiple infrared tags in a package of a single node [18], [19]. The latter studies closely relate to the development of our device in terms of multiple infrared beacons, if they consider it for event use. From the descriptions in those papers, it is not clear that they are applicable to this type of event.

Without the event usage, our system resembles the Relate system [20], [21] for estimating position and orientation both locally and globally. The Relate system provides a capability to measure spatial relations in a peer-to-peer fashion and to visualize the whole target in a map for mobile devices such as laptop PCs. The system uses a special USB-connected dongle that emits ultrasound and simultaneously senses neighbors’ respective distances and angles using ultrasonic time-of-flight (ToF) and AoA. To estimate the global position, however, the Relate system uses accurate distance information among objects, although the assumption of distance in our system is extremely low. In fact, the distance is a fixed value that is derived from the maximum infrared effective range in our simulation in Section III-B and the experiment described in Section IV-C. A detailed estimation algorithm is explained in [21], but from the system behavior it is not clear if the distance becomes inaccurate. In contrast to that system, we have proposed a method for estimating position and orientation using simple proximity infrared tags.

VI. CONCLUSIONS AND FUTURE WORK

This paper presents a topological estimation method that can estimate the global position and orientation from local relationships between two objects collected by infrared tags. A computation simulation and an experiment with the prototype implementation showed that context estimation is possible using the proposed extended spring dynamics model. Particularly, the proposed method represented good estimation precision for orientation. Under the conditions and parameters given in Section III, it achieved the estimation error of an angle of less than 10 deg for more than 12 objects. The position estimation error also decreases as the number of objects increases: that is an advantage of our topological approach, which connects many surrounding objects.

A new issue was highlighted during experimentation. A divergence of computation arose if only a few static objects existed, representing anchoring points, for total objects. That divergence is reducible using object allocation examined through experimentation. As a subject for future work, we would like to tackle this issue using temporal change. Resolving that issue would show topological estimation to be a more important practical method.

In addition to the issue described above, we are examining a method for collecting local relationships without using infrared
tags, along with applications to situations other than conferences, and derivation of the high abstraction level context from the estimation results within the group using data-mining techniques.

**ACKNOWLEDGMENT**

This research has been supported in part by Industrial Technology Research Grant Program (project ID 04A11502a) from NEDO (New Energy and Industrial Technology Development Organization of Japan).

**REFERENCES**


