View-Based Imitation Learning by Conflict Resolution with Epipolar Geometry

Yuichiro Yoshikawa and Minoru Asada
Dept. of Adaptive Machine Systems
Graduate School of Engineering
Osaka University
Suita, Osaka 565-0871, Japan
e-mail: yoshikawa@er.ams.eng.osaka-u.ac.jp, asada@ams.eng.osaka-u.ac.jp

Abstract

Imitation Learning is not simply one of the most promising ways to accelerate the behavior acquisition for humanoid robots but also one of the most interesting cognitive issues to model how we human beings learn to acquire various kinds of behaviors. However, the existing robotic approaches have focused on the behavior generation assuming the observation of the internal model of the demonstrator, but have not paid any attention how to build such a model from the learner's perception. This paper presents a computational model of view-based imitation learning without any internal model of the demonstrator. Instead, based on opt-geometric constraint (stereo epipolar constraint), the robot learns to imitate the demonstrator's motion by applying adaptive visual servoing that minimizes the residual between the recovered demonstrator's body parts supposed to be viewed by the demonstrator and the learner's ones in the learner's stereo image planes, and then reproducing the recovered demonstrator's trajectories without any reconstruction of 3-D trajectories. Computer simulation and real experiment are shown and discussion is given.

1 Introduction

Imitation Learning is one of the most promising ways to accelerate the behavior acquisition for humanoid robots [1]. Because, machine learning theories seems difficult to directly apply to real robot tasks as they are due to the huge search space caused by multimodal sensor space and many DOFs which also add much more uncertainties than computer simulations.

Another aspect of the imitation learning is that it is also one of the most interesting cognitive issues to model how we human beings learn to acquire various kinds of behaviors by building real robots capable of imitation learning [2]. However, the existing robotic approaches have focused on the behavior generation assuming the observation of the internal model of the demonstrator given the 3-D geometrical parameters (ex. [3, 4, 5]) or the coordinate transformation from the demonstrator's to the learner's (ex. [6, 7]). Hence, they have not paid any attention how to build such an internal model from the learner's perception from a viewpoint of the internal observer.

This paper presents a computational model of view-based imitation learning without any internal model of the demonstrator. Instead, based on an opt-geometric constraint (stereo epipolar constraint [8]), the robot learns to imitate the demonstrator's motion. Unlike the previous work [9], we do not need the initial posture assumption that the demonstrator's initial posture is the same as the learner's to estimate the epipolar constraint. Adaptive visual servoing (hereafter, AVS) [10] resolves this limitation by minimizing the residual between the recovered demonstrator's body parts supposed to be viewed by the demonstrator and the learner's one on the learner's stereo image plane, and then reproducing the recovered demonstrator's trajectories without any reconstruction of 3-D trajectories.

The rest of the paper is organized as follows. First, the principle of the demonstrator's view recovery based on the stereo epipolar constraint is given. Next, conflict resolution with stereo geometry is introduced to estimate the epipolar constraint. Then, the experimental results by computer simulation and real experiment are shown. Finally, discussion is given.

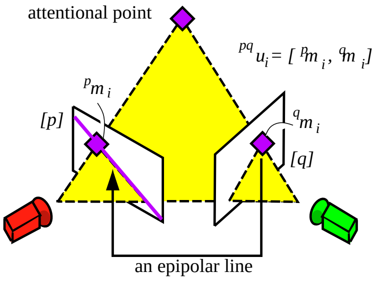


Figure 1: Epipolar geometry

2 Demonstrator's view recovery based on epipolar geometry

2.1 Epipolar geometry

Fig.1 shows an epipolar constraint between a pair of stereo images $[p]$ and $[q]$. Given a point $^p\boldsymbol{m}_i$ in the left (right) image $[p]$ ($[q]$), its corresponding point $^q\boldsymbol{m}_i$ ($^p\boldsymbol{m}_i$) in the right (left) image$[q]$ ($[p]$) is constrained to lie on a line called *epipolar line*. This relationship (constraint) between two cameras is called *epipolar geometry*.

If these two cameras can be approximated by affine camera model [11], epipolar geometry is given by

$$^{pq}\boldsymbol{u}_i^T\, {}^{pq}\boldsymbol{f} + {}^{pq}f_{33} = 0, \tag{1}$$

where $^{pq}\boldsymbol{u}_i = [^pm_i, {}^qm_i]^T$ is a vector which consists of the i-th matched image point between $[p]$ and $[q]$, and $^{pq}\boldsymbol{f} = [^{pq}f_{13}, {}^{pq}f_{23}, {}^{pq}f_{31}, {}^{pq}f_{32}]^T$ consists of nonzero elements in the affine fundamental matrix for the epipolar geometry between $[p]$ and $[q]$.

2.2 Estimation of affine fundamental matrix [11]

In general, a minimum of 4 pairs of matched points are required to uniquely determine the affine fundamental matrix. It can be determined by minimizing the sum of distances of each point $^{pq}\boldsymbol{u}_i$ to the hyperplane (eq.1) in the 4-dimensional space.

By this method, the affine fundamental matrix $^{pq}\boldsymbol{f}$ is determined as an eigenvector associated with the minimal eigenvalue of $^{pq}\boldsymbol{W}$, where

$$^{pq}\boldsymbol{W} = \sum_{i=1}^{N}(^{pq}\boldsymbol{u}_i - \frac{1}{n}\sum_{j=1}^{N}{}^{pq}\boldsymbol{u}_j)(^{pq}\boldsymbol{u}_i - \frac{1}{n}\sum_{j=1}^{N}{}^{pq}\boldsymbol{u}_j)^T, \tag{2}$$

then,

$$^{pq}f_{33} = -{}^{pq}\boldsymbol{u}_0^T\, {}^{pq}\boldsymbol{f}. \tag{3}$$

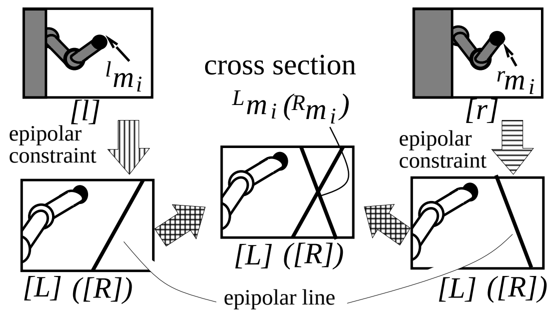


Figure 2: An Overview of the method to find a corresponding point in an added view [L] ([R]) from the two views [l] and [r].

2.3 Finding a corresponding point in a different view

We add one more camera [L] ([R]) observing a point which is also observed in [l] and [r]. The problem is how to find the corresponding points in the view [L] ([R]) with ones in the views [l] and [r].

Based on epipolar constraints, the matched points $^l\boldsymbol{m}_i$ ($^r\boldsymbol{m}_i$) are constrained to lie on the epipolar lines on [L] ([R]). We can find the matched points $^L\boldsymbol{m}_i$ ($^R\boldsymbol{m}_i$) on the cross sections of epipolar lines (see Fig.2). And $^L\boldsymbol{m}_i$ and $^R\boldsymbol{m}_i$ must also satisfy epipolar constraint between [L] and [R].

Thus, there are five epipolar equations ([l,L], [r,L], [l,R], [r,R], and [L,R]) to be satisfied. Developing their formulations algebraically, we obtain

$$\boldsymbol{A}\,{}^{LR}\boldsymbol{u}_i = -\boldsymbol{B}\boldsymbol{c}_i, \tag{4}$$

where

$$\boldsymbol{A} = \begin{bmatrix} ^{lL}f_{31} & ^{lL}f_{32} & 0 & 0 \\ ^{rL}f_{31} & ^{rL}f_{32} & 0 & 0 \\ 0 & 0 & ^{lR}f_{31} & ^{lR}f_{32} \\ 0 & 0 & ^{rR}f_{31} & ^{rR}f_{32} \\ ^{LR}f_{31} & ^{LR}f_{32} & ^{LR}f_{13} & ^{LR}f_{23} \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} ^{lL}f_{13} & ^{lL}f_{23} & ^{lL}f_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & ^{rL}f_{13} & ^{rL}f_{23} & ^{rL}f_{33} \\ ^{lR}f_{13} & ^{lR}f_{23} & ^{lR}f_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & ^{rR}f_{13} & ^{rR}f_{23} & ^{rR}f_{33} \\ 0 & 0 & 0 & 0 & 0 & ^{LR}f_{33} \end{bmatrix}$$

$$\boldsymbol{c}_i = [\,^lx_i \quad ^ly_i \quad 1 \quad ^rx_i \quad ^ry_i \quad 1\,]^T$$

If $\boldsymbol{A}^T\boldsymbol{A}$ is nonsingular, $^{LR}\boldsymbol{u}_i$ is determined from the matched points of the other two stereo images, such as,

$$^{LR}\boldsymbol{u}_i = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{B}\boldsymbol{c}_i. \tag{5}$$

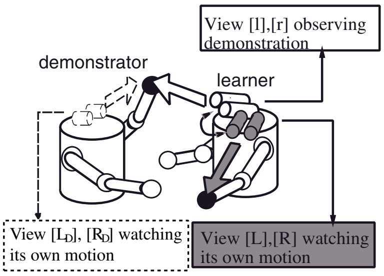


Figure 3: The learner's view at observing ([l], [r]) and imitating ([L], [R]), and the demonstrator's view ([L_D], [R_D]) at watching itself.

2.4 Demonstrator's view recovery

Hereafter, we assume that the learner and the demonstrator have the same body structure, that is, the same link structure and the same camera parameters. Then, the demonstrator's stereo views ([L_D] and [R_D]) watching its self motion can be regarded as the learner's ones ([L] and [R]) watching its self motion if the learner succeeds in exactly imitating the demonstrator's motion. If we can recover the demonstrator's view, it means the learner's view to be realized. Then, the problem is how to recover the views ([L_D] and [R_D]) as the views ([L] and [R]) based on epipolar geometry (see Fig.3).

We can determine all affine fundamental matrices between [l],[r],[L] and [R] by assuming that initial postures of both the demonstrator and the learner are the same and the corresponding points on the body parts are given. Using these affine matrices, we can recover the demonstrator's view which shows the desired trajectories for the learner to realize based on the method described in 2.3. Then, we can apply the adaptive visual servoing method to imitate the demonstrator's motion [9].

3 Resolving Conflict with Epipolar Geometry

In order to recover the demonstrator's view based on the method described in the previous section, we need the true affine fundamental matrices between the views of observing the demonstration ([l] and [r]) and observing the learner itself ([L] and [R]). However, we cannot estimate the true fundamental matrix when the learner's posture is different from the demonstrator's one.

If the affine fundamental matrices are correctly estimated, the recovered points of the demonstrator's body parts are coincident with the corresponding learner's body parts. Else, the corresponding points are shifted from each other on the image plane. That is, the estimated affine fundamental matrix has a conflict with true epipolar geometry. In this section, we derive the evaluation function of this conflict, analyze its behavior, and resolve the conflict by minimizing the evaluation function.

3.1 Evaluation function

Let the evaluation function of the error caused by the estimated affine fundamental matrix be

$$E = \frac{1}{2N}\sum_{i=1}^{N}({}^{L}\boldsymbol{e}_i^{T}\,{}^{L}\boldsymbol{e}_i + {}^{R}\boldsymbol{e}_i^{T}\,{}^{R}\boldsymbol{e}_i), \tag{6}$$

where ${}^{j}\boldsymbol{e}_i (j = L, R)$ denotes a vector from a recovered point(${}^{j}\hat{\boldsymbol{m}}_i$) to its corresponding one(${}^{j}\boldsymbol{m}_i$) (see Fig.4). Note that this evaluation function consists of only the visual information of both observing the demonstration and monitoring the learner itself.

Since the projected points ${}^{LR}\boldsymbol{u}_i$ of the learner's body parts are the functions in terms of the learner's posture ($\boldsymbol{\theta} \in \Re^m$), they can be given as, ${}^{LR}\boldsymbol{u}_i = {}^{LR}\boldsymbol{u}_i(\boldsymbol{\theta})$. Supposing that the learner's posture is given by $\boldsymbol{\theta} = \boldsymbol{\theta}_D + \boldsymbol{\delta\theta}$, where $\boldsymbol{\theta}_D$ is the posture of the demonstrator, the projected point ${}^{LR}\boldsymbol{u}_i$ is given by the following equation including the perturbation ${}^{LR}\boldsymbol{\delta u}_i$,

$${}^{LR}\boldsymbol{u}_i = {}^{LR}\boldsymbol{u}_{Di} + {}^{LR}\boldsymbol{\delta u}_i, \tag{7}$$

where ${}^{LR}\boldsymbol{u}_{Di} = {}^{LR}\boldsymbol{u}_i(\boldsymbol{\theta}_D) = [{}^{L}\boldsymbol{m}_{Di}, {}^{R}\boldsymbol{m}_{Di}]^T$.

As long as ${}^{LR}\boldsymbol{\delta u}_i$ is small, the relationship between $\boldsymbol{\theta}$ and ${}^{LR}\boldsymbol{u}_i$ can be approximated by

$${}^{LR}\boldsymbol{\delta u}_i = \boldsymbol{J}_{\boldsymbol{u}i}\boldsymbol{\delta\theta}, \tag{8}$$

where $\boldsymbol{J}_{\boldsymbol{u}i} = \partial^{LR}\boldsymbol{u}_i/\partial\boldsymbol{\theta}^T \in \Re^{4\times m}$ is a Jacobian matrix of time-derivatives of the projected point vector with respect to the joint angle.

Based on the perturbation theory for eigenvalue, the estimated fundamental matrix ${}^{pq}\boldsymbol{f}$ is given by the following equation including a function of perturbation ${}^{pq}\boldsymbol{\delta f}$ in terms of ${}^{LR}\boldsymbol{\delta u}_i (i = 1, \cdots, n)$,

$${}^{pq}\boldsymbol{f} = {}^{pq}\boldsymbol{f}_{true} + {}^{pq}\boldsymbol{\delta f}({}^{LR}\boldsymbol{\delta u}_1, \cdots, {}^{LR}\boldsymbol{\delta u}_N), \tag{9}$$

where ${}^{pq}\boldsymbol{f}_{true}$ is the true fundamental matrix. From eq.(8), the second term of RHS in eq.(9) is given by a function in terms of $\boldsymbol{\delta\theta}$, such as ${}^{pq}\boldsymbol{g}(\boldsymbol{\delta\theta})$. Therefore ${}^{pq}\boldsymbol{f}$ is approximated by including it,

$${}^{pq}\boldsymbol{f} = {}^{pq}\boldsymbol{f}_{true} + {}^{pq}\boldsymbol{g}(\boldsymbol{\delta\theta}) \tag{10}$$

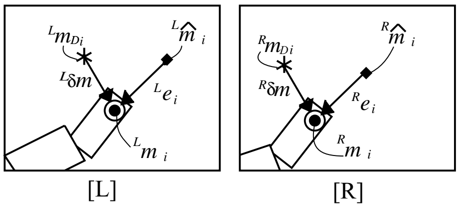


Figure 4: evaluation function

Substituting eq.(10) to eq.(5), and developing the above formulation algebraically in focusing on the dominant term, we obtain

$$E = \boldsymbol{\delta\theta}^T \boldsymbol{Q} \boldsymbol{\delta\theta}, \tag{11}$$

where $\boldsymbol{Q} \in \Re^{m \times m}$ is a positive-semidefinite matrix.

The positive-semidefinite matrix $\boldsymbol{Q}$ can be regarded as a positive-definite matrix because the evaluation function usually becomes zero only when the learner's posture corresponds to the demonstrator's one. Thus, the proposed evaluation function is expected to be a convex function and have a local minimum at which the learner's posture corresponds to the demonstrator's one.

4 Minimizing the evaluation function

Since it can be seen as a convex function in terms of joint angles, it is expected to be able to minimize the proposed evaluation function by a gradient method. In order to use the gradient method, we need the gradient vector, which is unknown in our case. Then, adaptive visual servoing method (AVS) [10] is applied to estimate the gradient vector of the unknown system.

The relationship between the learner's joint angle $\boldsymbol{\theta}$ and the evaluation function E is given by $E = E(\boldsymbol{\theta})$ from eq.(11). Differentiating it, we obtain a velocity relation,

$$\dot{E} = \boldsymbol{J}_E \dot{\boldsymbol{\theta}}, \tag{12}$$

where $\boldsymbol{J_E}$ is a Jacobian matrix of time-derivatives of the evaluation function with respect to joint angle velocity. Using AVS, the Jacobian matrix $\hat{\boldsymbol{J}}_{\boldsymbol{E}}$ can be estimated based on the recursive weighted least square method.

In order to minimize E, we can determine the control input vector $\boldsymbol{u}_{input} \in \Re^m$ in the following equation,

$$\boldsymbol{u}_{input} = -\boldsymbol{K} \hat{\boldsymbol{J}}_E E \tag{13}$$

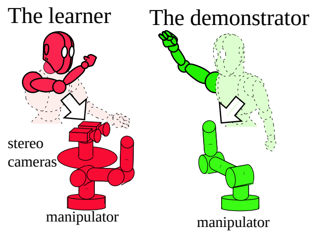


Figure 5: Assuming the learner and the demonstrator as two identical manipulator.

where $\boldsymbol{K} \in \Re^{m \times m}$ is a positive-definite gain matrix.

5 Experiments

To show the validity of the proposed method, some experimental results are given in this section. Two identical manipulators are assumed as bodies of the learner and the demonstrator, respectively. A pair of stereo cameras are assumed as learner's view points (see Fig.5).

5.1 Behaviors of the evaluation function by computer simulation

To examine the behavior of the evaluation function, the computer simulation is performed. When the demonstrator's joint angle are $(\theta_2, \theta_4) = (45°, -90°)$, the learner observes the demonstrator's posture and know the position of the ten feature points on the demonstrator's body in its stereo views. Then the learner changes the gaze direction and watchs its own matched feature points. Fig.7 shows the calculated evaluation function when moving its joint angles, $\theta_2 = 0° \sim 90°, and \theta_4 = 0° \sim -180°$.

We can regard that the calculated evaluation function is almost a convex function which has a local minimum at $(\theta_2, \theta_4) = (41°, -90°)$. The local minimum is very close to the point which the learner's posture corresponds to demonstrator's one.

5.2 Behaviors of the evaluation function in real experiment

To examine the behavior of the evaluation function, the result in the real experiment is shown in this section. In this experiment, the right and left manipulators are assumed as the learner's body and the demonstrator's one, respectively (see Fig.6). And the learner's view is a pair of center stereo cameras of

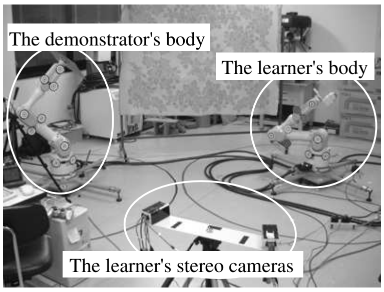
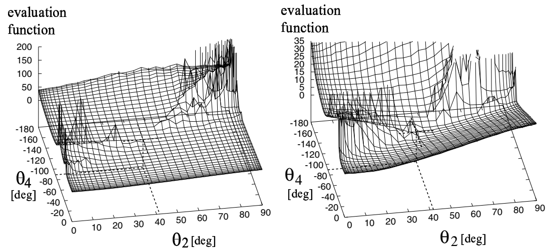


Figure 6: Overview of the experimental setup



(a) Global view (b) Close up

Figure 7: Evaluation function in the computer simulation; global view (left) and close up (right) of the evaluation function.

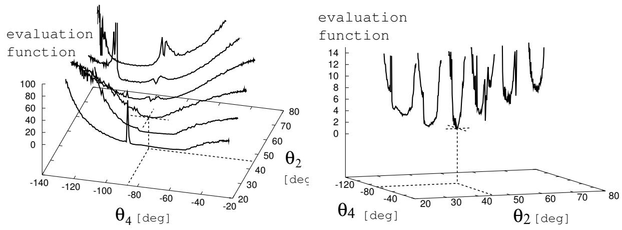
which image size and the baseline are 640 × 240 and 60cm, respectively. Fig.8 shows the calculated evaluation function in the same manner as in the computer simulation. In the real experiment, the the learner moves its joint angles, $\theta_2 = 25°, 35°, 45°, 55°, 65°$ and $75°$, and $\theta_4 = -25° \sim -135°$.

Similar to the computer simulation, the calculated evaluation function can be regarded as a convex one which has a local minimum at $(\theta_2, \theta_4) = (45°, -86°)$.

From the results of the computer simulation and the real experiment, we may conclude that the proposed evaluation function is a convex one and has a local minimum at which the learner's posture corresponds to the demonstrator's one.

5.3 Posture imitation by resolving conflict

To show the validity of the method to resolve the conflict of the estimated affine fundamental matrix with epipolar geometry, the result of the real experiment is shown in this section. Adaptive visual ser-



(a) Global view (b) Close up

Figure 8: Evaluation function in the real experiment; global view (left) and close up (right) of the evaluation function.

voiung (hereafter, AVS) is applied as one of a gradient method to estimate the true epipolar geometry and then to realize the imitation. The control input is determined by eq.(13). The initial values of the Jacobian matrix to be estimated are arbitrarily chosen as

$$\hat{\boldsymbol{J}}_{eval} = [\,-1.0 \quad 1.0\,], \tag{14}$$

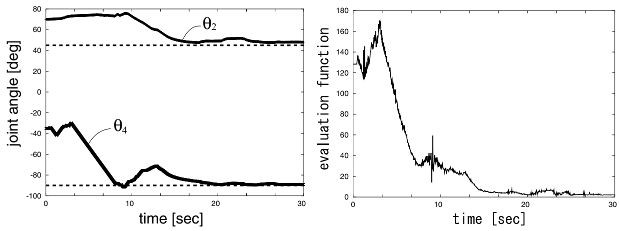
and forgetting factor ρ and the positive-definite gain matrix $\boldsymbol{K}$ are 0.95, and $diag(0.5 \times 10^{-2}, 0.5 \times 10^{-2})$, respectively.

After observing ten feature points on the demonstrator's body, the learner minimizes the evaluation function by using AVS. Fig.9 (a) shows the trajectories of two joint angles θ_2 and θ_4, respectively, during the control. They are evidently converged to the demonstrator's posture (broken lines). Fig.9 (b) shows that the evaluation function also converges to zero by the method.

5.4 Trajectory imitation based on estimated epipolar geometry

To show the correctness of estimated affine fundamental matrix through minimization process in the real experiment, the learner imitates the demonstration using it. The learner stores the trajectory of the demonstrator's endeffector. Then, it recovers the observed trajectory on its view of monitoring the self motion using the method described in 2.4, and the learner imitates by reproducing the recovered trajectory based on feedback control using AVS.

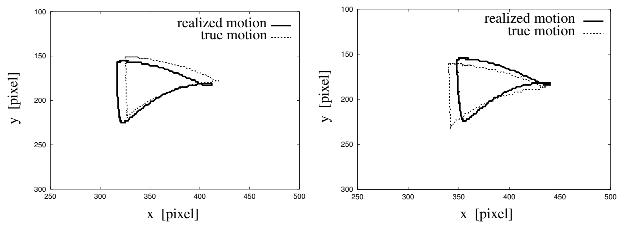
Fig.10 shows the images of the learner's view during the imitation. Each figure contains two trajectories; one is learner's and the other is the true one. Since two



(a) The changes of the learner's joint angles

(b) The change of the evaluation function

Figure 9: Experiment: [left] the changes of the learner joint angles (solid line) and the desired one (broken line); [right] the change of evaluation function.



(a) $[L]$

(b) $[R]$

Figure 10: The trajectories of imitated motion (solid lines) and those (broken lines) of desired in the learner's view obtained by the same control input as the demonstrator.

trajectories are almost overlapped with each other, it can be regarded that the estimation of affine epipolar geometry is sufficient to imitate the demonstrator's motion.

6 Discussion and conclusion

In this paper, we proposed a computational model of view-based imitation learning without assuming that the initial postures of the demonstrator and the learner are the same. The evaluation function is defined to estimate the difference between the true epipolar geometry and the estimated one caused by the different initial postures. Adaptive visual servoing takes roles of minimization of this error and of realizing the imitation as well. The experimental results showed the validity of the method.

The current motion representation is the trajectory on the image plane, and no more abstraction in this level. To cope with environmental changes and/or task variations, more abstracted motion representation capable to generate various kind of behaviors combining the primitive motions should be developed through the imitation process based on the leraner's perception. This is our future work.

References

[1] S. Schaal. Is imitation learning the route to humanoid robots? *Trends in Cognitive Science*, 1999.

[2] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. In *Proc. of the 1st IEEE-RSA International Conference on Humanoid Robots*.

[3] K. Ikeuchi and T. Suehiro. Toward an assembly plan from observation. *IEEE Trans. on R & A*, pp. 368–385, 1994.

[4] T. Inamura, I. Toshima, H. Ezaki, and Y. Nakamura. Generation of whole body motion using mimesis loop and primitive symbols(in japanese). In *Proc. of the 18th Annual Conference of the Robotics Society of Japan*, pp. 801–802, 2000.

[5] A. Billard and Maja Mataric. Learning human arm movement by imitation: Evaluation of a biologically-ispired connectionist architecture. In *Proc. of the 1st IEEE-RAS International Conference on Humanoid Robots*, 2000.

[6] Y. Kuniyoshi, M. Inaba, and H. Inoue. Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. on R&A*, Vol. 10, No. 6, pp. 799–821, 1994.

[7] H. Miyamoto, S. Schaal, F. Gandolfo, H. Gomi, Y. Koike, R. Osu, E. Nakano, Y. Wada, and M. Kawato. A kendama learning robot based on bi-directional theory. *Neural Networks*, Vol. 9, pp. 1281–1302, 1996.

[8] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene of from two projections. *Nature*, Vol. 293, pp. 133–135, 1981.

[9] M. Asada, Y. Yoshikawa, and K. Hosoda. Learning by observation without three-dimensional reconstruction. In *Intelligent Autonomous Systems*, pp. 555–560, 2000.

[10] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *Proc. of IROS'94*, pp. 186–193, 1994.

[11] G. Xu and Z.Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach.* Kluwer Academic Publisher, 1996.