行政院國家科學委員會專題研究計畫 成果報告

具人機互動能力之自走式仿人類眼球運動機器人
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 98-2221-E-009-128-
執 行 期 間 ：98 年 08 月 01 日至 99 年 07 月 31 日
執 行 單 位 ：國立交通大學電機與控制工程學系（所）

計 畫 主 持 人 ：陳永平

計畫參與人員：碩士班研究生-兼任助理人員：劉澤翰
碩士班研究生-兼任助理人員：白文榜
碩士班研究生-兼任助理人員：黃文俊
博士班研究生-兼任助理人員：楊世宏

公 開 資 訊 ：本計畫可公開查詢

中 華 民 國 99 年 10 月 31 日

行政院國家科學委員會專題研究計畫成果報告

具人機互動能力之自走式仿人類眼球運動機器人

計畫編號：NSC 98-2221-E-009-128-
執行期限：98 年 8 月 1 日至 99 年 7 月 31 日

主持人：陳永平　　　國立交通大學電控工程研究所
共同主持人：
計畫參與人員：楊世宏、黃文俊、白文榜、劉澤翰
國立交通大學電控工程研究所

中文摘要

本計劃嘗試結合仿人類眼球運動機器人以及自走車，成為自走式仿人眼機器人，設計智慧型法則，以視覺為基礎執行人機互動。本計劃已完成智慧型手勢辨識系統設計以及智慧型多軸追蹤控制，其中手勢辨識系統根據人腦所認知手姿態狀態來識別不同的手勢，總共有九種手勢可被此系統描述。而多軸追蹤控制利用倒傳遞演算學習法則的離線學習來訓練控制器，以學習目標位置與對應速度之間的關係，經由實驗證明確實可成功地完成物體的定點追蹤控制和水平追蹤控制。

關鍵詞：仿人眼機器人、手勢辨識、追蹤控制

Abstract

This project proposes a mobile humanoid eye robot, or MHER for simplicity, constructed from combining a humanoid eye-ball robot and a mobile robot. The MHER is designed to interact with human based on vision processing and intelligent algorithm. This project has developed an intelligent hand gesture recognition system and an intelligent multiaxial tracking control. The hand gesture recognition system can recognize nine hand gestures according to cognitive posture states of human brain. Furthermore, the neural networks learn the relationship of object position and corresponding velocity by the back-propagation algorithm in off-line way. The success of the set-point control and the horizontal trajectory tracking control of the MHER can be concluded from the experiment results.

Keywords: eye robot, hand gesture recognition, tracking control

I. Motivation and Goal

The field of human-robot interaction with a robot is a major challenge. Using hand gestures is a natural way for interaction between people. Hand detection and hand gesture recognition could be essential to human-robot interaction. Considering the global hand pose and each finger joint, the human hand motion has roughly 27 degrees of freedom (DOFs) [1]. This project makes the image processing easier for the vision-based hand gesture recognition. The color-based algorithm is implemented to meet the real-time performance, accuracy and robustness requirements. In recent years, many researchers have been devoted to developing artificial intelligence system for high-level vision [2], representation, and reasoning [3]. In general, the spatial temporal reasoning adopts the concept of state machine to describe a hand gesture represented by a specified state sequence. However, it is not intelligent to design a state machine manually. In order to solve above problems, this project proposes a system combined recurrent neural network and spatial temporal reasoning together.

On the other hand, computer vision and

image processing are important researches owing to their wide varieties of application. Among all of these applications, visual tracking has become a great interest in this domain. The speed of capturing an image must be as fast as possible to detect the movement of the object in real-time. However, the camera has a constraint on the number of capturing image per second. The purpose of this project is to deal with this problem and implement the visual servo control system. This project introduces an efficient approach of a real-time moving object tracking by the two eyes of the MHER and describes the design of an algorithm based on the neural network structure.

II. Hardware

The MHER tracking system is built with two cameras and five motors to emulate human eyeballs as shown in Fig. 1. The MHER adopts five FAULHABER DC servomotors to steer the MHER in tracking system. With RS-232 interface, the controller of DC servomotors is executed by the motion control card, MCDC 3006S, in a positioning resolution of 0.18°. Therefore, the MHER has two pan-direction video cameras, a conjugated tilt motor and a pan-tilt neck. The range of pan is approximately 120 degrees, and tilt is approximately 60 degrees. The size of the head is about 25 cm width and 10 cm height for the eye part.

III. Intelligent Hand Gesture Recognition

The system structure of the proposed hand gesture recognition system [4] includes image data retriever, posture data generator, posture state encoder, and gesture analyzer. The main goal of the proposed hand gesture recognition system is to recognize hand gestures captured by a CCD camera. This project will focus on nine gestures, named upward, downward, turn left, turn right, left around, right around, follow, learning and warming as shown in Fig. 2.

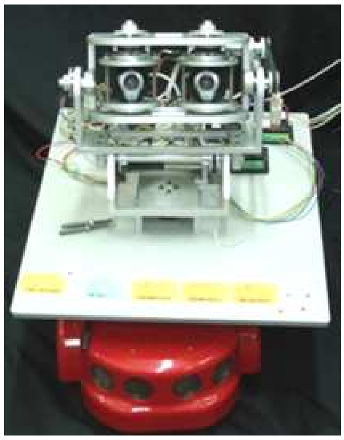

Fig. 1. HMER.

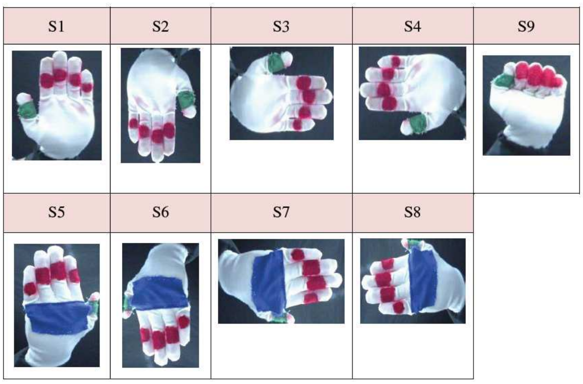

Fig. 2. Hand gestures.

To describe a hand gesture by a posture state sequence, it is often based on a fixed-length sequence of posture states. However, a hand gesture usually happens during an uncertain time; for example one hand posture can change to another hand posture quickly or slowly and thus it is difficult to represent a hand gesture by a fixed-length state sequence. To solve this problem, the hand gesture analyzer contains a trigger net and a gesture classifier, developed to learn a posture state sequence unfixed length. The single state eliminator and the repeated state processor are implemented by the feedforward neural network, while the trigger net is implemented by the recurrent neural network.

There are two kinds of clock mode, synchronous and asynchronous, used in this gesture recognition system. If repeated state retriever is applied for feed-forward classifier, then the gesture recognition system adopts synchronous clock mode for recurrent classifier; otherwise, it adopts asynchronous clock mode.

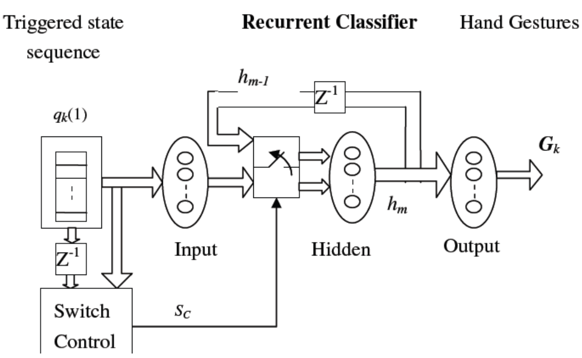

Fig. 3. Recurrent classifier.

Table I. Performance comparison.

	Size of neural network	Complexity of Training	Motion event Extendable	Test error
Feed-forward Classifier	Bigger	Fast	Bad	Good
Recurrent Classifier	Smaller	Slow	Batter	Good

In synchronous clock mode, the output of feed-forward classifier will be processed according to each coming clock. In asynchronous clock mode, however, the output of recurrent classifier shown in Fig. 3 will be processed according to the output of trigger net which is shifted down.

The hand gestures can be recognized by the gesture classifier according to different triggered state sequences by the recurrent neural network. As a result, the performance comparison between feedforward and recurrent neural network is shown in Table I.

The experimental results show that this structure can achieve satisfactory real-time performance and high classification accuracy. The proposed intelligent hand gesture recognition system contained several advantages described below. Another application of the hand gesture can be implemented by increasing hand posture states. There is no limit of length of state sequence for a recognized hand gesture (trigger net). This system would not be influenced by undefined hand gesture (recurrent classifier).

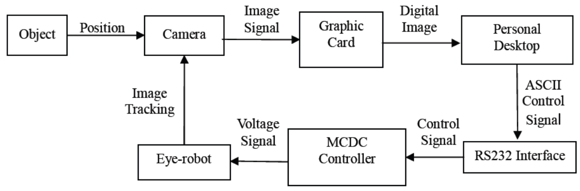

Fig. 4. System framework.

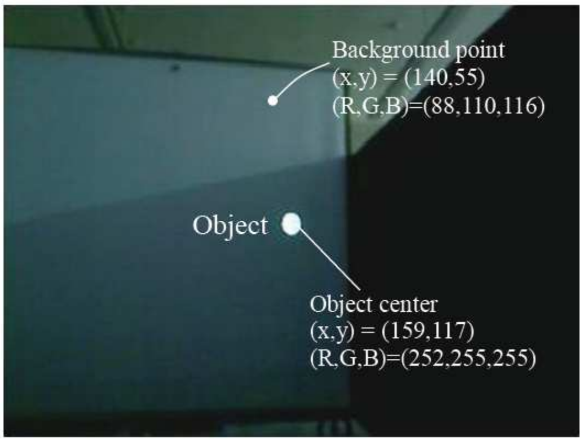

Fig. 5. Experimental scene.

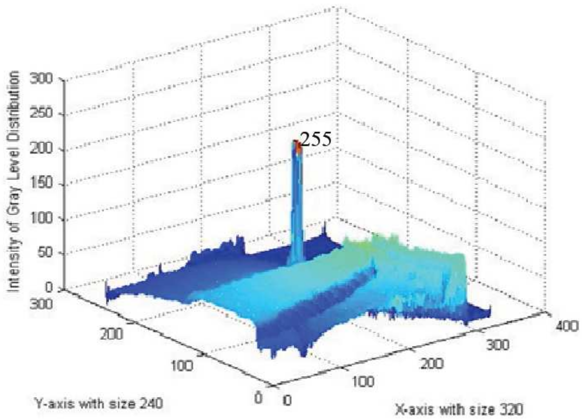

Fig. 6. Intensity of gray level distribution.

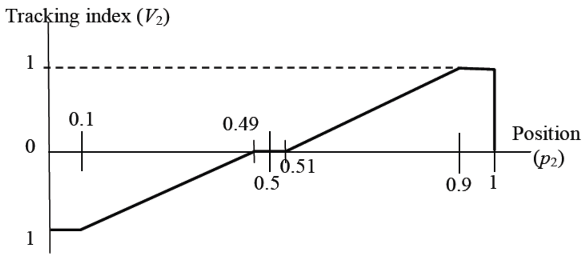

Fig. 7. Neck control training patterns.

IV. Intelligent Multiaxial Tracking Control

The tracking control system shown in Fig. 4 is the main part of this project. The Eye-robot tracking control is based on neural network by the off-line training [5] . There are four neural networks used in this project, including NNneck, NNeye, NNeye1-neck, and NNeye2-neck.

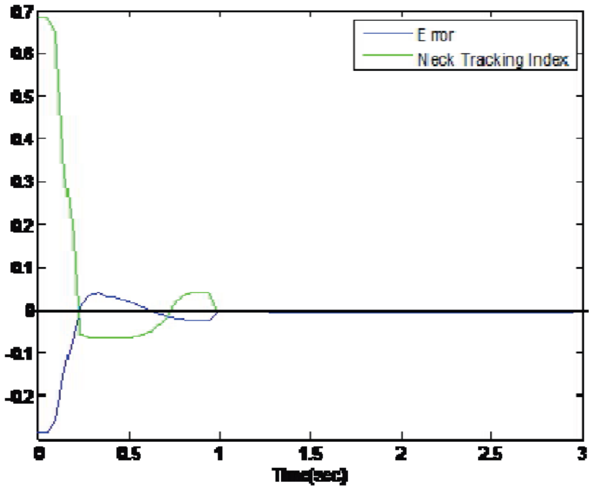

Fig. 8. Set-point control of the neck.

The object tracked by the Eye-robot is projected by an overhead projector in a screen. The speed of the object motion velocity is 28 cm/sec. The distance between the Eye-robot and screen is set as 240 cm and the proportion of the vision object in the Eye-robot is designed as 10cm x 10cm shown in Fig. 5.

To detect the object on the screen, the image of the left-eye camera is first transformed into a gray level image with gray level intensity distribution shown in Fig. 6. Then the object can be extracted.

The neural-network-based controller is achieved via offline training. The input of the neural network is the object center *p* and the output is the tracking index *V* related to the driving voltage of the Eye-robot. The neural network controller receives the error *e* between the current horizontal position x_c and the desired position x_d of the target and provides a tracking index *V* to drive the Eye-robot. The neural network controller design will emphasize on the training pattern driving the Eye-robot to trace the object.

With the patterns shown in Fig. 7, the human visual motion can be mimicked. The left eye is selected as the dominant eye in this project and is designed to have the ability to find the object in searching mode. Once the object is detected, the eyes and the neck will work together to trace the object. It is clear that the tracking motion is mainly executed by the neck when the object appears at the position far away the vision center x_c, while the dominant eye and right eye play the roles of concentration on the object.

After the offline training, the trained NN will be applied to the Eye-robot tracking, i.e., the object is kept in the image center. During the object tracking, the maximal velocity of the neck is set to be v_{max}=1500rpm/min. It is clear from Fig. 8 that the motors employed to steer the two cameras have also reached the control goal to locate the object around the visual center with an error near to zero.

Reference:

[1] Y. Wu and T. S. Huang, “Hand modeling analysis and recognition for vision-based human computer interaction,” *IEEE Signal Process Mag.—Special Issue on Immersive Interactive Technology*, vol. 18, no. 3, pp. 51–60, 2001.

[2] N. Badler, “Temporal Scene Analysis - Conceptual Descriptions of Object Movements,” *Report TR* 80, 1975.

[3] R. Brooks, “Symbolic reasoning among 3D models and 2D images,” *Artificial Intelligence*, pp. 285-348, 1981.

[4] 洪新光，智慧型手勢辨識系統設計，交通大學碩士論文，2009。

[5] 徐傳源，應用於眼球機器人之智慧型多軸追蹤控制，交通大學碩士論文，2010。

附件：封面格式

行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※※
※ ※
※ 具人機互動能力之自走式仿人類眼球運動機器人 ※
※ ※
※※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：■個別型計畫　□整合型計畫
計畫編號：NSC　98-2221-E-009-128-
執行期間：98 年 8 月 1 日至 99 年 7 月 31 日

計畫主持人：陳永平
共同主持人：
計畫參與人員：楊世宏、梁瑋哲、徐傳源、黃文俊、白文榜、劉澤翰

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學電控工程研究所

中　華　民　國　99 年　10 月　15 日

無衍生研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：陳永平			計畫編號：98-2221-E-009-128-				
計畫名稱：具人機互動能力之自走式仿人類眼球運動機器人							
成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		National Computer Symposium: Workshop on ICM
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	3	3	100%	人次	
		博士生	1	1	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		36th Annual Conference of the IEEE Industrial Electronics Society
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估 ■達成目標 □未達成目標（請說明，以 100 字為限） 　　□實驗失敗 　　□因故實驗中斷 　　□其他原因 說明：
2. 研究成果在學術期刊發表或申請專利等情形： 論文：■已發表 □未發表之文稿 □撰寫中 □無 專利：□已獲得 □申請中 ■無 技轉：□已技轉 □洽談中 ■無 其他：（以 100 字為限）
3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限） 學術成就: 本計劃所開發的系統利用類神經網路學習空間–時間序列，將來可進一步利用演化式計算改進學習的法則。 技術創新: 本計劃之智慧型手勢辨識及追蹤控制系統具有學習功能，為智慧型機器人之重要一環。 社會影響: 因為高齡人口比例與日俱增，使得服務性機器人日趨重要，本計劃所開發之智慧型手勢辨識提供人機一項方便的互動介面，不管年齡層，皆可適用，此外，追蹤控制系統可添增人機互動的樂趣，是服務性機器人不可或缺的元素。