Mental Image Expression by Media Integration
~ COMICS (Computer Organized Media Integration & Communication System) ~

Seiki Inoue

Media Integration & Communications Research Laboratories, ATR
2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02, Japan
Tel: +81-774-95-1460, Fax: +81-774-95-1408
Email: sinoue@mic.atr.co.jp

Abstract

We introduce the outline of our research plans concerning mental image expression for the purpose of communication. We are studying the Computer Organized Media Integration & Communication System (COMICS) as a flexible media handling environment. The system can expand the potential of mental image expression by freely handling and integrating various kinds of media. An experimental COMICS is also introduced.

1. Introduction

In conventional communication, logical information is mainly transmitted through voice. But recent progress in video processing and the expansion of the communication bandwidth have made it possible to communicate with each other using a wider range of media including video. Such a system can also transmit nonlogical information, such as feeling or emotion using video and sounds.

Against such a background, ATR Media Integration & Communications Research Laboratories (MIC) was founded in March 1995. Its major objective is to study new types of communication based on multimedia. In MIC, the following 3 themes are now being studied.

(1) Communication environment technology
The reconstruction and creation of communication environments

(2) Communication support technology
Creative thinking support, agents for the communication and transmission of mental images

(3) Science of human communication
Studies of human communication processes

In our department, we are studying how to express the mental images in a human's head as one form of communication support. By the way, in this paper we use the word "image" as something which exists in a human's head and cannot be expressed easily. For image expression, we are now developing a unified media handling system, COMICS (Computer Organized Media Integration & Communication System)[1] which can provide a flexible and efficient media handling environment. Users can use the necessary media tools easily, even though they are not familiar with each piece of hardware. Various kinds of media are combined with some meaning to express an image. In this paper, the concept and an experimental system of COMICS are described.

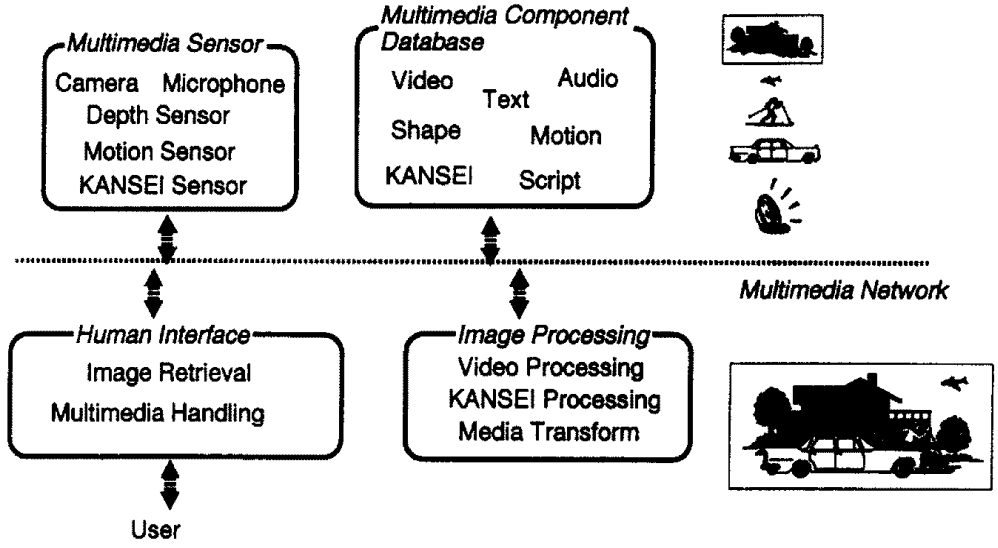

Fig.1 System configuration for mental image expression

2. System Configuration for Mental Image Expression

Fig. 1 shows the total system for mental image expression which we are going to study. It mainly consists of 4 parts.

Before explaining each part, we will explain about KANSEI. KANSEI is a very vague Japanese word which cannot be defined exactly. KANSEI is like emotion but not exactly. For example, there are persons who can give an audience a greater impression than other people, even though they are using the same set of video and sound materials. At that time we often say that he or she has KANSEI. KANSEI is some kind of ability. KANSEI as an ability contains the following three kinds of abilities: feeling ability, expanding ability and expressing ability. As you can see in Fig. 2, there are many kinds of feelings or impressions. We intend to analyze the reason why expression can give an audience such an impression, and use it in the next expression as research into KANSEI processing.

Now, we will explain each part in Fig. 1.

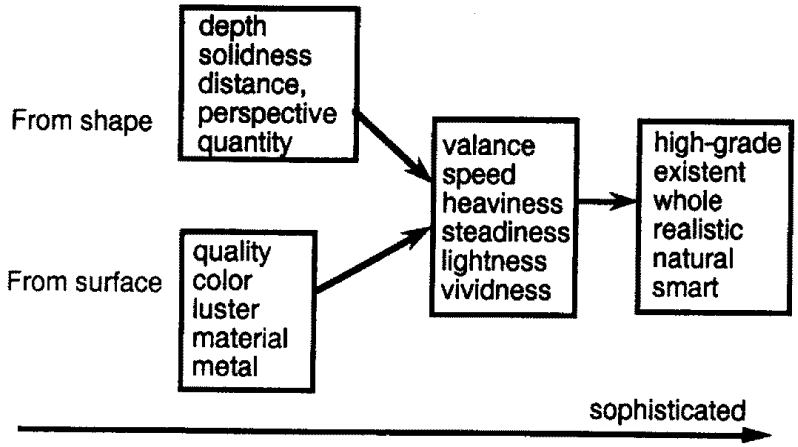

Fig. 2 Examples of impression

(1) Multimedia component database

Not only video and audio components but also additional or extracted data, such as 3-D shapes or the motion of objects, camera work and lighting conditions, are treated as media components. They are stored in a database and combined to form an expression. The information concerning this combination is called a script. Scripts are also stored in a database as components.

(2) Image processing

This includes not only video and audio processing but also KANSEI processing. In video processing, for example, more intelligent processing is achieved using KANSEI data. In addition to conventional preprocessing methods, such as edge detection and segmentation, methods which can make videos more vivid or more speedy are employed with the help of KANSEI processing.

(3) Human interface

With this part, users can retrieve images which are similar to theirs and know how and what components are combined. They can change components and customize the retrieved images by handling the components freely. They don't have to know each piece of hardware related to each medium.

(4) Multimedia Sensor

To store media components, they have to be extracted from ordinary input devices. Depth and motion sensors are necessary to acquire depth and motion information in the same way that cameras and microphones are used to input of video images and sounds. Sensors don't always mean hardware sensors. They sometimes mean software algorithms. We will study software sensors in order to acquire medium information.

3. COMICS

3.1 Concept of COMICS

We have started studying media handling environments. There is a handling tool for each medium and it is being developed separately. Musical instruments are for sounds, and video processing equipment is for video images. They require special knowledge and skills to handle. Our COMICS will provide users with an environment where they can handle various kinds of media freely without special knowledge or skills about media tools, and with which they can combine several media to express their images. Fig. 3 shows the concept of COMICS.

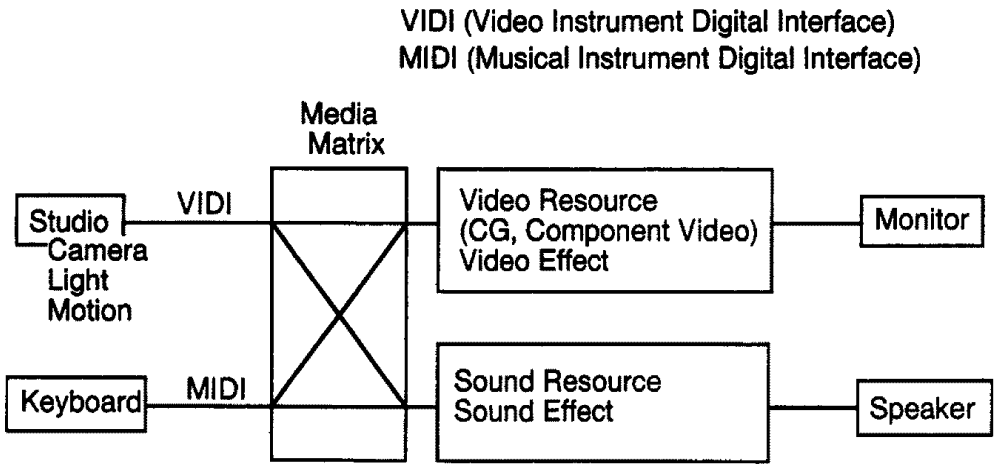

Fig. 3 The concept of COMICS

In computer music, like a MIDI (Musical Instrument Digital Interface) system, we think the input device is a musical instrument such as a MIDI keyboard and the output device is a speaker. The system has a digital interface MIDI which can control sound resources. As an analogy of computer music, we'd like to develop computer video. The input device for computer video is a camera and the output device is a video monitor. The system also has a digital video interface VIDI (Video Instrument Digital Interface) which can control video resources such as CG (Computer Graphics) or video components. There are video effects just as there are sound effects. Furthermore, if we can combine computer music, computer video and other computer media and change input and output devices, the range of expression will expand drastically. So in COMICS, users can connect each input to each output freely and, for example, play video instruments like musical instruments. The input device doesn't always mean a hardware device, and can sometimes be information extracted from ordinary video images or sound. For example, light information or motion can be connected to sound.

3.2 An experimental system

Fig. 4, Fig. 5 and Fig. 6 show the configuration of an experimental COMICS system, an operating scene and an example of an interface screen.

The left side of the screen stands for input devices: MIDI keyboard, motion, GUI (Graphical User Interface), and the right one for output devices: sound, video effects, CG. In the following three demonstrations, inputs are connected to outputs.

(1) From the MIDI keyboard to sound and CG (Fig. 6 and Fig. 7)

When you play the MIDI keyboard, you can hear its sound and see the CG of characters rotate in relation to the keys. The strength with which the keys are pressed controls the rotation speed and color.

(2) From GUI to video and sound

You can paint a picture using a mouse pointer and hear sound played simultaneously according to its movement.

(3) From movement to video effects and sound (Fig. 8)

The hi-lighted area around the ball moves according to the motion of the ball. At the same time, a heart beat can be heard. The nearer the ball comes to the keeper, the higher the tone of the heart beat becomes. So this sound can express the keeper's mind.

These are only demonstrations and are not the purpose of our research. The purpose of our research is to develop technologies to express the images in humans' heads easily. Through these demonstrations, we can confirm that the concept of COMICS is efficient for image expression. Up to now, connections have required special knowledge about media and are limited. We are now developing data models and human interfaces to handle an experimental COMICS easily.

3.3 Data model for COMICS

From the viewpoint of computer software, we have to establish a data model to handle the system easily. The task of expression has to be divided into small software modules. They have to be classified in a systematic way where some, for example, depend on hardware and others do not. They are expected to be manageable and to be easy to link to each other. If a data model which can manage modules is made and users

handle media along this model, the system can describe how and what media was used. And we can analyze this description for KANSEI processing. With this purpose in mind, we are now studying a three-layer model as shown in Fig. 9.

(1) Raw data processing layer

This provides basic functions, such as the input and output of media data and the management of different media. Management, for example, means methods which can extract attribute data, such as the motion of objects from videos.

(2) Tool operating layer

This provides functions for controlling media

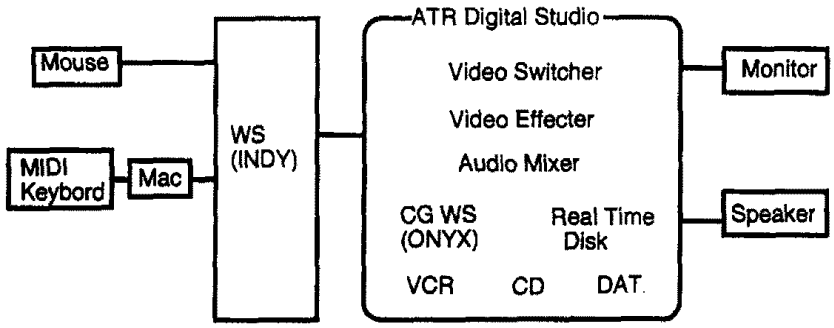

Fig.4 Hardware configuration of an experimental COMICS

Fig.5 An operating scene

Fig.6 An example of an interface screen

Fig. 7 An example of expression (Rotating characters)

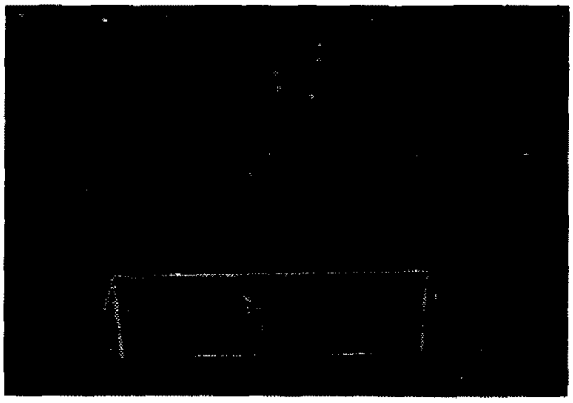

Fig. 8 An example of expression (Video effect)

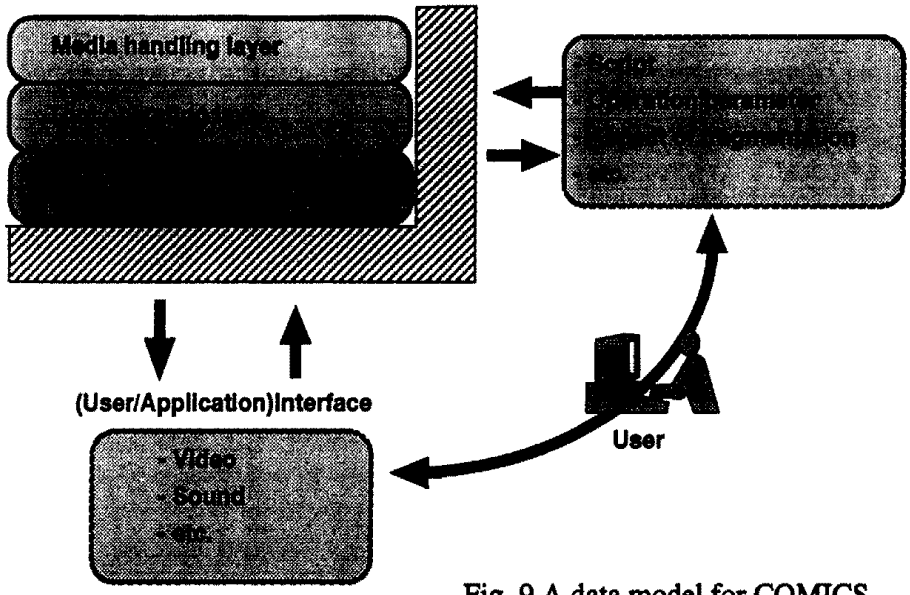

Fig. 9 A data model for COMICS

tools such as a video effecter or CG generator, and the management of such tools. Management, for example, means methods which can transfer sound control data to the parameters of a CG generator.

(3) Media handling layer

This provides abstract functions for handling media and the management of expression descriptions. It has a sophisticated interface and is easy to use without any special knowledge.

This model has not been finished. We will develop it by reordering and rearranging the programs of an experimental COMICS and, of course, adding new programs.

4. Conclusion

In this paper, we introduce our research plans for image expression in communication, and propose the concept of COMICS as a media handling environment. An experimental COMICS shows the validity of the concept.

To improve the system, we will study ways to attain easy operation. For that purpose, software models and friendly human interfaces will be studied.

In addition to media handling, we have also started to study the detection of 3 dimensional motion in sports scenes such as soccer[2], scripts of dramas[3] and KANSEI processing. We are now studying color reproduction as a part of the KANSEI processing of color. We also plan to study about a media server as an efficient and flexible database.

Research of mental image expression is being promoted by the following members:
Seiki Inoue, Masayuki Inoue, Tsuneki Haizuka, Hisashi Miyama, Michitoshi Ishiwaka.

References

(1) M. Ishiwaka, et al., "A Framework and a Prototype of Integrated Media Handling Environment COMICS" Conv. IPS Japan, March 1996.

(2) H. Miyama, et al., "Moving Object Tracking by Autonomous Distributed Architecture" Nat. Conf. of IEICE Japan, March 1996.

(3) T. Haizuka, et al., "Representation method in Noh video" Nat. Conf. of IEICE Japan, March 1996.