Proceedings of the 13th
KACC, October 1998

Vision Module to Estimate Robot's Self Moving Variation

Yoshiyasu Goto, Hiroshi Mizoguchi, Ken-ichi Hidai,
Takaomi Shigehara, and Taketoshi Mishima

Department of Information and Computer Sciences,
Faculty of Engineering, Saitama University
255 Shimo Okubo, Urawa 338-8570, Japan
Tel.: +81-48-858-3723
Fax.: +81-48-858-3750
E-mail: goto@me.ics.saitama-u.ac.jp

Abstract

In this paper the authors propose a model about interaction of inner modules of autonomous robot which is possible to learn walking action without external and explicit supervisor signal. A main feature of the model is that completed and fixed module for estimating robot's motion parameter by utilizing binocular parallax can be a supervisor for the module to learn the walking action.

1. Introduction

Humans and animals flexibly adapt to changeful environment by utilizing experience or acquiring new skill or action. And skills acquired previously is improved on efficiency. According to connectionism, these useful properties of a biological system are realized with plasticity of each neurons and interaction of modules composed of neurons in brain. Our ultimate goal is to reproduce these property with real world robots, for more detailed understanding of learning and adaptation system of a biological system.

So that, modeling such systems and making sure of its validity (by implementation) is firstly required. Difficulty about modeling life's learning system is derived from that it does not need and accept external and explicit supervising signal. The authors endorse assertion which says that such learning is realized with self organization and interaction of modules composed of neurons. We now attend to latter, interaction of modules. Example of interaction we are considering is one between vision and walking module (too much general module division, we know, this is only for explanation of framework). Assuming that vision module has been almost converged and fixed, and walking module has not yet, we think walking module can use vision module as supervisor instead of external one. There continues a little notes about our detailed assumption. In this case, main role of vision module is to estimate self moving direction and quantity as output, from image series (temporally stored via camera) as input. Walking module translates a directive (previously planed by higher modules) about self movement which was given as input to angles or torque series of legs as output. We accent that walking module can use difference between previously planed movement directive and self moving variation guessed by vision module as error signal for training.

As previous stage before detailed argument and implementation of such relation of two modules, in this paper the authors aim construction of reliable and ideal visual module. In this stage, regardless of problem about how such module be acquired, we treat only its converged and fixed form, i.e. what calculation completed vision module do to estimate self moving variation.

In the following, first we describe our model of inter development between vision and motion modules. Next a calculation method in the model is described. The method visually estimates motion parameters. In the method robot's moving variation is estimated utilizing binocular stereo parallax. The binocular stereo is supposed as a subsystem in the vision module. Then we summarize this work and mention some concluding remarks. Future works are also mentioned.

2. Model of Inter-Development between Vision and Motion Modules

Operation from Vision Module to Motion Module

The proposed model explains a process in which the vision module trains the motion module to obtain walking capability by visually estimating robot's motion variation.

Biology and physiology have accumulated various knowledge and observations about relationship of visual module to eye movements and head motions in biological systems, such as human beings and animals.

This work intends to apply these knowledge and hypothesis to relationship between artificial visual module and mobile robot's walking function, instead of the eye or head motion.

Fig. 1 shows a block diagram of the proposed model. In the figure, the upper left block denotes a trajectory generation module. It receives motion commands, such as "go forward" or "rotate right", from higher level module and generates trajectories of legs to realize the commanded motion. This trajectory generation module treats difference between the motion command from the higher level module and self motion estimated by vision module as input error signal. That is the vision module plays a role of supervisor in this learning model.

As for the supervised learning of motion, discussion in section 5.6 of a book by Dr. Kawato of ATR[1] is quite useful and applicable. It treats general motion and discusses in detail

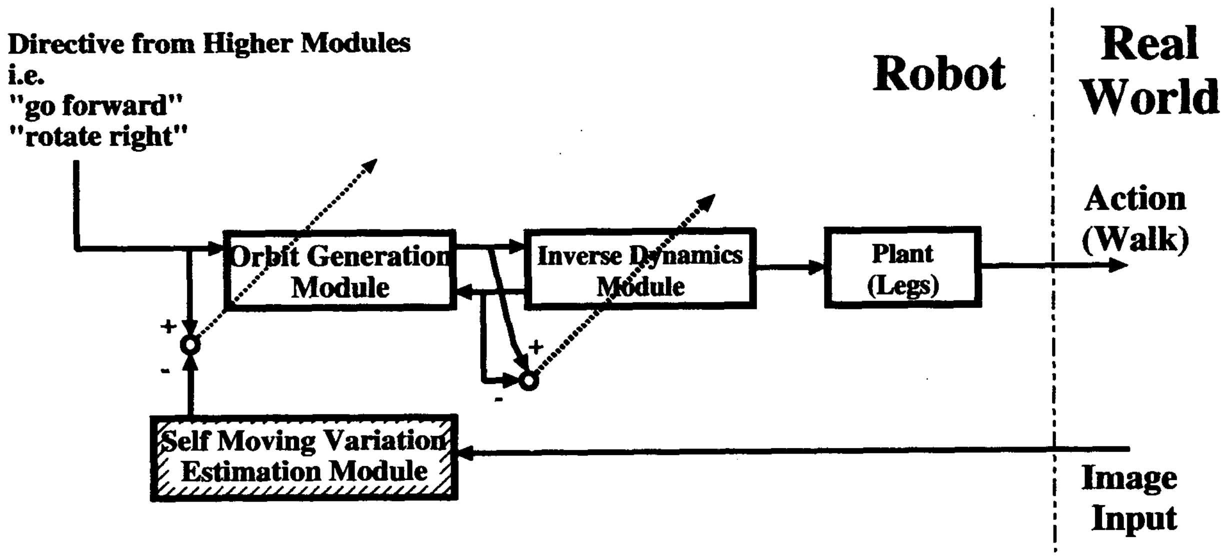

Fig. 1 Block Diagram of Proposed Inter-Development Model

transformation from difference between intended motion and estimated one in real world to error of the motion plan. As for self acquiring of gait control in legged robot, learning with classifier system is studied[3]. However it has not been revealed what is supervisor in the learning. The proposed model in this paper is intended to augment this problem utilizing Kawato's theory on learning of motion.

Feedback from Motion Module to Vision Module

Once the motion module obtains walking capability through the learning by vision module, then it can be expected that development of the vision module arises in turn through the motion. If the robot is possible to move as it plans, utilizing its motion parameters, it can actively obtain information of its environment. In this stage, binding conditions to solve *shape from X* utilized by early active vision research can be obtained. It is impossible to obtain the reliable binding condition, while the motion module cannot control body motion as planned.

3. Motion Parameter Estimation by Vision Module Utilizing Binocular Stereo

In this section, we describe a calculation method to estimate motion parameter by the vision module. The method is intended to estimate robustly even in dynamic environment.

To simplify the discussion, coordinate system of the robot is assumed to be identical to that of onboard camera.

It is also assumed that at every sampling time the robot can calculate a three dimensional coordinate $\boldsymbol{P}=[X,\ Y,\ Z]'$ by searching matched pair of coordinates $\boldsymbol{p}_L$ and $\boldsymbol{p}_R$ that are projected points of the $\boldsymbol{P}$ onto left and right images correspondingly.
Both rotation and translation are happened from the time t_i to t_{i+1}.
As shown in fig. 2, $\boldsymbol{\theta}$ denotes the rotational parameters around X, Y, and Z axes, and $\boldsymbol{T}$ denotes the translational parameters along these axes.

$$\boldsymbol{\theta}=[\theta_x,\theta_y,\theta_z]' \quad (1)$$

$$\boldsymbol{T}=[T_x,\ T_y,\ T_z]' \quad (2)$$

In this situation, if a point $\boldsymbol{P}$ observed at the time t_i moves to $\boldsymbol{P'}$ at t_{i+1}, the transformation can be represented as follows.

$$R_z(-\theta_z)R_y(-\theta_y)R_x(-\theta_x)\boldsymbol{P}-\boldsymbol{T}=\boldsymbol{P'} \quad (3)$$

Where, R_z, R_y, and R_x denote matrix of rotation around X, Y, and Z axes respectively.
If the rotational angles, θ_i, are sufficiently small, the equation (3) can be approximated by equation (4) as follows.

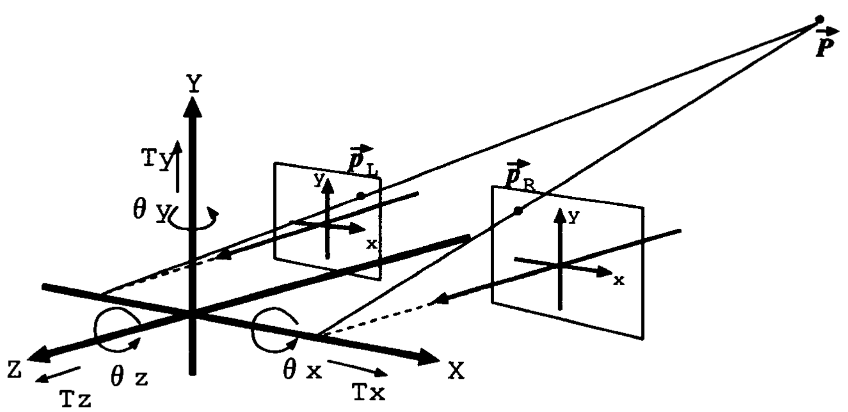

Fig. 2 Motion Parameters and Coordinate System of Onboard Camera

$$\begin{bmatrix} 0 & Z & -Y \\ -Z & 0 & X \\ Y & -X & 0 \end{bmatrix} \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} X - X' \\ Y - Y' \\ Z - Z' \end{bmatrix} \quad (4)$$

Thus it is ideally possible to calculate $\boldsymbol{T}$ and $\boldsymbol{\theta}$ utilizing observed coordinates of points corresponding to $\boldsymbol{P}$ and $\boldsymbol{P'}$. Sufficient number of the pair is theoretically three. However, in real situation, it is not possible to obtain precise motion parameters by simply observing arbitrary three pairs of points, because there are the following difficulties.

* Stereo measurement error in coordinates of $\boldsymbol{P}$ and $\boldsymbol{P'}$

* Disappearance of points by occlusion

* Disappearance of points from visual field according the robot's motion

To cope with these difficulties, this paper introduces the following four *similarity* in order to calculate and obtain a precise and reliable solution to the equation (4).

* Similarity between right and left image frames at the time, t_i.

* Similarity between right and left image frames at the time, t_{i+1}

* Similarity between a point in left image at the time, t_i, and one at t_{i+1}.

* Similarity between a point in right image at the time, t_i, and one at t_{i+1}

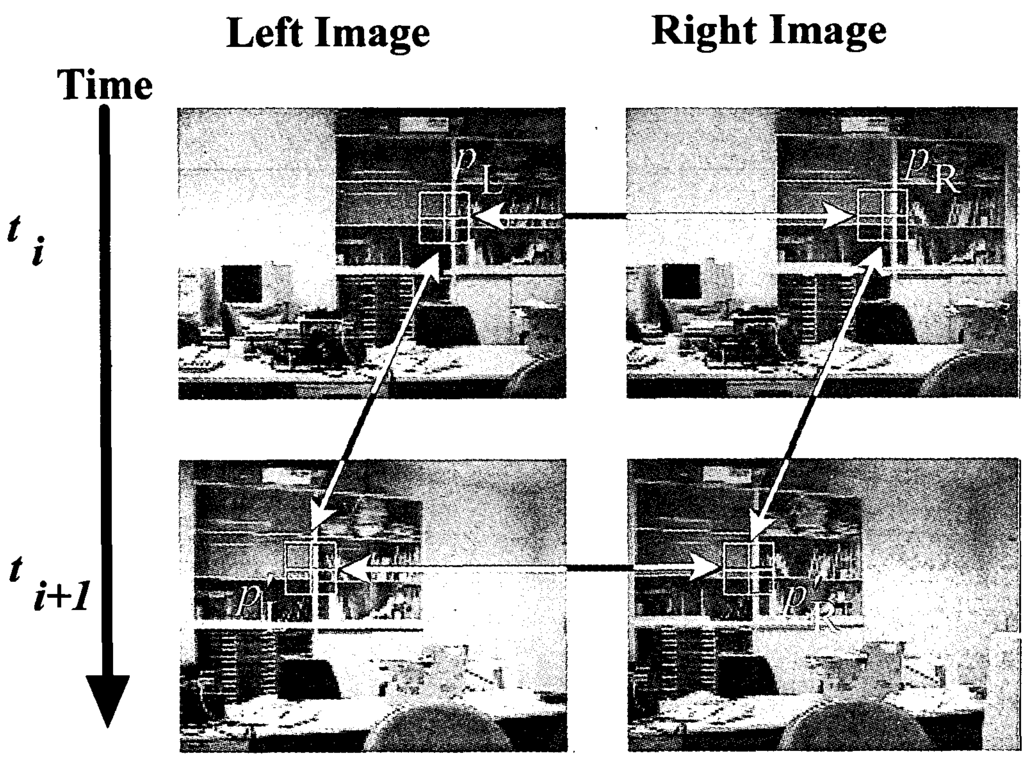

Fig. 3 Spatio-temporal Matching Scheme

Based upon these similarity, a reliable and precise solution of the equation (4) can be calculated, only if the following two conditions are satisfied at the same time. One condition is that a multiplied product of all these similarity values is higher than a threshold. The other is that a variance of pixel values within a block area around the observed point in the image is higher than another threshold, i.e. brightness within the block is not uniform. Because such points that satisfy these conditions are regarded as well matched in both left and right images, and inter frames. In other words, those points are well matched both spatially and temporally. Fig. 3 illustrates this spatio-temporal matching scheme.

To prove the reliability and precision of the above described method to estimate motion parameters, , an experiment and evaluation of the method should be done. To conduct such experiment accurately and quantitatively, we must utilize TV cameras whose various parameters are clearly opened and motion stage whose dynamic characteristics are disclosed. According to preliminary experiments utilizing computer controllable moving camera head, we have confirmed possibility to detect motion of rotation and translation.

4. Concluding Remarks

This paper has proposed a model on inter-development between vision and motion modules. A main feature of the model is the vision module to estimate robot's motion parameters can be a supervisor for the motion module to learn walking action. A calculation method to estimate the motion parameters by the vision module is also described. Preliminary experiment supports feasibility and possibility of the method.

Future works include computer simulation of the proposed model, application of the model to a research purpose quadruped mobile robot, evaluation of the model and the calculation method and improvement of them. Currently binocular stereo utilizes a simple block matching method. Thus the stereo matching tends to fail if image size changes or image rotates. When the camera moves forward or back along its optical axis, the image size is enlarged or shrunk. When the camera rotates around the optical axis, the rotation of the image happens. It is also a future work to improve the matching method so that it is robust against the image size change or the rotation of the image utilizing Log-Polar conversion.

References

[1] Mitsuo Kawato, *Computational Theory of Brain*, Sangyo Tosho, Tokyo, Japan, 1996.

[2] Yoshinori Kuno, "Active Vision -Retrospective and Perspective- ", *Journal of Japanese Society of Artificial Intelligence*, Vol.10, No.4, pp.493-499, 1995.

[3] Shuichi Sato, et al., "Study on Acquiring Gait of An Autonomous Quadruped Robot", *Proceedings of the 1998 JSME Conference on Robotics Mechatronics*, No.98-4, 1B1-8, 1998

[4] B. Irie and S. Miyake, "Capabilities of Thee-layered Perceptrons", *Proceedings of International Conference on Neural Networks*, pp.641-648, 1988.

[5] C. M. Brown and D. Terzopoulos, *Real-time Computer Vision*, Cambridge University Press, Cambridge, UK, 1994.

[6] M. Gini, D. F. Hougen and J. Slagle, "Rapid Unsupervised Connectionist Learning for Backing a Robot with Two Trailers", *Proceedings of ICRA'97*, pp.2950-2957, 1997.

[7] K. Hosoda, M. Kamado and M. Asada, "Vision-Based Servoing Control for Legged Robots", *Proceedings of ICRA'97*, pp.3154-3159, 1997.

[8] J. Batista, et al., "Real-Time Visual Behaviors with a Binocular Active Vision System", *Proceedings of ICRA'97*, pp.3391-3396, 1997.

[9] N.Mitsumoto, et al., "Control of the Distributed Autonomous Robotic System Based on the Biologically Inspired Immunological Architecture", *Proceedings of ICRA'97*, pp.3551-3556, 1997.

[10]James J. Clark, "Spatial Attention and Saccadic Motion", *Proceedings of ICRA'98*, pp.3247-3252, 1998.

[11]T. Nakamura, "Development of Self-Learning Vision-Based Mobile Robots for Acquiring Soccer Robots Behavior", *Proceedings of ICRA'98, pp.2592-2598*, 1998.

[12]S. Yamada and M. M. Murota, "Unsupervised Learning to Recognize Environments from Behavior Sequencess in a Mobile Robot, *Proceedings of ICRA'98*, pp.1871-1878, 1998.

[13]F. Mura and I. Shimoyama, "Visual Guidance of a Small Mobile Robot using Active, Biologically-Inspired, Eye Movements", *Proceedings of ICRA'98*, pp.1859-1864, 1998.

[14]M. Buehler, R. Battaglia, A. Cocosco, G. Hawker, J. Sarkis, and K. Yamazaki, "SCOUT: A Simple Quadruped that Walks, Climbs and Runs, *Proceedings of ICRA'98*, pp.1707-1712, 1998.

[15]E.Uchibe, M. Asada and K. Hosoda, "Environmental Complexity Control for Vision-Based Learning Mobile Robot", *Proceedings of ICRA'98*, pp.1865-1870, 1998.

[16]S. Kagami, M. Kabasawa, K. Okada, T. Matsuki, Y. Matsumoto, A. Konno, M. Inaba and H. Inoue, "Design and Development of a Legged Robot Research Platform JROB-1", *Proceedings of ICRA'98*, pp.146-151, 1998.