Virtual Collaborative Workspace

Gen SUZUKI and Takashi KOUNO

NTT Human Interface Laboratories Visual Media Laboratory
1-2356 Take, Yokosuka-shi, Kanagawa-ken, 238-03 Japan

e-mail: gen@ntthcs.ntt.jp, kouno@ntthcs.ntt.jp

ABSTRACT
This paper proposes the concept of a "Virtual Collaborative Workspace" as a widely acceptable Telecommunication service. A model of the collaborative workspace and the basic functions needed for this service are described. Based on these descriptions, the concept of an intelligent handset that integrates the basic functions needed for visual telecommunication is proposed. System configuration of the handset is discussed with the view of realizing a "Virtual Collaborative Workspace." In order to confirm this concept, a psychological evaluation is conducted on face masking to confirm the usefulness of facial expressions.
The prototype of the intelligent handset (ScopeHand) is implemented. Use of the prototype system confirms the effectiveness of the concept.

Keywords: Videophone, Virtual Reality, Visual Telephone, Human Interface.

1. INTRODUCTION

Visual telecommunication has the potential to be a powerful tool that supports human collaboration over long distances. The role of visual telecommunication can be thought of as the exchange of what one wants to see or what one wants to show the other side. Videophone and video conference systems have already been introduced as electronic environments in which we can exchange facial expressions[1]. However, compared with real face-to-face collaboration, these existing visual teleconferencing tools have extremely limited human interface capabilities, because they employ a fixed camera and a fixed display. Fixed video systems are unsuitable for collaborative work which involves various topical objects such as equipment, furniture, and dresses. Many objects can not be easily placed into the field of view of a fixed camera. A freely moving camera is necessary to support dynamic collaborative activities such as the operation of equipment, explanation of furniture, and the management of building construction.
On the other hand, a freely moving display is necessary to support practical collaboration in the various fields that need considerable body movement.

HMD (Head Mounted Display) is a freely moving display that can create a virtual space[2]. It uses a face direction sensor and head mounted displays which cover both eyes. A stereoscopic sensation is created by the parallax of the pictures seen by both eyes. A computer senses the position and direction of the user's head, and draws pictures that match the view point of the user's head position. HMD can create various virtual spaces that are very effective for human-computer interfaces.

Before applying HMD to the communication interface between humans, the following problems must be considered.
(1) Facial expressions can not be shown because the display covers both eyes.
(2) The immediate environment can not be noticed, for the same reason.
(3) HMD is difficult to put on and take off because of its totally enclosed style.

From the viewpoint of human communication or collaboration, stereoscopic sensation is not always necessary. A simpler display might be able to create virtual workspaces that are very effective for supporting communication, even if the display is for a single eye with no stereoscopic sensation [3]. Facial expressions can be recognized if just one eye is covered by a miniature display.

A combined face camera and display for a single eye is thought to be a practical and effective communication tool because it allows the use of facial expressions, freely moving video input/output, and the creation of a truly effective virtual workspace.

This paper proposes the concept of a virtual collaborative workspace that uses a combined face camera and a display that covers a single eye.

2. Model of Visual Space for Human Collaboration

2.1 Virtual Collaborative Workspace

Collaboration with face to face conversation is the fundamental activity characterizing human society. Almost all social activities are collaborative and are accompanied by interactive conversation. In order to support collaboration with face to face conversation, special physical spaces such as meeting rooms and work rooms are usually prepared. In the case of field work, the actual field location such as the operation room or machine room is the collaborative workspace. People gather in these collaborative workspaces and bring various topical materials such as documents, photographs, design models, equipments and goods that help explain the topical subjects.

To realize effective collaboration, it is indispensable to show and see these topical materials. These materials play a major role in explaining and recognizing the topical subjects, in solving the problems of the topical subjects, and in creating new ideas.

In order to support effective collaboration over long distances, various information related to the workspace should be combined in a visual environment. We have named this kind of artificially created environment the "Virtual Collaborative Workspace."

High capacity networks will allow the realization of a "Virtual Collaborative Workspace" that supports collaboration over long distances.

2.2 Topical materials

A model of visual objects in the collaborative workspace is schematically shown in Figure 1. The collaborative workspace contains of two kinds of visual objects :
(a) information related to people
(b) information related to topical materials

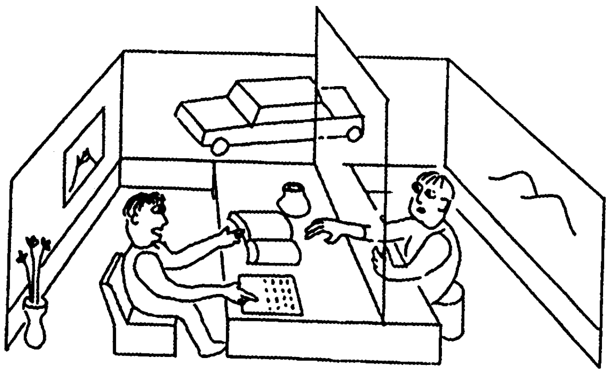

Figure 1. Model of a Collaborative Workspace

Visual information related to people includes human figures and their backgrounds. This information supports the understanding of the other party's situation and is the main service provided by videophones. This information plays some role in creating the feeling of realism experienced in a video conference.

Visual information related to topical materials contains figures of materials and the arrangement of these materials. This information supports the understanding of the topic subjects. This information plays the main role in promoting collaboration.

Because various goods can be used as topical materials, the usual fixed camera is unsuitable. A freely moving camera is essential to capture the topical materials from various view points in various work spaces. In particular, a head mounted camera that permits the user to move freely might be effective and powerful in dynamic collaborations [4].

2.3 Coexistence

In order to support collaboration over long distances by using visual telecommunication, the transmission of visual information related to the participants and topical materials is necessary, but is not sufficient. A human interface that creates a sensation of coexistence is essential to realize an effective environment for interactions.

When people collaborate in the same workspace, they can use visual gestures such as facial expressions, hand gestures, and eye movements, as well as conversation. In order to realize coexistence in a virtual collaborative workspace, support of to such visual clues is essential [5].

As basic visual interactions, in addition to the facial expressions, hand gestures and eye movements play important roles in expressing information and increasing mutual understanding. Usually, as shown in Figure 2, hands and eyes are linked to topical materials. Topical materials are referenced by hands and eyes. These movements strongly influence the scope and efficiency of the conversation.

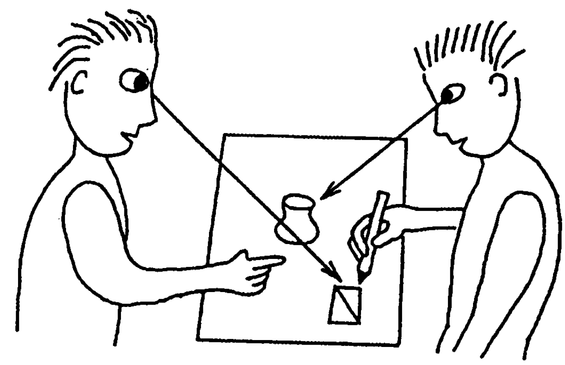

Figure 2. Linkage of topical Materials with Hands and Eyes

3. Proposal of Intelligent Handset

The concept of the intelligent handset called ScopeHand is shown in Figures 3 and 4.
A camera, a small display, a microphone, a speaker, and a position and angle sensor are all integrated into a small handset. The user covers one eye and one ear with the handset. The visual image from the view point of the covered eye is taken by the color camera in the handset. Visual information is displayed on the small display that directly covers the eye. The displayed picture can be controlled by the direction of the user's face. Realistic virtual spaces can be created by the displayed pictures which match the user's facial movement. Facial expressions can be captured and clearly recognized by the partner because only one eye is covered; the other eye and other facial features are uncovered.

Elementary functions for human visual telecommunication: speaking, listening, looking, and showing are all supported by this handy tool. It uses the following basic functions to support collaborative activities over long distances.

Figure 3. The Concept of Intelligent Handset (ScopeHand)

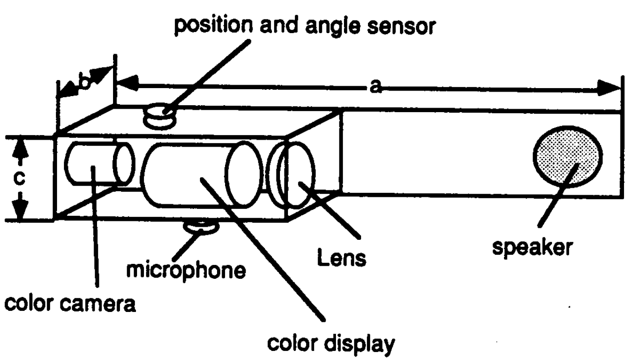

Figure 4. Structure of Intelligent Handset (ScopeHand)

(1) Virtual workspace
A face display that covers a single eye can create virtual environments similar to those realized by the HMD. The current virtual environment created by the usual HMD provides a stereoscopic sensation using the parallax of pictures seen by both eyes. However, a stereoscopic sensation is not essential to create a practical virtual space.

If the picture seen by the covered eye tracks the movements of the user's face, the user feels as if he is using a large display or multiple displays. For collaborative tasks, a wide workspace such as a conference table is necessary to layout a number of topical materials and compare them. By using a face attached display, wide workspaces such as virtual tables or virtual desks can be realized visually without any physical restriction.

(2) Sharing an eye
The face camera included in the handset can be moved freely because it follows the movements of the user's face. The camera takes pictures from the user's view point and shows everything seen by the user.

By viewing pictures taken from the user's view point we create the feeling of "sharing an eye" with the partner. The viewer feels as if he is occupying the user's location.

For collaborative activities, the freely moving camera and shared view are extremely effective in promoting understanding.

(3) Shared Workspace with coexistence feelings
In order to realize a shared workspace that offers the feeling of coexistence, the superimpose technique is introduced as shown in Figure 5. Pictures of topical materials and gestures of the user's hands are taken by his face camera. Video images taken by the face cameras of both sides are superimposed by a video mixer. The superimposed picture is an overlay of the images of both sides. This superimposed image is distributed to both sides.

The coexistence feeling is created from this superimposed image which makes it seem as if both parties are together and pointing at the same topical materials.

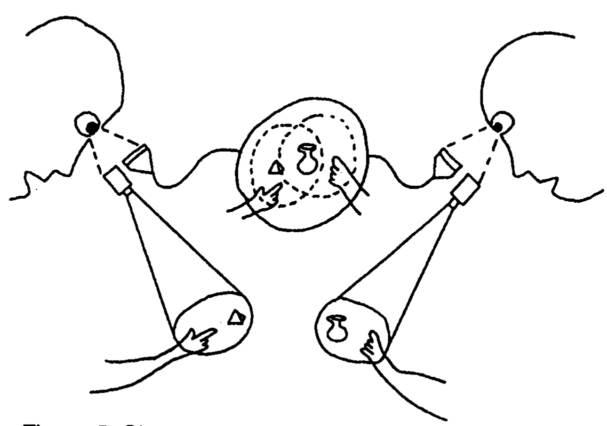

Figure 5. Shared Workspace Promoting Coexistence Feelings

(4) Eye contact
The head mounted display allows virtual images to be seen in any location. Using this flexibility, eye contact function can be realized as follows.

In order to take pictures of your own face, an external camera is placed at the position where you would expect the partner' face to be. The video picture of the partner's face taken by the external camera of the partner's side is shown when you directly look at the external camera. If the partner wants to see your face, he simply looks directly at the external camera. This permits eye contact to be established very easily.

(5) Awareness of both real and virtual spaces
The usage of HMD strongly immerses people into a virtual space if both eyes are covered because the user can not see

his immediate environment.
However, awareness concerning the environment situation is essential for most collaboration activities such as the operation of equipment.

The proposed intelligent handset (ScopeHand) covers only one eye, and the another eye can be used to observe the immediate surroundings. Such semi immersion or simultaneous observation of both the virtual space and the real world is very useful for most regular business activities.

4. Effectiveness of Facial expressions

As described above, the proposed face display and face camera encourage collaboration. These merits are offset to some extent because ScopeHand partially hides facial expressions.

The effects of facial masking were evaluated by psychological opinion tests.

The structure of the experimental system used in the opinion test is shown in Figure 6. Experimental conditions are shown in Table 1. Opinion scores were recorded while masking two facial expressions: happy and angry. The masking positions: both eyes, a single eye, a single eye and mouth, the mouth, mid-point between eye and mouth and no masking were tested. The opinion score scale is shown in Table 2.

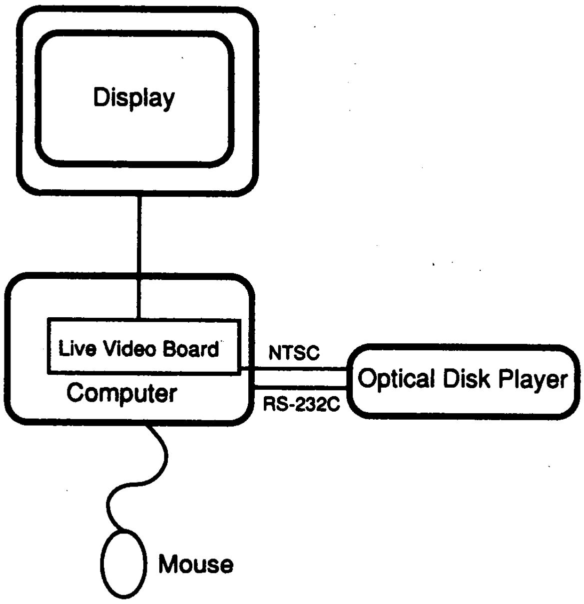

Figure 6. Structure of the Experimental System

Table 1. Experimental Conditions

Display Size	13 inch(W200mm x H145mm)
Distance from Display to Person	90cm(6H)
Facial Expression	Happy and angry
Mask Pattern	Postion:Mouth,Eye,Eyes, Eye and Mouth,Between Eye and Mouth Size:8 variations
Number of Repetitions	2
Display Time	2 seconds
Age of Subjects	23-42 years old
Number of Subjects	10 (5 men and 5 women)

Table 2. Scale of Opinion Score

Scale	Obstructive Feeling
1	No strangeness
2	Slightly strange, but no obstructive feeling
3	Slightly strange and weak obstructive feeling
4	Some obstructive feeling
5	Strong obstructive feeling

Results are shown in Figures 7 and 8. The following points are derived from these results.
(1) MOS (Mean Opinion Score) increases with the size of masking area.
(2) MOS saturates at some masking size
(3) MOS increases with masking position in the following order:
Both eyes > single eye and mouth > single eye > mouth > mid-point
(4) MOS for a single eye and MOS for the mouth are saturated at the same value, MOS=3.

From these results, the masking of a single eye is found to be slightly obstructive to the recognition of facial expressions. This level is almost the same as the masking of the mouth.

These results, indicate that covering a single eye is acceptable in practical use.

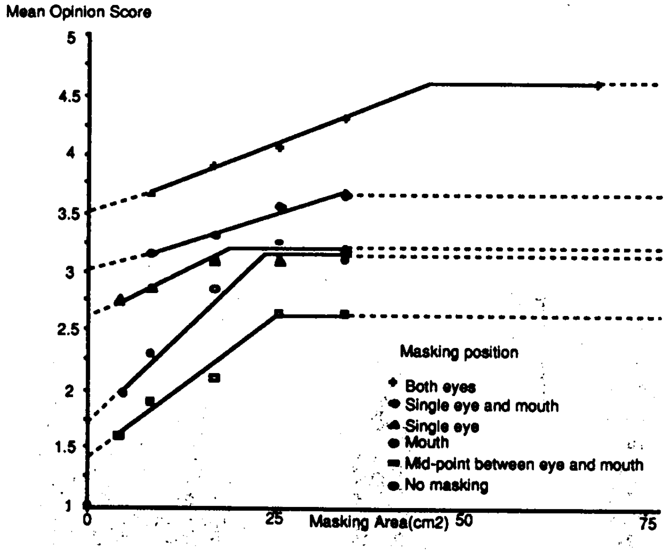

Figure 7. Relation between Size of Masking Area and MOS for Happy Face

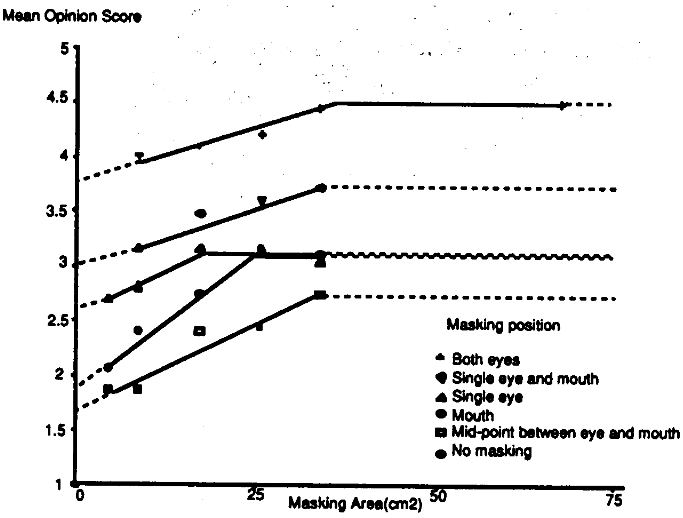

Figure 8. Relation between Size of Masking Area and MOS for Angry Face

5. Prototype System

In order to evaluate the usability of the concept described above, a prototype terminal system (named ScopeHand) was implemented. Specifications and the configuration of the prototype system are shown in Table 3 and Figures 9 and 10. Appearance of the prototype is shown in Figures 11,12,13, and 14.

A liquid crystal display, a CCD camera, a microphone, a speaker, and an angle and position tracker were integrated into a small handset.

The handset is symmetric so that it can be used with either the right or left eye.

Table 3. Specifications of the Prototype System

Handset	Size a x b x c:	300mm x 110mm x 40mm
	Weight	450g
	Video Output	Color LCD
		Picture Size : 21mm x16mm
		Resolution : 55000 pixels
	Video Input	Color CCD Camera
		Resolution : 410000 pixels
	Audio Output	Dynamic Speaker
	Audio Input	Condenser Microphone
Control System	Video and Audio Switcher	Video Signal : 1 Vp-p VBS/VS 75Ω
		Audio Signal : -10dB 50K Ω
		Interface : RS-232C 9600 Baud
	Angle and Position Tracker	Three dimensional digitizer with electro magnetic coupling
		Static Accuracy : 0.1" RMS for the X,Y or Z sensor position 0.1" RMS for sensor orientation
		Output Update Rate : 30 times/sec
		Interface : RS-232C 19200 Baud

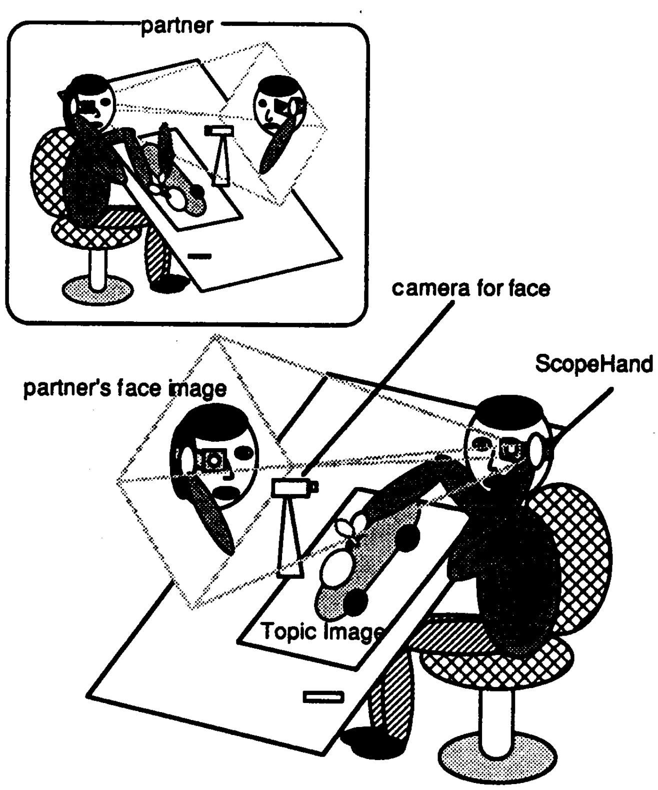

Figure 9. System Image using ScopeHand

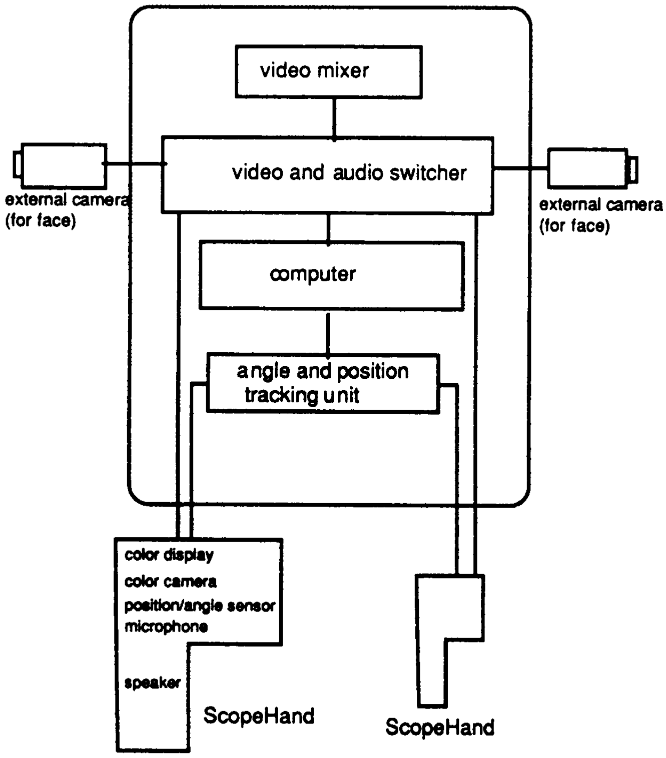

Figure 10. Structure of the System

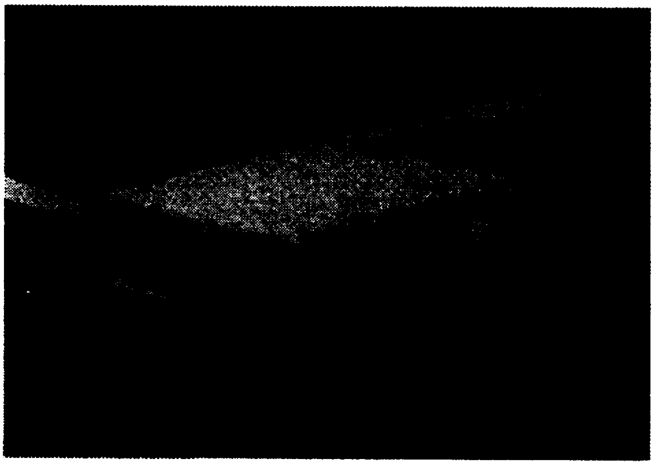

Figure 11. Appearance of ScopeHand

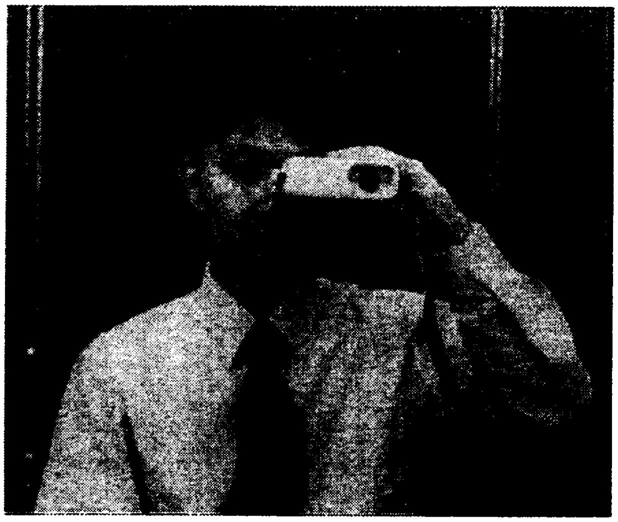

Figure 12. Facial Masking with ScopeHand

Figure 13. Appearance using ScopeHand

Figure 14. Shared Space of ScopeHand

6. Evaluation

The prototype system was tested in our laboratory. The following points were noted during the evaluation.

(1) The face camera is very effective in explaining topical materials.

(2) The face display can create a wide virtual workspace that corresponds to multiple displays.

(3) Superimposing the camera views creates a feeling of virtual coexistence.

(4) The all-in-one type handset is very easy and comfortable to use.

(5) It is not a big problem for each eye to see a different picture.

7. CONCLUSION

(1) The concept of the "Virtual Collaborative Workspace" has been proposed as a widely acceptable visual communication service. The model of the collaborative workspace and basic functions needed for this service were introduced.

(2) The concept of an intelligent handset that integrates the functions needed for visual telecommunication was proposed. System configuration was discussed from the view point of realizing a "Virtual Collaborative Workspace."

(3) A psychological evaluation of face masking was conducted, and confirmed the usefulness of facial expressions.

(4) A prototype of the intelligent handset (ScopeHand) has been implemented. The performance of the prototype proved the effectiveness of this concept.

The intelligent handset is able to support various effective communication functions, in various communication fields such as collaborative environment, wireless video telecommunication and entertainment.

REFERENCES

[1] T.V.Crater: "The Picturephone System: Service Standards", B.S.T.J, vol.50, no.2, pp.235-269 (1971)

[2] S.S.Fisher:" New Ineraction Media: Interacting with Three-Dimensional Environments", CHI'90 Tutorial p.p.190, (1990)

[3] Reflection Technology: "PC Private Eye: User Manual"(1990)

[4] Kuzuoka:"Spatial Workspace Collaboration: A Sharedview Video Support System For Remote Collaboration Capability", CHI'92, pp533-540, (1992)

[5] H.Ishii:"TeamWorkStation: Towards A Seamless Shared Workspace", CSCW'90(1990)