2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Visual Feedback Pose Synchronization with A Generalized Camera Model

Tatsuya Ibuki, Takeshi Hatanaka, Masayuki Fujita and Mark W. Spong

***Abstract*— This paper investigates visual feedback pose synchronization in leader-follower type visibility structures on the Special Euclidean group $SE(3)$. We first define visual robotic networks with a generalized camera model. We then propose a visual feedback pose synchronization law combining a vision-based observer with the pose synchronization law presented in our previous works. We then prove that for a static leader, the network with the control law achieves visual feedback pose synchronization. Moreover, for a moving leader, we evaluate the tracking performance of the network. Finally, the validity of the proposed control law and analysis is demonstrated through experiments.**

I. Introduction

Mobile sensor networks [1] are collections of interconnected multiple mobile robots with sensing devices and computing capability. Mobile sensor networks have potential advantages in performances, robustness against sensor failures and various sensor-driven tasks especially in dynamical environments. In operation, each sensor is required to act cooperatively using only limited information so that the total system attains specified behaviors. Cooperative control gives methods for such distributed control [2], [3].

Cooperative control problems for mobile sensor networks are formulated as pose (position and attitude) coordination problems [1]. In this paper, we tackle pose synchronization as one of such problems whose objective is to lead agents' poses to a desired one by utilizing distributed control strategies.

In the stage of implementation, it is unavoidable to consider how to acquire neccesary information for cooperative control laws. In multi-agent systems, agents might be capable of communicating with neighboring agents, where measurements on global information might be assumed, or measuring relative information with respect to neighbors via relative sensors without communication. In this paper, we address the latter scenario for ease of implementation and cost reduction. Among such relative measurements, a visual sensor brings one rich information including three-dimensional pose of the other agents compressed into two-dimensional image plane.

While numerous research works have been devoted to the combination of control techniques with vision [4],[5], vision-based cooperative control is also tackled [6]-[8]. In one of our previous works [6], we presented a leader-following visual feedback pose synchronization law. However, we have proved only convergence with a not moving leader and not analyzed the tracking performance of the network to a moving leader. Moreover, since we have considered pinhole cameras as vision sensors, a field of view may not be sufficiently wide.

T. Ibuki, T. Hatanaka and M. Fujita are with the Department of Mechanical and Control Engineering, Tokyo Institute of Technology, Tokyo 152-8552, JAPAN fujita@ctrl.titech.ac.jp

M. W. Spong is with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080 mspong@utdallas.edu

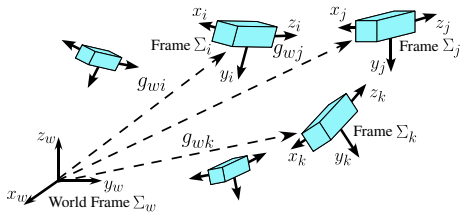

Fig. 1: Rigid Body Motion

In this paper, we investigate a leader-following pose synchronization problem for a network of rigid bodies with a generalized camera model. We first introduce a notion of visual robotic networks to be controlled. Here, we introduce not only pinhole camera models [6] but also generalized panoramic camera models consisting of a pinhole camera and a hyperbolic mirror in order to get a wider field of view. After defining visual feedback pose synchronization, we present a synchronization law consisting of a vision-based observer and synchronization law. We then prove synchronization for a static leader. Furthermore, we analyze the tracking performance of the network for a moving leader. The effectiveness of the control scheme is demonstrated through experiments on a planar testbed.

The main contributions of this paper are follows: (i) We deal with a wider class of camera models than usual pinhole camea models. (ii) We give the tracking performance analysis of the network for a moving leader. (iii) We perform the experiment in order to confirm the effectiveness of the proposed control law.

II. Visual Robotic Network

A. Rigid Body Motion

In this paper, we consider a network of n rigid bodies in three-dimensional space (see Fig. 1). Let Σ_w be an inertial coordinate frame and $\Sigma_i,\ i \in \mathcal{V} := \{1, \cdots, n\}$ body-fixed coordinate frames. We denote the pose of body i in Σ_w by $(p_{wi}, e^{\hat{\xi}_{wi}\theta_{wi}}) \in SE(3)$ or homogeneous representation

$$g_{wi} = \begin{bmatrix} e^{\hat{\xi}_{wi}\theta_{wi}} & p_{wi} \\ 0 & 1 \end{bmatrix} \in SE(3),\ i \in \mathcal{V}.$$

Here, $\xi_{wi} \in \mathcal{R}^3$ $(\xi_{wi}^{\mathrm{T}}\xi_{wi} = 1)$ and $\theta_{wi} \in \mathcal{R}$ specify the direction and angle of rotation, respectively. For simplicity, we use $\hat{\xi}\theta_{wi}$ to denote $\hat{\xi}_{wi}\theta_{wi}$.

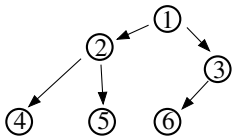

Fig. 2: Leader-follower Type Visibility Structure

Let us now introduce the velocity of each rigid body to represent rigid body motion of Σ_i relative to Σ_w. We define the body velocity of body i relative to Σ_w as $V_{wi}^b = [(v_{wi}^b)^{\mathrm{T}}\ (\omega_{wi}^b)^{\mathrm{T}}]^{\mathrm{T}} := (g_{wi}^{-1}\dot{g}_{wi})^{\vee} \in \mathcal{R}^6$, where $v_{wi}^b \in \mathcal{R}^3$ and $\omega_{wi}^b \in \mathcal{R}^3$ represent the linear and angular velocities. Then, rigid body motion is represented by the kinematic model

$$\dot{g}_{wi} = g_{wi}\hat{V}_{wi}^b,\ i \in \mathcal{V}. \tag{1}$$

We denote the pose of Σ_j relative to Σ_i as $g_{ij} = (p_{ij}, e^{\hat{\xi}\theta_{ij}}) := g_{wi}^{-1}g_{wj} \in SE(3)$. Then, differentiating g_{ij} with respect to time yields relative rigid body motion

$$V_{ij}^b := (g_{ij}^{-1}\dot{g}_{ij})^{\vee} = -\mathrm{Ad}_{(g_{ij}^{-1})}V_{wi}^b + V_{wj}^b, \tag{2}$$

where $\mathrm{Ad}_{(g_{ij})} \in \mathcal{R}^{6\times 6}$ is the adjoint transformation associated with g_{ij} [9].

B. Visibility Structure

We describe visibility structures among rigid bodies. Throughout this paper, we assume each body has vision to capture other visible bodies. A set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is defined so that $(j,i) \in \mathcal{E}$ means body j is visible from body i. We next define the set of visible bodies from body i as

$$\mathcal{N}_i := \{j \in \mathcal{V} \mid (j,i) \in \mathcal{E}\},\ i \in \mathcal{V}. \tag{3}$$

Let us now make the following assumptions on the visibility structure.

Assumption 1: Visibility structures have leader-follower structures (see Fig. 2 [6]).

C. Visual Measurement

Suppose that each rigid body j has s ($s \geq 4$) feature points, whose positions relative to Σ_j are denoted by $p_{jj_k} \in \mathcal{R}^3,\ k \in \{1, \cdots, s\}$. A coordinate transformation yields the positions of feature points relative to frame Σ_i as $p_{ij_k} = g_{ij}p_{jj_k}$, where p_{ij_k} and p_{jj_k} should be regarded as $[p_{ij_k}^{\mathrm{T}}\ 1]^{\mathrm{T}}$ and $[p_{jj_k}^{\mathrm{T}}\ 1]^{\mathrm{T}}$, respectively [9].

Let us now define visual measurements of rigid bodies with panoramic camera models. We first introduce feature points obtained by pinhole camera models. We denote the k-th feature point on the image plane as $f_{ij_k} \in \mathcal{R}^2$. Then, by perspective projection [9], f_{ij_k} is given by

$$f_{ij_k} = \frac{\lambda_i}{z_{ij_k}}\begin{bmatrix} x_{ij_k} \\ y_{ij_k} \end{bmatrix}, \tag{4}$$

where $p_{ij_k} = [x_{ij_k}\ y_{ij_k}\ z_{ij_k}]^{\mathrm{T}}$ and $\lambda_i \in \mathcal{R}$ is a focal length.

We next define feature points obtained by panoramic camera models [5]. A panoramic camera model consists of the pinhole camera model and a hyperbolic mirror (see Fig. 3). We denote the pose of rigid body i's mirror coordiate

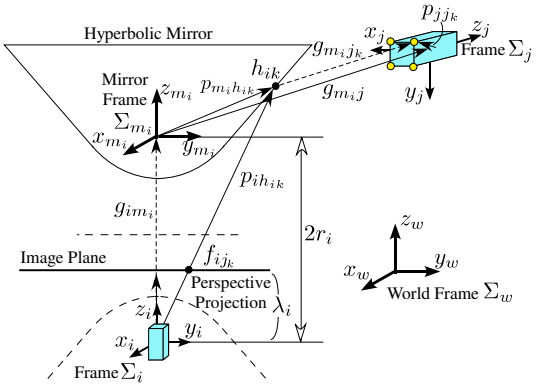

Fig. 3: Panoramic Camera Model

Σ_{m_i} relative to Σ_w by $g_{wm_i} = (p_{wm_i}, e^{\hat{\xi}\theta_{wm_i}}) \in SE(3)$, pose and position of k-th feature point of body j relative to Σ_{m_i} by $g_{m_ij} = (p_{m_ij}, e^{\hat{\xi}\theta_{m_ij}}) \in SE(3)$ and $p_{m_ij_k} \in \mathcal{R}^3$, respectively. Then, similarly to (2), relative rigid body motion of body j relative to the mirror of body i is represented by

$$V_{m_ij}^b = -\mathrm{Ad}_{(g_{m_ij}^{-1})}\mathrm{Ad}_{(g_{im_i}^{-1})}V_{wi}^b + V_{wj}^b, \tag{5}$$

where $g_{im_i} = (p_{im_i}, e^{\hat{\xi}\theta_{im_i}}) \in SE(3)$ is the pose of Σ_{m_i} relative to Σ_i.

We denote the point on body i's mirror by h_{ik} as shown in Fig. 3 and the position of the point relative to Σ_{m_i} by $p_{m_ih_{ik}} := [x_{m_ih_{ik}}\ y_{m_ih_{ik}}\ z_{m_ih_{ik}}]^{\mathrm{T}} \in \mathcal{R}^3$. Then, body i gets body j's feature point f_{ij_k} on its image plane which is p_{ij_k} projected onto the plane through h_{ik}.

Let a_i, b_i and $r_i := \sqrt{a_i^2 + b_i^2}$ be the hyperbolic mirror parameters satisfying

$$\frac{(z_{m_ih_{ik}} + r_i)^2}{a_i^2} - \frac{x_{m_ih_{ik}}^2 + y_{m_ih_{ik}}^2}{b_i^2} = 1. \tag{6}$$

We denote the position of point h_{ik} relative to Σ_i by $p_{ih_{ik}} := [x_{ih_{ik}}\ y_{ih_iik}\ z_{ih_{ik}}]^{\mathrm{T}} \in \mathcal{R}^3$. Then, by perspective projection (4), $f_{ij_k} = (\lambda_i/z_{ih_{ik}})[x_{ih_{ik}}\ y_{ih_{ik}}]^{\mathrm{T}}$ holds. We moreover set $p_{im_i} = [0\ 0\ 2r_i]^{\mathrm{T}}$, $e^{\hat{\xi}\theta_{im_i}} = I_3$ and $p_{m_ih_{ik}} = cp_{m_ij_k}$ ($0 < c < 1$) (see Fig. 3). Then, the following equation holds by substituting $cp_{m_ij_k}$ into (6).

$$c(p_{m_ij_k}) = \frac{b_i^2(r_iz_{m_ij_k} + a_i\|p_{m_ij_k}\|_2)}{a_i^2x_{m_ij_k}^2 + a_i^2y_{m_ij_k}^2 - b_i^2z_{m_ij_k}^2}.$$

Finally, since $z_{ih_{ik}} = 2r_i + c(p_{m_ij_k})z_{m_ij_k}$ holds,

$$f_{ij_k} = \frac{\lambda_i c(p_{m_ij_k})}{2r_i + c(p_{m_ij_k})z_{m_ij_k}}\begin{bmatrix} x_{m_ij_k} \\ y_{m_ij_k} \end{bmatrix}. \tag{7}$$

If we select $a_i = 1$ and $b_i = \sqrt{-1}$, then the panoramic camera's feature point (7) is equivalent to the pinhole camera's one (4). Namely, a panoramic camera model includes a pinhole comera model as a special case.

We assume each rigid body can extract the feature points (7) of visible bodies from its image data. We thus define visual measurements of body i as

$$f_i := \{f_{ij}\}_{j\in\mathcal{N}_i},\ i \in \mathcal{V},\ f_{ij} := [f_{ij_1}^{\mathrm{T}}\ \cdots\ f_{ij_s}^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{R}^{2s}. \tag{8}$$

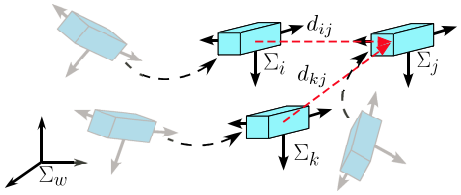

Fig. 4: Pose Synchronization

Hereafter, the aggregate system consisting of n rigid bodies with kinematic model (1), visibility structures (3) satisfying Assumption 1 and visual measurements (8) is called visual robotic network Σ.

III. Visual Feedback Pose Synchronization

A. Definition of Visual Feedback Pose Synchronization

We first define the virtual relative pose $\tilde{g}_{ij} \in SE(3)$ as

$$\tilde{g}_{ij} := \begin{bmatrix} e^{\hat{\xi}\theta_{ij}} & p_{ij} - d_{ij} \\ 0 & 1 \end{bmatrix}, \ i, j \in \mathcal{V},$$

where $d_{ij} \in \mathcal{R}^3, \ i, j \in \mathcal{V}$ are constant biases such that each rigid body guarantees collision avoidance and visibility to neighbors in the final configuration. We assume each body has biases relative to its neighbors $d_{ij}, \ j \in \mathcal{N}_i$ (see Fig. 4).

Then, the goal of this paper is to design a body velocity input V^b_{wi} so that the visual robotic network Σ achieves visual feedback pose synchronization defined below.

Definition 1: On the visual robotic network Σ, a control input V^b_{wi} is said to achieve visual feedback pose synchronization if V^b_{wi} depends only on visual measurement (8) and

$$\lim_{t\to\infty} \Pi(\tilde{g}_{ij}) = 0 \ {}^\forall i, j \in \mathcal{V}. \tag{9}$$

Here, $\Pi(g_{wi}) := (1/2)\|p_{wi}\|_2^2 + \phi(e^{\hat{\xi}\theta_{wi}}) \geq 0$, $\phi(e^{\hat{\xi}\theta_{wi}}) := (1/2)\mathrm{tr}(I_3 - e^{\hat{\xi}\theta_{wi}}) \geq 0$ is the energy of pose errors. By the definition, $\Pi(g_{wi}) = 0$ if and only if $g_{wi} = I_4$.

Equation (9) means that relative positions among rigid bodies converge to desired ones and orientations to a common value (see Fig. 4). Unlike [3] premising the measurement of g_{ij}, the objective of this paper is to present a velocity law using only visual measurements (8).

B. Visual Feedback Pose Synchronization Law

We introduce the structure of visual feedback pose synchronization law. First of all, each rigid body has to estimate relative pose g_{m_ij} by a nonlinear observer since visual measurements (8) are two-dimensional. Hereafter, we denote the estimate of g_{m_ij} by $\bar{g}_{m_ij} = (\bar{p}_{m_ij}, e^{\hat{\bar{\theta}}_{m_ij}}) \in SE(3)$.

We next define the desired relative pose $g_{dij} = (d_{ij}, I_3) \in SE(3)$, control error $g_{cij} = (p_{cij}, e^{\hat{\xi}\theta_{cij}}) \in SE(3)$ and control error vector $e_{cij} \in \mathcal{R}^6$ as

$$g_{cij} := g^{-1}_{dij} g_{im_i} \bar{g}_{m_ij}, \ e_{cij} := \begin{bmatrix} p_{cij} \\ \mathrm{sk}(e^{\hat{\xi}\theta_{cij}})^\vee \end{bmatrix}.$$

Here, $\mathrm{sk}(e^{\hat{\xi}\theta_{ij}}) \in \mathcal{R}^{3\times 3}$ is the skew-symmetric part of matrix $e^{\hat{\xi}\theta_{ij}}$. Note that $e_{cij} = 0$ if and only if $g_{cij} = I_4$ as long as $|\theta_{cij}| < \pi$ and hence $g_{im_i}\bar{g}_{m_ij} = g_{dij}$.

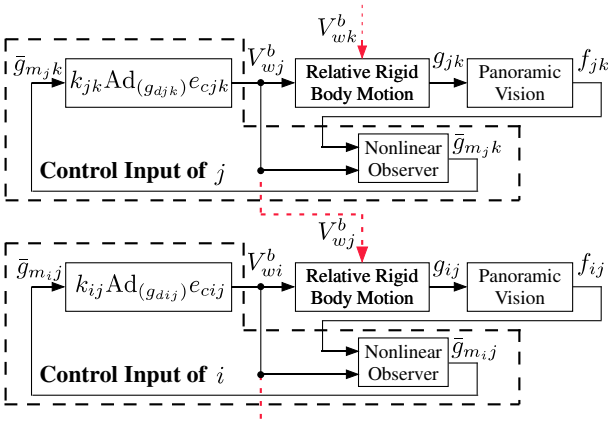

Fig. 5: Total Control System

We also define the estimation error $g_{eij} = (p_{eij}, e^{\hat{\xi}\theta_{eij}}) \in SE(3)$ between the actual relative pose g_{m_ij} and its estimate $\bar{g}_{m_ij}$ and its vector representation $e_{eij} \in \mathcal{R}^6$ as

$$g_{eij} := \bar{g}^{-1}_{m_ij} g_{m_ij}, \ e_{eij} := \begin{bmatrix} p_{eij} \\ \mathrm{sk}(e^{\hat{\xi}\theta_{eij}})^\vee \end{bmatrix}.$$

Note that $e_{eij} = 0$ if and only if $g_{eij} = I_4$ as long as $|\theta_{eij}| < \pi$, and hence $\bar{g}_{m_ij} = g_{m_ij}$. Therefore, both $e_{cij} = 0$ and $e_{eij} = 0$ mean $g_{ij} = g_{dij}$ as long as $|\theta_{cij}| < \pi, \ |\theta_{eij}| < \pi$. Namely, equation (9) is achieved if and only if $e_{cij} = 0, \ e_{eij} = 0, \ j \in \mathcal{N}_i \ {}^\forall i \in \mathcal{V}$.

By using above notations, we propose the following control law.

$$\textit{Controller}: \ V^b_{wi} = k_{ij}\mathrm{Ad}_{(g_{dij})}e_{cij}, \tag{10a}$$

$$\textit{Observer} \begin{cases} \bar{V}^b_{m_ij} = -\mathrm{Ad}_{(\bar{g}^{-1}_{m_ij})}\mathrm{Ad}_{(g^{-1}_{im_i})}V^b_{wi} + u_{ij} & \text{(10b)} \\ u_{ij} = k_{eij}\left(e_{eij} - \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{cij}})}e_{cij}\right), \\ \qquad\qquad j \in \mathcal{N}_i, \ i \in \mathcal{V}. & \text{(10c)} \end{cases}$$

where $k_{ij}, \ k_{eij} \in \mathcal{R}$ are positive gains. The total control system is shown in Fig. 5.

Velocity input (10a) is the same as that in [3] except for using $g_{im_i}\bar{g}_{m_ij}$ instead of g_{ij}. Equation (10b) simulates relative rigid body motion (5) by using the estimate $\bar{g}_{m_ij}$ as its state. Here, $u_{ij} \in \mathcal{R}^6$ is an external input to be determined so that the estimated values $\bar{g}_{m_ij}$ and $\bar{V}^b_{m_ij} := (\bar{g}^{-1}_{m_ij}\dot{\bar{g}}_{m_ij})^\vee \in \mathcal{R}^6$ are driven to their actual values. Differentiating g_{eij} with respect to time and utilizing (5) and (10b), we get the following estimation error system.

$$V^b_{eij} := (g^{-1}_{eij}\dot{g}_{eij})^\vee = -\mathrm{Ad}_{(g^{-1}_{eij})}u_{ij} + V^b_{wj}. \tag{11}$$

In (10c), e_{eij} can be reconstructed by visual measurements f_{ij} [5]. This means that the present control law (10) can be calculated only by visual measurements (8) in the absence of communication or any measurements of own states. It is thus sufficient for visual feedback pose synchronization to prove (9).

C. Convergence Analysis

We prove that control law (10) on the visual robotic network Σ achieves (9). Differentiating g_{cij} with respect to

time gives

$$V_{cij}^b := (g_{cij}^{-1}\dot{g}_{cij})^\vee = -\mathrm{Ad}_{(g_{cij}^{-1})}\mathrm{Ad}_{(g_{dij}^{-1})}V_{wi}^b + u_{ij}. \quad (12)$$

This system is the same as (10b) and called control error system.

We next consider the total system combining control error system (12) with estimation error system (11) as

$$\begin{bmatrix} V_{cij}^b \\ V_{eij}^b \end{bmatrix} = \begin{bmatrix} -\mathrm{Ad}_{(g_{cij}^{-1})} & I_6 \\ 0 & -\mathrm{Ad}_{(g_{eij}^{-1})} \end{bmatrix} \begin{bmatrix} \mathrm{Ad}_{(g_{dij}^{-1})}V_{wi}^b \\ u_{ij} \end{bmatrix} + \begin{bmatrix} 0 \\ V_{wj}^b \end{bmatrix}. \quad (13)$$

In this paper, the collection of the combining system (13) for $j \in \mathcal{N}_i,\ i \in \mathcal{V}$ with control law (10) is called collective error system Σ_{col}, whose state, denoted by $x_e \in \mathcal{R}^{12(n-1)}$, is given by the stuck vector of $e_{ij} := [e_{cij}^{\mathrm{T}} \quad e_{eij}^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{R}^{12},\ j \in \mathcal{N}_i,\ i \in \mathcal{V}$.

Let us now show that control law (10) on the visual robotic network Σ achieves visual feedback pose synchronization.

Theorem 1: Suppose that the leader does not move ($V_{w1}^b = 0$). Then, control law (10) on the visual robotic network Σ achieves visual feedback pose synchronization at least locally if

$$\begin{cases} I_6 - D_{i1} > 0,\ i \in \mathcal{V}_p \\ k_{jk} < \frac{2k_{ij}k_{eij}}{k_{ij}+k_{eij}},\ k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_q \\ \begin{cases} k_{jk} < 2k_{eij} \\ k_{jk}(k_{eij}I_6 + k_{ij}(I_6 - D_{ij})) < 2k_{ij}k_{eij}(I_6 - D_{ij}) \end{cases}, \\ \qquad\qquad k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_r \end{cases} \quad (14)$$

where $D_{ij} := (1/2)\mathrm{Ad}_{(g_{dij})}^{\mathrm{T}}\mathrm{Ad}_{(g_{dij})}$ (refer to [6] for the definitions of V_*).

Proof: From the definition of x_e, if the equilibrium point $x_e = 0$ is asymptotically stable, then local visual feedback pose synchronization is achieved. This can be proved by differentiating the following Lyapunov function candidate with respect to time and utilizing completing square.

$$U := \sum_{i=2}^{n}\sum_{j\in\mathcal{N}_i} q_i \left(\Pi(g_{cij}) + \Pi(g_{eij})\right) \geq 0.$$

Here, $q_i \in \{1, 2, 3, \cdots\}$ is corresponding natural numbers [6]. Note that $U = 0$ if and only if $g_{cij} = I_4,\ g_{eij} = I_4$ (i.e. $e_{cij} = 0,\ e_{ij} = 0$ as long as $|\theta_{cij}| < \pi,\ |\theta_{eij}| < \pi),\ j \in \mathcal{N}_i\ ^\forall i \in \mathcal{V}$ and otherwise $U > 0$. ■

Gain conditions (14) imply that if the backward rigid bodies move fast, then visual feedback pose synchronization is achieved. For example, if we set $d_{ij} = 0\ ^\forall i, j \in \mathcal{V}$, conditions (14) are represented by

$$\begin{cases} k_{jk} < \frac{2k_{ij}k_{eij}}{k_{ij}+k_{eij}},\ k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_q \\ k_{jk} < \frac{2k_{ij}k_{eij}}{k_{ij}+2k_{eij}},\ k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_r \end{cases}.$$

These conditions mean that the forward body moves more slowly than the backward one [6]. This explains the intuition that motion of forward bodies has large influences on group motion while that of backward ones has small impact. It is noted that conditions (14) are reduced to linear matrix inequalities on control gains. Thus, we can find gains by using existing solvers if it is feasible.

It should be noted that Theorem 1 proves synchronization for the system integrating the observers instead employing certainly equivalence principle. It is well known in robot control that proving stability for the integrated system in observer-based control strategies is much more difficult than the individual control and estimation problems even for a single passive system [10]. It should be also true or might be much harder for synchronization since it is required to estimate not their own but the other individuals' information only from relative measurements.

D. Pose Synchronization with Desired Velocities

In the previous subsection, we have presented a control law to achieve visual feedback pose synchronization in the sense of (9). Notice now that all rigid bodies would stop in the final configuration though it is sometimes required for bodies to move in the desired direction while achieving pose synchronization. We thus add a common desired velocity to all bodies in this subsection.

Suppose that all rigid bodies have a common desired velocity $V_d \in \mathcal{R}^6$ and each body knows the velocity in its own coordinate frame $\mathrm{Ad}_{(e^{-\hat{\xi}\theta_{wi}})}V_d$. Let us fix the form of each body velocity as

$$V_{wi}^b = \tilde{V}_{wi}^b + \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{wi}})}V_d$$

for some $\tilde{V}_{wi}^b$. Then, relative rigid body motion (2) can be represented by

$$V_{ij}^b = -\mathrm{Ad}_{(g_{ij}^{-1})}\tilde{V}_{wi}^b + \tilde{V}_{wj}^b. \quad (15)$$

Also, estimation error system (11) is derived as

$$V_{eij}^b = -\mathrm{Ad}_{(g_{eij}^{-1})}u_{ij} + \tilde{V}_{wj}^b. \quad (16)$$

Equations (15) and (16) mean that control and estimation error systems do not change except for using $\tilde{V}_{wi}^b$ instead of V_{wi}^b. Therefore, we propose the following control law.

$$\textit{Controller}:\ V_{wi}^b = k_{ij}\mathrm{Ad}_{(g_{dij})}e_{cij} + \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{wi}})}V_d$$
$$\textit{Observer}\ \begin{cases} \bar{V}_{m_ij}^b = -\mathrm{Ad}_{(\bar{g}_{m_ij}^{-1})}\mathrm{Ad}_{(g_{im_i}^{-1})}\tilde{V}_{wi}^b + u_{ij}\ , \\ u_{ij} = k_{eij}\left(e_{eij} - \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{cij}})}e_{cij}\right) \\ \qquad\qquad j \in \mathcal{N}_i,\ i \in \mathcal{V}. \end{cases} \quad (17)$$

Then, we have the following corollary which can be proved in the same way as Theorem 1.

Corollary 1: Suppose the leader's body velocity is $\mathrm{Ad}_{(e^{-\hat{\xi}\theta_{w1}})}V_d$. Then, control law (17) on the visual robotic network Σ achieves visual feedaback pose synchronization at least locally if gain conditions (14) are satisfied.

We use the assumptions that all rigid bodies have a common velocity V_d in Corollary 1. However, even without such common knowledge (the only leader has its own velocity), it is expected for followers to track the leader within a bounded error and achieve flocking-like behaviors. We thus analyze the tracking performance in the presense of V_{w1}^b based on the theory of $\mathcal{L}_2$-gain analysis in the following section.

IV. Tracking Performance Analysis

In this section, for the moving leader ($V_{w1}^b \neq 0$), we analyze the tracking performance of the other rigid bodies based on the theory of $\mathcal{L}_2$-gain analysis by regarding the leader's velocity as an external disturbance.

Theorem 2: Suppose the leader has its own velocity ($V_{w1}^b \neq 0$). Then, for any positive scalars $\epsilon,\ \gamma_i,\ i \in \mathcal{V}_q$, control law (10) on the visual robotic network Σ achieves

$$\|x_e\|_{\mathcal{L}_2} \leq \gamma \|V_{w1}^b\|_{\mathcal{L}_2} + \delta,\ \gamma := \sqrt{\sum_{i\in\mathcal{V}_q} \frac{\gamma_i}{2\epsilon}} \tag{18}$$

with a nonnegative scalar δ if

$$\begin{cases} \begin{cases} k_{ei1} - \frac{1}{2\gamma_i} - \epsilon > 0 \\ \left(k_{i1} + k_{ei1} - \epsilon - \frac{2\gamma_i k_{ei1}^2}{2\gamma_i(k_{ei1}-\epsilon)-1}\right) I_6 - k_{i1} D_{i1} > 0 \end{cases}, \\ \qquad\qquad i \in \mathcal{V}_p \\ \begin{cases} k_{jk} < 2(k_{eij} - \epsilon) \\ k_{jk} < \frac{2((k_{ij}-\epsilon)(k_{eij}-\epsilon) - k_{eij}\epsilon)}{k_{ij}+k_{eij}-\epsilon} \end{cases}, \\ \qquad\qquad k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_q \\ \begin{cases} k_{jk} < 2(k_{eij} - \epsilon) \\ \left(k_{ij} + k_{eij} - \epsilon - \frac{2k_{eij}^2}{2k_{eij}-k_{jk}-2\epsilon}\right) I_6 - k_{ij} D_{ij} > 0 \end{cases}, \\ \qquad\qquad k \in \mathcal{N}_j,\ j \in \mathcal{N}_i,\ i \in \mathcal{V}_r \end{cases} . \tag{19}$$

Proof: The time derivative of U is given by

$$\dot{U} = \sum_{i=2}^{n} \sum_{j\in\mathcal{N}_i} q_i \left(-e_{ij}^{\mathrm{T}} Q_{ij} e_{ij} + e_{eij}^{\mathrm{T}} \mathrm{Ad}_{(e^{\hat{\xi}\theta_{eij}})} V_{wj}^b \right). \tag{20}$$

Here, $Q_{ij} := \begin{bmatrix} (k_{ij}+k_{eij}) I_6 & -k_{eij}\mathrm{Ad}_{(e^{\hat{\xi}\theta_{cij}})} \\ -k_{eij}\mathrm{Ad}_{(e^{-\hat{\xi}\theta_{cij}})} & k_{eij} I_6 \end{bmatrix}$. Completing square for $e_{ei1}^{\mathrm{T}} \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ei1}})} V_{w1}^b,\ i \in \mathcal{V}_p$ yields

$$\begin{aligned} e_{ei1}^{\mathrm{T}} \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ei1}})} V_{w1}^b &= -\frac{\gamma_i}{2} \left\| \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ei1}})} V_{w1}^b - \frac{1}{\gamma_i} e_{ei1} \right\|_2^2 \\ &\quad + \frac{\gamma_i}{2} \|V_{w1}^b\|_2^2 + \frac{1}{2\gamma_i} \|e_{ei1}\|_2^2 \\ &\leq \frac{\gamma_i}{2} \|V_{w1}^b\|_2^2 + \frac{1}{2\gamma_i} \|e_{ei1}\|_2^2 \end{aligned}$$

for any positive scalars $\gamma_i,\ i \in \mathcal{V}_p$. Therefore, if gain conditions (19) are satisfied, we get

$$\dot{U} \leq \sum_{i\in\mathcal{V}_p} \frac{\gamma_i}{2} \|V_{w1}^b\|_2^2 - \epsilon \|x_e\|_2^2.$$

Integrating the above inequality from 0 to T with respect to time yields

$$U(T) - U(0) \leq \sum_{i\in\mathcal{V}_q} \frac{\gamma_i}{2} \int_0^T \|V_{w1}^b(t)\|_2^2 dt - \epsilon \int_0^T \|x_e(t)\|_2^2 dt.$$

Thus, the following inequality is derived.

$$\|x_e\|_{\mathcal{L}_2} \leq \sqrt{\sum_{i\in\mathcal{V}_q} \frac{\gamma_i}{2\epsilon}} \|V_{w1}^b\|_{\mathcal{L}_2} + \sqrt{\frac{1}{\epsilon} U(0)}.$$

Then, by defining $\gamma := \sqrt{\sum_{i\in\mathcal{V}_q} \frac{\gamma_i}{2\epsilon}},\ \delta := \sqrt{\frac{1}{\epsilon} U(0)}$, we get inequality (18). ■

Although γ_i appears in only conditions of rigid body $i \in \mathcal{V}_p$, the arguments k_{i1} also appear in the other constraints. This fact implicitly means that γ influences all bodies' gains.

Theorem 2 means that if we regard V_{w1}^b as the disturbance input and x_e as the output of collective error system Σ_{col}, then Σ_{col} has $\mathcal{L}_2$-gain less than or equal to γ. Since γ evaluates the control and estimation errors for the leader's velocity, it can be regarded as an indication of the tracking performance of the group. Therefore, by setting control gains making γ small, we can achieve a high tracking performance.

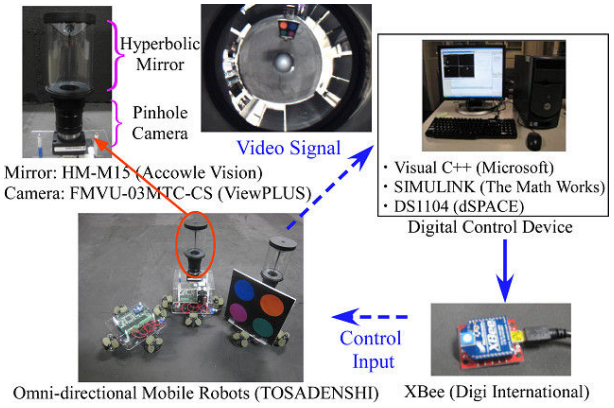

Fig. 6: Experimental Environment

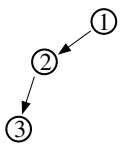

Fig. 7: Visibility Structure in Experiment

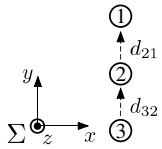

Fig. 8: Final Configuration in Experiment

V. Experiment

In this section, we demonstrate the effectiveness of the proposed control laws through experiments on a planar testbed.

We use three omnidirectional mobile robots with four wheels as rigid bodies. Each robot has a pinhole camera with a panoramic mirror. We attach a plate with four colored circles to each robot in order to improve accuracy of extracting feature points. We also use an overhead camera attached above the robots to measure the actual pose of robots. Transmitted video signals are loaded into PC and the control law is calculated in real time. Then the control inputs are sent to robots via an embedded wireless communication device. The sampling period of the controller is 20 [ms]. This experimental schematic is shown in Fig. 6.

We use the visibility structure depicted in Fig. 7. We let gains be

$$\text{GainA}: \begin{cases} k_{21} = 5.0,\ k_{e21} = 8.0 \\ k_{32} = 5.7,\ k_{e32} = 9.3 \end{cases},\ \gamma = 2.70,$$

$$\text{GainB}: \begin{cases} k_{21} = 0.37,\ k_{e21} = 3.0 \\ k_{32} = 0.5,\ k_{e32} = 3.0 \end{cases},\ \gamma = 18.1,$$

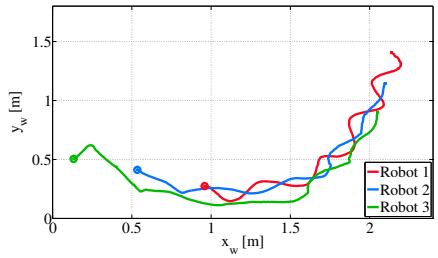

Fig. 9: Positions in Σ_w

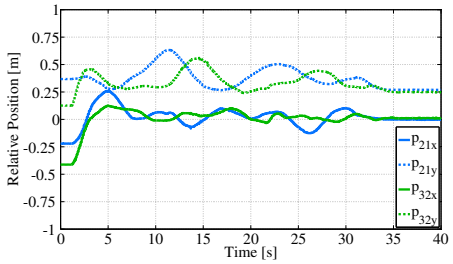

Fig. 10: Relative Positions

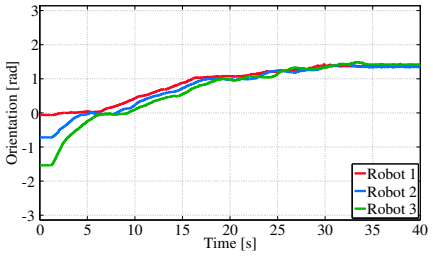

Fig. 11: Rotation Angles in Σ_w

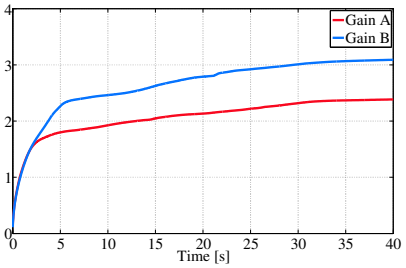

Fig. 12: Tracking Performance

where each gain setup satisfies conditions (14) and (19) ($\epsilon = 0.05$). Position biases are $d_{21} = d_{32} = [0\ 0.25\ 0]^{\mathrm{T}}$ [m] (Fig. 8). Initial conditions are set as

$$p_{w1}(0) = [0.96\ 0.27\ 0]^{\mathrm{T}}, \quad \xi\theta_{w1}(0) = [0\ 0\ -0.06]^{\mathrm{T}},$$
$$p_{w2}(0) = [0.53\ 0.41\ 0]^{\mathrm{T}}, \quad \xi\theta_{w2}(0) = [0\ 0\ -0.72]^{\mathrm{T}},$$
$$p_{w3}(0) = [0.13\ 0.51\ 0]^{\mathrm{T}}, \quad \xi\theta_{w3}(0) = [0\ 0\ -1.54]^{\mathrm{T}}.$$

Finally, we set the desired common velocity as 0 and the leader's body velocity V_{w1}^b as

$$V_{w1}^b = \begin{cases} [0.1\sin t\ 0.1\ 0\ 0\ 0\ 0]^{\mathrm{T}} & t \in [0,5) \\ [0.1\sin t\ 0.1\ 0\ 0\ 0\ 0.15]^{\mathrm{T}} & t \in [5,15) \\ [0.1\sin t\ 0.1\ 0\ 0\ 0\ 0]^{\mathrm{T}} & t \in [15,30) \\ 0 & t \in [30,40] \end{cases}.$$

The experimental results are shown in Figs. 9-12. Fig. 9 illustrates the trajectories of the robots on 2-dimensional plane for Gain A, Fig. 10 time resposes of relative positions and Fig. 11 orientations in Σ_w. Figs. 12 shows $\sqrt{\int_0^t \|x_e(\tau))\|_2^2 d\tau}$ for Gain A and Gain B, respectively. We see from Figs. 9 and 10 that when the leader moves, the other robots track it successfully, and the desired relative positions are almost achieved at around 35 [s] when the leader is static. Moreover, Fig. 11 shows that all orientations converge to almost a common value (robot 1's value) at that time. The results mean that the proposed control law (10) achieves visual feedback pose synchronization and thus the synchronization law works successfully.

Figs. 12 shows that the tracking performance is improved for the smaller values of γ. Therefore, γ is adequate for the performance indicate of the visual feedback pose synchronization.

VI. Conclusions

In this paper, we have investigated pose synchronization by using visual information as measured output of each rigid body. We have first introduced visual robotic networks. Then after defining visual feedback pose synchronization, we have proposed a visual feedback pose synchronization law combining a vision-based observer with the pose synchronization law. We have then proved that the network with the control law achieves visual feedback pose synchronization in the absence of communication or any other measurements of the states. Moreover, we have analyzed the tracking performance of the network for the moving leader. Finally, the experimental results have demonstrated the validity of our results.

References

[1] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratatoni and R. E. Davis, "Collective Motion, Sensor Networks and Ocean Sampling," *Proc. of the IEEE*, Vol. 95, No. 1, pp. 48–74, 2007.

[2] F. Bullo, J. Cortes and S. Martinez, *Distributed Control of Robotic Networks*, Princeton Series in Applied Mathematics, 2009.

[3] Y. Igarashi, T. Hatanaka, M. Fujita and M. W. Spong, "Passivity-based Output Synchronization and Flocking Algorithm in SE(3)," *Proc. of the 47th IEEE Conference on Decision and Control*, pp. 1024–1029, 2008.

[4] F. Chaumette and S. A. Hutchinson, "Visual Servoing and Visual Tracking," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, eds., Springer-Verlag, pp. 563–583, 2008.

[5] H. Kawai, T. Murao and M. Fujita, "Visual Motion Observer-based Pose Control with Panoramic Camera via Passivity Approach," *Proc. of the 2010 American Control Conference*, pp. 4534–4539, 2010.

[6] T. Ibuki, T. Hatanaka, M. Fujita and M. W. Spong, "Visual Feedback Leader Following Pose Synchronization: Convergence Analysis," *Proc. of the 2011 American Control Conference*, pp. 493–498, 2011.

[7] F. Morbidi, G. L. Mariottini and D. Prattichizzo, "Observer Design via Immersion and Invariance for Vision-based Leader-follower Formation Control," *Automatica*, Vol. 46, No. 1, pp. 148–154, 2010.

[8] P. Vela, A. Betser, J. Malcolm and A. Tannenbaum, "Vision-based Range Regulation of A Leader-follower Formation," *IEEE Trans. on Control Systems Technology*, Vol. 17, No. 2, pp. 442–448, 2009.

[9] Y. Ma, S. Soatto, J. Kosecka and S. S. Sastry, *An Invitation to 3-D Vision,* Springer, Chapter 2, 2003.

[10] H. Berghuis and H. Nijmeijer, "A Passivity Approach to Controller-Observer Design for Robots," *IEEE Trans. on Robotics and Automation*, Vol. 9, No. 6, pp. 740–754, 1993.