Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Frame Rate-based Discrete Visual Feedback Pose Regulation: A Passivity Approach

Tatsuay Ibuki * **Johannes R. Walter** ** **Takeshi Hatanaka** *
Masayuki Fujita *

* *Tokyo Institute of Technology, Tokyo, JAPAN (e-mail: fujita@ctrl.titech.ac.jp)*
** *University of Stuttgart, Stuttgart, GERMANY (e-mail: tky57436@stud.uni-stuttgart.de)*

Abstract: This paper studies a visual feedback 3D pose regulation problem explicitly handling camera frame rates. Although numerous works have already tackled vision-based estimation/control problems by focusing on the limitation of measured output (2D visual information), almost all of them have assumed that visual measurements of a camera are continuously available. However, camera frame rates and image processing time are often non-negligible compared with other computation time. In view of this fact, we newly propose a discrete visual feedback pose regulation law under the situation that visual measurements are sampled due to frame rates. We first give a sufficient condition of frame rates to achieve a desired relative pose to a static target object. Then, the tracking performance for a moving target is analyzed via the notion of ultimate boundedness. The present analysis provides a guideline for the design of estimation/control gains. We finally show the significance and validity of this work through 3D simulation.

1. INTRODUCTION

Fusion of control theory and computer vision has been in wide spread use thanks to rich visual information (Chaumette and Hutchinson [2006, 2007]). This technique is originally motivated by robot control as in Hager and Hutchinson [1996] and recently handles various problems such as surveillance or security, image-based medical procedures or comprehension of biological perceptual information processing (Ding et al. [2012], Gao et al. [2010], Han et al. [2010]). One of main issues here is how to handle 2D visual information in order to estimate/control a 3D relative pose (position and orientation) to a target object (Dani et al. [2012], Cunha et al. [2011], Ayazoglu et al. [2011], Karasev et al. [2011], Luca et al. [2008]).

This paper addresses a visual feedback 3D pose regulation problem whose objective is to drive the camera pose relative to a target object to a desired one by using only visual information extracted by a single camera. To meet this goal, our previous works (Fujita et al. [2007], Kawai et al. [2011]) have proposed a visual feedback 3D pose estimation mechanism, called *visual motion observer*, and observer-based control mechanisms. Here, passivity of rigid body motion plays a central role and energy-based approaches have been taken for convergence and performance analysis. However, these works assume that visual measurements are continuously available. This assumption permits any positive estimation and control gains though we know that high gain estimation or control does not work in experiments. Moreover, most of works on these problems focus only on the limitation of measured output and do not consider the fact that frame rates and image processing time are not small enough to be negligible compared with other computation time.

In the case that frame rates or image processing time are non-negligible, we have to consider estimation/control input sampled by discrete visual measurements. When we handle such sampled input, an event-based control technique (Tabuada [2007], Bemporad et al. [2010], Garcia and Antsaklis [2013]) becomes a good help for convergence analysis. The appealing point of event-based control is the possibility to reduce the number of re-computing input and transmissions while guaranteeing desired performances. One of the main topics here is to investigate inter-event time guaranteeing monotonic decrease of potential functions for convergence. This technique is suited to our previous works (Fujita et al. [2007], Kawai et al. [2011]) since they also use non-increasing properties of potential functions for convergence/performance analysis.

In view of these facts, this paper studies a visual feedback 3D pose regulation problem explicitly handling camera frame rates and image processing time. We newly propose a discrete pose regulation mechanism, where the present input is sampled due to frame rates. We next provide the relation between frame rates and estimation/control gains to achieve the desired relative pose for a static target object, which is our main contribution. The tracking performance analysis is then conducted for a moving target. Here, it is shown that estimation and control errors are ultimately bounded by a function of the camera frame rate, the estimation/control gains and the target object velocity. We also give how to deal with the case that frame rates and image processing time are variable. The second contribution is that the convergence and performance analysis provides the guideline for gain settings. We finally demonstrate the validity and importance of this work via 3D simulation.

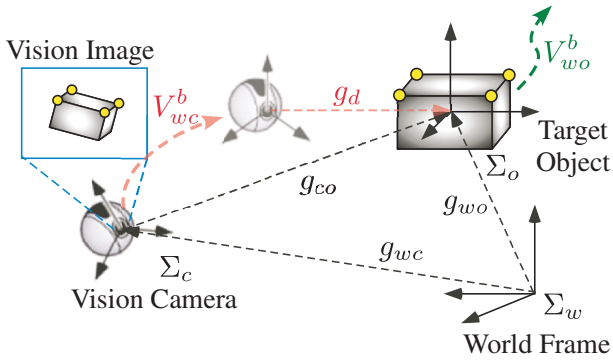

Fig. 1. Coordinate frames for visual feedback system

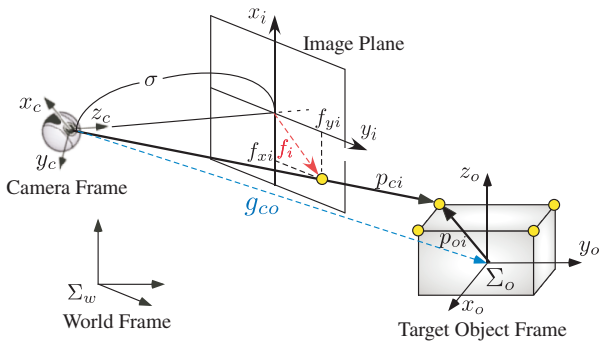

Fig. 2. Pinhole camera model

2. VISUAL FEEDBACK POSE REGULATION

We first give the problem formulation of visual feedback 3D pose regulation and introduce the main results of Fujita et al. [2007] as the preliminary of this work.

2.1 Rigid Body Motion

We consider a visual feedback pose regulation system shown in Fig. 1 throughout this work. Here, we respectively denote the world frame, the camera frame and the object frame as Σ_w, Σ_c and Σ_o. Then, the pose of the origin of the camera frame Σ_c relative to the world frame Σ_w is represented by $g_{wc} = (p_{wc}, e^{\hat{\xi}_{wc}\theta_{wc}}) \in SE(3)$, where $\xi_{wc} \in \mathcal{R}^3$ ($\|\xi_{wc}\| = 1$) and $\theta_{wc} \in (-\pi, \pi]$ are respectively the direction and the angle of rotation. For notational simplicity, we hereafter represent $\xi_{wc}\theta_{wc}$ by $\xi\theta_{wc}$. The notation '$\wedge$' gives $\hat{a}b = a \times b$, $a, b \in \mathcal{R}^3$ for the vector cross-product $\times$ and its inverse is represented by '$\vee$'. We similarly represent the pose of the object frame Σ_o relative to the world frame Σ_w by $g_{wo} = (p_{wo}, e^{\hat{\xi}\theta_{wo}}) \in SE(3)$.

We also denote the body velocities of the camera frame Σ_c and the object frame Σ_o relative to the world frame Σ_w by $V^b_{wc} = [v^{\mathrm{T}}_{wc}\ \omega^{\mathrm{T}}_{wc}]^{\mathrm{T}}$, $V^b_{wo} = [v^{\mathrm{T}}_{wo}\ \omega^{\mathrm{T}}_{wo}]^{\mathrm{T}} \in \mathcal{R}^6$, where $v \in \mathcal{R}^3$ and $\omega \in \mathcal{R}^3$ are respectively the linear and angular body velocities (Ma et al. [2003]).

We now introduce the homogeneous representations of g and V^b for calculations as follows (note here that we use another definition of '$\wedge$' for $V^b \in \mathcal{R}^6$).

$$g = \begin{bmatrix} e^{\hat{\xi}\theta} & p \\ 0 & 1 \end{bmatrix} \in \mathcal{R}^{4\times4},\ \hat{V}^b = \begin{bmatrix} \hat{\omega} & v \\ 0 & 0 \end{bmatrix} \in \mathcal{R}^{4\times4}.$$

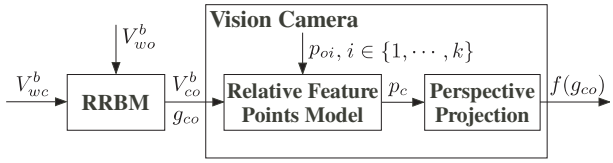

Fig. 3. Block diagram of camera model with RRBM (RRBM means relative rigid body motion)

Then, the body velocities V^b_{wc} and V^b_{wo} are respectively defined as $\hat{V}^b_{wc} := g^{-1}_{wc}\dot{g}_{wc}$ and $\hat{V}^b_{wo} := g^{-1}_{wo}\dot{g}_{wo}$ (Ma et al. [2003]).

Finally, the pose of the object frame Σ_o relative to the camera frame Σ_c and its body velocity are respectively denoted by $g_{co} = (p_{co}, e^{\hat{\xi}\theta_{co}}) \in SE(3)$ and $V^b_{co} := (g^{-1}_{co}\dot{g}_{co})^{\vee} \in \mathcal{R}^6$. Then, $g_{co} = g^{-1}_{wc}g_{wo}$ holds, and the following *relative rigid body motion* is obtained from $\hat{V}^b_{co} = g^{-1}_{co}\dot{g}_{co}$ (Ma et al. [2003]).

$$\dot{g}_{co} = -\hat{V}^b_{wc}g_{co} + g_{co}\hat{V}^b_{wo}. \tag{1}$$

2.2 Visual Measurement

We next introduce 2D visual measurements extracted by a vision camera as measured output for 3D pose regulation. Although we only introduce the extraction by perspective projection in this paper, we can also introduce panoramic camera models for the subsequent discussions as in Kawai et al. [2011].

We consider the target object with k ($k \geq 4$) feature points. We represent the positions of the feature points relative to the object frame Σ_o by $p_{oi} \in \mathcal{R}^3$, $i \in \{1, \cdots, k\}$. Then, the positions relative to the camera frame Σ_c are given by $p_{ci} = g_{co}p_{oi}$ from the coordinate transformation. Here, we use the homogeneous representations $[p^{\mathrm{T}}_{ci}\ 1]^{\mathrm{T}}$ and $[p^{\mathrm{T}}_{oi}\ 1]^{\mathrm{T}}$. We next denote the k feature points on the image plane by $f = [f^{\mathrm{T}}_1\ \cdots\ f^{\mathrm{T}}_k]^{\mathrm{T}} \in \mathcal{R}^{2k}$. Then, well known perspective projection (Ma et al. [2003]) yields the following relation for each $f_i \in \mathcal{R}^2$ (Fig. 2).

$$f_i = \frac{\sigma}{z_{ci}} \begin{bmatrix} x_{ci} \\ y_{ci} \end{bmatrix},\ p_{ci} = [x_{ci}\ y_{ci}\ z_{ci}]^{\mathrm{T}}. \tag{2}$$

Here, $\sigma > 0$ is the focal length of the camera. Suppose that the visual measurements f is only available for pose regulation and the feature points p_{oi} are known *a priori*. Then, the visual measurement is the function only of the relative pose g_{co} (i.e. $f(g_{co})$). Fig. 3 illustrates the block diagram of the relative rigid body motion with perspective projection.

The objective of this work is then to propose a discrete pose regulation mechanism to drive g_{co} to the fixed desired pose, denoted by $g_d = (p_d, e^{\hat{\xi}\theta_d}) \in SE(3)$, only from the visual measurement $f(g_{co})$ under the situation that the frame rate and image processing time are non-negligible (Fig. 1).

2.3 Review of Visual Feedback Pose Regulation

We next review the visual feedback pose regulation presented in Fujita et al. [2007] as the preliminary. Note here that the authors in Fujita et al. [2007] assume that visual measurements are continuously available.

Since available visual measurements (2) are two dimensional, it is necessary for 3D pose regulation to estimate the relative pose g_{co} by a nonlinear observer. We now represent the estimate of g_{co} by $\bar{g}_{co} = (\bar{p}_{co}, e^{\hat{\bar{\xi}}\bar{\theta}_{co}}) \in SE(3)$. Similarly to the Luenberger-type observer (Luenberger [1971]), we first build the copy model of relative rigid body motion (1) as follows.

$$\dot{\bar{g}}_{co} = -\hat{V}^b_{wc}\bar{g}_{co} + \bar{g}_{co}\hat{u}_e. \tag{3}$$

Here, $u_e = [u_{ep}^{\mathrm{T}}\ u_{eR}^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{R}^6$ is observer input for the estimation of g_{co}, and notice that the model (3) does not include V^b_{wo} information since it is unknown (the pose regulation problem with target velocity estimation has been tackled in Ibuki et al. [2013]). It should be also noted that the estimated visual measurements $\bar{f}$ can be computed by $\bar{g}_{co}$ and (2).

We next introduce the estimation error $g_{ee} = (p_{ee}, e^{\hat{\xi}\theta_{ee}}) \in SE(3)$ between g_{co} and $\bar{g}_{co}$ and the estimation error vector $e_e \in \mathcal{R}^6$ defined as follows.

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

Here, $\mathrm{sk}(e^{\hat{\xi}\theta}) := (1/2)(e^{\hat{\xi}\theta} - e^{-\hat{\xi}\theta}) \in so(3)$. We note the important property that for $\theta_{ee} \in (-\pi, \pi)$, $e_e = 0$ holds if and only if $g_{ee} = I_4$, i.e. $\bar{g}_{co} = g_{co}$ ($I_n \in \mathcal{R}^{n\times n}$ represents the n-dimensional identity matrix). It should be also noted that e_e can be approximately reconstructed by the measurement error $f_e \in \mathcal{R}^{2k}$defined as $f_e := f - \bar{f}$ (refer to Fujita et al. [2007]). Then, the time differentiation of g_{ee} with (1) and (3) produces the following *estimation error system*.

$$\dot{g}_{ee} = -\hat{u}_e g_{ee} + g_{ee}\hat{V}^b_{wo}. \tag{4}$$

This is given in the vector form by

$$V^b_{ee} := (g_{ee}^{-1}\dot{g}_{ee})^\vee = -\mathrm{Ad}_{(g_{ee}^{-1})}u_e + V^b_{wo}.$$

Here, $\mathrm{Ad}_{(g)} \in \mathcal{R}^{6\times 6}$ represents the adjoint transformation associated with g (Ma et al. [2003]).

We also build the control error system similarly to the estimation error system (4). We respectively define the control error $g_{ce} = (p_{ce}, e^{\hat{\xi}\theta_{ce}}) \in SE(3)$ and its vector $e_c \in \mathcal{R}^6$ as

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

Notice again that for $\theta_{ce} \in (-\pi, \pi)$, $e_c = 0$ holds if and only if $g_{ce} = I_4$ (i.e. $\bar{g}_{co} = g_d$). Then, the time differentiation of g_{ce} with (3) gives the following *control error system*.

$$\dot{g}_{ce} = -g_d^{-1}\hat{V}^b_{wc}g_d g_{ce} + g_{ce}\hat{u}_e. \tag{5}$$

This is written in the vector form by

$$V^b_{ce} := (g_{ce}^{-1}\dot{g}_{ce})^\vee = -\mathrm{Ad}_{(g_{ce}^{-1})}\mathrm{Ad}_{(g_d^{-1})}V^b_{wc} + u_e.$$

In summary, combining the estimation error system (4) and the control error system (5) yields the following *total error system*.

$$\begin{bmatrix} V^b_{ce} \\ V^b_{ee} \end{bmatrix} = -\begin{bmatrix} \mathrm{Ad}_{(g_{ce}^{-1})} & -I_6 \\ 0 & \mathrm{Ad}_{(g_{ee}^{-1})} \end{bmatrix} u_{ce} + \begin{bmatrix} 0 \\ V^b_{wo} \end{bmatrix}, \tag{6}$$

where $u_{ce} := [(\mathrm{Ad}_{(g_d^{-1})}V^b_{wc})^{\mathrm{T}}\ u_e^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{R}^{12}$. Then, the authors in Fujita et al. [2007] have shown that if $V^b_{wo} \equiv 0$ holds, the total error system (6) is *passive* from the input u_{ce} to the output $\nu_{ce} \in \mathcal{R}^{12}$ defined as

$$\nu_{ce} := Ne_{ce},\ N := \begin{bmatrix} -I_6 & 0 \\ \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{ce}})} & -I_6 \end{bmatrix} \in \mathcal{R}^{12\times 12}.$$

Here, $e_{ce} \in \mathcal{R}^{12}$ is the total control and estimation errors defined as $e_{ce} := [e_c^{\mathrm{T}}\ e_e^{\mathrm{T}}]^{\mathrm{T}}$. Note then that for $\theta_{ce}, \theta_{ee} \in (-\pi, \pi)$, $e_{ce} = 0$ means $g_{co} = g_d$ (i.e. the goal is achieved). The corresponding storage function $U \geq 0$ for this passivity is given by

$$U := \frac{1}{2}\|p_{ce}\|^2 + \phi(e^{\hat{\xi}\theta_{ce}}) + \frac{1}{2}\|p_{ee}\|^2 + \phi(e^{\hat{\xi}\theta_{ee}}),$$
$$\phi(e^{\hat{\xi}\theta}) := \frac{1}{4}\|I_3 - e^{\hat{\xi}\theta}\|_F^2 = \frac{1}{2}\mathrm{tr}(I_3 - e^{\hat{\xi}\theta}) \geq 0.$$

Note here that $U = 0$ means $e_{ce} = 0$, i.e. visual feedback pose regulation works.

The paper Fujita et al. [2007] proposes the following pose regulation input from this passivity.

$$u_{ce} = -K\nu_{ce},\ K := \begin{bmatrix} k_c I_6 & 0 \\ 0 & k_e I_6 \end{bmatrix} \in \mathcal{R}^{12\times 12}. \tag{7}$$

Here, $k_c, k_e > 0$ and it has the following fact.

Fact 1. If $V^b_{wo} \equiv 0$ holds, then the equilibrium point $e_{ce} = 0$ for the closed-loop system (6) and (7) is asymptotically stable. In addition, given a positive scalar α, if K satisfies

$$\begin{cases} k_e - \dfrac{1}{2\alpha} - \dfrac{1}{2} > 0 \\ k_c - \dfrac{1}{2} - \dfrac{k_e(\alpha+1)}{\alpha(2k_e - 1) - 1} > 0 \end{cases},$$

then the system (6) and (7) with the input V^b_{wo} and the output ν_{ce} has $\mathcal{L}_2$-gain $\leq \alpha$.

The first statement means that the visual feedback pose regulation works successfully when the target object is static. Also, the second claim says that even for a moving object, the regulation works well with the tracking performance indicator α. The block diagram of the present visual feedback system is depicted in Fig. 4 without *Sampler* blocks.

We now note that Fact 1 does not consider camera frame rates or image processing time though they are not negligible compared with other computation time (general cameras have 15, 30, 60 or less fps when the image processing time is included). Therefore, we tackle a new visual feedback pose regulation problem explicitly taking into consideration camera frame rates in the next section.

3. FRAME RATE-BASED DISCRETE VISUAL FEEDBACK POSE REGULATION

In this section, we reconsider a visual feedback pose regulation problem by explicitly handling the camera frame rate and the time for extracting feature points via image processing. Throughout this section, we give a little complicated analysis or conditions in order to make conservativeness as less as possible.

3.1 Convergence Analysis

Suppose that the camera has the frame rate $\tau > 0$ [fps] and, for notational simplicity, this includes the image processing time to extract feature points. Assume also that

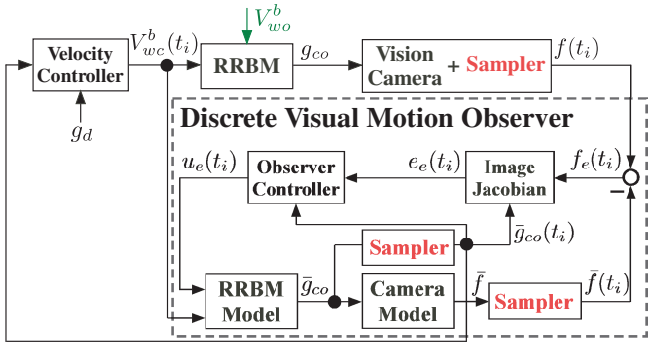

Fig. 4. Block diagram of discrete visual feedback system

the computation time to calculate the control/estimation input is negligible compared with $1/\tau$. Based on the frame rate τ, we now introduce sampling time sequence $\{t_0, t_1, t_2, \cdots\}$ such that $t_{i+1} - t_i = 1/\tau$ holds for all $i \in \mathbb{N}_0$ ($\mathbb{N}_0$ represents the union of natural numbers and $\{0\}$). Then, the visual measurements (2) are extracted at each time instant t_i.

We thus newly propose frame rate-based discrete input as follows.

$$u_{ce}(t) = -K\nu_{ce}(t_i) = -KN(t_i)e_{ce}(t_i), \quad k_e \geq k_c > 0,\ t \in [t_i, t_{i+1}). \tag{8}$$

Note here that the present pose regulation input is constant till the next sampling time (the velocity input is also sampled because it is built by the estimated relative pose $\bar{g}_{co}$ from sampled visual information). Moreover, we newly utilize the gain condition $k_e \geq k_c$ with hope for good transient responses. The block diagram of the present discrete visual feedback system is depicted in Fig. 4. Since the image Jacobian in the observer is the function of not only f but $\bar{f}$ and $\bar{g}_{co}$ (Fujita et al. [2007]) and the velocity law consists of $\bar{g}_{co}$, we apply the same samplers to these estimates.

Let us now introduce the error $e \in \mathcal{R}^{12}$ caused by the samplers as

$$e(t) := N(t)e_{ce}(t) - N(t_i)e_{ce}(t_i),\ t \in [t_i, t_{i+1}),\ {}^\forall i \in \mathbb{N}_0.$$

We then have the following result (this theorem is proved by the similar approach in Ibuki et al. [2013]).

Theorem 2. Suppose that the target object is static ($V_{wo}^b \equiv 0$). Then, if the camera frame rate satisfies the condition

$$\tau \geq \frac{\sqrt{k_e(16k_c + 31k_e)}}{2\arctan\left(\frac{(\sqrt{5}-2)k_c\delta\sqrt{k_e(16k_c+31k_e)}}{2k_ck_e+(17+5\sqrt{5})k_e^2+(10-3\sqrt{5})k_ck_e\delta}\right)} \tag{9}$$

for $\delta \in (0,1)$, there exist finite time $t_e > 0$ and a positive scalar β such that

$$U(t) \leq U(t_e)e^{-\beta(t-t_e)}\ {}^\forall t \geq t_e.$$

Namely the equilibrium point $e_{ce} = 0$ for the closed-loop system (6) and (8) is exponentially stable after time t_e.

Claim 3. The condition (9) implies that large feedback gains k_c and k_e require fast frame rates τ (small sampling intervals). This property is intuitive because large gains increase the influence of e, which might be a poor impact on estimation and control. Said differently, after choosing a camera with a certain rate, it is not free to let gains large. Although the condition (9) is only sufficient and much conservative so far, we believe that this analysis provides the significant insight that camera frame rates or image processing time is not negligible (Section 4 gives an example).

One of our future works is to reduce the conservativeness of the sufficient condition (9).

3.2 Tracking Performance Analysis for Moving Target

We next give tracking performance analysis for a moving target by introducing the theory of ultimate boundedness (Khalil [2002]).

We first suppose that $\|V_{wo}^b(t)\| \leq \kappa\ {}^\forall t \geq t_0$ holds for a positive scalar κ (i.e. the target velocity is bounded). Then, the time differentiation of the potential error function U along the trajectories of (6) and (8) yields

$$\begin{aligned}\dot{U} =& e_{ce}^{\mathrm{T}}N^{\mathrm{T}}u_{ce} + e_e^{\mathrm{T}}\mathrm{Ad}_{(e^{\hat{\xi}\theta_{ee}})}V_{wo}^b \\ \leq& -ke_{ce}^{\mathrm{T}}N^{\mathrm{T}}KN(t_i)e_{ce}(t_i) + \|e_e\|\|V_{wo}^b\| \\ \leq& -\frac{(3-\sqrt{5})k_c}{2}\|e_{ce}\|^2 + \frac{(1+\sqrt{5})k_e}{2}\|e_{ce}\|\|e\| + \kappa\|e_{ce}\|,\end{aligned}$$

where the first term is given by the similar approach in Ibuki et al. [2013]. Thus, if the error e satisfies

$$\|e\| \leq \gamma$$

for a positive scalar γ, we get

$$\dot{U} \leq -\frac{(3-\sqrt{5})k_c}{2}\|e_{ce}\|^2 + \frac{(1+\sqrt{5})k_e\gamma + 2\kappa}{2}\|e_{ce}\|. \tag{10}$$

Then, since the right-hand side of (10) consists only of $\|e_{ce}\|$, we can employ ultimate boundedness analysis, and we have the following result from the similar analysis in Ibuki et al. [2013].

Theorem 4. Suppose that the target object velocity is upper-bounded by κ. Then, for every $e_{ce}(t_0)$, there exists $t_b \geq 0$ such that the solution $e_{ce}(t)$ of the closed-loop system (6) and (8) meets

$$\|e_{ce}(t)\| \leq \frac{(1+\sqrt{5})k_e\gamma + 2\kappa}{\sqrt{2}(3-\sqrt{5})\delta k_c}\ {}^\forall t \geq t_0 + t_b \tag{11}$$

if $\theta_{ce}(t), \theta_{ee}(t) \in (-\pi/2, \pi/2)$ holds and the frame rate satisfies

$$\begin{aligned}\tau \geq& \frac{(1+\sqrt{5})k_e}{2\ln\left(1 + \frac{(1+\sqrt{5})k_ck_e\delta\gamma}{4k_c\delta\kappa+\sqrt{2}(23+11\sqrt{5})((1+\sqrt{5})k_e\gamma+2\kappa)k_e}\right)}, \\ & \text{for } \|e_{ce}(t_0)\| \leq \frac{(1+\sqrt{5})k_e\gamma + 2\kappa}{(3-\sqrt{5})\delta k_c}, \\ \tau \geq& \frac{(1+\sqrt{5})k_e}{2\ln\left(1 + \frac{(1+\sqrt{5})k_e\gamma}{4\kappa+(7+5\sqrt{5})k_e\sqrt{2U(t_0)}}\right)}, \\ & \textit{otherwise}.\end{aligned} \tag{12}$$

Claim 5. It can be seen from (11), (12) and $k_e \geq k_c$ that large gains achieve good performances but require fast frame rates. It is also intuitive that the target object velocity directly influences the tracking performance. In addition, we can consider γ as the indicator of the performance and the sufficiently allowable gains. For instance, choosing a camera with a certain rate enables us to design k_c and k_e for a desired performance related to γ from (11) and (12) (Section 4 provides an example).

Theorem 4 gives two frame rate conditions depending on the initial estimation/control errors which might be confusing. However, if we first run the present estimation/control law (8) before the target moves, we can consider only the first condition. Also, we see from simulation experiences that the condition $\theta_{ce}(t), \theta_{ee}(t) \in (-\pi/2, \pi/2)$ is generally satisfied unless the initial estimation/control errors are too big or κ is too large. Eliminating this condition is a future work.

3.3 Variable Camera Frame Rate Case

So far, we have considered the situation that the visual measurement (2) can be extracted every fixed sampling time $(1/\tau)$. However, actual sampling time for frame rates and extracting feature points by image processing is much variable. To tackle this issue, we handle the worst case (the maximum sampling time).

Let us now consider the case that the visual measurement (2) is extracted at time instants $\{t_0, t_1, t_2, \cdots\}$, where $t_{i+1} - t_i = 1/\tau_i$, $\tau_i > 0$, $i \in \mathbb{N}_0$. Namely, we do not assume that every sampling interval is the same. We now suppose that the worst frame rate, denoted by $\tau_{min} > 0$, is known *a priori* from advance image processing tests. Then, note from the convergence/performance analysis approaches that by simply replacing τ for τ_{min} in conditions (9) and (12), the non-increasing property of U is always guaranteed. We thus get the following corollary.

Corollary 6. Suppose that τ_i, $i \in \mathbb{N}_0$ satisfy $\tau_i \geq \tau_{min}$. Then, the same statements as in Theorems 2 and 4 hold by replacing τ for τ_{min}.

4. VERIFICATIONS

We finally show the validity of the present discrete visual feedback pose regulation mechanism (8) and the significance to handle frame rates via 3D simulation.

We consider a 60fps camera pointing in the direction of the z-axis of Σ_c with $\sigma = 0.003$ [m]. The initial poses of the camera and the target object are respectively set as $g_{wc}(0) = I_4$, $p_{wo}(0) = [1\ 0\ 2]^{\mathrm{T}}$ [m] and $\xi\theta_{wo}(0) = [0\ -\pi/3\ 0]^{\mathrm{T}}$ [rad]. We also set $p_{o1} = [0.5\ 0.5\ 0.5]^{\mathrm{T}}$, $p_{o2} = [0.5\ -0.5\ 0.5]^{\mathrm{T}}$, $p_{o3} = [-0.5\ 0.5\ 0.5]^{\mathrm{T}}$, $p_{o4} = [-0.5\ -0.5\ 0.5]^{\mathrm{T}}$ [m] and

$$V_{wo}^b = \begin{cases} 0 & t \in [0,5) \\ [0\ 0\ 0.15\ 0\ 0\ 0]^{\mathrm{T}} & t \in [5,10) \\ [0.1\ 0\ 0\ 0\ 0\ -0.01]^{\mathrm{T}} & t \in [10,15) \\ [-0.1\ 0\ 0\ 0\ 0\ 0.01]^{\mathrm{T}} & t \in [15,20) \\ [0\ 0\ -0.15\ 0\ 0\ 0]^{\mathrm{T}} & t \in [20,25) \\ 0 & t \in [25,35) \end{cases}.$$

We now apply the present discrete input with $k_c = k_e = 0.8$, $\bar{p}_{co} = [0\ 0\ 1]^{\mathrm{T}}$ [m] and $\bar{\xi}\bar{\theta}_{co}(0) = 0$ [rad] to the camera in order to achieve the desired relative pose $p_d = [0\ 0\ 0.5]^{\mathrm{T}}$ [m] and $\xi\theta_d = 0$ [rad]. In this setting with $\delta = 0.95$, the sufficient condition (9) gives $\tau \geq 59.45$ (i.e. satisfied). Moreover, we get from (12) the best performance indicator $\gamma = 0.2$ for $k_c = k_e = 0.8$ and 60fps. As a result, Theorem 4 gives the fact that there exists $t_b \geq 0$ such that $\|e_{ce}(t)\| \leq 0.996\ {}^{\forall}t \geq t_b$.

The results are depicted in Figs. 5-7. We first show the time response of the discrete input $u_{ce}(t)$ in Fig. 5. This figure provides the confirmation that the input is actually sampled due to the frame rate. The time responses of the potential error function $U(t)$ and the norm of the estimation/control errors $\|e_{ce}(t)\|$ are respectively illustrated in Figs. 6 and 7. These figures give the validity of the convergence (after 25s) and performance analysis, i.e. the present discrete visual feedback pose regulation mechanism works successfully.

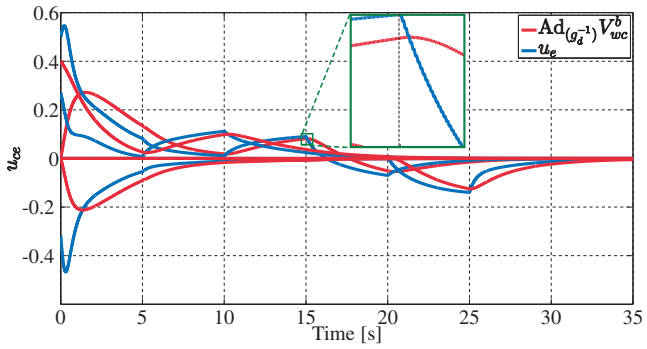

Fig. 5. Discrete estimation and control input

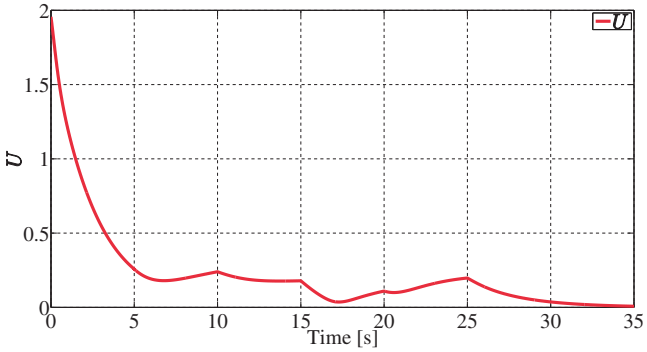

Fig. 6. Potential error function

In order to clarify the significance of our claim for handling frame rates, we finally give the example that the high gain setting for a slow frame rate causes poor estimation and control. We now set the frame rate including image processing time as 5fps. This rate is obtained by the worst case of our experimental system, where we use FMVU-03MTC CS camera (ViewPLUS) with 30fps and image processing software OpenCV (Willow Garage) to extract feature points. Then, by setting $k_c = 3$ and $k_e = 4$, we get the oscillating response of $\|e_{ce}(t)\|$ as shown in Fig. 8. Here, the period of oscillation is 0.2s which is given by the current frame rate. This result means that although the sufficient conditions (9) and (12) are much conservative so far, there must exist the upper-bound of gains as we claim in this work.

5. CONCLUSIONS

This paper has investigated a visual feedback 3D pose regulation problem in the case that camera frame rates and image processing time are non-negligible. We have first introduced the visual feedback pose regulation mechanism proposed in Fujita et al. [2007] where visual information is supposed to be continuously available. Then, a new pose regulation scheme explicitly handling frame rates has been proposed and it has been shown that the present

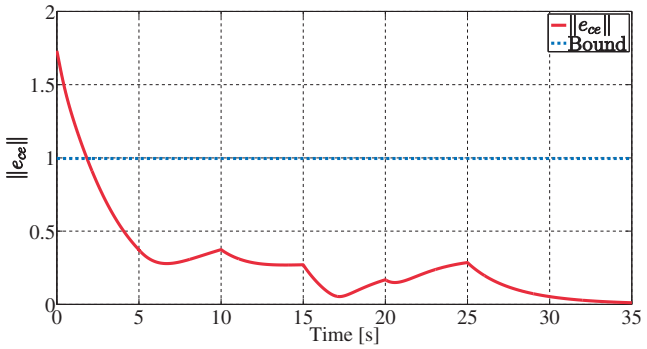

Fig. 7. Norm of estimation and control errors

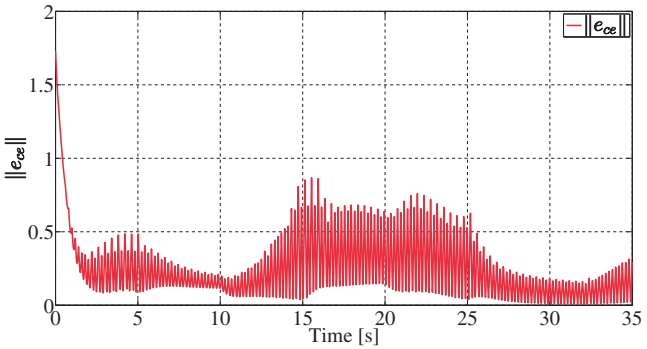

Fig. 8. Norm of estimation and control errors (5fps)

law works successfully for a static target object. We have also analyzed the tracking performance for a moving target based on the theory of ultimate boundedness. Additionally, we have given how to deal with the case that frame rates and image processing time are variable. The present analysis provides a guideline for the design of estimation and control gains. Finally, the effectiveness of the present scheme and analysis has been shown via 3D simulation.

The future directions of this work are to reduce the conservativeness of the sufficient frame rate conditions, to introduce the dynamics of rigid body motion, and to employ the theory of $\mathcal{L}_2$-gain stability for performance analysis as in Fujita et al. [2007], Kawai et al. [2011].

REFERENCES

F. Chaumette and S. Hutchinson. Visual servo control, part I: Basic approaches. *IEEE Robotics and Automation Magazine*, 13(4):82–90, 2006.

F. Chaumette and S. Hutchinson. Visual servo control, part II: Advanced approaches. *IEEE Robotics and Automation Magazine*, 14(1):109–118, 2007.

G.D. Hager and S. Hutchinson Eds. Special section on vision-based control of robot manipulators. *IEEE Trans. on Robotics and Automation*, 12(5):649–650, 1996.

C. Ding, B. Song, A. Morye, J.A. Farrell and A.K. Roy-Chowdhury. Collaborative sensing in a distributed PTZ camera network. *IEEE Trans. on Image Processing*, 21(7):3282–3295, 2012.

Y. Gao, R. Sandhu, G. Fichtinger and A. Tannenbaum. A coupled global registration and segmentation framework with application to magnetic resonance prostate imagery. *IEEE Trans. on Medical Imaging*, 29(10):1781–1794, 2010.

S. Han, A. Censi, A.D. Straw and R.M. Murray. A bioplausible design for visual pose stabilization. *Proc. of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5679–5686, 2010.

A.P. Dani, N.R. Fischer and W.E. Dixon. Single camera structure and motion. *IEEE Trans. on Automatic Control*, 57(1):238–243, 2012.

R. Cunha, C. Silvestre, J. Hespanha and A.P. Aguiar. Vision-based control for rigid body stabilization. *Automatica*, 47(5):1020–1027, 2011.

M. Ayazoglu, B. Li, C. Dicle, M. Sznaier and O.I. Camps. Dynamic subspace-based coordinated multicamera tracking. *Proc. of the 2011 IEEE International Conference on Computer Vision*, pages 2462–2469, 2011.

P.A. Karasev, M.M. Serrano, P.A. Vela and A. Tannenbaum. Depth invariant visual servoing. *Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4992–4998, 2011.

A.D. Luca, G. Oriolo and P.R. Giordano. Feature depth observation for image-based visual servoing: Theory and experiments. *The International Journal of Robotics Research*, 27(10):1093–1116, 2008.

M. Fujita, H. Kawai and M.W. Spong. Passivity-based dynamic visual feedback control for three dimensional target tracking: Stability and $\mathcal{L}_2$-gain performance analysis. *IEEE Trans. on Control Systems Technology*, 15(1):40–52, 2007.

H. Kawai, T. Murao and M. Fujita. Passivity-based visual motion observer with panoramic camera for pose control. *Journal of Intelligent and Robotic Systems*, 64(3-4):561–583, 2011.

P. Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Trans. on Automatic Control*, 52(9):1680–1685, 2007.

A. Bemporad, M. Heemels and M. Johansson. *Networked Control Systems (Lecture Notes in Control and Information Sciences), Vol. 406*. Springer, 2010.

E. Garcia and P.J. Antsaklis. Model-based event-triggered control for systems with quantization and time-varying network delays. *IEEE Trans. on Automatic Control*, 58(2):422–434, 2013.

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

D.G. Luenberger. An introduction to observers. *IEEE Trans. on Automatic Control*, AC-16(6):596–602, 1971.

T. Ibuki, T. Hatanaka and M. Fujita. Passivity-based visual feedback pose regulation integrating a target motion model in three dimensions. *SICE Journal of Control, Measurement, and System Integration*, 6(5):322–330, 2013.

T. Ibuki, Y. Namba, T. Hatanaka and M. Fujita. Passivity-based discrete visual motion observer taking account of camera frame rates. *Proc. of the 52nd IEEE Conference on Decision and Control*, pages 7660-7665, 2013.

H.K. Khalil. *Nonlinear Systems, Third Edition*. Prentice Hall, 2002.