

身体的メタファーとしてのメディアスペース

○三樹 弘之

沖電気工業(株) マルチメディア研究所

山崎 敬一

埼玉大学教養学部

葛岡 英明

筑波大学構造工学系

山崎 晶子

同愛病院

加藤 浩, 鈴木 栄幸

NEC 情報メディア研究所

共同作業の支援を目的としてビデオを利用することは、医療や監視システムといった専門家用のシステムで成功を収めている以外は、実験的な研究にとどまったり、あまり実用的な成功をおさめていなかったりする。専門家用ではなく、遠隔教育や遠隔指示のような実際の社会生活への応用を考える場合、どのように設計したら良いのであろうか。この疑問に答えるため、我々はまず、画面の反転以外は特別な処理を施さないローテクで、カメラとモニターの配置を様々試すことにより、ビデオ利用の問題点、ならびに問題点を解消する技術の方向性を探った。この結果、身体的なメタファーと呼ぶ概念の重要性を認識した。本論文では、この身体的なメタファーと呼ぶ概念を紹介し、議論する。

Media Space as a Reconstructed Body

Hiroyuki Miki

Oki Electric Industry Co., Ltd.

Keiichi Yamazaki

Saitama University

Hideaki Kuzuoka

University of Tsukuba

Akiko Yamazaki

Nurse School of Douai Hospital

Hideyuki Suzuki, Hiroshi Kato

NEC Corporation

Although utilization of video as a support for collaborative work has had a success in some practical fields, say, surgery room or plant control, most of such efforts have resulted in a working prototype or in a little success. How can we design a system for remote instruction or remote operation? As a first attempt to answer this question, we used only audio-video connections and built up some systems which had different allocations of cameras and monitors. We conducted experiments on each allocation pattern, and have explored problems of each setting. Out of these efforts, we conceived of a novel concept named 'Body Metaphor'. This paper introduces the 'Body Metaphor' concept and discusses it.

1. 序論

TV会議のようなビデオを媒介にしたコミュニケーションでは、対面的なコミュニケーションにおいて用いられる視線を含めたジェスチャーや他の身体行動は、あまり有効に働かないという問題点が指摘されてきた(Heath and Luff 1991, 1992)。

この問題を解決するために、メディアスペースと呼ばれる、ビデオとコンピュータを用いたコミュニケーションの研究において、これまで様々な研究が行われてきた。VideoDraw(Tang et. al 1990)やClearBoard(石井・小林 1992)といった、協同的なデザインの支援システムは、実験システムではあるが、限られた対象において、上記の問題を解決した。これらの実験システムでは、遠く離れた人があたかも向き合っただ対面的な対話をしているかのような環境を再現した。他方、実用的なシステム(手術室

におけるビデオ映像の利用(Nardi et. al 1993)や、プラント制御におけるビデオ映像の利用(Tani et. al 1992)等)では、対面的な環境を提供しないものもある。

本論文では、ビデオとコンピュータを用いた協同作業の支援システムを、人工呼吸法や作業の指示(Kuzuoka 1994, 1995)のような遠隔指示の医療や教育に適用する場合を考える。この場合は、前述のようなシステムと違い、指示する側か指示される側か、指示者か作業者かといった立場の違い、作業環境の実体と同じ場所にいるか遠隔から参加するかに応じた適切なインターフェースを提供する必要がある。特に、共有する作業対象が3次元的な形状をもち、3次元空間に分散して設置されていて、さらに作業者が空間内を動き回る様な作業を支援できなければならない。この意味で、対面や同一環境、またはモニターがキーワードであった前記のシス

テムとは、対象が異なる。

我々は、アルゴブロックというコンピュータを用いた協同学習の道具を用いて、複数のカメラやモニターを空間的にさまざまに配置し、遠隔から協同作業を支援する実験を繰り返した。

複数のカメラやモニターを用いた点では、我々はMTV (Multiple Target Video) の研究 (Gavor et.al 1993, Heath et. al 1995) の影響を受けている。Gaver や Heath らは、どの映像が重要であるか、人々がどのような映像の映ったモニターを用いたかという点に着目して分析をおこなった。一方、我々は、モニターに表示された身体の一部を作業空間にどのように配置すべきかという点に着目した。これは、Buxton (1992, 1995) の提唱するサロゲート (「代理人」) の考え方に影響され、身体の一部と同じように考えた結果である。

我々の目的とするのは、一つは、人々が、空間にさまざまに配置された身体像を用いて、エスノメソドロジー (Goddwin 1991, 山崎 1994) が詳説しているような緻密な相互行為を達成できるのか、達成できるのならばどのようにしてかを調べることである。二つ目は、緻密な相互行為を達成できるようにする環境を設計するための設計指針を確立することである。そして第3の目的は、そのようなシステムを実際に構築することである。

2. 実験対象と実験概要

2.1. 協同作業とアルゴブロック

アルゴブロック (Suzuki et. al. 1993) は、実体を持つプログラム言語である (図1)。すなわち、物理的実体を持つブロックの一つ一つにプログラム言語のコマンドが割り当ててあり、利用者は、これらのブロックを手で相互に接続することによってプログラムを作ることができる。ブロックとしては、「前進」「回転」「繰り返し」「条件分岐」などが用意されている。ブロックのパラメータ入力スイッチによって、距離、角度、繰り返しの回数を設定で

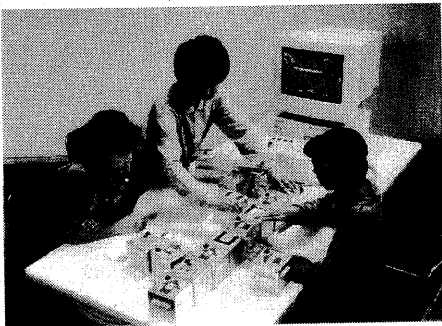


図1 アルゴブロックの使用

きる。現在、CRT 画面上の潜水艦のキャラクタを動かしてゴールまで到達させる子供向けのシステムが実現されている。

通常のアルゴブロックの利用においては、作業者 (作業者) も指示者 (指示者) も同じ部屋にいるが、遠隔指示の環境を作るために、指示者は別の場所から遠隔で指示するようにした。

この研究にアルゴブロックを利用する理由は2つある。1つは、わかりやすさである。ブロックの操作により個々の作業者の考えていることが見えやすく、他の協同作業だけでなく指示者にとっても、作業がどのように進行しているかが分かりやすくなる。2つめの理由は、比較的固定した方向性である。ブロックを接続していく方向 (図1では奥から手前) やモニタの位置が固定している (図1では奥) ために、作業の方向性があまり変化せず、カメラとモニターによる遠隔指示といった位置変化に弱いローテクでも、遠隔から最低限の指示は行える。

2.2. 概要

- 実験：予備実験2回、(本) 実験3回、のべ約30時間 (カメラやモニターの配置を変えるなどの改良を加えながら行った。)
- 被験者：合計14組。複数の指示者と複数の作業者で一組。小学生と大学生が参加 (混在した組は無し)。指示者にはあらかじめアルゴブロックの使い方を教えた。
- 機材 (図2, 図4を参照):
 - 指示者側:
 - 1) カメラ：合計2個
 - a) 指示画面を接写するカメラ:
 - b) 指示者の顔を映すカメラ
 - 2) モニター：指示者の前に合計3台。作業者側のカメラからの音声・映像を移す。
 - a) 左側のモニター：パーソナルコンピュータのディスプレイで、遠隔制御カメラをコントロールするためのユーザ・インタフェースとそのカメラからの映像が表示される。
 - b) 中央のモニター：文脈提供カメラからの映像が表示される。
 - c) 右側のモニター：文脈提供カメラ、遠隔制御カメラ、天井カメラの映像、そして、アルゴブロック画面に表示されているものと同じ画面とをセレクターによって選択できる。この画面は指示カメラによって接写されているため、この画面の上で指示者が手振りを使うと、作業者側の指示画面には手振りが入った映像が表示されることになる。
 - 作業者側：

1) カメラ：合計3台

a) 前方1：リモートコントロールカメラ
遠隔側の指示者の操作によってパン、チルト、ズームを変えられる。

b) 前方2：文脈提供カメラ
アルゴブロックを組み立てる机の上、部屋後方に設置されたモニター、そして作業者を全て撮影できるように、設置された。

c) 天井：天井カメラ
作業者を含めた作業空間全体が撮影できる。

2) モニター：合計3台

a) 指示者の顔画面：

b) アルゴブロックの画面：

c) 指示画面：作業環境の映像に初期の実験ではマウスカーソルがスーパーインポーズされた画面を用いたが、その後改良を加え指示者の手がスーパーインポーズされた画面を用いた。これによって、指示者は手を利用して指さしやその他のジェスチャーを見せることができる。

ここでは次の章で記述される、(本)実験1、2の3種類の配置パターンを解説する。

ア) 作業側第1パターン(実験1)：従来のテレビ会議モデルの対面メタファーから発想したもの。全てのモニターを部屋の前側に、指示側と同一に並べた。

イ) 作業側第2パターン(実験2)：部屋前方に、顔画面とその左側に指示画面を並べて、後方にアルゴブロック画面を配置した。

ウ) 作業側第3パターン(実験2、図2を参照)：部屋前方に顔画面を置き、指示画面とアルゴブロック画面を後方に並べた。

3.実験データの分析

3.1.実験1：指示側と同一の作業側の画面配置

我々はまず、指示側と同一の画面配置である第1パターンについて実験を行った。作業側は、指示者の様々なモニターを見たり、さまざまなモニターに手を伸ばしたりしながら、共同作業やコミュニケーションを行っていた。しかし、作業者がどういう映像が映っているどの画面(顔画面、指示画面、アルゴブロック画面)をみて共同作業やコミュニケーションをしているかが、指示者には簡単には分からなかった(注：GaverやHeathたちのMTVの実験では、作業者が対面カメラの正面で、しかもカメラのすぐ近くにいたために、作業者の指向がわかりやすかった)。このため、指示画面による指示者の指示は、必ずしもうまく行われなかった。

また、指示者は、言葉だけでなく視線の動きやジェスチャーによって作業者に対する指示を行お

うとしていたが、これも困難だった。ジェスチャーの誇張や、言葉による確認作業が頻繁に行われた。こうした一種の補償作業によって、実験そのものは何とか続けられた。だが特に小学生を対象にした実験の場合には、指示者や作業者の一部が途中で協同作業に加わらなくなったりした。また、作業者が指示者の言うことを無視し、自分たちだけで作業を進めてしまう場合もあった。

この実験から我々は、以下のことを認識した。

- 指示者が自分の行動に対してどう評価しているかを作業者が知ろうとする時や、作業者が指示者に対して質問する場合には、指示者の顔を見る場合が多い。それゆえ、指示者の顔を見ることができ、また指示者の顔への作業者の指向が指示者に分かることが重要である。
- 指示者が、まず自分の指示している映像を確認でき、その上で、指示している対象を作業者が見ていることを、指示者が見ることができ、またそのことを指示者の言葉や動作で作業者が確認できることが重要である。またそうした作業が、お互いの空間における身体的な配置と相互行為を通して成し遂げられる。

3.2.身体メタファーという概念

上記の認識は、実は通常のインストラクションの場面では、あたりまえのことである。しかしこれがパターン1でうまくいかないのは、指示側と同一にした作業側の画面配置や、カメラの配置が主たる問題であった。ここで我々は、Buxton (1992, 1995)のサロゲート(「代理人」)の概念を思い起こした。Buxtonらはカメラ、スピーカー、マイクを目、口、耳の「代理人」として考え、さらに social protocol をメディアスペース設計に適用することにより、顔

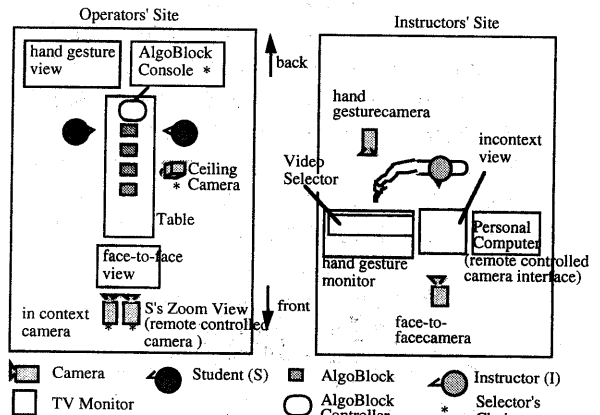


図2：作業側第3パターン時の全体の配置図

画面のためのモニターの配置方法に関するガイドラインを提案した。遠隔指示という我々の対象では、これをそのままは当てはめられないが、ほぼ同様の路線を取ることが出来るのは想像に難くなかった。我々は、まず、Buxtonの「代理人」(目、口、耳、等)に加えて、指示画面を指示者の手の「代理」であると考えることによって、手や作業空間、そして、作業者の位置関係をも考慮したシステムの配置を、身体メタファーと呼び、これを基本的な設計指針とすることにした。指示に置く身体メタファー(図3)とは、例えば、指示者の指先が指示者と作業者の目の前方に配置されていること、作業者が指示者の視野内に配置されていること、さらに、作業者が指先を見ているのか指示者の顔を見ているのか分かる程度に離れて指と顔が配置されていること、等を最低限要求する。

実験は、この身体メタファーに即しているパターン3のシステムを組み立てることから始まった。この時、我々が考慮したのは次の二つの点である。

(1) 文脈提供カメラによって、指示者が、作業側側の作業空間の文脈の中で、指示画面を見えるようにする。これによって、作業者が指示画面に対してどのように指向しているか、また自分の指さしによる指示に対して作業者がどのように応答しているかが指示者に見えるようになる。

(2) 指示者の顔画像を指示画面やアルゴブロックの画面と反対の位置におく。また、文脈提供カメラを顔画像の方向から撮り、指示者の声を顔画像からだすようにする。これによって作業者が指示画面を見ているか、指示者の顔画像を見ているかが指示者側に分かるようになる。

我々は、身体メタファーに即しているこうしたパターン3の特徴と限界を調べるために、パターン2と比較するという実験を行った。パターン3のシステムとパターン2のシステムの主な違いは、指示者が、指示画面を相手の作業空間のコンテキストで見えるかどうか、指示画面と顔画像に対する作業者の指向を明確に区別することができるかどうかである。なお、パターン2の指示を身体メタファーで考えると、顔を下に向けないまま、指示者が右脇の下の画面を、左手で指さすか、あるいは右手で手首を

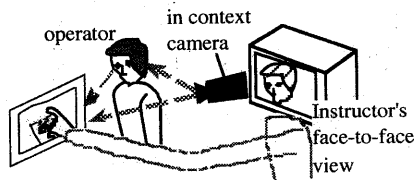


図3: 指示における身体メタファーの概念

返して指さすような配置になっている。

3.3.実験2: パターン2, パターン3の比較

実験は、指示者1名、作業者2名で4組行った。被験者は大学生で、1時間の実験を、それぞれパターン2, パターン3という配置を変えながら行った。

この実験において、次のことが確認された。

(1) 作業者が作業側の指示画面を見ながら作業をしていることを、指示者は文脈提供画面において確認しながら指示を行っていた(図4)。指示者が特に細かい指示を行うときには、指示者は指示者側の指示画面と文脈提供画面を交互に見ていた。そして指示者は作業者が指示者の指先を見ていることを確認しながら指示作業を行っていた(テーブル1)。

指示者が指示を行う前に、指示者は作業者が指示者の指先が映る指示画面を見る体勢をしているかどうかを確認していた。そのような体勢ができていないときは、'この'という指示代名詞を使って作業側側の注意を喚起して、指示画面を見るようにさせた。また、指示者が指示を行っているときに、作業者が指示画面をみていないことがモニターされたときは指示を途中でやめた。

要約すると、このような配置(パターン3)では、作業者が指示画面を見るということは、受け手性(Goodwin 1981)の表示や、自分が聞き手であることを示すことになっている。指示者は作業者が受け手性を示すことを確認しながら指示を行える。

これに対してパターン2の配置では、パターン1と違いアルゴブロックの画面は文脈提供画面で確認できるが、指示画面と顔画面は文脈提供画面で確認できない。また、指示画面と顔画面は同じ側にある両者の区別は付きにくい。結果として、実験1の問題と同様の現象が依然として確認された。だが原因は少なくなり、指示画面を文脈提供画面で見れないことによって、自分の手の動きが自分で確認出来ないことが最大の問題だった。相手がどう見ているかよりも、これでまずつまづいた。特に実験が始まってすぐのときに顕著であり、指示者が、「手見えています?」「手映っていますか?」という言葉をとえ作業者が「見えています」と言葉で確認した後も何度も繰り返している例が象徴的だった。

(2) パターン3では、指示者の指先の動きと、指示者の指先が画面に置かれ続けていることが、相互的な行為の資源として使われている。トランスクリプト(テーブル1)で示されたように、作業者が指示画面を見ながら「はい」と言った時に、指示者は指先を離している。そして別の対象を指さした。

つまり指示者の指先の動きは、指示者が作業者の行動を理解したことを示しているのである。また、

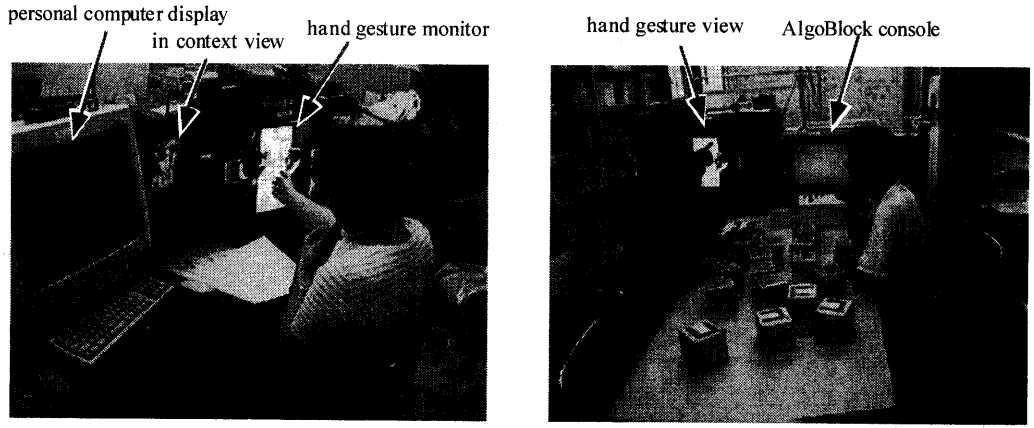


図4 パターン3の実験の様子(左:指示者側, 右:作業側(文脈提供画面の映像より))

この指示者の指先の動きは、指示者が作業者の行動をモニターしていることを示す相互的な道具にもなっている。

このようなことは、指示者の顔画面と指示画面の配置や指示者の応答とあわせて、パターン3においては作業者が作業者の行動が指示者にみられていることをモニターする重要な資源となっている。作業者は指示者の指先を示す画面である指示画面をみることで、作業者が指示者によってモニターされていることをモニターすることができるのである。

一方、パターン2においては、指示者が指先をうごかす速度が作業者にとって早すぎたりゆっくりすぎたりした。指示者は、作業者が指示者の指先をみているかどうかモニターできなかった。すなわち、「指示画面と顔画像に対する作業者の指向を明確に区別することができない。」パターン2のレイアウトの結果と言える。

自分の手の動きが見えないという(1)で指摘した問題と合わせて、パターン2のレイアウトは指示には不適切と思われる。(2)での問題だけならば、例えば指示画面を前方右端に、その他の画面やカメ

ラを前方左端に設置して、指示画面(身体メタファアの「手」にあたる)を前方右端に設置して両者のできるだけ(身体メタファアの「顔(頭)」から)離せば、作業者による指示者の「手」への指向と「顔」への指向は区別できるだろう。(1)と(2)の両方の問題を解決するには、例えば、少なくとも指示画面が指示者の「顔」から見えるように少し後方に下げてさらに90度くらい回転させるといった事が考えられるだろう。

(3)パターン3においては、作業者が質問をする時には指示者の顔画面をみる事が多かった。またアルゴブロックの操作が成功して終わったときには、指示者の反応や評価をモニターするために、指示者の顔画面を振り返ってみることもしばしばあった。パターン2においては、顔画面が同じ側にあることによる問題は、顔画面が作業方向(アルゴブロックの画面のある方向)とは逆に位置するために比較的少なかった。また、指示者からの指示を受ける必要がなくなった状態(作業者だけで課題を解ける)になると、パターン3との違いは特になかった。

テーブルI パターン3の実験からのトランスクリプト

	P-----	P-----	-----			
I2	HM-----	C-----	HM-----	C-----	HM-----	C-----
	tunagerutokiwa	konoblockwo	konodengenwo	kitte kirankya	ikenain	desuyone
O5	AL-----	H-----		-----		
		hai				
O6	H-----	-----				

注:このトランスクリプトでは、楽譜のように全体の進行が左から右に時間的にすすんでいる。この場面には、作業者側にO5とO6と記号化した二人の作業者がいて、指示者側にはI2がいる。会話の上の行には視線の方向を示している。---はそれが継続していることを示している。視線を示す行での、記号は次のようなものである。ALとHは作業者側のAlgo Blockと指示画面であり、HGとCは指示画面と文脈提供画面を示す。さらにPは指示画面に指示者が指をおいたことを示している。Pの連続は、指先が継続して置かれていることを示す。

作業者の作業にとっては、総じて言えば、パターン3の方が良い。作業中に向くのはほとんどアルゴ画面の方向であり、この方が指示者の唐突なアドバイスや指示を見損なうことが減る。また、作業中で振り返れないときでも、質問に対する答えの指示は、顔画面方向を振り返らなくても見られる。

しかしパターン2が必ずしも悪いわけではない。パターン2だと、質疑応答のターンがはっきり切れる時が多い。作業者が、質問を作るときはアルゴブロックの画面やアルゴブロック等を見ているが、指示者の回答を見るには指示画面を見るために振り返る必要があるからである。これは、特に一つ一つの発話が確実に相手に聞こえることが重要なタスクの場合はメリットとなるだろう。

4. 議論

パターン3で十分なのか、言い換えれば、視線やジェスチャーや指示はこれで良いのかについて、もう少し議論してみたい。例えば一見便利に見える指差したが、便利さゆえに、インタラクションの中でこの資源の重要性が高まり、指差しできない画面でも指さしをしてしまう。画面やモニターの位置だけでは解決できない、場の相互的な作られ方といった課題がそこにはある。また、作業者は複数だが、指示者が誰を見ているかがわかるかまでは求めている（これは難しい。C.f. 森川（1996））。

さらには、身体メタファーが指示者にしか提供されておらず、作業者には提供されていないという事もある。このために、作業者は指示者の志向の構えを認識できず、例えば、自分達が身を乗り出して凝視している対象を、教師が同じような大きさで同じように見ていることを仮定し、よもや教師が見ていないとは思わずに、「これってこの向きでいいの?」とか聞いて、指示者が「えっ? どれ?」と言って慌ててセレクタを切り換えたりリモートコントロールカメラをズームアップしたりする場面が見受けられた。Heath等（1991）の言う、Communicative Asymmetryの問題の典型的な例である。これは、利用可能な資源の非対象性という問題である。例えば自分が行うジェスチャーは、遠隔でも、相手が自然な環境と同じように見えていると想定してしまう錯覚である。緻密な分析が必要な大きな問題である。例えば前述の例では、文脈提供カメラを指示者側（の後方）にも設置し、作業者に指示者のコンテキスト（指示者がどの画面をいかに見ているか）を提供することも出来る。しかし、作業者の作業中は指示者は観察に専念することが出来るのに対して、作業者は指示者の観察よりも作業の遂行に主たる注意が向いている。難しいだろう。そもそも、kendon（1991）やエスノメソドロジ

ストが指摘しているように、視線やジェスチャーは、対話開始の機能だけでなく対話中にも重要な機能を果たす。被指示者の参加を得ることだけでなく、その維持、修復、終了等に対して重要な役割を担う。ジェスチャーの認識には基本的に周辺知覚が重要という指摘がある（Heath et.al 1991）が、ビデオを介したコミュニケーションにおいては、視線やジェスチャーを周辺知覚で伝えることはきわめて困難である事が指摘されている（森川 1996）。

これまで議論してきたように、指針としての身体メタファーは、それほど多くの問題を解決出来るわけではない。しかしBuxtonが指摘しているように、作成とテストの繰り返し設計がメディアスペースの設計には重要である。まず、身体メタファーに従ったプロトタイプで実験をしてデータを集め、その上で特に強化したい部分を改善するという段階を取った方が良いのではないだろうか。身体メタファーを使って考えれば、例えばパターン2のように身体メタファーを考えると変な物については、問題がある程度予測できる。また、身体メタファーは、遠隔指示システムの使い勝手の評価にも指針として利用出来るだろう。目の機能、顔の機能、手の機能、胴体の機能、あるいはこれらの組み合わせが、自然と行えるかどうかが出発点となる。

我々は今後、緻密な相互行為を可能とする設計に向けて、葛岡（1994,1995）が提唱しているようなテクノロジーを導入して、この遠隔指示システムを改良していく予定である。

References

- Buxton, W. (1992): "Telepresence", *Graphics Interface'92*.
- Buxton, W. (1995): "Ubiquitous Video", in *Nikkei Electronics*, no. 632, 1995.3.27.
- Gaver, W., Sellen, A., Heath, C, and Luff, P. (1993). "One is not enough", *INTERCHI'93*.
- Goodwin, C. (1981): *Conversational Organization*, Academic Press.
- Heath, C., & Luff, P. (1991): "Disembodied Conduct", *CHI'91*.
- Heath, C., Luff, P., and Sellen, A. (1995): "Reconsidering the Virtual Workplace", *ECSCW'95*.
- Ishii, H. & Kobayashi, M. (1992): "ClearBoardt", *CHI'92*.
- Kendon, A. (1990): *Conducting Interaction*, Cambridge Univ. Press.
- Kuzuoka, H., et al (1994): "GestureCam", *CSCW'94*.
- Kuzuoka, H., et al (1995): "Can the GestureCam be a surrogate?", *ECSCW'95*.
- Miki, H. (1995): "Cultural Differences in an Assembling Task", *IPSI SIG-GW Vol.95, No.38*.
- Morikawa, M. (1996): "Difference of Gaze Perceptions between on face-to-face and video communication", 12th Symposium on Human Interface
- Nardi, B.A., et al (1993): "Turning Away from Talking Heads", *INTERCHI'93*.
- Suzuki, H. and Kato, H. (1995): "Interaction-Level Support for Collaborative Learning: AlgoBlock", *CSCW'95*.
- Tani, M., et al (1992): "Object-oriented video", *CHI'92*.
- Tang, J., and Minneman, S. (1990): "VideoDraw", *CHI'90*.
- Yamazaki, K. (1994): *The Pitfall of Beautiful Face*, Harvest-sha.