The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Task selection for control of active vision systems

Yasushi Iwatani

Abstract— **This paper discusses the task selection problem for active vision systems, that is, what tasks should be selected to control active vision systems and which order of priority should be set for the selected tasks. One possible task to determine an optimal camera placement is to obtain high resolvability, or equivalently, to optimize a perceptibility measure which is a quantitative scaling measure from error of the measured velocity of image features in the image plane to error of the corresponding velocity computed in the camera coordinate frame. This paper first shows that optimization of the perceptibility measure may produce unreasonable motion responses for active vision systems, and it should not be selected for a primary task to control active vision systems. This paper then proposes target tracking as a primary task and optimization of a perceptibility measure as a secondary task. The perceptibility measure for the secondary task is induced by certain Jacobian matrices, not the image Jacobian matrix, to produce cooperative behavior for active vision systems with multiple cameras.**

I. Introduction

An active vision system is a robotic system with a camera (or cameras) mounted on the robot end effector. The camera position and/or the direction of the optical axis can be controlled, and the system provides a variable and wide field of view. Active vision systems may give wide tracking area or accurate 3D shape estimation due to the variable field of view. One typical class of active vision systems is the class of pan-tilt camera systems, and the standard objective is to track a given target object [1].

This paper discusses what tasks should be selected to control active vision systems and which order of priority should be set for the selected tasks, since the task selection problem or the task sequencing problem is one of major and significant problems in the area of robot control including vision-based control [2], [3]. Possible tasks to determine an optimal camera placement for a given target object are to obtain high resolvability and to keep visibility of image features, the field of view and the depth of field in a certain range [4]. In particular, several quantitative measures of resolvability have been proposed in [5], [6], [7] where resolvability is called motion perceptibility. The perceptibility measures are quantitative scaling measures from error of the measured velocity of image features in the image plane to error of the corresponding velocity computed in the camera coordinate frame. The measures are calculated by using the image Jacobian matrix that relates the camera velocity in the camera coordinate frame to the velocity of image features in the image plane. In particular, the perceptibility ellipsoid is defined by a product of all the singular values of the image Jacobian matrix. Therefore the measure can be easily computed for vision-based control systems, although quantization error analysis can be achieved only for a specific active vision system [8]. The motion perceptibility is successfully integrated with visual servoing [9], since visual servo techniques basically use the image Jacobian matrix to derive the control input signal.

This paper first shows that optimization of perceptibility measures should not be selected for a primary task to control active vision systems. This follows from the fact that the perceptibility measures are basically optimized when the tracked target is captured at the edge of the camera's field of view. Hence optimization of a perceptibility measure may produce unreasonable motion responses for active vision systems, when it is selected as the primary task. Based on the discussion, this paper proposes optimization of a perceptibility measure as a secondary task, while target tracking is selected as the primary task. The primary task moves the center of gravity of image features to the center of the image plane. The secondary task gives high perceptibility of motion under the primary task. The perceptibility measures for the secondary task is induced by certain Jacobian matrices, not the image Jacobian matrix, since the image Jacobian matrix does not have any information for control of active vision systems with multiple cameras. Several examples demonstrate that the proposed task selection produces cost functions to obtain a reasonable camera placement and cooperative behavior for active vision systems with multiple cameras.

II. Motion Perceptibility

A. Motion Perceptibility Induced by Image Jacobian

This section introduces the quantitative measures of motion perceptibility proposed in [7].

Let us first define the image Jacobian matrix. A 3D point with coordinates (X_i, Y_i, Z_i) in the camera coordinate frame is projected onto the image plane as a 2D point with coordinates

$$\begin{aligned} \boldsymbol{s}_i &= \begin{bmatrix} u_i & v_i \end{bmatrix}^\top \\ &= \begin{bmatrix} \dfrac{X_i}{Z_i} & \dfrac{Y_i}{Z_i} \end{bmatrix}^\top \end{aligned} \tag{1}$$

where the focal length is set to 1. For given m image features, the image feature vector $\boldsymbol{s}$ is defined by

$$\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_1^\top & \boldsymbol{s}_2^\top & \dots & \boldsymbol{s}_m^\top \end{bmatrix}^\top . \tag{2}$$

Let (v_x, v_y, v_z) and $(\omega_x, \omega_y, \omega_z)$ denote the camera velocity and angular velocity in the camera coordinate frame, respec-

Y. Iwatani is with the Department of System Information Sciences, Tohoku University, Aramaki Aza Aoba 6-6-01, Sendai, Japan iwatani@ic.is.tohoku.ac.jp

tively. We set

$$\boldsymbol{v} = \begin{bmatrix} v_x & v_y & v_z & \omega_x & \omega_y & \omega_z \end{bmatrix}^\top. \tag{3}$$

The relationship between $\dot{\boldsymbol{s}}$ and $\boldsymbol{v}$ is given by

$$\dot{\boldsymbol{s}} = \boldsymbol{L}\boldsymbol{v} \tag{4}$$

where

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_1^\top & \boldsymbol{L}_2^\top & \dots & \boldsymbol{L}_m^\top \end{bmatrix}^\top, \tag{5}$$

$$\boldsymbol{L}_i = \begin{bmatrix} -\frac{1}{Z_i} & 0 & \frac{u_i}{Z_i} & u_i v_i & -(1+u_i^2) & v_i \\ 0 & -\frac{1}{Z_i} & \frac{v_i}{Z_i} & 1+v_i^2 & -u_i v_i & -u_i \end{bmatrix}. \tag{6}$$

The matrix $\boldsymbol{L}$ is called the image Jacobian matrix or the interaction matrix. This paper sometimes consider motion in the (X, Z) plane to simplify our discussion. The image Jacobian matrix in the (X, Z) plane is given by (5) with

$$\boldsymbol{L}_i = \begin{bmatrix} -\frac{1}{Z_i} & \frac{u_i}{Z_i} & -(1+u_i^2) \end{bmatrix}. \tag{7}$$

One measure of motion perceptibility is the minimum singular value of $\boldsymbol{L}$, since the error bound is given by

$$\|\Delta \boldsymbol{v}\| \le \frac{\|\Delta \dot{\boldsymbol{s}}\|}{\sigma_{\boldsymbol{L}\min}} \tag{8}$$

where $\Delta \boldsymbol{v}$ is error in the computed camera velocity, $\Delta \dot{\boldsymbol{s}}$ is error in the measured visual feature velocity and $\sigma_{\boldsymbol{L}\min}$ is the minimum singular value of $\boldsymbol{L}$. It is seen from (8) that $\sigma_{\boldsymbol{L}\min}$ should be large to reduce the error $\Delta \boldsymbol{v}$. Thus maximization of $\sigma_{\boldsymbol{L}\min}$ is a candidate for a task to control active vision systems.

The perceptibility ellipsoid defined by

$$w_i = \begin{cases} \sqrt{\det(\boldsymbol{L}^\top \boldsymbol{L})}, & \text{if } k \ge \ell, \\ \sqrt{\det(\boldsymbol{L}\boldsymbol{L}^\top)}, & \text{if } k < \ell \end{cases} \tag{9}$$

for the $k \times \ell$ matrix $\boldsymbol{L}$ also provides a quantitative measure of motion perceptibility. Another alternative measure of motion perceptibility is the condition number for $\boldsymbol{L}$, that is, $\sigma_{\boldsymbol{L}\max}/\sigma_{\boldsymbol{L}\min}$ where $\sigma_{\boldsymbol{L}\max}$ is the maximum singular value of $\boldsymbol{L}$ [5]. The condition number should be small to reduce the error $\Delta \boldsymbol{v}$.

Remark 1: The image Jacobian matrix $\boldsymbol{L}$ defined by (5) is valid only for single-camera systems. For multi-camera systems, the image Jacobian matrix is given by the block diagonal matrix whose diagonal parts are the image Jacobian matrices of the corresponding cameras. Hence the optimal motion perceptibility induced by image Jacobian for multi-camera systems is determined by the individual optimization for each camera.

B. Motion Perceptibility Induced by Target Jacobian

The measures described above represent motion perceptibility between the image feature velocity $\dot{\boldsymbol{s}}$ and the camera velocity $\boldsymbol{v}$ in the camera frame. The goal for active vision systems is to obtain the target state $\boldsymbol{q}_t \in \mathbb{R}^n$ in the world coordinate frame or in the robot coordinate frame when the target is also one of end effectors in the system. Perceptibility measures for $\dot{\boldsymbol{q}}_t$ can be also defined in a similar manner to the measures induced by the image Jacobian matrix. To this end, we use

$$\dot{\boldsymbol{s}} = \boldsymbol{J}_t \dot{\boldsymbol{q}}_t + \boldsymbol{J}_a \dot{\boldsymbol{q}}_a \tag{10}$$

where $\boldsymbol{q}_a$ is the vector of generalized coordinates of active cameras in the system and

$$\boldsymbol{J}_t = \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{q}_t}, \tag{11}$$

$$\boldsymbol{J}_a = \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{q}_a}. \tag{12}$$

The matrix $\boldsymbol{J}_t$ is called the target Jacobian matrix in this paper.

One perceptibility measure between $(\dot{\boldsymbol{s}}, \dot{\boldsymbol{q}}_a)$ and $\dot{\boldsymbol{q}}_t$ is given by

$$\rho := \frac{\sigma_{\boldsymbol{J}_t \min}}{\sigma_{\boldsymbol{J}_a \max}} \tag{13}$$

where $\sigma_{\boldsymbol{J}_t \min}$ is the minimum singular value of $\boldsymbol{J}_t$, and $\sigma_{\boldsymbol{J}_a \max}$ is the maximum singular value of

$$\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix}. \tag{14}$$

The error bound is given by

$$\|\Delta \dot{\boldsymbol{q}}_t\| \le \frac{\sigma_{\boldsymbol{J}_a \max}}{\sigma_{\boldsymbol{J}_t \min}} \left\| \Delta \begin{bmatrix} \dot{\boldsymbol{s}} \\ \dot{\boldsymbol{q}}_a \end{bmatrix} \right\|. \tag{15}$$

The camera state $\boldsymbol{q}_a$ usually can be obtained from sensors other than cameras. Thus (15) implies that the computational error $\Delta \dot{\boldsymbol{q}}_t$ is bounded by sensing error. The measure ρ defined by (13) should be large to reduce the error $\Delta \dot{\boldsymbol{q}}_t$. When $k \ge n$ for the $k \times n$ matrix $\boldsymbol{J}_t$, we can use the minimum singular value of

$$\boldsymbol{J}_c = \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix}^+ \boldsymbol{J}_t \tag{16}$$

as a perceptibility measure, where $\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix}^+$ is the Moore-Penrose inverse of $\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix}$. When $k < n$, $\boldsymbol{J}_c$ is not full rank and the minimum singular value of $\boldsymbol{J}_c$ does not provide an appropriate measure. Meanwhile (13) is always available.

An alternative and simpler measure is the minimum singular value of $\boldsymbol{J}_t$. The error bound is represented by

$$\|\Delta \dot{\boldsymbol{q}}_t\| \le \frac{\|\Delta(\dot{\boldsymbol{s}} - \boldsymbol{J}_a \dot{\boldsymbol{q}}_a)\|}{\sigma_{\boldsymbol{J}_t \min}}. \tag{17}$$

Furthermore, we can use the perceptibility ellipsoid

$$w_t = \begin{cases} \sqrt{\det(\boldsymbol{J}_t^\top \boldsymbol{J}_t)}, & \text{if } k \ge n, \\ \sqrt{\det(\boldsymbol{J}_t \boldsymbol{J}_t^\top)}, & \text{if } k < n \end{cases} \tag{18}$$

as a perceptibility measure induced by target Jacobian. The advantage to use only $\boldsymbol{J}_t$ is that the closed form solution of the perceptibility ellipsoid can be computed easier than the minimum singular value. Moreover, the perceptibility ellipsoid w_t may be differentiable even when the minimum singular value is not differentiable as shown later in Example 2.

III. Motivating Examples

A. Control of an Active Vision System Using Motion Perceptibility Induced by Image Jacobian

It is seen from (5) and (6) that the minimum singular value $\sigma_{\boldsymbol{L}\,\min}$ is large when $\|\boldsymbol{s}_i\|$ is large or when the depth Z_i is small. In particular, $\sigma_{\boldsymbol{L}\,\min}$ is sometimes maximized when $\|\boldsymbol{s}_i\| \to \infty$ or when $Z_i = 0$. Thus maximization of $\sigma_{\boldsymbol{L}\,\min}$ may produce unreasonable motion responses as shown below.

Let us consider an active vision system with an active camera, a linear slider and a revolute joint in the (X, Z) plane as illustrated in Fig. 1. The camera's X position x_a and the angle θ can be controlled. The state vector of the camera is represented by

$$\boldsymbol{q}_a = \begin{bmatrix} x_a & \theta \end{bmatrix}^\top . \tag{19}$$

The Z position of the camera is always set at 0. Let a point be given as a target captured in the image plane. The target position is denoted by

$$\boldsymbol{q}_t = \begin{bmatrix} x_t & z_t \end{bmatrix}^\top \tag{20}$$

in the world coordinate. Then we have the image Jacobian (5) with (7) and $m = 1$ where

$$Z_1 = -(x_t - x_a)\sin\theta + z_t \cos\theta, \tag{21}$$

$$u_1 = \frac{(x_t - x_a)\cos\theta + z_t \sin\theta}{-(x_t - x_a)\sin\theta + z_t\cos\theta}. \tag{22}$$

Suppose first that $\theta = 0$, that is, the optical axis of the camera is parallel to the Z axis. Then the image Jacobian matrix is given by

$$\boldsymbol{L} = \begin{bmatrix} -\frac{1}{z_t} & \frac{x_t - x_a}{z_t^2} & -\left(1 + \frac{(x_t - x_a)^2}{z_t^2}\right) \end{bmatrix}. \tag{23}$$

From the above equation, $|x_t - x_a|$ goes to infinity as $\sigma_{\boldsymbol{L}\,\min}$ increases. This implies that the target should be captured at the edge of the camera field of view from the viewpoint of motion perceptibility induced by the image Jacobian. Thus maximization of $\sigma_{\boldsymbol{L}\,\min}$ does not yield a reasonable motion response in this case.

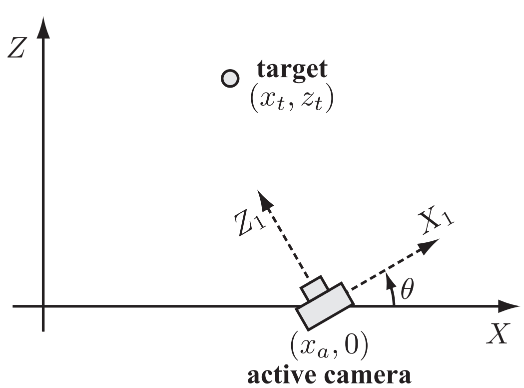

Fig. 1. Active vision system with an active camera, a linear slider and a revolute joint.

Suppose next that $x_a = x_t$. Then we obtain

$$\boldsymbol{L} = \begin{bmatrix} -\frac{1}{z_t\cos\theta} & \frac{\tan\theta}{z_t\cos\theta} & \frac{-1}{\cos^2\theta} \end{bmatrix} \tag{24}$$

and

$$\sigma_{\boldsymbol{L}\,\min} = \frac{1 + z_t^2}{z_t^2 \cos^4\theta}. \tag{25}$$

Maximization of $\sigma_{\boldsymbol{L}\,\min}$ yields that $|\theta| \to \pi/2$. This also means that the target should be captured at the edge of the camera field of view.

B. Control of Active Vision Systems Using Motion Perceptibility Induced by Target Jacobian

Consider again the system discussed in the previous section. Then we have

$$\boldsymbol{J}_t = \frac{1}{Z_1^2}\begin{bmatrix} z_t & -(x_t - x_a) \end{bmatrix}, \tag{26}$$

$$\boldsymbol{J}_a = \frac{1}{Z_1^2}\begin{bmatrix} -z_t & (x_t - x_a)^2 + z_t^2 \end{bmatrix} \tag{27}$$

where Z_1 is given by (21).

The measure $\sigma_{\boldsymbol{J}_t\,\min}$ is written by

$$\sigma_{\boldsymbol{J}_t\,\min} = \frac{\sqrt{z_t^2 + (x_t - x_a)^2}}{Z_1^2} \tag{28}$$

The measure $\sigma_{\boldsymbol{J}_t\,\min}$ goes to infinity as $|x_a| \to \infty$ when $\theta = 0$ or as $|\theta| \to \pi/2$ when $x_a = x_t$. Hence maximization of $\sigma_{\boldsymbol{J}_t\,\min}$ is not suitable for a primary task to control active vision systems.

Moreover, straightforward calculations lead to

$$\rho = \sqrt{\frac{z_t^2 + (x_t - x_a)^2}{Z_1^2 + z_t^2 + \{(x_t - x_a)^2 + z_t^2\}^2}}. \tag{29}$$

It is seen from the optimality condition that ρ has a local maximum at $x_t = x_a$ and $\theta = \pi/2$. Thus Z_1 may approach to zero when a greedy algorithm determines the local best. This also implies that the target should be captured at the edge of the camera field of view from the viewpoint of motion perceptibility induced by target Jacobian.

IV. Task Selection for Control of Active Vision Systems

The previous section shows that the measures of motion perceptibility are not appropriate for a primary task to control active vision systems. This paper proposes target tracking as a primary task and maximization of a perceptibility measure induced by target Jacobian as a secondary task to control of active vision systems.

The target tracking task for systems with a single active camera is to minimize

$$\left\| \sum_{i=1}^{m} \boldsymbol{s}_i \right\| \tag{30}$$

where the center of the image plane is set to $\mathbf{0}$. The measure (30) is not replaced by $\|\boldsymbol{s}\|$ or $\|\boldsymbol{s} - \boldsymbol{s}_g\|$ for a goal position $\boldsymbol{s}_g$, since (30) leads to redundancy but minimization of $\|\boldsymbol{s}\|$ or

$\|\boldsymbol{s}-\boldsymbol{s}_g\|$ does not allow performing a secondary task for an active vision system that tracks multiple targets. In particular, minimization of $\|\boldsymbol{s}\|$ may yield unreasonable behavior for multi-target systems as will shown in Example 2. Note that minimization of $\|\boldsymbol{s}-\boldsymbol{s}_g\|$ is the visual servoing problem. Let us then consider a system with r active cameras. The measure for the tracking task in active camera $j \in \{1,\dots,r\}$ is defined by

$$\left\| \sum_{i=\mathbb{I}_j} \boldsymbol{s}_i \right\| \tag{31}$$

where $\mathbb{I}_j$ is the index set that contains indexes of image features captured by camera j. The measure defined by (31) is minimized for each active camera. Note that

$$\bigcup_{j=1}^{r} \mathbb{I}_j \neq \{1,2,\dots,m\} \tag{32}$$

where m is the number of the image features, when the system has a stationary camera.

The secondary task uses one of the perceptibility measures induced by target Jacobian not image Jacobian, since the image Jacobian matrix does not have any information for control of multi-camera systems as discussed at Remark 1 in Section II-A. Hence the measures induced by image Jacobian do not make cooperative behavior of active cameras. On the other hand, the measures induced by target Jacobian are valid even for multi-camera systems as shown later in Example 3.

The following three examples demonstrate that the proposed task selection provides appropriate cost functions that produces reasonable motion responses for active vision systems.

Example 1: Consider again the active vision system discussed in Section III-A. Suppose that the primary task is completed. Then it holds that

$$x_t - x_a = -z_t \tan\theta \tag{33}$$

and we have

$$\boldsymbol{J}_t = \frac{\cos\theta}{z_t}\begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}, \tag{34}$$

$$\boldsymbol{J}_a = \frac{1}{z_t}\begin{bmatrix}-\cos^2\theta & z_t\sin^2\theta\end{bmatrix}. \tag{35}$$

From (33) and (34), the minimum singular value $\sigma_{\boldsymbol{J}_t\,\min}$ is maximized at

$$x_a = x_t, \quad \theta = 0. \tag{36}$$

The measure ρ is given by

$$\rho = \sqrt{\frac{\cos^2\theta}{z_t^2 + \cos^4\theta + z_t^2\sin^4\theta}} \tag{37}$$

when (33) holds. It is seen from the optimality condition that ρ also has a local maximum at (36). The solution (36) gives a reasonable camera placement, and it demonstrates the validity of the proposed task selection.

Example 2: Consider the system discussed in Example 1 again, but it tracks two target points represented by (x_{t1}, z_{t1}) and (x_{t2}, z_{t2}) in the (X, Z) plane. The state vector of the target points is denoted by

$$\boldsymbol{q}_t = \begin{bmatrix}x_{t1} & z_{t1} & x_{t2} & z_{t2}\end{bmatrix}^\top. \tag{38}$$

Let us first chose $\|\boldsymbol{s}\|$ as the measure of the primary task. Minimization of $\|\boldsymbol{s}\|$ is achieved at

$$x_a = \frac{z_{t1}x_{t2} - z_{t2}x_{t1}}{z_{t1} - z_{t2}}, \tag{39}$$

$$\theta = \tan^{-1}\left(\frac{x_{t_1} - x_{t2}}{z_{t1} - z_{t2}}\right). \tag{40}$$

It is seen that $|\theta| \to \pi/2$ as $z_{t1} \to z_{t2}$. This is not appropriate behavior to perform tracking.

Let us use $\|\boldsymbol{s}_1 + \boldsymbol{s}_2\|$ as the measure. The primary task is completed when

$$\theta = \frac{1}{2}\tan^{-1}\frac{(x_{t1}-x_a)z_{t2} + (x_{t2}-x_a)z_{t1}}{(x_{t1}-x_a)(x_{t2}-x_a) - z_{t1}z_{t2}} \tag{41}$$

holds. Suppose that $z_{t1} = z_{t2}$ to compare the result obtained above with the proposed approach. Then it can be seen from the optimality condition that the minimum singular value $\sigma_{\boldsymbol{J}_t\,\min}$ is maximized at

$$x_a = \frac{x_{t1} + x_{t2}}{2}, \quad \theta = 0 \tag{42}$$

when

$$|x_{t1} - x_{t2}| \neq 2|z_t| \tag{43}$$

holds. The condition (43) is a sufficient condition to derive the optimal solution (42), since it guarantees the differentiability of $\sigma_{\boldsymbol{J}_t\,\min}$ at (42). Fig. 2 shows that (42) gives the optimal solution that maximizes $\sigma_{\boldsymbol{J}_t\,\min}$ when (43) is not satisfied. The ellipsoid measure defined by (18) is always maximized at (42), and (43) is not necessary for the ellipsoid measure.

The proposed task selection is also available when $z_{t1} \neq z_{t2}$. Let the two target points be set at $(x_{t1}, z_{t1}) = (0.1, 1.1)$ and $(x_{t2}, z_{t2}) = (-0.1, 0.9)$. The minimum singular value $\sigma_{\boldsymbol{J}\,\min}$ and the measure ρ defined by (13) are maximized at

$$x_a = 0.2780, \quad \theta = 0.2790. \tag{44}$$

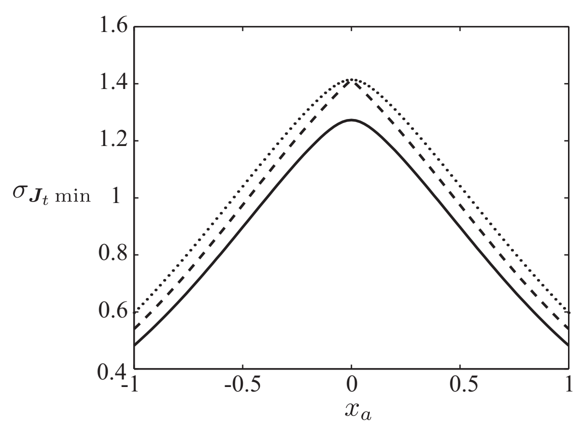

Fig. 2. The minimum singular value $\sigma_{\boldsymbol{J}_t\,\min}$ versus the angle θ. Solid line: $x_{t1} = -x_{t2} = 0.9$ and $z_{t1} = z_{t2} = 1$. Dashed line: $x_{t1} = -x_{t2} = 1$ and $z_{t1} = z_{t2} = 1$. Dotted line: $x_{t1} = -x_{t2} = 1.1$ and $z_{t1} = z_{t2} = 1$.

and

$$x_a = 0.3407, \quad \theta = 0.3354 \tag{45}$$

respectively, as illustrated in Fig. 3. The optimal image plane is not oriented parallel to the plane on the tracked targets. In fact, the optimal angles in (44) and (45) are less than

$$\frac{\pi}{4} = \tan^{-1} \frac{z_{t1} - z_{t2}}{x_{t1} - x_{t2}}. \tag{46}$$

Example 3: Consider an active vision system with two active cameras that share a linear slider and each camera has a revolute joint as depicted in Fig. 4. The state of the two active cameras in the world coordinate is denoted by

$$\boldsymbol{q}_a = \begin{bmatrix} x_{a1} & \theta_1 & x_{a2} & \theta_2 \end{bmatrix}^\top. \tag{47}$$

Suppose that the primary task is accomplished. Then we have

$$\boldsymbol{J}_t = \frac{1}{z_t} \begin{bmatrix} \cos^2\theta_1 & \cos\theta_1 \sin\theta_1 \\ \cos^2\theta_2 & \cos\theta_2 \sin\theta_2 \end{bmatrix}, \tag{48}$$

$$\boldsymbol{J}_a = \frac{1}{z_t} \begin{bmatrix} -\cos^2\theta_1 & z_t \sin^2\theta_1 & 0 & 0 \\ 0 & 0 & -\cos^2\theta_2 & z_t \sin^2\theta_2 \end{bmatrix}. \tag{49}$$

The minimum singular value $\sigma_{\boldsymbol{J}_t \min}$ is maximized at

$$\boldsymbol{q}_a = \pm \begin{bmatrix} x_t + z_t & \pi/4 & -x_t - z_t & -\pi/4 \end{bmatrix}^\top. \tag{50}$$

The two cameras are symmetry with respect to the X axis at (50). The optimal camera placement is natural, since $\sigma_{\boldsymbol{J}_t \min} = \sigma_{\boldsymbol{J}_t \max}$ and the resolution in the two directions is same there. The resulting two shapes of $\sigma_{\boldsymbol{J}_t \min}$ and ρ are similar to each other as shown in Figs. 5 and 6.

Remark 2: If a perceptibility measure induced by image Jacobian is selected for the secondary task in Example 3, the measure is optimized at

$$x_{a1} = x_{a2} = x_t, \quad \theta_1 = \theta_2 = 0. \tag{51}$$

The optimal solution (51) is same as (36) for the single-camera system, and no cooperative behavior is found. For example, (51) is not valid for 3D reconstruction of the target

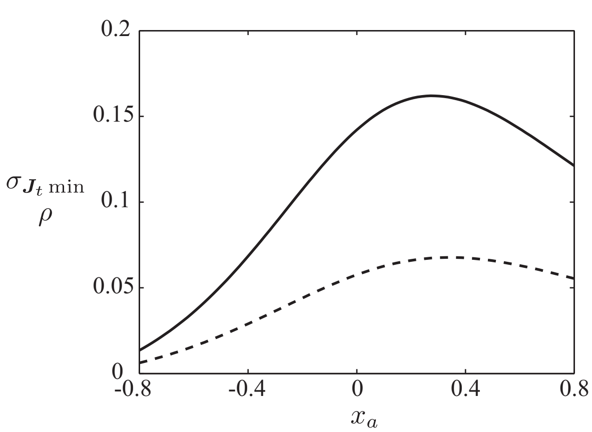

Fig. 3. Solid line: the minimum singular value $\sigma_{\boldsymbol{J}_t \min}$ versus the angle θ. Dashed line: the measure ρ versus the angle θ. The two targets are set at $(x_{t1}, z_{t1}) = (0.1, 1.1)$ and $(x_{t2}, z_{t2}) = (-0.1, 0.9)$ for the both cases.

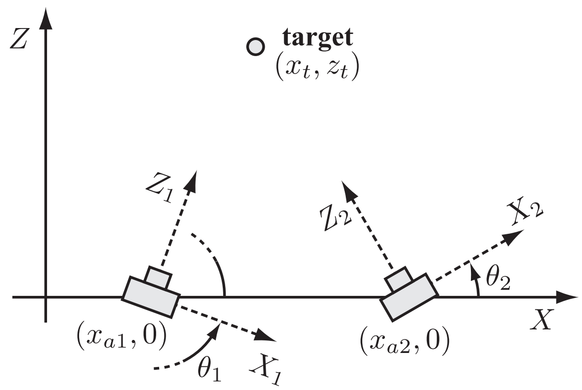

Fig. 4. Active vision system with two active cameras that share a linear slider and each camera has a revolute joint.

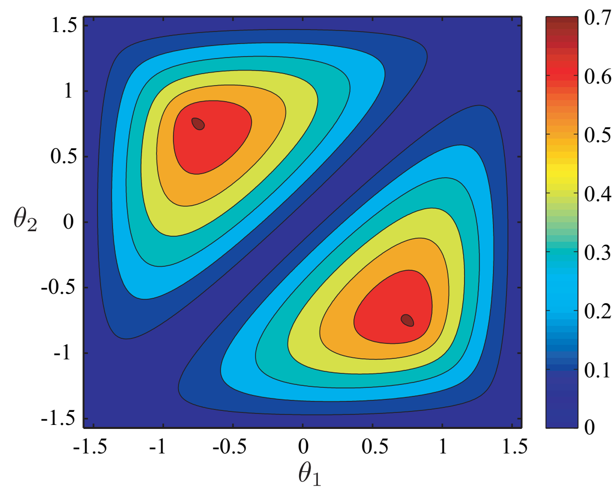

Fig. 5. The minimum singular value $\sigma_{\boldsymbol{J}_t \min}$

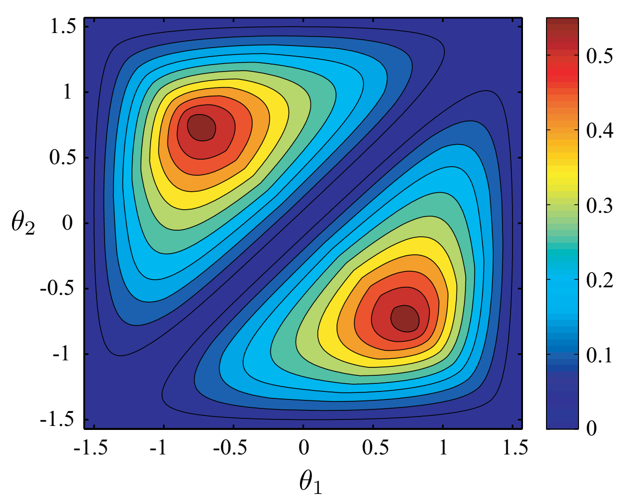

Fig. 6. The measure ρ

position. On the other hand, perceptibility measures induced by target Jacobian provide reasonable solutions for both single-camera systems and multi-camera systems.

V. Controller Design

Controllers for active vision systems can be designed by using the task function approach for redundant robot systems [10]. Suppose that the input signal $\boldsymbol{u}$ is given by $\boldsymbol{u} = \dot{\boldsymbol{q}}_a$. The error function for the primary task can be written by

$$\boldsymbol{e} = \begin{bmatrix} \boldsymbol{e}_1^\top & \boldsymbol{e}_2^\top & \dots & \boldsymbol{e}_r^\top \end{bmatrix}^\top, \tag{52}$$

$$\boldsymbol{e}_j = \sum_{i=\mathbb{I}_j} \boldsymbol{s}_i. \tag{53}$$

The controller based on the task function approach is represented by

$$\boldsymbol{u} = -k_1 \boldsymbol{J}_e^+ \boldsymbol{e} + k_2 (\boldsymbol{I} - \boldsymbol{J}_e^+ \boldsymbol{J}_e) \frac{\partial \phi}{\partial \boldsymbol{q}_a} \tag{54}$$

where k_1 and k_2 are positive constants, ϕ is a selected perceptibility measure and

$$\boldsymbol{J}_e = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{q}_a}. \tag{55}$$

The selected perceptibility measure should be differentiable to implement (54). Note that (54) also can be used in practice, if the selected measure is differentiable for all $\boldsymbol{q}_a$ except on a set of measure zero. Thus the measures ρ and $\sigma_{\boldsymbol{J}_t \min}$ are available in practice, although they may be undifferentiable at a point as illustrated in Fig. 2. If differentiability is still desired everywhere, the ellipsoid measure w_t is useful.

Example 4: Consider the system discussed in Example 3 again. The minimum singular value $\sigma_{J_t \min}$ is selected as a perceptibility measure. The controller (54) with

$$k_1 = 1, \quad k_2 = 2 \tag{56}$$

is implemented. Fig. 7 confirms the validity of (54) where the target is fixed at $(x_t, z_t) = (0, 1)$. The primary task is accomplished, since $\|\boldsymbol{e}\|$ goes to zero. The secondary task is also achieved, since the state $\boldsymbol{q}_a$ converges to (50).

VI. Conclusion

This paper has discussed the task selection problem for active vision systems. Optimization of a motion perceptibility measure is not suitable for a primary task to control active vision systems, since it may produce unreasonable motion responses. This paper has proposed target tracking as a primary task and optimization of a perceptibility measure induced by target Jacobian as a secondary task. The secondary task gives high perceptibility of motion under the primary task.

To balance the rates of change of $\dot{\boldsymbol{s}}$ and $\dot{\boldsymbol{q}}_a$, a weighting matrix $\boldsymbol{W}$ may be introduced. If the weight is required, the perceptibility measured are computed with

$$\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix} \boldsymbol{W}, \tag{57}$$

instead of $\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{J}_a \end{bmatrix}$.

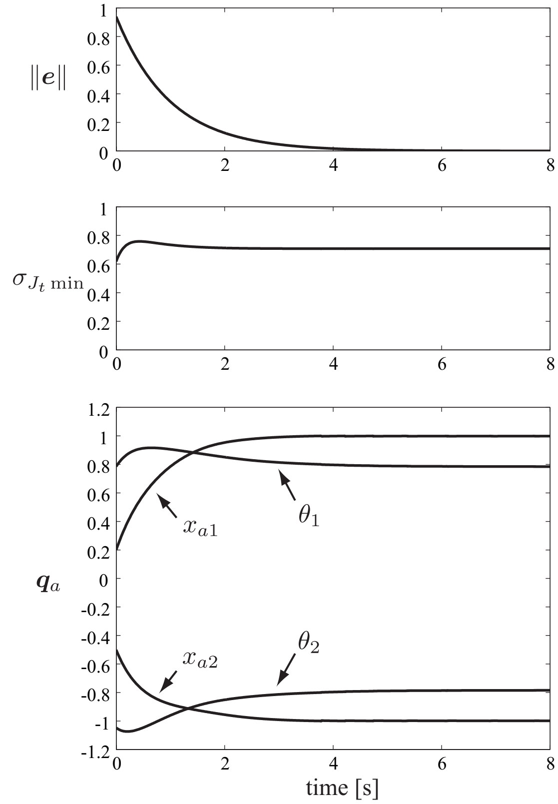

Fig. 7. Time responses of Example 4

References

[1] Y. Nakabo, M. Ishikawa, H. Toyoda, and S. Mizuno, "1ms column parallel vision system and it's application of high speed target tracking," in *Proc. of IEEE International Conference on Robotics and Automation*, 2000, pp. 650–655.

[2] E. Marchand and G. D. Hager, "Dynamic sensor planning in visual servoing," in *Proc. of IEEE International Conference on Robotics and Automation*, 1998, pp. 1988–1993.

[3] N. Mansard and F. Chaumette, "Task sequencing for high-level sensor-based control," *IEEE Trans. on Robotics*, vol. 23, no. 1, pp. 60–72, 2007.

[4] K. A. Tarabanis, R. Y. Tsai, and P. K. Allen, "The MVP sensor planning system for robotic vision tasks," *IEEE Trans. on Robotics and Automation*, vol. 11, no. 1, pp. 72–85, 1995.

[5] J. T. Feddema, C. S. G. Lee, and O. R. Mitchell, "Weighted selection of image features for resolved rate visual feedback control," *IEEE Trans. on Robotics and Automation*, vol. 7, no. 1, pp. 31–47, 1991.

[6] B. J. Nelson and P. K. Khosla, "Force and vision resolvability for assimilating disparate sensor feedback," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 5, pp. 714–731, 1996.

[7] R. Sharma and S. Hutchinson, "Motion perceptibility and its application to active vision-based servo control," *IEEE Trans. on Robotics and Automation*, vol. 13, no. 4, pp. 607–617, 1997.

[8] Y. Nakabo, T. Mukai, Y. Hattori, Y. Takeuchi, and N. Ohnishi—, "Variable baseline stereo tracking vision system using high-speed linear slider," in *Proc. of IEEE International Conference on Robotics and Automation*, 2005, pp. 1567–1572.

[9] L. Deng, F. Janabi-Sharifi, and W. J. Wilson, "Hybrid motion control and planning strategies for visual servoing," *IEEE Trans. on Industrial Electronics*, vol. 52, no. 4, pp. 1024–1040, 2005.

[10] C. Samson, M. L. Borgne, and B. Espiau, *Robot Control: the Task Function Approach*. Oxford University Press, 1991.