Dynamic Learning-Based Jacobian Estimation for Pan-Tilt-Verge Head Control

SILA SORN SUJITRA and ARTHIT SRIKAEW
Robotics and Automation Research Unit for Real-World Applications (RARA)
School of Electrical Engineering
Suranaree University of Technology
111 University Ave., Suranaree Subdistrict, Muang District, Nakhon Ratchasima 30000
THAILAND

Abstract: - Pan-tilt-verge (PTV) vision system is one of the most widely used in active vision. The main advantage of using such system is its 4 DOF which allows tracking of moving objects efficiently. Besides a physical design of the head, an overall tracking performance of the system depends on its controller. This paper presents a development of controlling PTV head to achieve one of human-like eye movement behaviors, i.e. saccade. The PTV head is driven directly from a controller using visual feedback. The dynamic Jacobian estimation is obtained by using a self-organizing map network with unsupervised learning scheme. The estimated Jacobian is used in PTV head controller and results are desirable for both performance and speed of learning. Moreover, the system can eventually perform tracking without a priori knowledge of the head structure, e.g. mathematical model of the head and hardware calibration. Thus the system can conveniently be implemented.

Key-Words: - Jacobian estimation, self-organizing map, active vision system, saccade, pan-tilt-verge, tracking moving object.

1 Introduction
Pan-tilt-verge (PTV) vision system is one of the most widely used in active vision. The main advantage of using such system is its 4 DOF which allows tracking of moving objects efficiently. Besides a physical design of the head, an overall performance of the tracking system depends on its controller. There are various researches that focus on designing a PTV controller to achieve the human-like eye movements. Saccade is one of such movements. It allows the PTV to turn to the target immediately such that the object in the image plane is brought to the center of the image at once. Sharma and Shrinivasa [4] applied a self-organizing invertible map (SOIM) for a saccadic controller without any calibration. Two cameras were used and the network was off-line training. Peng et al. [1] built a PTV head called ISAC. They have deployed an ANN to achieve nonlinear controller with saccade and smooth-pursuit movement behaviors. Their system needs no calibration and the training process can be both off-line and on-line. Lee and Geliana [2] created bilateral model to simulate the saccade and smooth-pursuit behaviors. The resulting oculomotor control system model can be used to control a robot head with high speed object tracking and very small error.

There are also many researches about robot head controller. Most of them requires calibration or a model of the robot head [3][6][7][8][9]. However, calibration processes and model calculation could be significantly complicated. Moreover, the assembly of robot head must be efficiently accurate. Otherwise, the model of the robot head might inherit some degree of error which can cause undesired effects in the controller.

This work has applied a self-organizing map (SOM)[5] for controlling a PTV head to track moving object with saccade movement. SOM allows the system to learn using competitive learning scheme which is unsupervised learning. Each weight of the network stores motor command corresponding to the distance of the target from the center of the image plane. SOM dynamically adjusts itself to achieve better motor command values. This learning scheme does not require any calibration or model calculation in order to obtain better accuracy for motor commands. The adjustments of weights use image Jacobian which provides a simple implementation and a fast calculation. The following section describes more details about image Jacobian-based controller for the PVT head.
2 Image Jacobian-Based PTV Head Controller

Generally, when PTV joints change, robot Jacobian [11] relates those changes. In the mean while, PTV joint changes also yield image parameters to change. This change is called image Jacobian [10]. Let \( \hat{\theta} \) be a PTV joint position vector and \( \hat{r} \) be a vector that indicates a movement of the cameras. Let \( \hat{f} \) be image-parameters vector. The relationship between joints velocity \( \dot{\theta} \) and camera velocity \( \dot{r} \) can be described as follows:

\[
\dot{r} = J \cdot \dot{\theta} \tag{1}
\]

In additions, the relationship between \( \dot{f} \) and \( \dot{r} \) is

\[
\dot{f} = J_r \cdot \dot{r} \tag{2}
\]

Consequently, the relationship between \( \dot{f} \) and \( \dot{\theta} \) can be related as image Jacobian \( J_q \) which is

\[
\dot{f} = J_q \cdot \dot{\theta} \tag{3}
\]

where \( J_q = J_r \cdot J \).

2.1 Dynamic Image Jacobian Estimation

Image Jacobian can be dynamically estimated from changes of PTV joints and changes of image parameters using linear relationship in (3). Thus, the PTV joint error can be computed by

\[
\Delta \theta = J_q^{-1} \Delta f \tag{4}
\]

In order to calculate image Jacobian \( J_q \), PTV joint parameters and image parameters must be known. The estimated image Jacobian can then be computed by \( J_q = \Delta f / \Delta \theta \).

2.2 Learning-based Image Jacobian Estimation

Consider PTV joint parameters \( \Delta \theta \) as the amount of joint motors movement and image parameters \( \Delta f \) as the change of the target image when joint motors change by \( \Delta \theta \). The next motor command \( \Delta \theta^{\text{new}} \) can be computed by

\[
\Delta \theta^{\text{new}} = \frac{\Delta f^{\text{new}}}{J_q} \tag{5}
\]

where \( \Delta f^{\text{new}} \) is the image parameter of error distance in pixels between the target image and center of the image plane.

Equation (5) shows that the accuracy of the motor command \( \Delta \theta^{\text{new}} \) to bring the target to the center of the image at once depends on the image Jacobian \( J_q \). In order to reduce such error on image Jacobian calculation, a self-organizing map network (SOM) is applied to estimate the image Jacobian. Using a competitive learning scheme, the \( i^{th} \) weight of SOM is updated by

\[
w_i^{\text{new}} = w_i^{\text{old}} + \eta \Delta w \tag{6}
\]

where \( \eta \) is learning rate and \( \Delta w \) is the amount of the next motor command. Each weight of the network is corresponding to the amount of motor command to bring the target toward the center of the image. The new learning rule can then be defined as follows:

\[
w_i^{\text{new}} = w_i^{\text{old}} + \eta \Delta \theta^{\text{new}} \tag{7}
\]

3 System Implementation

The estimated image Jacobian is applied in the TWINETR system built-in house at Robotics and Automation Research Unit for Real-World Applications (RARA), Suranaree University of Technology. The system is composed of 2 color CCD cameras mounted on a 4-DOF pan-tilt-verge head as shown in Figure 1.

Figure 2 shows example of stereo images with size of 320x240 pixels from the TWINETR vision system. Each image is divided into 8x6 subimages having size of 40x40 pixels each. Each subimage contains individual weight of SOM corresponding to the motor command which can bring target image located in particular subimage to center of the image for both cameras. The target can be segmented and located using simple color segmentation.

The resulting online training of TWINETR vision system to achieve saccade movement is presented in the next section. Note that the SOM network can be trained both off-line and on-line. The system control diagram is displayed in Figure 3 and a flowchart of training process is shown in Figure 4.
4 Experimental Results

The system is trained in 3 set of experiments. The first experiment is to test the convergence of the network on left and right motors only. The target is moved and stopped around the workspace. Each time it stops, the system updates corresponding weight and drives motor to move until the target is brought to the center of the image. The results are demonstrated in Figure 5. The final weight update is guaranteed such that the corresponding weight can bring the target to the center of the image in merely one motor command. Figure 6 shows the result of tracking after the training is complete. The results show that the system can bring the target to the center of the image with saccade characteristics.

The second experiment is to train the system with random position of the target with the left, right and tilt motors. While the target moves around the work space, the network is trained corresponding to the current position of the target. The results in Figure 7 shows that the training is eventually converge.

The final experiment is to train the system with the target moving around in circular pattern. The results shown in Figure 8 demonstrate the convergence of the network after some period of time in which the system is able to track and keep the target at the center of the image.

The results from all experiments show that the system can learn to track the moving object with unsupervised characteristics. Desirable efficiency and speed are also obtained. The system can adapt itself without using any mathematical model calculation or any hardware calibration.
5 Conclusions and Future Work

This paper presents the development of controlling PTV head to achieve one of human-like eye movement behaviors, i.e. saccade. The visual feedback controller using image Jacobian has been implemented. The dynamic Jacobian estimation is obtained by using unsupervised learning scheme and self-organizing map network. The results show desirable efficiency and speed of learning. The system can perform tracking without using any mathematical model or any hardware calibration. Thus the system can conveniently be implemented. Due to the fact that no model or calibrations are not required. The system can be extended to apply for a verge motor. In this case, the movement of such motor can causes all other motor to change their physical positions. The controller can become more complicated in design. The proposed method is hence very promising to be able to take care of such case efficiently.

Fig. 4 Training process.

Fig. 5 Target positions from center of images: (a) left image and (b) right image.

Fig. 6 Saccadic movement in object tracking after training.
Fig. 7 Convergence of training on (a) left motor (b) right motor and (c) tilt motor.

Fig. 8 On-line training with the moving target in circular pattern.
References: