Cooperative reference frame estimation for multi-agent systems via formation control

Asai, Ryo; Sakurama, Kazunori

FULL PAPER

Cooperative Reference Frame Estimation for Multi-Agent Systems via Formation Control

R. Asai[a] and K. Sakurama[a]

[a]Graduate School of Informatics, Kyoto University, Kyoto, Japan

ARTICLE HISTORY
Compiled June 6, 2023

KEYWORDS
multi-agent systems; distributed estimation; formation control; reference frame

ABSTRACT
In this study, we propose a distributed method for multi-agent systems to estimate a global reference frame by constructing a common reference frame for all agents with only local measurements. Our method is a combination of formation control and similarity evaluation of feature point matching. Here, we design a distributed controller to achieve a prescribed configuration and a distributed estimator to construct a common reference frame, which calculates the similarity between the measured relative positions and the prescribed configuration. Then, we quantify the error range between the common reference frame and the global reference frame and show that the constructed common reference frame is the best. The proposed method requires only the information on the relative positions of neighbor robots through local measurements and does not require inter-robot communication or relative posture observation. Finally, we conducted simulations with 12 agents and an experiment with 8 two-wheeled robots to verify the effectiveness of the proposed method.

1. Introduction

Simultaneous operation of multiple autonomous robots is expected to improve fault-tolerance and work efficiency. For example, autonomous underwater vehicles (AUVs) can survey topography [1], explore mineral deposits [2], and inspect pipeline and cables [3] at the bottom of the ocean, where humans cannot reach. For another example, micro air vehicles (MAVs) can survey a large area [4] and transport heavy load cooperatively [5].

When a group of robots work cooperatively, constructing a common reference frame is important to obtain coordinates consistent with all robots [6,7]. In addition, the common reference frame is preferably identical to the global reference frame defined by geographic features such as geomagnetism and landmarks to interact with objects in surroundings [8]. Unless GPS (Global Positioning System) or motion capture systems are available, agents need to estimate their own states (e.g., positions and postures) using their own sensors, such as IMUs (Inertial Measurement Units), LiDAR (Light

CONTACT R. Asai. Email: asai.ryo.83w@st.kyoto-u.ac.jp

Detection And Ranging) sensors and cameras. Information based on such sensors can be inconsistent with each other, which leads to the mismatch of the local reference frames of robots. For this reason, it is required to construct a common reference frame using such inconsistent information. Construction of a common reference frame enables agents to share positional information without accurate absolute measurements.

The problems of constructing a common reference frame are studied in the field of multi-agent systems. Paper [9] has proposed a distributed algorithm for static sensor networks in two-dimensional space, based on a gossip communication scheme, i.e., asynchronous and random pairwise communication. In [10], a consensus-based algorithm for multi-robot systems in two-dimensional space has been proposed. This method uses a consensus protocol of the estimations of a common reference frame by communication, exploiting measured relative positions. In [11], two distributed algorithms for multi-robot systems in three-dimensional space have been proposed; one has an asymptotic convergence property and the other has a finite-time convergence property. This method employs a consensus protocol of auxiliary matrices to derive a common reference frame. These existing methods construct a common reference frame by exchanging the information on relative positions (and orientations) with wireless communication.However, wireless communication can be degraded by packet loss and shielding effects [12] and is unavailable underwater due to the high attenuation of electromagnetic waves. Also, robots having payload and battery limitations such as MAVs cannot be equipped with extra devices [13]. In these cases, constructing a common reference frame without communication is expected. To the best knowledge of the authors, such a method has not been developed so far.

In this study, we propose a method to construct a common reference frame close to the global reference frame in a distributed and cooperative manner with only local measurements. The proposed method is a combination of formation control and similarity evaluation of feature point matching. First, we introduce a prescribed configuration of formation as a clue to common reference frame construction. Second, to achieve the prescribed configuration, we use the optimal controller in terms of the smallest ambiguity for the configuration. Third, we develop a method to construct a common reference frame, which calculates the similarity between the prescribed configuration and measured relative positions of neighbor agents. Further, we quantify the error range between the common and global reference frames and show the proposed method is optimal in terms of volume of the error range set. Finally, we demonstrate the effectiveness of the proposed method by simulations with 12 agents and an experiment with 8 two-wheeled robots.

The proposed method does not require communication between agents and can be implemented only by onboard sensing devices. Furthermore, it can be applied in both two- and three-dimensional spaces.

Notation: Let $\mathbb{R}$ be the set of real numbers. $I_d \in \mathbb{R}^{d\times d}$ denotes identity matrix of size d. For a finite set $\mathcal{A}$, let $|\mathcal{A}|$ denote the number of elements in $\mathcal{A}$. For n elements $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$ and their indices set $\mathcal{C} = \{i_1, \ldots, i_{|\mathcal{C}|}\}$, let $x_\mathcal{C}$ represents the following matrix

$$x_\mathcal{C} = [x_{i_1}, x_{i_2}, \ldots, x_{i_{|\mathcal{C}|}}] \in \mathbb{R}^{d\times|\mathcal{C}|}. \tag{1}$$

Frobenius norm is denoted as $\|\cdot\|$. For matrix $X \in \mathbb{R}^{d\times n}$ and the set $\mathcal{A} \subset \mathbb{R}^{d\times n}$, the

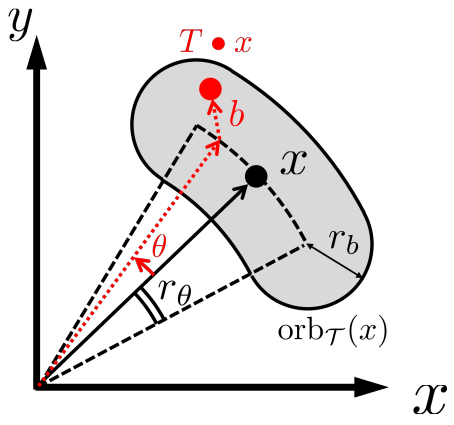


Figure 1. Operation $T \bullet x$ the region of $\mathrm{orb}_{\mathcal{T}}(x)$.

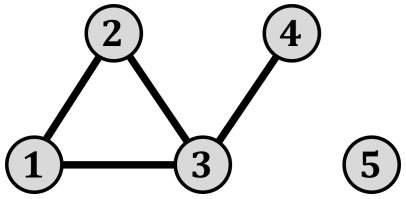


Figure 2. An example of graph $\mathcal{G}$.

distance between X and $\mathcal{A}$ is defined as

$$\mathrm{dist}(X, \mathcal{A}) = \inf_{Y \in \mathcal{A}} \|X - Y\|. \tag{2}$$

2. Mathematical Preliminaries

2.1. Special Euclidean Group

Let $\mathrm{SO}(d)$ be the set of d-dimensional orthogonal matrices whose determinant is 1. Let $\mathrm{SE}(d)$ denote the d-dimensional special Euclidean group as

$$\mathrm{SE}(d) = \{T \in \mathbb{R}^{(d+1)\times(d+1)} : T = \begin{bmatrix} M & b \\ \mathbf{0}_d^\top & 1 \end{bmatrix}, M \in \mathrm{SO}(d), b \in \mathbb{R}^d\}, \tag{3}$$

where $\mathbf{0}_d \in \mathbb{R}^d$ is the d-dimensional vector which has all components equal to 0. The following matrices are the examples of $\mathrm{SO}(d)$ and $\mathrm{SE}(d)$ for $d = 2$:

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathrm{SO}(2), \tag{4}$$

$$T = \begin{bmatrix} \cos\theta & -\sin\theta & b_x \\ \sin\theta & \cos\theta & b_y \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{SE}(2). \tag{5}$$

The action of $\mathrm{SE}(d)$ on $\mathbb{R}^d$ is denoted using the operator $\bullet$ as follows:

$$T \bullet x = Mx + b \tag{6}$$

for $T = \begin{bmatrix} M & b \\ \mathbf{0}_d^\top & 1 \end{bmatrix} \in \mathrm{SE}(d)$ and $x \in \mathbb{R}^d$. In Figure 1, the red dotted lines and the red point represent the transformation of $x \in \mathbb{R}^2$ by the operation in (6) with T in (5), i.e. x is rotated by an angle θ and then translated by $b = [b_x, b_y]^\top \in \mathbb{R}^2$. Besides, the

action of $\mathrm{SE}(d)$ on $\mathbb{R}^{d\times n}$ is denoted as follows:

$$T \bullet x_{\mathcal{N}} = [T \bullet x_1, T \bullet x_2, \ldots, T \bullet x_n], \tag{7}$$

where $x_{\mathcal{N}} = [x_1, \ldots, x_n] \in \mathbb{R}^{d\times n}$.

For $\mathcal{T} \subset \mathrm{SE}(d)$, the product of $\mathcal{T}$ with an element $T' \in \mathrm{SE}(d)$ is defined as follows:

$$T'\mathcal{T} = \{T'T \in \mathrm{SE}(d) : T \in \mathcal{T}\}. \tag{8}$$

Let $\mathcal{T}^{-1} \subset \mathrm{SE}(d)$ be the set of inverse elements of $\mathcal{T}$. $\langle\mathcal{T}\rangle$ denotes subgroup generated by $\mathcal{T}$, that is the smallest subgroup of $\mathrm{SE}(d)$ containing every element of $\mathcal{T}$. The $\mathcal{T}$-orbit of $x_{\mathcal{N}}$ is defined as follows:

$$\mathrm{orb}_{\mathcal{T}}(x_{\mathcal{N}}) = \{T \bullet x_{\mathcal{N}} \in \mathbb{R}^{d\times n} : T \in \mathcal{T}\}, \tag{9}$$

which represents the collection of $T \bullet x_{\mathcal{N}}$ for all $T \in \mathcal{T}$. The gray area in Figure 1 represents $\mathrm{orb}_{\mathcal{T}}(x)$ for $x \in \mathbb{R}^2$ and

$$\mathcal{T} = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta & b_x \\ \sin\theta & \cos\theta & b_y \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{SE}(d) : b_x, b_y, \theta \in \mathbb{R}, |\theta| \le r_\theta, \sqrt{b_x^2 + b_y^2} \le r_b \right\}. \tag{10}$$

2.2. Graph Theory

Consider an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with a node set $\mathcal{N} = \{1, \ldots, n\}$ and an edge set $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$. The neighbor set of node i is denoted as

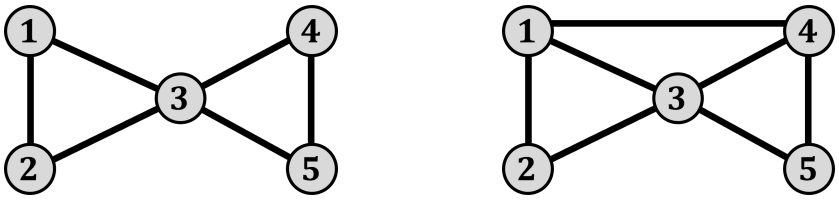
$$\mathcal{N}_i = \{j \in \mathcal{N} : \{i, j\} \in \mathcal{E}\} \cup \{i\}. \tag{11}$$

Note that $\mathcal{N}_i$ includes node i itself. For a node subset $\mathcal{C} \subset \mathcal{N}$, the subgraph $\mathcal{G}|_{\mathcal{C}} = (\mathcal{C}, \mathcal{E}|_{\mathcal{C}})$ is the *induced subgraph* of $\mathcal{C}$ with the node set $\mathcal{C}$ and the edge set $\mathcal{E}|_{\mathcal{C}}$ consisting of all of the edges in $\mathcal{E}$ that contains the pairs of the nodes in $\mathcal{C}$. A node subset $\mathcal{C}$ is called a *clique* if the induced subgraph $\mathcal{G}|_{\mathcal{C}}$ is complete. A clique $\mathcal{C}$ is said to be *maximal clique* if $\mathcal{C}$ is contained by no other cliques. Let us consider the graph $\mathcal{G}$ in Figure 2. The induced graph $\mathcal{G}|_{\{1,2,3\}} = (\{1,2,3\}, \{(1,2),(1,3),(2,3)\})$ is complete, which implies $\{1,2,3\}$ is a clique. On the other hand, $\mathcal{G}|_{\{2,3,4\}} = (\{2,3,4\}, \{(2,3),(3,4)\})$ is incomplete, which implies $\{2,3,4\}$ is not a clique.

For graph $\mathcal{G}$ and matrix $X \in \mathbb{R}^{d\times n}$, a pair $(\mathcal{G}, X)$ is called a *framework*. Let $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_q \subset \mathcal{N}$ be the maximal cliques in the graph $\mathcal{G}$ and $\mathcal{Q} = \{1, 2, \ldots q\}$ be the set of its indices. For a given graph $\mathcal{G}$ and matrix $x^*_{\mathcal{N}} = [x^*_1, \ldots, x^*_n] \in \mathbb{R}^{d\times n}$, framework $(G, x^*_{\mathcal{N}})$ is said to be *clique-rigid* [14] if the following holds for each matrix $x_{\mathcal{N}} = [x_1, \cdots, x_n] \in \mathbb{R}^{d\times n}$

$$x_{\mathcal{C}_k} \in \mathrm{orb}_{\mathrm{SE}(d)}(x_{\mathcal{C}^*_k}) \quad \forall k \in \mathcal{Q} \Rightarrow x_{\mathcal{N}} \in \mathrm{orb}_{\mathrm{SE}(d)}(x^*_{\mathcal{N}}). \tag{12}$$

Figure 3 shows a clique-rigid framework and a non-clique-rigid framework.



(a) Non-clique-rigid. (b) Clique-rigid.

Figure 3. Examples of frameworks.

3. Problem Formulation

Consider n agents in a d-dimensional space. The agents are numbered from 1 to n, and the set of the numbers is denoted as $\mathcal{N} = \{1, \ldots, n\}$.

Let Σ be the *global reference frame* defined by environmental features such as a geomagnetic field and landmarks. Each agent measures its own position and orientation in Σ. One example of how to measure the own absolute position is self-localization with a map, which can be implemented with relative sensors, e.g, LiDAR. We should note that this method involves localization errors on agents, which are different from each other. Especially, we focus on the bias error. Due to the bias, a reference frame $\Sigma^{[i]}$ constructed by agent i, called a *local reference frame*, is different from Σ. Figure 4 shows the relation between Σ and $\Sigma^{[i]}$. The vector b_i and the matrix M_i represent the origin and the rotation of $\Sigma^{[i]}$ from Σ, respectively. Let $T_i = \begin{bmatrix} M_i & b_i \\ \mathbf{0}_d^\top & 1 \end{bmatrix} \in \mathcal{T} \subset \mathrm{SE}(d)$ according to (3), which is called a *transformation matrix*.Then, the following relationship holds between the position $z \in \mathbb{R}^d$ defined in Σ and the corresponding position $z^{[i]} \in \mathbb{R}^d$ in $\Sigma^{[i]}$

$$z = T_i \bullet z^{[i]}. \tag{13}$$

Suppose that the value of T_i is unknown, but the set $\mathcal{T}$ that T_i belongs to is known to all the agents.

Remark 1. As mentioned previously, $\mathcal{T}$ represents the error range of self-localization. Hence, it can be estimated from the accuracy of the relative sensor. For example, it was reported in [15] that the self-localization scheme using Hokuyo UST-20LX in a corridor environment caused a position error of up to 0.15 m and an attitude error of up to 1 deg. Hence, we set $\mathcal{T}$ in (10) with $r_b = 0.15\,\mathrm{m}$ and $r_\theta = 1\,\mathrm{deg}$ in this case.

The position of agent i in Σ is represented by $x_i(t) \in \mathbb{R}^d$. The control input $u_i(t) \in \mathbb{R}^d$ is given by the velocity in $\Sigma^{[i]}$. Then, the kinematic model of agent i is denoted as follows:

$$\dot{x}_i(t) = M_i u_i(t). \tag{14}$$

The structure of the sensing network is denoted by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with an edge set $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$. Let $\mathcal{N}_i \subset \mathcal{N}$ be the neighbor set of agent i defined as (11). Let $x_j^{[i]} \in \mathbb{R}^d$ be the relative position of agent j from agent i, given as

$$x_j^{[i]} = T_i^{-1} \bullet x_j. \tag{15}$$

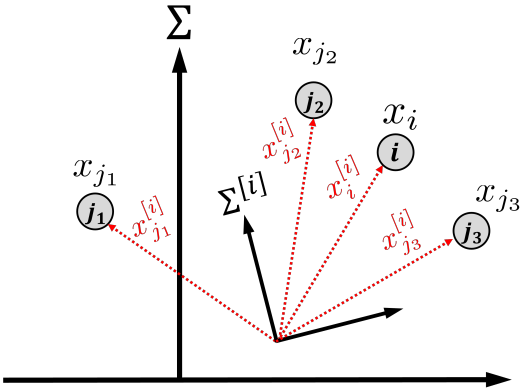


Figure 4. Information that agent i obtains by local measurements.

The information that agent i obtains by local measurement is the collection $x_{\mathcal{N}_i}^{[i]}$ of the relative positions of agent i's neighbors, which is expressed from (1) as

$$x_{\mathcal{N}_i}^{[i]} = [x_{j_1}^{[i]}, x_{j_2}^{[i]}, \ldots, x_{j_{|\mathcal{N}_i|}}^{[i]}] \tag{16}$$

for $\mathcal{N}_i = \{j_1, j_2, \ldots, j_{|\mathcal{N}_i|}\}$. Figure 4 depicts the position of agent i, $x_i^{[i]}$, and those of the neighbors, x_{j_1}, x_{j_2}, and x_{j_3}, which form $x_{\mathcal{N}_i}^{[i]}$ for $\mathcal{N}_i = \{i, j_1, j_2, j_3\}$.

For the agents to cooperate with each other, they need to construct a *common reference frame* Σ_{c}, whose origin and orientation must be identified by all agents locally. The common reference frame is preferably identical to Σ, but it cannot be realized in general because of the different measurement biases. For this reason, the objective of this study is to construct a common reference frame that is as close to Σ as possible. Let T_{c} be the transformation matrix of Σ_{c} from Σ. Note that T_{c} corresponds to the error bias common in the agents, and Σ_{c} coincides with Σ if $T_{\mathrm{c}} = I_{d+1}$. Figure 5 shows the relation among Σ, $\Sigma^{[i]}$, and Σ_{c}. By constructing a common reference frame, each agent can use the coordinate in Σ_{c}, which is equivalent to knowing the transformation matrix $T_{\mathrm{c}}^{-1}T_i$ of $\Sigma^{[i]}$ from Σ_{c}. The reason is as follows. Let z_{c} be the coordinate corresponding to z in Σ_{c}, satisfying

$$z = T_{\mathrm{c}} \bullet z_{\mathrm{c}}. \tag{17}$$

Information that agent i can obtain by local measurement is $z^{[i]}$. From (13) and (17), the following is obtained

$$z_{\mathrm{c}} = (T_{\mathrm{c}}^{-1}T_i) \bullet z^{[i]}. \tag{18}$$

Hence if agent i knows the value of $T_{\mathrm{c}}^{-1}T_i$, it can achieve the value of z_{c} from $z^{[i]}$. For this reason, we just have to estimate $T_{\mathrm{c}}^{-1}T_i$.

Let $\hat{T}_i(t) \in \mathrm{SE}(d)$ be the agent i's estimation of $T_{\mathrm{c}}^{-1}T_i$ for a certain $T_{\mathrm{c}} \in \mathrm{SE}(d)$. Then, the estimation is expected to satisfy

$$\exists T_{\mathrm{c}} \in \mathcal{T}_{\mathrm{c}} \ \text{ s.t. } \lim_{t\to\infty} \hat{T}_i(t) = T_{\mathrm{c}}^{-1}T_i \ \ \forall i \in \mathcal{N}, \tag{19}$$

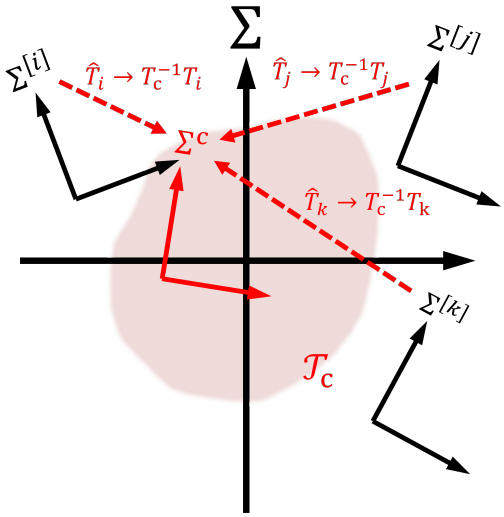


Figure 5. Global reference frame Σ, local reference frame $\Sigma^{[i]}$ and common reference frame Σ. Light red range expresses the error bias range $\mathcal{T}_c$.

where $\mathcal{T}_c \subset \mathrm{SE}(d)$ corresponds to the error bias range of the estimation. Hence, $\mathcal{T}_c$ must contain I_{d+1}, corresponding to the global reference frame Σ, and its volume should be as small as possible. Figure 5 explains (19), where the light red area $\mathcal{T}_c$ (with the orientation bias ignored) contains Σ_c.

Because agent i can only obtain the information (16), we design an estimator and a controller of the forms

$$\hat{T}_i(t) = f_i(x^{[i]}_{\mathcal{N}_i}(t)), \tag{20}$$

$$u_i(t) = g_i(x^{[i]}_{\mathcal{N}_i}(t)), \tag{21}$$

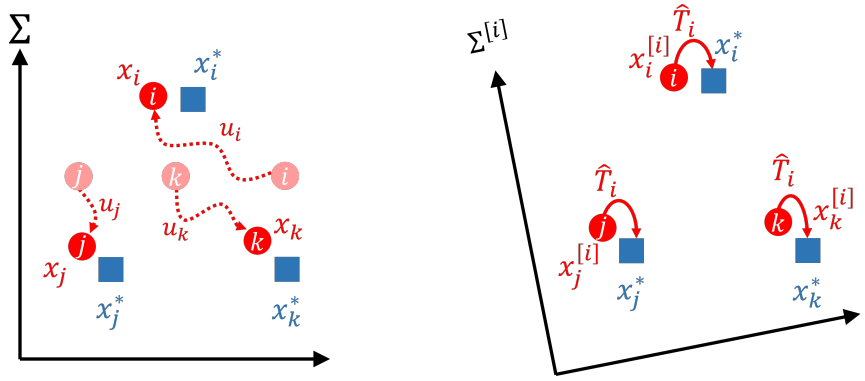
where $f_i : \mathbb{R}^{d\times|\mathcal{N}_i|} \to \mathrm{SE}(d)$, $g_i : \mathbb{R}^{d\times|\mathcal{N}_i|} \to \mathbb{R}^d$. We call (20) and (21) a distributed estimator and controller with relative measurements, respectively.

In summary, the target problem of this paper is given as follows.

Problem 1. Design a distributed estimator and controller (20), (21) with relative measurements that realize the objective (19) with a bias range $\mathcal{T}_c$, which contains I_{d+1} and has the smallest volume. Furthermore, quantify such $\mathcal{T}_c$.

4. Main Result

The main idea to solve Problem 1 is illustrated in Figure 6, where the solid circles represent the agent positions $x_i(t)$ and the blue squares represent a prescribed configuration, denoted by x_i^*. First, we control each agent to achieve a configuration as close to a prescribed one x_i^* as possible via a distributed controller (21), as shown in Figure 6(a). Next, each agent updates the estimation value $\hat{T}_i(t)$ via a distributed estimator (20) to estimate the transformation matrix $T_c^{-1}T_i$ from the common reference frame Σ_c. This is successful by comparing the measured positions $x^{[i]}_{\mathcal{N}_i}(t)$ of neighbors with the prescribed configuration $x^*_{\mathcal{N}_i}$, as shown in Figure 6(b). Then, a common



(a) Controller. (b) Estimator.

Figure 6. Illustration of main idea.

reference frame Σ_c is constructed.

Let $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_q \subset \mathcal{N}$ be the maximal cliques in $\mathcal{G}$ and let $\mathcal{Q}_i$ be the indices set of maximal cliques which agent i belongs to. We propose the following estimator and controller:

$$f_i(x^{[i]}_{\mathcal{N}_i}) = \arg\min_{T \in \mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - T \bullet x^{[i]}_{\mathcal{N}_i}\|, \tag{22}$$

$$g_i(x^{[i]}_{\mathcal{N}_i}) = -(x^{[i]}_i - \Phi_i(x^{[i]}_i) \bullet x^*_i) - \sum_{k \in \mathcal{Q}_i} (x^{[i]}_i - \Psi_{ik}(x^{[i]}_{\mathcal{C}_k}) \bullet x^*_i) \tag{23}$$

with functions $\Phi_i : \mathbb{R}^d \to \mathcal{T}^{-1}$ and $\Psi_{ik} : \mathbb{R}^{d \times |\mathcal{C}_k|} \to \langle \mathcal{T} \rangle$ satisfying

$$\Phi_i(x^{[i]}_i) = \arg\min_{T \in \mathcal{T}^{-1}} \|x^{[i]}_i - T \bullet x^*_i\|, \tag{24}$$

$$\Psi_{ik}(x^{[i]}_{\mathcal{C}_k}) = \arg\min_{T \in \langle \mathcal{T} \rangle} \|x^{[i]}_{\mathcal{C}_k} - T \bullet x^*_{\mathcal{C}_k}\|. \tag{25}$$

The estimator (22) computes the similarity between the prescribed configuration $x^*_{\mathcal{N}_i}$ and the measured relative positions $x^{[i]}_{\mathcal{N}_i}(t)$ with a transformation T, and the controller (23) moves the agent position $x_i(t)$ to the prescribed configuration x^*_i with a transformation T. Consequently, the solution T of the optimization problem (22) coincides to the transformation matrix $T_c^{-1} T_i$.

We consider the case for $d = 2$ and $\mathcal{T}$ in (10) to illustrate the way to solve (24) and (25). Then, explicit solutions can be derived as follows. First, (25) is associated with the pattern matching technique of image data [16] because $\langle \mathcal{T} \rangle = \mathrm{SE}(d)$ for $\mathcal{T}$ in (10) holds, and an explicit solution can be obtained. Second, (24) is reduced to

$$\begin{aligned} &\min_{(\theta, b) \in \mathbb{R} \times \mathbb{R}^2} \left\| x^{[i]}_i - (R(\theta) x^*_i + b) \right\| \\ &\text{subject to} \quad |\theta| \leq r_\theta, \quad \|b\| \leq r_b, \end{aligned} \tag{26}$$

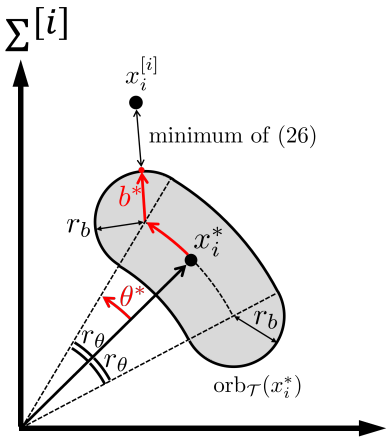


Figure 7. Solutions θ^* and b^* of (26).

where $R : \mathbb{R} \to \mathrm{SO}(2)$ is defined as

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{27}$$

The solutions $\theta^* \in \mathbb{R}$ and $b^* \in \mathbb{R}^2$ of the optimization problem (26) are derived of explicit forms as follows:

$$\theta^* = \begin{cases} \bar{\theta} & (|\bar{\theta}| < r_\theta) \\ r_\theta \frac{\bar{\theta}}{|\bar{\theta}|} & (|\bar{\theta}| \geq r_\theta) \end{cases} \tag{28}$$

$$b^* = \begin{cases} x_i^{[i]} - R(\theta^*)x_i^* & (\|x_i^{[i]} - R(\theta^*)x_i^*\| < r_b) \\ r_b \frac{x_i^{[i]} - R(\theta^*)x_i^*}{\|x_i^{[i]} - R(\theta^*)x_i^*\|} & (\|x_i^{[i]} - R(\theta^*)x_i^*\| \geq r_b) \end{cases}, \tag{29}$$

where $\bar{\theta} \in \mathbb{R}$ is denoted as

$$\bar{\theta} = \frac{x_i^* \times x_i^{[i]}}{|x_i^* \times x_i^{[i]}|} \arccos\left(\frac{(x_i^*)^\top x_i^{[i]}}{\|x_i^*\| \|x_i^{[i]}\|} \right) \tag{30}$$

with cross product $\times$ and inverse cosine function $\arccos : [-1, 1] \to [0, \pi]$. Figure 7 shows a geometric explanation of the solutions θ^* and b^* of the optimization problem (26). In this figure, the gray area represents the set of all $R(\theta)x_i^* + b$ satisfying $|\theta| \leq r_\theta$ and $\|b\| \leq r_b$, from which we must find the point nearest to $x_i^{[i]}$ to solve (26). In this case, the line connecting $x_i^{[i]}$ and the nearest point is orthogonal to the circle with radius r_b and center $R(r_\theta)x_i^*$. This point is represented as $R(\theta^*)x_i^* + b^*$ with θ^* and b^* in (28) and (29), respectively, in the latter cases.

The following theorem is the main result of this paper.

Theorem 1. *Suppose that framework $(\mathcal{G}, x_{\mathcal{N}}^*)$ is clique-rigid, and consider the system (14). The objective (19) is realized with the distributed estimator (22) and controller*

(23) with relative measurements for the following $\mathcal{T}_c \in \mathrm{SE}(d)$

$$\mathcal{T}_c = \{T \in \mathrm{SE}(d) : \exists x_{\mathcal{N}} \in \mathbb{R}^{d\times n} \ \text{s.t.} \ x^*_{\mathcal{N}} = T \bullet x_{\mathcal{N}},\ x_i \in \mathrm{orb}_{T_i\mathcal{T}^{-1}}(x^*_i) \ \ \forall i \in \mathcal{N}\}. \quad (31)$$

Proof. The controller (23) is derived from the gradient-based controller

$$g_i(x^{[i]}_{\mathcal{N}_i}(t)) = -M_i^{-1}\frac{\partial v}{\partial x_i}(x_{\mathcal{N}}) \quad (32)$$

with the objective function $v : \mathbb{R}^{d\times n} \to \mathbb{R}$

$$v(x_{\mathcal{N}}) = \sum_{k\in\mathcal{Q}} v_{ck}(x_{\mathcal{C}_k}) + \sum_{i\in\mathcal{N}} v_{ai}(x_i), \quad (33)$$

$$v_{ck}(x_{\mathcal{C}_k}) = \frac{1}{2}(\mathrm{dist}(x_{\mathcal{C}_k}, \mathrm{orb}_{\langle\mathcal{T}\rangle}(x^*_{\mathcal{C}_k})))^2, \quad (34)$$

$$v_{ai}(x_i) = \frac{1}{2}\det(M_i)^{\frac{2}{d}}(\mathrm{dist}(x_i, \mathrm{orb}_{\mathcal{T}^{-1}}(x^*_i)))^2. \quad (35)$$

By the controller (32), the state $x_{\mathcal{N}}(t)$ converges to a point where the function $v(x_{\mathcal{N}})$ takes the minimum. Actually,

$$\lim_{t\to\infty} \mathrm{dist}(x_{\mathcal{N}}(t), v^{-1}(0)) = 0, \quad (36)$$

is realized [17,18]. Moreover, the zero set $v^{-1}(0)$ of $v(x_{\mathcal{N}})$ is calculated from (24) and (25) as

$$\begin{aligned} v^{-1}(0) =& \{x_{\mathcal{N}} \in \mathbb{R}^{d\times n} : \exists T \in \mathrm{SE}(d)\, \text{s.t.}\, x_{\mathcal{N}} = T \bullet x^*_{\mathcal{N}}\} \\ & \cap \bigcap_{i\in\mathcal{N}} \{x_i \in \mathbb{R}^d : x_i \in \mathrm{orb}_{T_i\mathcal{T}^{-1}}(x^*_i)\} \end{aligned} \quad (37)$$

if framework $(G, x^*_{\mathcal{N}})$ is clique-rigid [17,18]. From (37), (36) is reduced to

$$\exists T_{\mathrm{c}} \in \mathcal{T}_{\mathrm{c}} \ \ \text{s.t.} \ \ \lim_{t\to\infty} x_{\mathcal{N}}(t) = T_{\mathrm{c}} \bullet x^*_{\mathcal{N}} \ \ \forall i \in \mathcal{N} \quad (38)$$

for $\mathcal{T}_{\mathrm{c}}$ in (31).

Next, we consider the effect of the estimator (22). The following relationship is obtained from (16)

$$\|x^*_{\mathcal{N}_i} - T \bullet x^{[i]}_{\mathcal{N}_i}(t)\| = \|x^*_{\mathcal{N}_i} - (TT_i^{-1}) \bullet x_{\mathcal{N}_i}(t)\|. \quad (39)$$

From (20), (22), and (39),

$$\begin{aligned} \hat{T}_i(t) &= \arg\min_{T\in\mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - T \bullet x^{[i]}_{\mathcal{N}_i}(t)\| \\ &= \arg\min_{T\in\mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - (TT_i^{-1}) \bullet x_{\mathcal{N}_i}(t)\| \\ &= \left(\arg\min_{\bar{T}\in\mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - \bar{T} \bullet x_{\mathcal{N}_i}(t)\|\right) T_i \end{aligned} \quad (40)$$

is achieved where $\bar{T} = TT_i^{-1}$. From (38) and (40), we obtain

$$\begin{aligned}\lim_{t\to\infty} \hat{T}_i(t) &= \left(\arg\min_{\bar{T}\in \mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - \bar{T} \bullet (T_\mathrm{c} \bullet x^*_{\mathcal{N}_i})\| \right) T_i \\ &= \left(\arg\min_{\bar{T}\in \mathrm{SE}(d)} \|x^*_{\mathcal{N}_i} - (\bar{T}T_\mathrm{c}) \bullet x^*_{\mathcal{N}_i}\| \right) T_i \\ &= T_\mathrm{c}^{-1} T_i. \end{aligned} \tag{41}$$

Here, $\bar{T} = T_\mathrm{c}^{-1}$ is only the solution of the second term of (41) because of the clique-rigidity. From (41) and $T_\mathrm{c} \in \mathcal{T}_\mathrm{c}$, (19) is obtained for $\mathcal{T}_\mathrm{c}$ in (31).

□

Theorem 1 guarantees that a common reference frame is constructed with the proposed estimator (22) and controller (23) with some bias T_c which belongs to the set $\mathcal{T}_\mathrm{c}$ in (31). This bias range $\mathcal{T}_\mathrm{c}$ with the proposed estimator and controller is the best one in the sense of the volume as follows.

Proposition 1. *The set $\mathcal{T}_c$ of (31) contains I_{d+1}. Moreover, under the estimation law with (22), the controller (23) provides the smallest volume of the sets for which the objective (19) is realized by distributed controllers with relative measurements.*

Proof. From the definition of $\mathcal{T}$, $T_i \in \mathcal{T}$ holds, yielding $I_{d+1} \in T_i\mathcal{T}^{-1}$. Hence, $\mathcal{T}_\mathrm{c}$ in (31) contains I_{d+1}.

From (38) and (41), the set $\mathcal{T}_\mathrm{c} \in \mathrm{SE}(d)$ in (38) is equivalent to (19). Hence, the volume of the set $\mathcal{T}_\mathrm{c}$ in (38) is equal to the volume of $\mathcal{T}_\mathrm{c}$ in (19). It is known that the controller (23) provides the smallest volume of the sets for which (38) is realized by distributed controllers with relative measurement [17]. Consequently, the controller (23) provides the smallest volume of the sets for which the objective (19) is realized by distributed controllers under the estimation law with (22). □

Remark 2. The main purpose of this study is to construct a common reference frame, expressed in (19), not formation control or attitude control. In this study, we assume that communication is unavailable, and thus (19) cannot be realized in common ways, including consensus protocol. Instead, we propose a method using formation control to realize (19) without communication. Formation control (attitude control) itself is not a control objective. Even if we expected to achieve attitude synchronization, it would not be easily achievable under our setting because we additionally assume that the attitude of other robots are unobservable.

5. Simulation

Simulations are conducted to evaluate the effectiveness of the proposed method. Consider a system of 12 agents in a two-dimensional space. The transformation matrix set $\mathcal{T}$ of local reference frames $\Sigma^{[i]}$ is (10) with $r_\theta = \pi/8$, $r_b = 0.5$. Each T_i is chosen from $\mathcal{T}$ randomly. The kinematic model of agents is given as (14), and the proposed estimator (22) and controller (23) are employed. The prescribed configuration x_i^* and

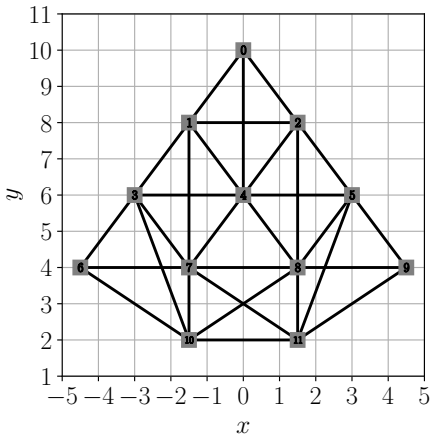
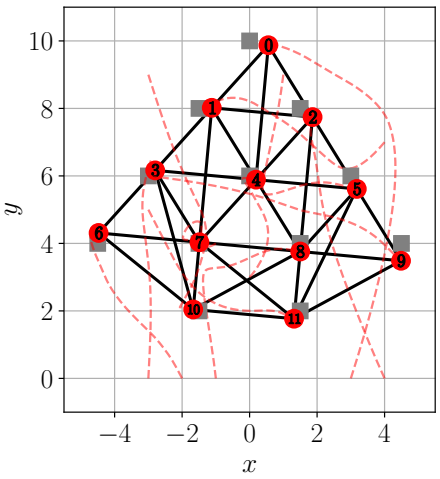
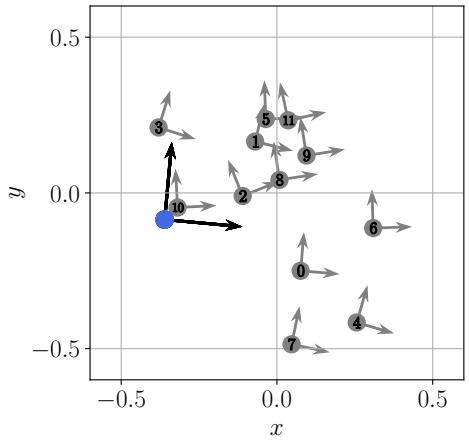


Figure 8. Prescribed configuration in simulations.



(a) Trajectories of agent positions in simulations



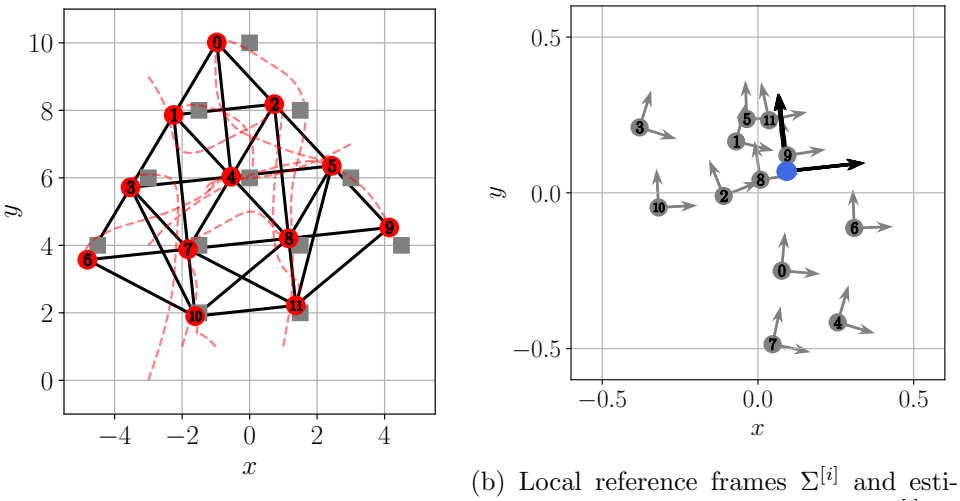
(b) Local reference frames $\Sigma^{[i]}$ and estimated common reference frames $\Sigma_{\mathrm{c}}^{[i]}$ in simulations.

Figure 9. Results of Simulation 1.

edges are depicted in Figure 8. The origins, the x-, and y- axes in this and the following figures correspond to the global frame Σ.

The result of simulation 1 is shown in Figure 9. The trajectories of agent positions are depicted in Figure 9(a). The gray squares represent the prescribed configurations x_i^*, and the black line segments represent edges. The red numbered circles represent the terminal positions of the agents, and the dashed lines represent the trajectories. The agents achieve the prescribed configuration with some translation and rotation. Figure 9(b) depicts local reference frames $\Sigma^{[i]}$ with the gray circles and arrows and the estimation of common reference frames $\Sigma_{\mathrm{c}}^{[i]}$ with the blue circles and arrows, corresponding to the transformation T_i and $T_{\mathrm{c}} = \lim_{t\to\infty} T_i \hat{T}_i^{-1}(t)$ from (19). It can be observed that the common reference frames Σ_{c} are successfully obtained and they are close to the global reference frame Σ.

Simulation 2 is conducted with the same T_i with different initial positions from Simulation 1. The result is shown in Figure 10. As in the case of simulation 1, the



(a) Trajectories of agent positions in simulations

(b) Local reference frames $\Sigma^{[i]}$ and estimated common reference frames $\Sigma_{\mathrm{c}}^{[i]}$ in simulations.

Figure 10. Results of Simulation 2.

agents achieve the prescribed configuration with some translation and rotation, but the terminal positions of the agents differ from simulation 1 because of the different initial positions. Moreover, it can be observed that different initial positions of the agents result in a different common reference frame Σ_{c}.

Next, 100 simulations are conducted under the same conditions from different initial positions. Figure 11 depicts the common reference frames constructed in 100 simulations. Each circle and the attaching arrow represent the origin and y-axis of the common reference frame Σ_{c}, respectively, given in each simulation. It is observed that the constructed common reference frames differ from each other according to the initial positions, and their origins are within a certain range. The same trend is confirmed for the orientations. These results imply that common reference frames are constructed with $T_{\mathrm{c}} \in \mathcal{T}_{\mathrm{c}}$ for $\mathcal{T}_{\mathrm{c}}$ in (31), as shown in Theorem 1.

6. Experiment

An experiment was conducted to confirm the practicability of the proposed method. Consider a system of 8 robots, TurtleBot 3 Burger [19] (Figure 12) running ROS (Robot Operating System), in a two-dimensional space.

Each robot was equipped with a single board computer (Raspberry Pi 3 Model B+) and a 360-degree two-dimensional laser range scanner (LDS-01), and could move forward, backward, and rotate with two wheels. This type of robot is under a non-holonomic constraint whereas the target system in (14) is not. Hence the actual input to this robot, the translation velocity $v_i(t) \in \mathbb{R}$ and rotation velocity $\omega_i(t) \in \mathbb{R}$, were derived from the input $u_i(t)$ in (21) as follows [20]:

$$v_i(t) = \|u_i(t)\| \cos\theta_i(t), \tag{42}$$
$$\omega_i(t) = \|u_i(t)\| \sin\theta_i(t), \tag{43}$$

where $\theta_i(t) \in \mathbb{R}$ is the angle of $u_i(t)$ to the forward direction of robot i in $\Sigma^{[i]}$ (Fig-

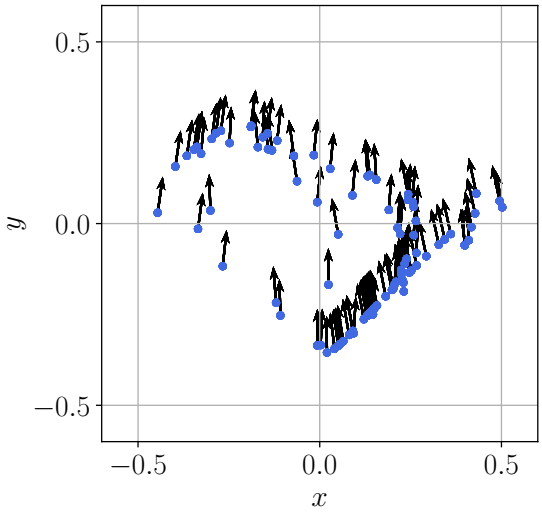


Figure 11. Constructed common reference frames in 100 simulations.

Figure 12. Robots used in the experiments.

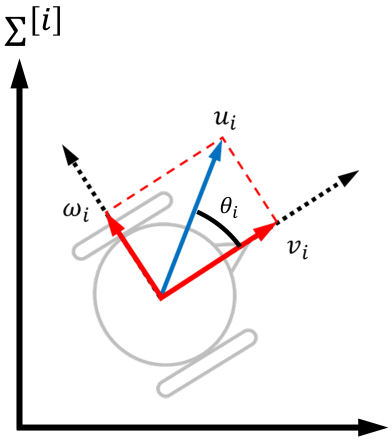


Figure 13. Two-wheeled robot model.

ure 13). Here, (42) represents the projection of $u_i(t)$ with respect to the forward direction, and (43) represents the projection of $u_i(t)$ with respect to the normal to the forward direction. To avoid collision between robots, repulsive force was added to the control input. Besides, the input was multiplied by some gain.

The transformation matrix set of the local reference frame $\mathcal{T}$ is given as (10) with $r_\theta = \pi/8$, $r_b = 0.4$. The global position x_i is measured with a motion capture system, and a random bias $T_i \in \mathcal{T}$ is applied as an artificial noise. Then, the resultant is regarded as the measured position $x_i^{[i]}$. This is to show the validity of the method more clearly. The relative positions of neighbor robots $x_{\mathcal{N}}^{[i]}$ were measured with LiDAR in the following procedure: (i) the initial positions of neighbor robots were given with their numbers; (ii) their relative positions were tracked by local measurements with LiDAR; (iii) If LiDAR tracking failed due to occlusion, etc., lost positions were complemented through wireless communication. Figure 14 describes the prescribed configuration x_i^* and edges. The controller and estimator of a common reference frame were given by (23) and (22), respectively.

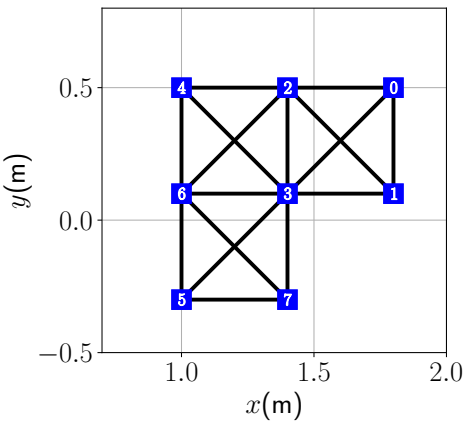


Figure 14. Prescribed configuration and edges.

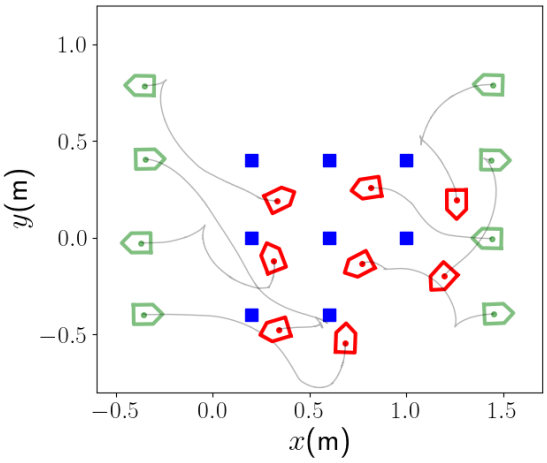


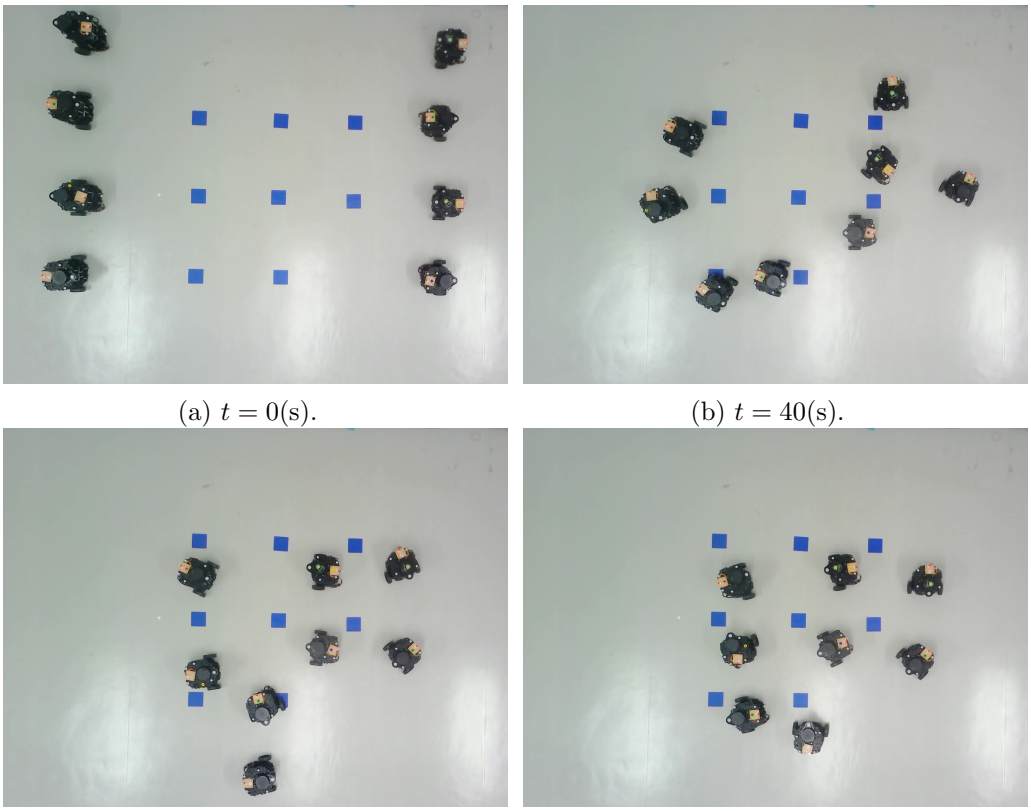
Figure 15. Trajectories of robot positions from $t = 0$(s) to 240(s).

The trajectories of the robot positions from $t = 0$(s) to 240(s) in the experiment are shown in Figure 15. The green pentagons represent the robot positions at $t = 0$(s), the red pentagons represent the positions at $t = 240$(s), and the blue squares represent the prescribed configuration in Figure 14. Also, the snapshots in the experiment are shown in Figure 16. Figures 15 and 16(d) show that the terminal positions of the robots and prescribed configuration have nearly a congruent relationship.

Next, the local and estimated common frames, $\Sigma^{[i]}$ and $\Sigma_c^{[i]}$, are shown in Figure 17(a), corresponding to the transformation matrices T_i and $T_i\hat{T}_i^{-1}(t)$, with the gray and light blue numbered circles, respectively. Figure 17(a) indicates that all $\Sigma_c^{[i]}$ nearly converge to the same frame close to Σ. The axis angles of $\Sigma^{[i]}$ and $\Sigma_c^{[i]}$ with Σ are shown in Figure 17(b), where the light blue lines represent the angles of $\Sigma_c^{[i]}$ and the gray lines do those of $\Sigma^{[i]}$. The x and y values of origin in each reference frame are depicted in Figures 17(c) and 17(d). These figures show that the estimated common reference frames $\Sigma_c^{[i]}$ approach a common reference frame as the prescribed configuration is achieved. (See also Figure 16.) Note that $\Sigma_c^{[i]}$ do not completely coincide because of sensor noise.

7. Conclusion

In this paper, we developed a distributed method to construct a common reference frame for all agents close to the global reference frame using only local measurements. The proposed method is composed of formation control and similarity evaluation of feature point matching. The error range between the constructed common reference frame and the global reference frame is quantified and shown to be optimal in terms of volume. Finally, simulations and an experiment were conducted to verify the effectiveness of the proposed method. The simulation result showed that the error of the constructed common reference frame was within the range around the global reference frame. The experiment result indicated that the system worked effectively in a real



(a) $t = 0$(s). (b) $t = 40$(s).

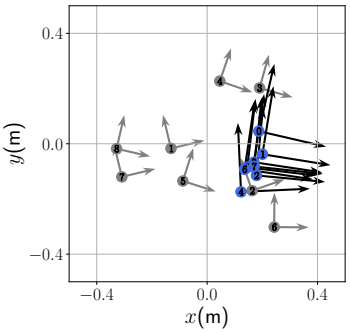
(c) $t = 120$(s). (d) $t = 240$(s).

Figure 16. Snap Shots of the experiment.

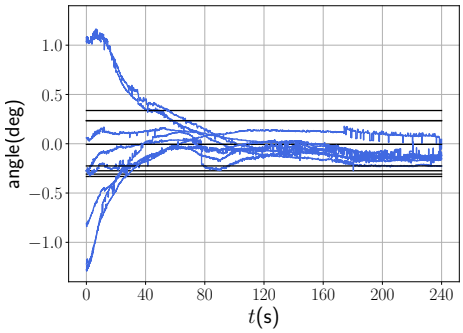
environment. Future work is to reduce the effect of the sensor noise to construct a completely common reference frame.

References

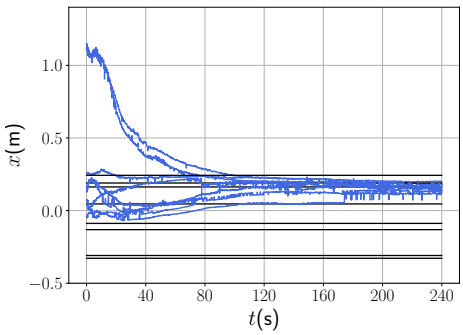
[1] Sahoo A, Dwivedy SK, Robi P. Advancements in the field of autonomous underwater vehicle. Ocean Eng. 2019;181:145–160.
[2] Suzuki K, Ishibashi J, Kato Y, et al. Preface: Front edge of submarine mineral resources research in japan. Geochem J. 2015;49(6):575–577.
[3] Sangekar MN, Thornton B, Bodenmann A, et al. Autonomous landing of underwater vehicles using high-resolution bathymetry. IEEE J Ocean Eng. 2020;45(4):1252–1267.
[4] Saska M, Vonásek V, Chudoba J, et al. Swarm distribution and deployment for cooperative surveillance by micro-aerial vehicles. J Intell Robot Syst. 2016;84:469–492.
[5] Loianno G, Kumar V. Cooperative transportation using small quadrotors using monocular vision and inertial sensing. IEEE Robot Autom Lett. 2018;3(2):680–687.
[6] Oh KK, Park MC, Ahn HS. A survey of multi-agent formation control. Automatica. 2015; 53:424–440.
[7] Flocchini P, Prencipe G, Santoro N, et al. Arbitrary pattern formation by asynchronous, anonymous, oblivious robots. Theor Comput Sci. 2008 11;407:412–447.
[8] Panigrahi PK, Bisoy SK. Localization strategies for autonomous mobile robots: A review. J King Saud Univ - Comput Inf Sci. 2021;.
[9] Franceschelli M, Gasparri A. Gossip-based centroid and common reference frame estima-



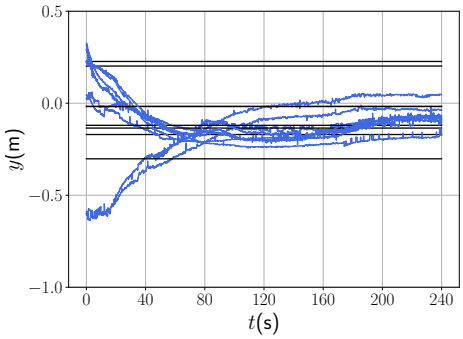
(a) Local reference frames $\Sigma^{[i]}$ and estimated common reference frames $\Sigma_c^{[i]}$.



(b) Axes angles of $\Sigma^{[i]}$ and $\Sigma_c^{[i]}$ with Σ.



(c) x values of origin of $\Sigma^{[i]}$ and $\Sigma_c^{[i]}$.



(d) y values of origin of $\Sigma^{[i]}$ and $\Sigma_c^{[i]}$.

Figure 17. Estimation result in the experiment.

tion in multiagent systems. IEEE Trans Robot. 2014;30(2):524–531.
[10] Lopez-Limon C, Franceschelli M, Seatzu C, et al. A consensus algorithm for common reference frame estimation in networked multi-agent systems. In: 53rd IEEE Conference on Decision and Control (CDC); 2014. p. 6117–6122.
[11] Tran QV, Ahn HS. Multi-agent localization of a common reference coordinate frame: An extrinsic approach. IFAC-PapersOnLine. 2019;52(20):67–72.
[12] Haxhibeqiri J, Jarchlo EA, Moerman I, et al. Flexible wi-fi communication among mobile robots in indoor industrial environments. Mob Inf Syst. 2018;2018.
[13] Zeng Y, Zhang R, Lim TJ. Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun Mag. 2016;54(5):36–42.
[14] Sakurama K, Azuma S, Sugie T. Multiagent coordination via distributed pattern matching. IEEE Trans Autom Control. 2019;64(8):3210–3225.
[15] Park G, Lee B, Sung S. Integrated pose estimation using 2d lidar and ins based on hybrid scan matching. Sensors. 2021;21(16).
[16] Kanatani K. Analysis of 3-d rotation fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1994;16(5):543–549.
[17] Sakurama K. Formation control of multi-agent systems with generalized relative measurements. In: 2020 59th IEEE Conference on Decision and Control (CDC); 2020. p. 2799–2804.
[18] Sakurama K. Unified formulation of multiagent coordination with relative measurements. IEEE Trans Autom Control. 2021;66(9):4101–4116.
[19] TurtleBot3 [Internet] California: Open Source Robotics Foundation I ; [cited 2022 March

26]. Available from: https://www.turtlebot.com/.
[20] Sakurama K. Formation control of non-holonomic multi-agent systems under relative measurements. In: 21st IFAC World Congress; 2020. p. 11006–11011.