Adaptive Visual Servoing for Legged Robots
– Vision-Cued Swaying of Legged Robots in Unknown Environments –

Koh Hosoda, Takahiro Miyashita, Susumu Takeuchi, and Minoru Asada

Dept. of Adaptive Machine Systems, Osaka University
hosoda@mech.eng.osaka-u.ac.jp

Abstract

This paper describes a method to achieve a vision-cued swaying task in unknown environments utilizing adaptive visual servoing. The proposed method has a hybrid structure consisting of a controller to keep the distances between feet constant (a stance servoing controller), and an adaptive visual servoing controller. Making use of the method, the motion of each joint need not be pre-programmed, but is generated by the method according to the motion of visual cues. An experimental result demonstrates how the proposed method realizes a vision-cued swaying behavior of the legged robot.

1 Introduction

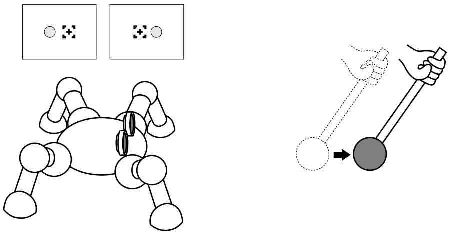
Among mobile abilities of robots, legged locomotion has an advantage to the others owing to its adaptivity/robustness against changes of terrain. The other reason why legged locomotion receives the attention is that most of natural living things such as human, animals, and insects utilize the ability. Because of these reasons, there have been numerous studies on legged locomotion [1].

Observing natural creatures, we can find that external sensors, such as vision, tactile and force sensors, play great roles to realize legged locomotion. However most of the previous work on legged robots are focusing on control utilizing *internal* sensors. They consider kinematics and/or dynamics of the legged robot and the environment, and calculate desired trajectories of the joints. Then, they apply the internal sensory feedback control based on the calculated trajectories, which cannot cope with changes of the environment. On the other hand, utilizing external sensors, the robot can be adaptive and robust to changes of the environment, as the natural creatures do.

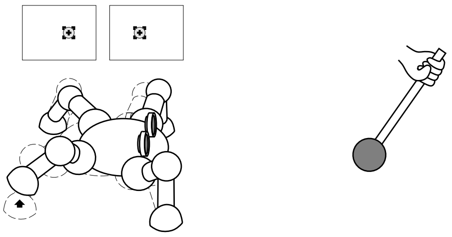
A vision sensor is one of the most powerful external sensors to make the legged robot robust against changes of the environment since the robot can capture global information without physically interacting with the environment. In the previous work on vision-based legged robots [2, 3], they make a hierarchical control structure of a lower servoing level and an upper navigating level. They utilize the vision sensors in the upper level to generate the desired inputs for the lower level servoing. However, vision sensors can also be utilized for generating leg motions in the servoing level, that is, visual servoing. Utilizing visual servoing, one can make the robot generate robust/adaptive leg motions against changes of the environment. Another interesting feature of applying visual servoing to a legged robot is that one can generate reactive swaying behaviors of the legged robot by showing some visual cues (see Figure 1). One can make the robot sway to track visual targets (Figure 1(a),(b)). By combining a certain gait strategy, one may even make the robot walk according to the movement of the visual targets (Figure 1(c)).

Recently, visual servoing has been received increasing attention in building robust/high-response robot systems [4–11]. The controller feeds the visual information back to control inputs directly, which makes the closed loop system robust against disturbance. Applying visual servoing to a legged robot, motions of the robot are generated by the movement of the visual cues, and therefore the resultant system becomes robust and adaptive against changes of the environment. There have been many studies on visual servoing applied to manipulators, but only one for legged robots [12] to the best of our knowledge. The main difficulty to apply visual servoing control to a legged robot is that it is not fixed on the ground unlike manipulators. In [12], they have proposed a vision-based servoing for legged robots utilizing visual servoing. However, they assume that the environment is known/calibrated, which makes the resultant system less adaptive.

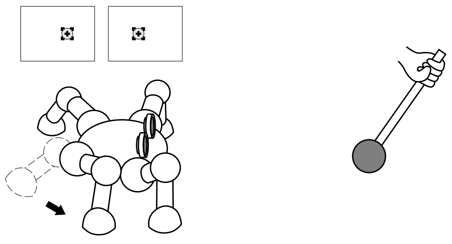
In this paper, we describe a method to achieve a vision-cued swaying task in unknown environments utilizing adaptive visual servoing [11]. The proposed method has a hybrid structure consisting of a controller to keep the distances between feet constant ("the stance servoing controller" in the following), and an adaptive visual servoing controller. Making use of the method, the motion of each joint does not need to be pre-programmed, but is generated by the method according to the motions of visual cues. An experimental result shows how the proposed method realizes



(a) If the taget moves,



(b) the proposed method makes the robot sway, and



(c) the robot walks.

Figure 1: Making a legged robot walk by showing visual cues

a vision-cued swaying behavior.

2 Vision-cued swaying utilizing adaptive visual servoing

2.1 Vision-cued swaying task

From the equipped camera(s), the legged robot can observe image features such as position, line length, contour length, and/or area of certain image patterns in the image plane(s). The task for the robot is a vision-cued swaying task to sway its body to make the image features converge to the desired values, that is, a visual servoing task. By moving the target that has image features, a swaying behavior of the robot can be realized (Figure 1(b)).

In [12], they assumed that all the parameters were known. To use such a method, however, the legged robot and the environment need to be calibrated carefully. Even if we have performed the tedious calibration process, the resultant system may become sensitive to modeling error and disturbance, and further, it cannot move in unknown environments. To make the robot accomplish a given task in unknown environments, we propose to apply the adaptive visual servoing [11]. The only assumption we need to make is that

- the relation between joint angles and positions of the feet with respect to the robot coordinate frame is known,

and nothing anymore. That is, all the intrinsic and extrinsic parameters of the camera(s), the shape of the terrain, and the relation between the legged robot and the image features are unknown.

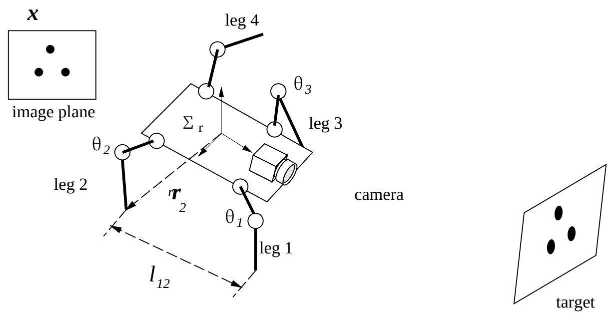


Figure 2: A 4-legged robot chasing the visual target

2.2 Controlling method for legged robots in unknown environments

There are a number of studies on controlling manipulators utilizing visual servoing, but there is only one study on controlling legged robots [12] to the best of our knowledge. One of the main difficulties to apply visual servoing to legged robots is that they are not fixed on the ground, whereas manipulators are fixed. If the position of a foot of the robot moves with respect to the ground by gait and/or slip while visual servoing is applied, one need to observe the amount of gait/slip and to re-calculate the Jacobian matrix which describes the relation between the velocities of the image features in the image planes and the velocities of the joints. In [12], to deal with such a problem, they have proposed a hybrid servoing structure consisting of a servoing controller to keep distances between feet constant (a stance servoing controller) and a visual servoing controller. However, this method can be applied only for calibrated robots and environments. To make legged robots act in unknown environments, we propose a hybrid servoing structure consisting of the stance servoing controller and an *adaptive* visual servoing controller.

In the rest of this paper, we deal with a 4-legged robot shown in Figure 2 whose legs 1, 2, and 3 contact with the ground at points with friction (Figure 2).

2.3 Stance servoing controller

First, we introduce the stance servoing controller to keep distances between feet constant. Let Σ_w and Σ_r

be a world coordinate frame fixed on the ground and a robot coordinate frame fixed to the robot body, respectively (see Figure 2). A vector from the origin of Σ_r to the i-th foot with respect to Σ_r, ${}^r\boldsymbol{r}_i$ $(i = 1, \cdots, 3)$, is a function of the joint angle vector of the i-th leg, $\boldsymbol{\theta}_i$:

$$ {}^r\boldsymbol{r}_i = {}^r\boldsymbol{r}_i(\boldsymbol{\theta}_i). \quad (1) $$

Differentiating eq.(1), we can obtain a velocity relation between $\dot{\boldsymbol{\theta}} = [\dot{\boldsymbol{\theta}}_1^T \ \dot{\boldsymbol{\theta}}_2^T \ \dot{\boldsymbol{\theta}}_3^T]^T \in \Re^9$ and ${}^r\dot{\boldsymbol{r}} = [{}^r\dot{\boldsymbol{r}}_1{}^T \ {}^r\dot{\boldsymbol{r}}_2{}^T \ {}^r\dot{\boldsymbol{r}}_3{}^T]^T \in \Re^9$,

$$ {}^r\dot{\boldsymbol{r}} = \boldsymbol{J}_{r\theta}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}}, \quad (2) $$

where $\boldsymbol{J}_{r\theta} = \partial^r\boldsymbol{r}/\partial\boldsymbol{\theta}$. Assume that each leg has necessary/sufficient degrees of freedom for positioning each foot, that is 3, and the Jacobian matrix $\boldsymbol{J}_{r\theta}$ becomes invertible without losing generality. In a case that a leg has more than 4 d.o.f., all we have to do is to find a certain subtask to deal with the redundancy.

Let a vector $\boldsymbol{l}$ denote a collection vector of the distances between feet,

$$ \boldsymbol{l} = \begin{bmatrix} \| {}^r\boldsymbol{r}_1 - {}^r\boldsymbol{r}_2 \| \\ \| {}^r\boldsymbol{r}_2 - {}^r\boldsymbol{r}_3 \| \\ \| {}^r\boldsymbol{r}_3 - {}^r\boldsymbol{r}_1 \| \end{bmatrix}. \quad (3) $$

Differentiating eq.(3), we can obtain a velocity relation

$$ \begin{aligned} \dot{\boldsymbol{l}} &= \boldsymbol{J}_{lr}({}^r\boldsymbol{r}){}^r\dot{\boldsymbol{r}} \\ &= \boldsymbol{J}_{l\theta}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}}, \end{aligned} \quad (4) $$

where $\boldsymbol{J}_{lr} = \partial\boldsymbol{l}/\partial\boldsymbol{r}^T$ and $\boldsymbol{J}_{l\theta} = \boldsymbol{J}_{lr}\boldsymbol{J}_{r\theta}$, respectively. From eq.(4), a feedback controller for joint velocities to keep the feet distance vector $\boldsymbol{l}$ as a constant desired vector $\boldsymbol{l}_d$ can be derived as

$$ \boldsymbol{u}_l = \boldsymbol{J}_{l\theta}{}^+\boldsymbol{K}_l(\boldsymbol{l}_d - \boldsymbol{l}) + (\boldsymbol{I}_9 - \boldsymbol{J}_{l\theta}{}^+\boldsymbol{J}_{l\theta})\boldsymbol{k}_1, \quad (5) $$

where $\boldsymbol{J}_{l\theta}{}^+$, $\boldsymbol{I}_9$, and $\boldsymbol{k}_1$ denote a pseudo inverse matrix of a matrix $\boldsymbol{J}_{l\theta}$, a 9×9 identity matrix, and an arbitrary vector that describes redundancy of the robot with respect to the stance servoing task, respectively. The matrix $\boldsymbol{K}_l \in \Re^{3\times3}$ is a gain matrix. Because the desired stance $\boldsymbol{l}_d$ may be the initial distance vector, one need not know the shape of the ground nor to sense the gait/slip displacement as far as all the legs of the robot keep touch with the ground at the initial configuration. Utilizing the null space of $\boldsymbol{J}_{l\theta}$, $(\boldsymbol{I}_9 - \boldsymbol{J}_{l\theta}{}^+\boldsymbol{J}_{l\theta})\boldsymbol{k}_1$, an adaptive visual servoing controller can be applied.

2.4 Adaptive visual servoing controller

From the camera(s) attached to the robot body, one can get some image features such as position, line length, contour length, and/or area of certain image patterns. Let a vector of the image features be $\boldsymbol{x} \in \Re^m$. Assume that the target is moving so slowly that one can neglect the velocity of the target comparing to the velocity of the robot. If the stance servoing controller (5) keeps the feet distance constant, we can apply an adaptive visual servoing controller [11].

As far as the stance servoing controller works, the image feature vector $\boldsymbol{x}$ is a function of the joint angle vector $\boldsymbol{\theta}$,

$$ \boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\theta}). \quad (6) $$

If the stance servoing controller does not work, $\boldsymbol{x}$ becomes also the function of gait/slip, therefore eq.(6) is not valid. Differentiating eq.(6), we get

$$ \dot{\boldsymbol{x}} = \boldsymbol{J}_{x\theta}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}}. \quad (7) $$

Here the Jacobian matrix $\boldsymbol{J}_{x\theta} = \partial\boldsymbol{x}/\partial\boldsymbol{\theta}^T$ consists not only of the kinematic parameters of the legged robot that are known, but also of intrinsic and extrinsic camera parameters and of the parameters of the environment. Therefore, by correcting $\boldsymbol{\theta}$ and $\boldsymbol{x}$ we need to estimate the matrix $\widehat{\boldsymbol{J}}_{x\theta}$ that satisfys eq.(7). We utilize a least squares method to identify the non-linear system in the discrete time domain [13]:

$$ \begin{aligned} &\widehat{\boldsymbol{j}}_i(k+1) - \widehat{\boldsymbol{j}}_i(k) = \\ &\frac{\{\boldsymbol{x}(k+1) - \boldsymbol{x}(k) - \widehat{\boldsymbol{J}}_{x\theta}(k)\boldsymbol{u}(k)\}_i}{\rho_i + \boldsymbol{u}(k)^T\boldsymbol{W}_i(k)\boldsymbol{u}(k)}\boldsymbol{W}_i(k)\boldsymbol{u}(k), \end{aligned} \quad (8) $$

where $\widehat{\boldsymbol{J}}_{x\theta}(k)$, $\widehat{\boldsymbol{j}}_i(k)$, $\boldsymbol{u}(k)(= T\dot{\boldsymbol{\theta}})$, ρ_i and $\boldsymbol{W}_i(k)$ denote a constant Jacobian matrix, its i-th row vector, a control input vector in the k-th step during sampling rate T, an appropriate positive constant and a weighting matrix, respectively. In a case that $\boldsymbol{W}$ is a covariance matrix and that ρ_i is in the range $0 < \rho \leq 1$, the proposed estimator is a well-known weighted recursive least squares estimator.

Using the estimated $\widehat{\boldsymbol{J}}_{x\theta}$, we can obtain a visual servoing control input $\boldsymbol{u}_v$ to track given desired trajectories $\boldsymbol{x}_d$,

$$ \boldsymbol{u}_v = \widehat{\boldsymbol{J}}_{x\theta}{}^+\boldsymbol{K}_v(\boldsymbol{x}_d - \boldsymbol{x}) + (\boldsymbol{I}_9 - \widehat{\boldsymbol{J}}_{x\theta}{}^+\widehat{\boldsymbol{J}}_{x\theta})\boldsymbol{k}_2, \quad (9) $$

where $\boldsymbol{K}_v$ and $\boldsymbol{k}_2$ denote a $m \times m$ gain matrix and an arbitrary vector that describes the redundancy of the robot with respect to the visual servoing task, respectively.

2.5 Hybrid structure of two servoing controllers

The legged robot has 3 legs contacting with the ground, therefore it has 9 degrees of freedom. The stance controller utilizes 3 degrees of freedom and the visual servoing controller uses 6 degrees of freedom if the number of image features is enough to determine the position/orientation of the robot. Therefore 9 degrees of freedom is necessary/sufficient to realize the visual tracking task.

As far as the stance controller keeps the distances between the feet, the Jacobian matrix $\widehat{\boldsymbol{J}}_{x\theta}$ is a function of $\boldsymbol{\theta}$, and the adaptive visual servoing controller works well. In this sense, the stance controller has an obvious priority over the visual servoing controller. According to the priority, from eqs.(5) and (9), we propose a hybrid structure to keep the stance and to make the image features converge to the desired ones,

$$\begin{aligned} \boldsymbol{u} &= \boldsymbol{J}_{l\theta}{}^{+}\boldsymbol{K}_l(\boldsymbol{l}_d - \boldsymbol{l}) \\ &\quad +(\boldsymbol{I}_9 - \boldsymbol{J}_{l\theta}{}^{+}\boldsymbol{J}_{l\theta})\{\widehat{\boldsymbol{J}}_{x\theta}(\boldsymbol{I}_9 - \boldsymbol{J}_{l\theta}{}^{+}\boldsymbol{J}_{l\theta})\}^{+} \\ &\quad \{\boldsymbol{K}_i(\boldsymbol{x}_d - \boldsymbol{x}) - \widehat{\boldsymbol{J}}_{x\theta}\boldsymbol{J}_{l\theta}{}^{+}\boldsymbol{K}_l(\boldsymbol{l}_d - \boldsymbol{l})\}. \end{aligned} \quad (10)$$

The diagram of the proposed method is shown in Figure 3. The proposed method changes the attitude of the legged robot to make the image features in the image planes converge to the desired ones. Therefore, by applying the method to a legged robot, one can realize a swaying behavior by moving the target.

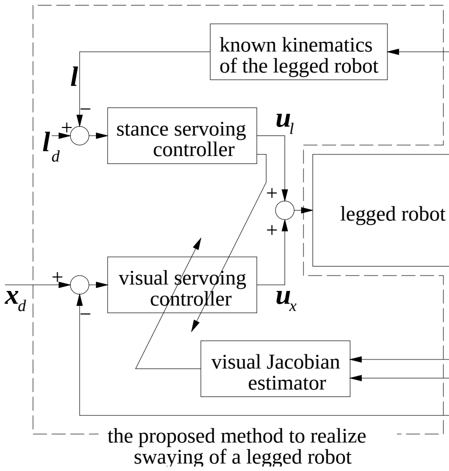


Figure 3: Block diagram of the proposed method

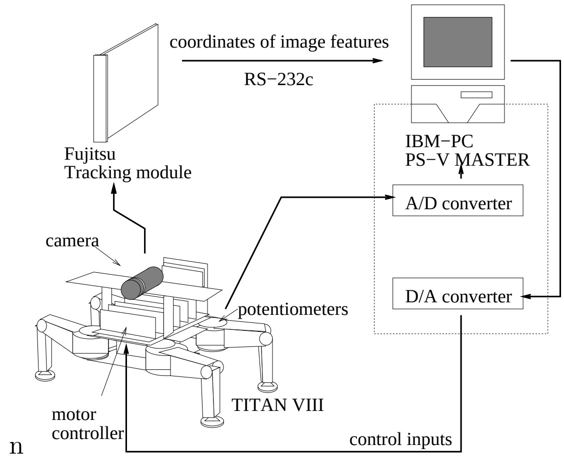


Figure 4: Experimental equipment

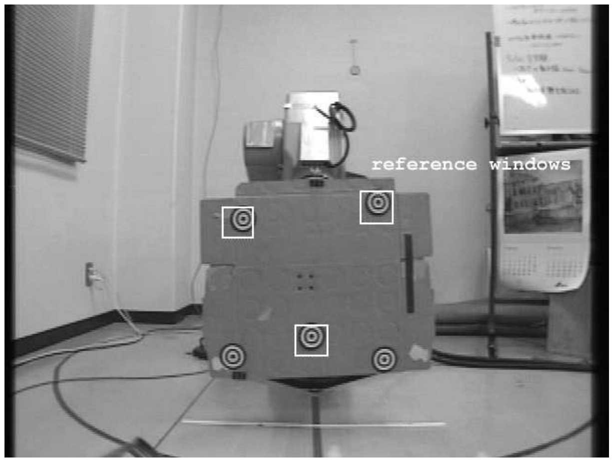


Figure 5: A sample image supplied for the legged robot

3 Experiments

To show how the proposed method generates a vision-cued swaying behavior of a legged robot, a preliminary experimental result is shown in this section.

3.1 Experimental setup

In Figure 4, a legged robot TITAN–VIII and its controller used for the experiment are shown. The quadruped walking robot TITAN–VIII is developed by Tokyo Institute of Technology [14], of which size is about 0.5(W)× 0.6(D)× 0.4(H)[m], and of which weight is about 25[kg]. The walking robot is equipped with one camera (UN401, ELMO). The image from the camera is sent to a tracking unit equipped with a high-speed correlation processor by Fujitsu Corp. utilizing a SAD (Sum of Absolute Difference) measure [15] of which image size is 512[pixel] × 512[pixel].

Before starting an experiment, we give three 16 × 16 certain patterns (called reference patterns) to be tracked. During the experiment the unit feeds coordinates where the correlation coefficient is the smallest with respect to the reference patterns to the host computer PS/V–MASTER (IBM Corp., CPU:Intel DX4 100MHz) through a RS–232C link in real–time (33[ms]). An example image is shown in Figure 5. Each joint of the walking robot is equipped with a potentiometer to observe its joint angle. The observed joint angles are sent to the computer through an A/D converter board. The computer calculates the desired joint velocities according to the proposed method and sends the commands to the velocity controllers of joints through a D/A converter board. A manipulator is used to move a board on which 3 target marks are drawn (see Figure 7). The distance between the target and the robot is about 1.0[m].

3.2 Experimental results

As the proposed method does not need any *a priori* knowledge on the Jacobian matrix $\boldsymbol{J}_{x\theta}$, we gave an arbitrary matrix as an intial matrix,

$$\boldsymbol{J}_{x\theta}(0) = \begin{bmatrix} 1.0 & 10 & 10 & 1.0 & -10 & -10 & 1.0 & -1.0 & -1.0 \\ -1.0 & -1.0 & -10 & 0.0 & 10 & 10 & 0.0 & -10 & 0.0 \\ -1.0 & -1.0 & 1.0 & -1.0 & 1.0 & -1.0 & -1.0 & -1.0 & 0.0 \\ 0.0 & -1.0 & 1.0 & 0.0 & -1.0 & -1.0 & 0.0 & 1.0 & 0.0 \\ 1.0 & 10 & 1.0 & -1.0 & -10 & -1.0 & 1.0 & -10 & -1.0 \\ -1.0 & -10 & -10 & -1.0 & 10 & 100 & -1.0 & -1.0 & 0.0 \end{bmatrix}.$$

We set the weighting matrix $\boldsymbol{W}_i$ in eq.(8) as an identical matrix. The positive constant ρ_i is selected not to make the system too sensitive to image noise, $\rho_i = 1.0$ in a trial and error manner.

We show a result to demonstrate how the proposed method works. At $t = 9.0$, the target begins to move leftward 0.2[m] in 8.0[s]. Then, it stays there for 5.0[s], moves rightward 0.2[m] in 8.0[s], stays for 5.0[s], moves rightward 0.2[m] in 8.0[s], stays for 5.0[s], moves leftward 0.2[m] in 8.0[s]. In Figure 7, we can see how the legged robot behaves to track the visual target. The left foreleg of the robot is the floating leg. The leftward and rightward translation of the robot is not so large in the figure, whereas the yawing motion can be observed. The reason seems that the controller cannot discriminate leftward and rightward translation from yawing rotation in the image planes, because the Jacobian matrix is estimated based on the image, not on the 3D information.

Error of the position of the target mark 1 (upper right mark in the image plane) with respect to the desired is shown in Figure 6. Along y-axis (vertically) and x-axis (horizontally) the robot can track the target remarkably.

4 Discussion and future work

Applying visual servoing to a legged robot, motions of the robot are generated by the movement of the visual cues, and therefore the resultant system becomes robust and adaptive against changes of the environment. The proposed method can make the robot sway according to the movement of the visual target in an unknown environment. From the point of view of visual servoing, the proposed method is a new application of visual servoing. This paper has proposed how to apply visual servoing to legged robots.

It is interesting to find that the robot cannot discriminate the leftward/rightward translation from the yawing rotation, when we apply the adaptive visual servoing to it. If the robot is fully calibrated it can discreminate [12]. But, in case that it is not calibrated it only yaws instead of translate motion, because the task given for the robot is not to move its body leftward/rightward, but to make the image features converge to the desired ones.

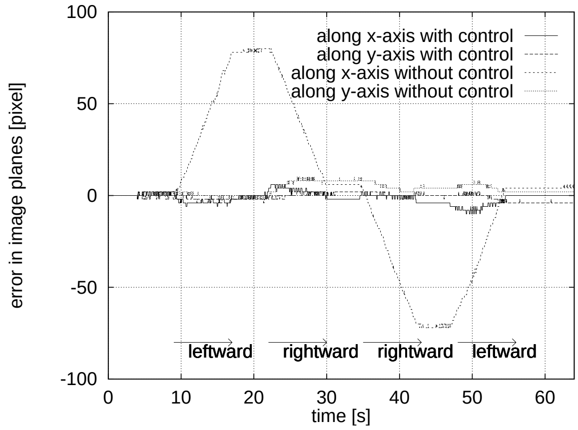


Figure 6: Error of the target mark 1 (upper right mark in the image plane)

When a human mother teaches her infant to walk, she may guide the infant by showing her hands in front of the infant. This fact may show that the visual cues have important roles in achieving a walking task in animals. The proposed method is one of the most powerful servoing strategies based on external sensors to achieve such a task.

Using the proposed scheme, we can realize a swaying behavior with three legs. If we can develop to design the gait of the another lifted leg based on sensors, we can realize sensor-based walking of legged robots. We can build an adaptive/robust walking robot that can walk in dynamic/unknown environments with its sensors. The next step to realize vision-based walking is how to design the gait of the another floating leg.

Acknowledgement

This work is supported in part by the Grant-in-Aid for Scientific Research on Priority Areas (Area number 266, Intelligent Robotics) by Ministry of Education, grant number 97245102.

References

[1] M. H. Raibert et al. Special issue on legged locomotion. *Int. J. of Robotics Research*, 3(2), 1984.

[2] M. Inaba, F. Kanehiro, S. Kagami, and H Inoue. Two-armed bipedal robot that can walk, roll over and stand up. In *Proc. of the 1995 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 297–302, 1995.

[3] D. J. Pack. Perception-based control for a quadruped walking robot. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2994–3001, 1996.

[4] P. I. Corke. Visual control of robot manipulators – a review. In *Visual Servoing*, pages 1–31. World Scientific, 1993.

[5] L. E. Weiss, A. C. Sanderson, and C. P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE J. of Robotics and Automation*, RA-3(5):404–417, 1987.

[6] J. T. Feddema and C. S. G. Lee. Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera. *IEEE Trans. on System, Man, and Cybernetics*, 20(5):1172–1183, 1990.

[7] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura. Manipulator control with image-based visual servo. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2267–2272, 1991.

[8] W. Jang and Z. Bien. Feature-based visual servoing of an eye-in-hand robot with improved tracking performance. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2254–2260, 1991.

[9] N. Maru, H. Kase, et al. Manipulator control by visual servoing with the stereo vision. In *Proc. of the 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1865–1870, 1993.

[10] N. P. Papanikolopoulos and P. K. Khosla. Adaptive robotic visual tracking: Theory and experiments. *IEEE Trans. on Automatic Control*, 38(3):429–445, 1993.

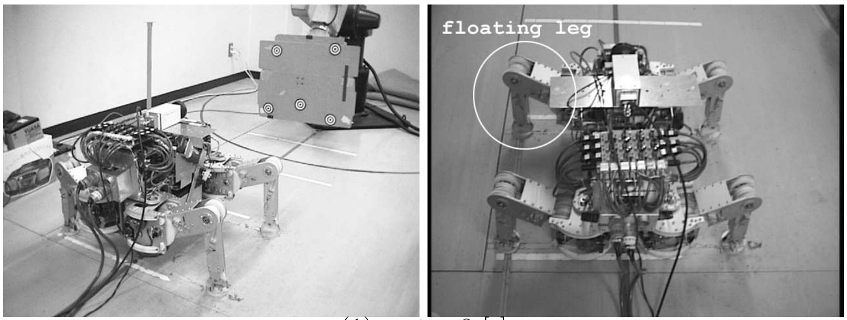
[11] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *Proc. of the 1994 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 186–193, 1994.

[12] K. Hosoda, M. Kamado, and M. Asada. Vision-based servoing control for legged robots. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 3154–3159, 1997.

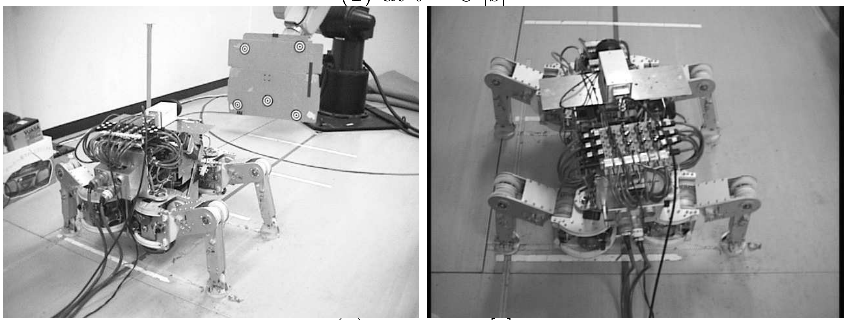
[13] P. Eykhoff. *System Identification*, chapter 7. John Wiley & Sons Ltd., 1974.

[14] K. Arikawa and S. Hirose. Development of quadruped walking robot TITAN–VIII. In *Proc. of the 1996 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 208–214, 1996.

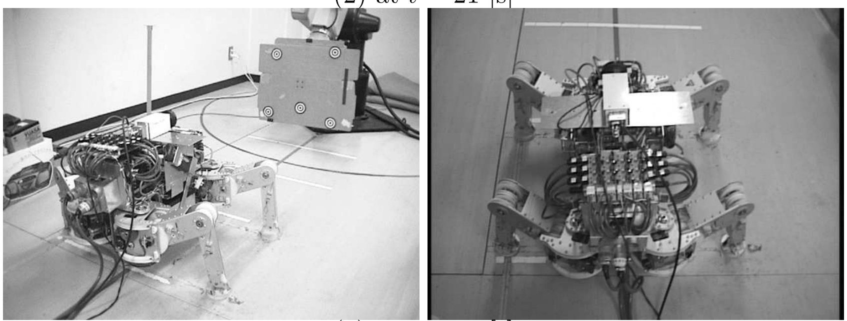
[15] M. Inaba, T. Kamata, and H. Inoue. Rope handling by mobile hand-eye robots. In *Proc. of Int. Conf. on Advanced Robotics*, pages 121–126, 1993.



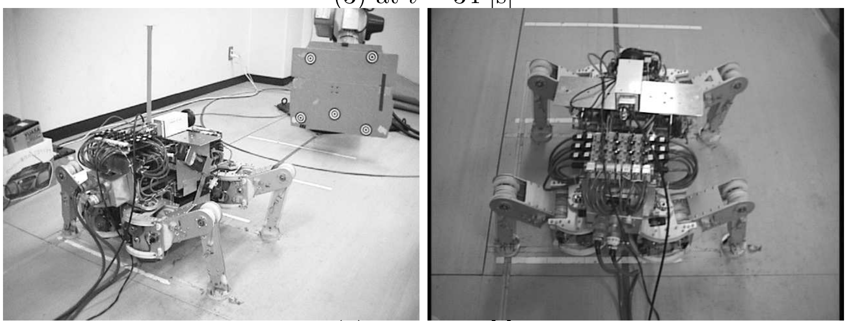
(1) at $t = 8$ [s]



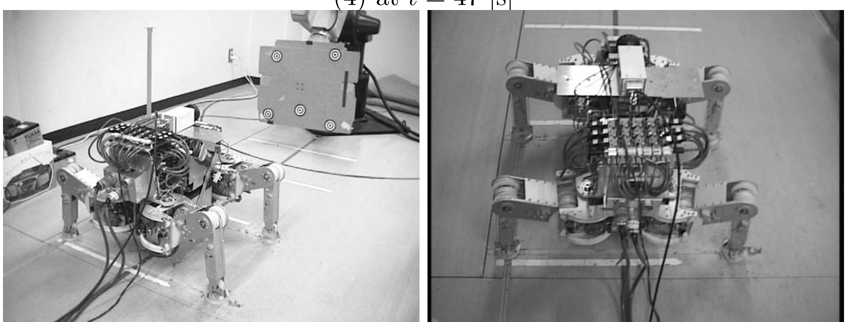
(2) at $t = 21$ [s]



(3) at $t = 34$ [s]



(4) at $t = 47$ [s]



(5) at $t = 60$ [s]

Figure 7: An experimental result: The legged robot sway to track the visual target