2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

The Distance Communication System by using Intelligent Space

Hiroo UMEDA , Toru YAMAGUCHI
Department of Electronic Systems Engineering, Tokyo Metropolitan Institute of Technology,
6-6 Asahigaoka , Hino , Tokyo 191-0065 , JAPAN
umeda@fml.ec.tmit.ac.jp , yamachan@fml.ec.tmit.ac.jp

Abstract **- We propose the Distance Communication System that is not only Making Distance Learning Contents but also controlling intellectual moving object. In order to make Distance Learning Contents (Video Contents), we must follow the motion of lecturer. In the former Systems and a person operates Video-Camera because it's not enough to follow the motion. In this research and we make the systems that can match the motion of lecturer naturally. The systems use Intelligent Space software and so the systems recognize lecturer's motion automatically and order Pan/Tilt-Type Camera to follow the motion. And we made possible to operate an intellectual moving object with application of this system.**

1. Introduction

A lot of distance learning contents have using Video On Demand style. This style is mainly using video multimedia with the lecturer's voice and picture, and HTML pages of the lecture's point.

In the making process of these contents, now we have many problems, so wasting time for making contents. In former process, first, we have to record the lecture to the videotape by following the motion of the lecturer, and this was done by human operation.

Secondary, we have to convert the recorded video data to some format that can use with the personal computer for adjusting to the Distance Learning System.

In final process, we divide the converted video data along with the lecture contents. When attending this kind of lecture, we have a merit that we can listen many time until you could understand. For those reasons, we have to divide video data into a paragraph of the lecture. And we can get above merit because we can search the point which we could not umderstand quickly. However in this process, a third person except lecturer cannot divide without listening the video data many time and understand the contents of the lecture. So this process takes time most. And this process needs exclusive software and skill that can manage the software. In the society which is changing quickly to age society, in early future, people who is unaccustomed to operate a personal computer is increasing, and the matter that those people giving a lecture is also increasing inevitably. For those people, the clicking mouse operation is to be a particular nervous operation, so it cannot say the easy operation that everyone can do. For the reasons, we need easy interface taking the place of mouse and keyboard. You can think that, for example, using human voice and movement. These interface that using compound interface with various methods are called Multi-Modal Interface, and which realize the natural interface for the human with Multi-Modal Interface is Natural Interface. We think the intention recognition with human movement is especially important for Natural Interface.

Then this paper describes Distance Learning System for making contents with few people, short time, and easily by using human movement of the hand and the face as interface. With this system, we think that making contents is becoming more easily and number of Distance Learning Contents increasing.

In Section 2, show the system outline and express the summary of using software. In 3rd Section, express the function and interface of the system, in Section 4, show the experiment and the result of this system. In Section 5, express the summary and future tasks.

2. System Outline

This system, it is shown in Fig.1, is constituted by intelligent interface agents that consist of the PC + camera + intelligent space software (in below, I-Space).

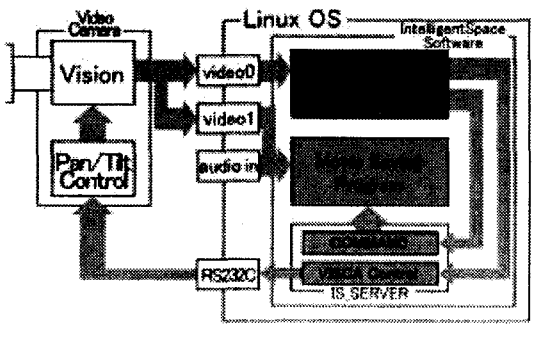

Fig.1. System Outline

This I-Space is open source software that is researched and developing at our laboratories and is working on the Linux operating system. This software is formed the following 2 large part.

1) The intelligent space core that detects the position of the object from color information.
   (In Fig.1: I-Space & Plugins)
2) The server that processes variously using 1)'s information.
   (In Fig.1: IS_SERVER)

The communication between above 2 part uses TCP/IP protocol. With former ISpace, this communication is done only 1 to 1. So in the process of making this system, we changed the program of I-Space itself for to enable to communicate between 1 I-Space and some servers. With this change, it became possible to divide I-Space's various process because each other process can work on each other PC when we need large-scale processing.

And this system uses the camera that can be controlled with VISCA (Video System Control Architecture) protocol. So the PC needs the Serial interface (RS232C) for outputting VISCA. Further the PC needs 2 video capture boards and 1 sound board, and input image of the video camera to each video capture boards, the sound of the lecturer to the sound board.

3. Intelligent Interface Agent

Intelligent Interface Agent recognize the command of vedeo recording START or STOP and follow the motion of the lecturer from the position or movement of lecturer's hands and face.

3.1. Input Information

I-Space fixes the position of the lecturer's hands and face by detecting those colors from the information of the video camera's picture and calculates the coordinates of the center of gravity. We call this methods tracking.

These tracking objects (The hand and the face) have two data that is X direction and Y direction as the coordinates of the center of gravity[x , y]. Taking this coordinate value as time series data (Examples in Fig.2), and set the extremum to the characteristic points. Taking the position of the hand, and from the relative positional relationship of the face position, the movement of the hands can be detected. In addition, characteristics of the movement are detected unless depending on the distance of the camera and the person by taking relative position and formalizing.

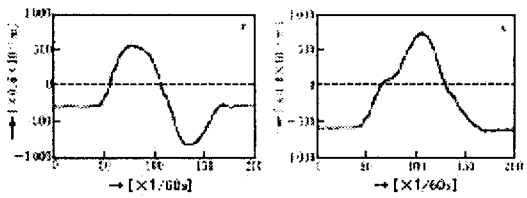

Fig.2 Example of time-series data

3.2. Fuzzy Associative Inferences

Movement of the human, the speed and size, is not fixed. So this system uses the fuzzy associative inference which has robustness to unstable noise for the command recognition. Fig.3 shows outline of the command recognition by I-Space.

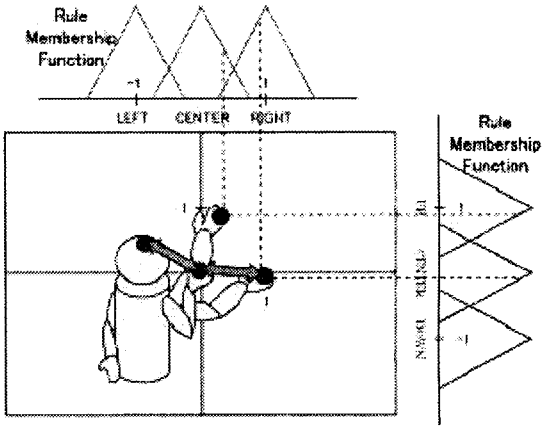

Fig.3 Command Recognition

In this recognition, characteristic points are set the fuzzy label by membership function, and processing propagation based on the rules in the network which is formed bidirectional associative memory (BAM).

3.3. Video Recording Control

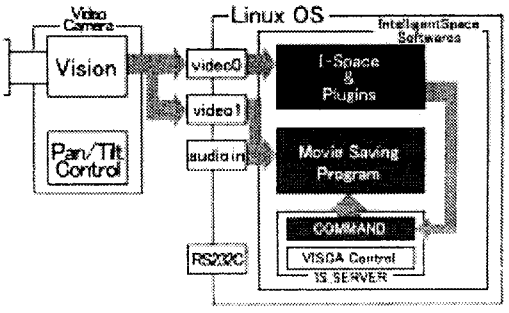

Fig.4 Data flow of Video Recording Control

Fig.4 shows the data flow of video recording control. Fig.5 is command recognition fuzzy rules that are expressed with network of associative memory.

The layer that represents the antecedent section of fuzzy rule is if-Layer, the layer which represents the consequent section is then-Layer. And the node that is inside of if-Layer is member ship function of the antecedent section, the node that is inside of then-Layer is member ship function or input/output function of the consequent section. In addition, correlation between the antecedent section and the consequent section is over the limit of BAM's remembrance generally, so the rule-Layer that one node represents one rule was set between if-Layer and then-Layer.

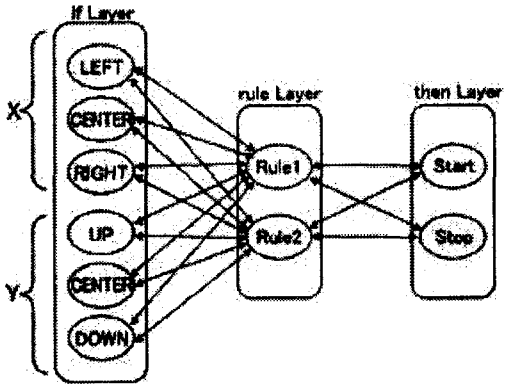

Fig.5 The fuzzy ru le of associative memory network

Therefore fuzzy rules are expressed by forming BAM between if-Layer and rule-Layer, rule-Layer and then-Layer.

Abovementioned processes control the video recording program. This video recording program saves the video data as AVI format. Because of these processes, themselves who lectures control START and STOP of video recording, it can make easy to divide the lecture contents, and the video data is saved as format that adjusts to PC.

3.4. Lecturer Following Control

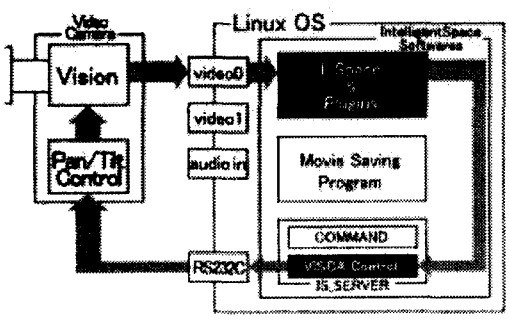

Fig.6 Data flow of Lecturer Following Control

Data flow of lecturer following control is shows in Fig.6. In order to make the contents from the video data, it is necessary for the face of the lecturer to be always on the picture. Fig.7 shows the range of lecturer's face. For example, the range of the lecturer's face is set 1/2 size of height and width as the center of Fig.7.

When the head of the lecturer goes outside from this range, the camera follows the movement in order that the head of the lecturer to be settled inside the range. As for this, camera control section (In Fig.1: VISCA Control) decides the direction of camera based on the rule of inside with input information (coordinates of the center of gravity) which is mentioned in Section 3.1. If the coordinates of the center of gravity is (x, y), height and width of window size that I-Space has recognized is H and W, this rule is shown in Chart.1.

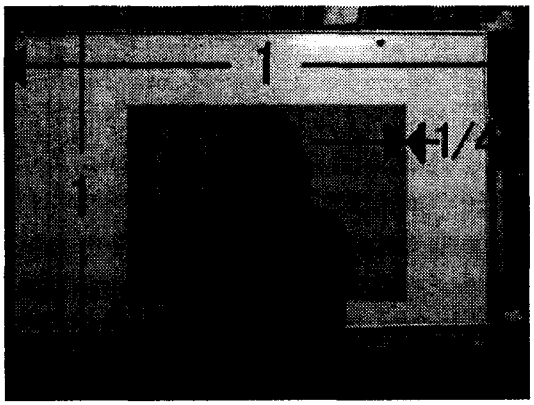

Fig.7 Range of Face

Chart.1 Rule of Camera Control

Rule	If	Then
1	$x > \frac{3}{4}W$	Pan+
	$x < \frac{1}{4}W$	Pan-
2	$y > \frac{3}{4}H$	Tilt+
	$y < \frac{1}{4}H$	Tilt-

The direction that is decided in this way is converted to the VISCA signal. And the signal outputs to camera through the serial interface (RS232C), so the camera can move the Pan and Tilt.

Because of this, the camera can move to follow the movement of the lecturer automatically, so the cameraman is not needed.

4. Experiments and Consideration

4.1. Command Recognition Experiment

The recording video data becomes the contents itself. So this command recognition part must operate securely. In below, the experiment of command recognition and the result are shown.

Tracking the face and hand of subject (the lecturer) who is 2 meter far from video camera by I-Space with 10 flames per second.

When raising the hand, video recording is START. And when moveing hand to the left, video recording is STOP.

Acting 100 times that Rasing the hand and moving hand to the left, and check correct recognition rate. The experimental result is shown in Chart.2.

Chart.2 Result of Command Recognition Experiment

Command	Rate	Frequency
START	**98 %**	98/100
STOP	**99 %**	99/100

4.2. Considerations

Both START and STOP recognition is high probability. There is no either thought that because presently is only two operational command, the recognition is high probability. However the system which applies this command recognition mechanism by I-Space can operate intellectual moving object with 7 types command input of natural human motion, the correct recognition rate is all high probability. The scene of experiment of garage inserting operation by the intellectual moving object is shown in Fig.8, and the correct recognition rate is shown in Chart.3.

Fig.8 Scene of Intellectual moving object

Chart.3 Recognition Rate of Intellectual moving Object

Command	**Rate**	**Command**	**Rate**
Forward	94%	Back	92%
Right	88%	Left	90%
Stop	100%	System Start	100%
System Stop	100%	Average	95%

As recognized from this result, the command recognition by I-Space is high probability.

In addition, the flame rate, 10 flames per second is really used to make Distance Learning contents at Distance Learning Center of Tokyo Metropolitan Institute of Technology, the flame rate does not give strange feeling to attendant who views video contents. Actually looking at recorded video data, especially, there is no strange feelings.

5. Conclusions

In this paper, we propose the Distance Communication System for making contents that needs less number of people, little time to making contents easily, and propose command recognition method with natural interface and the lecturer following function.

To improve this system, we plan to do the following research:

- To strength the recognition mechanism of I-Space by optimize filtering.
- Controlling number of camera by using attention mechanism.
- More naturalize of interface of command recognition by improving recognition of movement.

In addition, we will develop to the system include the vending and transmitting the contents.

References

[1] T.Yamaguchi , N.Ando, Intelligent Space and Human Centered Robotics, Proc. Of Euro Symposium on CI(2002.6)

[2] Distance Learning Center –Tokyo Metropolitan Institute of Technology-, http://www.dlc.tmit.ac.jp/

[3] Neural Network and Fuzzy Signal Processing, T.Yahagi , M.Hagiwara , T.Yamaguchi, CORONA Publishing, 1998