

超鏡対話における指差し行為

森川 治、山下樹里、福井幸男、佐藤 滋

生命工学工業技術研究所 {morikawa,juli,fukui,sato}@nibh.go.jp

概要 身振りなどの非言語情報の多くは、発話者と周辺事物との相対位置関係が重要な意味を持つことが多い。対話システム「超鏡」では、画面上の事物なら、指差しながら対話を行うことが可能である。本論文では、超鏡対話における指差しが利用者に受け入れられる事を確認する実験を行った。4つの指示課題を与え、自主的に超鏡指示を学習・習得する過程を観測した。その結果、4番目の課題では93%の利用者が、超鏡指差しを使用して課題遂行したことが確認された。

Pointing on HyperMirror video mediated communication

Osamu Morikawa, Juli Yamashita, Yukio Fukui, Sigeru Sato
National Institute of Bioscience and Human-Technology,

A lot of nonverbal information are understood by the relative positional relationship of peripheral objects in face-to-face conversation. In our HyperMirror system all participants and all objects have relative positional relationships of each other regardless of physical position. They can have a conversation with the relationship, for example, pointing an object on the HyperMirror screen. But if it is difficult for them to learn and use the action of pointing, they will not use the action. In this paper we confirm that people accept and use the action of HyperMirror pointing by experiment. In the experiment, we give to subjects the task that they instruct their partner to move predefined four objects using the HyperMirror system, and observe the process of their studying and accepting the HyperMirror pointing by themselves. We find that 93% of subjects execute the task by using the HyperMirror pointing at the forth object.

1、はじめに

一般にビデオ対話環境では、CRTやスクリーン上の対話相手と対話をする。その画面には、対話相手の映像だけでなく、対話相手のいる周辺の様子も同時に表示されている。同様に、相手のディスプレイには、自分の映像と周辺の様子が表示されている。しかし、両者のいる空間が異なることから、ディスプレイを境に自分の世界と相手の世界が分離されている。

一方、人間どうしのコミュニケーションでは、言葉だけでなく、身振り等の非言語情報を有効に使用している。一説によればメッセージの6割から9割をこれが担っていると言われている[1]。これらの非言語情報には、単独で意味を持つ場合と、周辺の事物との相対的位置関係から意味が確定する場合がある。例えば、壁にかかっている時計を見るという

ジェスチャーは、時計と話者の位置関係が理解されて初めて意味を持つ非言語情報の典型例である(図1)。このような後者の非言語情報は、従来のビデオ対話環境では世界が分離されているため正しく伝わらないことがある。しかし従来のビデオ対話システムでは、音声と映像が対面对話と酷似した形で提示され

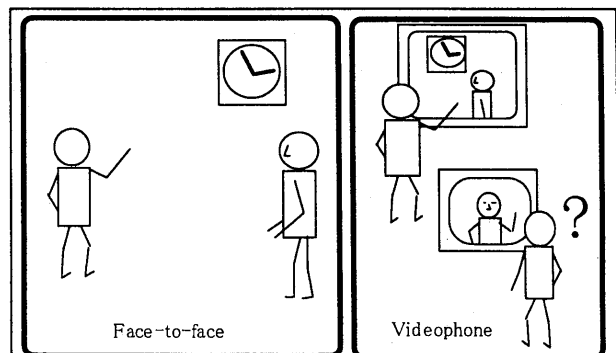


図1、ビデオ対話環境では、対話の世界が分離されているため伝わらない非言語情報がある。



図2、超鏡対話例

る。我々はこれらの情報の変形に気が付かず、対面对話と同じつもりで臨むことになる。その結果、ビデオ対話は使いにくい、疲れる、対話相手は見えるが遠い存在に感じる、といった感想を利用者に引き起こす結果になる。

ビデオ対話システムの多くは、これらの差異を減少することにより、対面と同じに対話できる事を目指して研究・改良されてきた。しかし本研究では逆に、ビデオ対話システムにより画像情報が変形を受けている事を明示的に利用者に示す事を考える。これにより利用者は、ビデオ対話システムの画像情報を、対面对話と同じ様に解釈する事を防ぐのである。すなわち、自己像を表示する新しいビデオ対話方式「超鏡：HyperMirror」を提案する[2][3][4][5]。

2、同室感覚のある超鏡システム

超鏡対話では、対話者全員が同一の画像を用いて対話を行う(図2)。対話は2地点に限らず、何地点でも構わない。利用者はシステムの前に立つだけで、装置の装着や操作を必要とせずに対話を行うことができる。対話に利用する映像は、各地にいる対話者全員があたかも同室にいる様子を鏡に写したような内容である。そこには自己像が含まれ、自己の近傍の事物と、対話相手の近傍の事物が混在している。対話者および近傍の事物は物理的に同室であるか否かに係わらず、画面上では等質となって表示される。対話者間を分断する壁は存在しない。全利用者はシステムの前を自由に動くことができ、対話相手との位置関係を自由に変えることができる。例えば画面上で、各地にいる任意の対話者のすぐ横に並ぶことができる。画面上でその人の肩をた



図3、机上型の超鏡対話システムを使って、リモコンの操作方法を教示している様子。

たいたり、握手をする事もできる。相手の持ち物を指差すこともできる(図3)。そのため、対話者全員が同じ場所にいるような感じが得られる。

超鏡対話には、映像の中で対話をしている「演者の自己」と、その映像を見ている「観客の自己」が共存する。対話は、同じ映像を見ている観客兼演者の対話相手と行う。このように超鏡は、対面对話および従来のビデオ対話とは全く異なる対話形式である。しかし超鏡のように演者と観客の両者が融合する場面が、他に無い訳ではない。実際、日常生活において鏡を見ている状況が正に、この両者が融合している典型例である。

2.1 話し手の視点と聞き手の視点の違い

実空間では話し手と聞き手は同じ事物を見て対話できる。しかし、お互いの視点が異なる為、厳密には、同じ対象物を見ているも同じ視覚情報に接している訳ではない。そこには、「自分の見ている事物を、相手も別の視点から見ている：図4左」状態と「相手の見ている事物を、自分も別の視点から見ている：図4右」状態があり、これらに対話進行状況に合わせて随時、切り替えている。

前者の場合、話し手は自分の視点で発話する。話し手を中心に事物は配置される。聞き手は傍観者になる。一方、後者の場合、聞き手の視点で事物が配置される。聞き手がメインになり、話し手は聞き手の立場に立った行動を強いられる。すなわち、聞き手の視点を推定し、その推定に基づいて発話する必要がある。自分が見ている視覚情報を、相手の受け取る視覚情報に翻訳して理解する必要があ

る。さらに従来のビデオ対話の場合には、翻訳する前の材料となる視覚情報自身を推測する必要がある。

超鏡対話では、全員が同じ視覚情報で対話するため、常に「観客の自己」の視点で対話を進められる。「観客の自己」に分かりやすいように行動すれば、結果的に対話相手にも分かりやすい行動となる。

すなわち、超鏡対話空間を「相手の見ている事物を自分も見ている (WYSIWIS :What you see is what I see)」を満たす空間と解釈するより、むしろ「自分の見ている事物を相手も見ている (WISIWYS: What I see is what you see)」を満たしていると考えた方がふさわしい。

2.2 超鏡対話の特徴

文献[4]で述べたように、超鏡システムを使用した対話を観察した結果、興味深い行動がいくつか観測されている。

- 同室にいるという社会性の発現：画面上で対話相手の前を横切りお互いの位置を交代する場合、「前を失礼」といいながら、少し小走りに、腰をかがめながら移動する事例。類似の現象は、対話者のシルエットを表示する Krueger の VideoPlace[6]でも観測されている。

- 身体による視線行動：対面对話ならば視線を話し手に向ける場面で、身体を画面上で相手の方に向け、画面上の相手を見つめる事例。

- 均一性：物理的に同室にいる相手と話す場合にも、実空間ではなく、超鏡空間で話す事例。

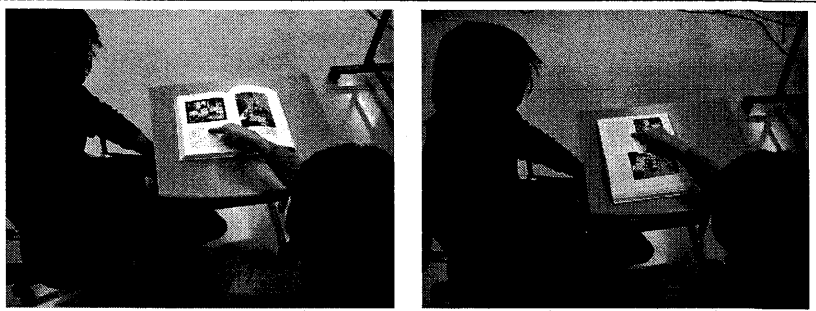


図4、実空間における事物配置の相違点。左では、「話し手の視点」で本が配置され、右では、「聞き手の視点」で本が配置されている。

- パーソナルスペース：超鏡画面上で対話相手と適切な間隔を保とうとする事例。

2.3 超鏡システムのハードウェア

超鏡システムの役割は単純である。各地のカメラで撮影された映像を合成して1枚の超鏡映像を作成することである。合成画像の作成方法には、光学的に行う方法(図5)と電子的に行う方法(図6)がある。光学的に行う場合には、2重写しが原則である。電子的におこなう場合には、例えば、位置センサーを用い、重なり部分の前後関係を自動判別し、優先順位を決定することにより、より自然な映像を得ることも可能である[7]。

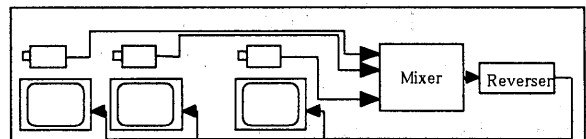


図6、電子的合成による超鏡システム結線例

我々が主に使用している超鏡システムは、2地点、クロマキー合成方法である(図7)。ビデオ信号はNTSCである。クロマキー合成をするために、一方の地点の壁1面をブル

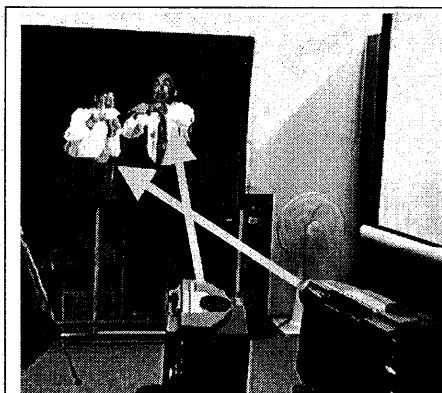


図5、光学的合成例

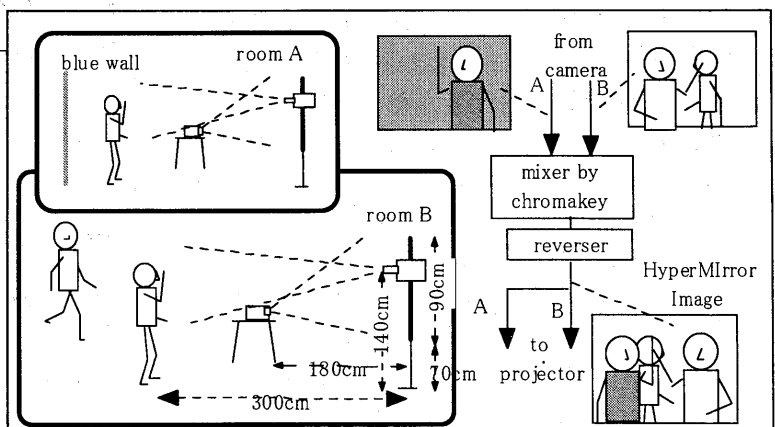


図7、クロマキー合成による超鏡システム例

バックにした。各地にはビデオカメラ、スピーカ、液晶プロジェクタによる床上 70cm、横 120cm 高さ 90cm のスクリーンがある。立って対話をする状況を想定し、利用者はスクリーン、カメラから 3 m 離れた位置に立った場合、ひざから上が撮影範囲になるように調整してある。

初期の超鏡システムでは、すべての利用者が鏡映像になっているため、地図や文字を使った対話等では不都合な場合もあった。そこで、対話地点が 2 地点に制約される代わりに、自己像は鏡映像、対話相手は正像という超鏡-II も作成した (図 8)。

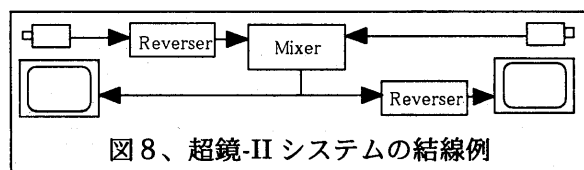


図 8、超鏡-II システムの結線例

3、指差し行為

対話において利用する非言語情報には、前述の「壁の時計を見る」様な消極的なジェスチャーだけでなく、周囲の事物を指差しながら話をする様な積極的なジェスチャーもある。以下、対話環境ごとの指差し行為について考察する。

3.1 対面対話の指差し行為

対面対話における指示行為は、話し手の視点で行われる。すなわち話し手は、近くにある場合には、近寄って行って、直接触って指差しをし、手に届かない場合には、事物を見つめる視線上に指先を置いて指差しをする。後者の場合、聞き手は、話し手の指先だけでなく、視線も理解する必要があり、その直線上から事物を検索する必要がある(図 9)。

3.2 ビデオ対話における指差し行為

従来のビデオ対話システムでは、指差せる事物は、話し手の周囲の事物だけである。聞き手の周囲の事物は、見えていても指差せない。従って、対話者が共通で指差しできる事物は存在しない。この性質は、ビデオ対話での対話内容を制約する一因になっている。

さらに指差せる事物であっても、対面対話における指差しを、そのままビデオ対話に適用すると多くの場合、失敗に終わる(図 9 中)。

それは、ビデオ対話では、対話相手のモニタに何が表示され、何が表示されていないかを知ることが困難だからである。指差している対象事物が相手のモニタの視野外にある場合は当然失敗する。対象事物がモニタに写っている場合でも、直接触らない限り確実に伝わらない。離れて指示している場合には、聞き手は対面対話の時と同様に、話し手の指先だけでなく、視線も理解する必要があり、その直線上から事物を検索する必要がある。しかし、ビデオ映像は平面であり奥行き情報が欠落し、画面の詳細さも実物より劣るために、聞き手が指示事物を同定するのは困難である。

3.3 超鏡対話の指差し行為

超鏡システムでは、WISIWYS のため、対面対話とは異なる、全く新しい指差し行為が可能になる。超鏡対話における新しい指差し行為は、話し手の「演者の視点」ではなく、「観客の視点」で行われる。超鏡画面上で、指差し対象事物と話し手の指を重ね合わせることで、指差しを行う(図 3、10)。これを超鏡指差しと呼ぶ。

画面上にある事物は、その事物の物理的位

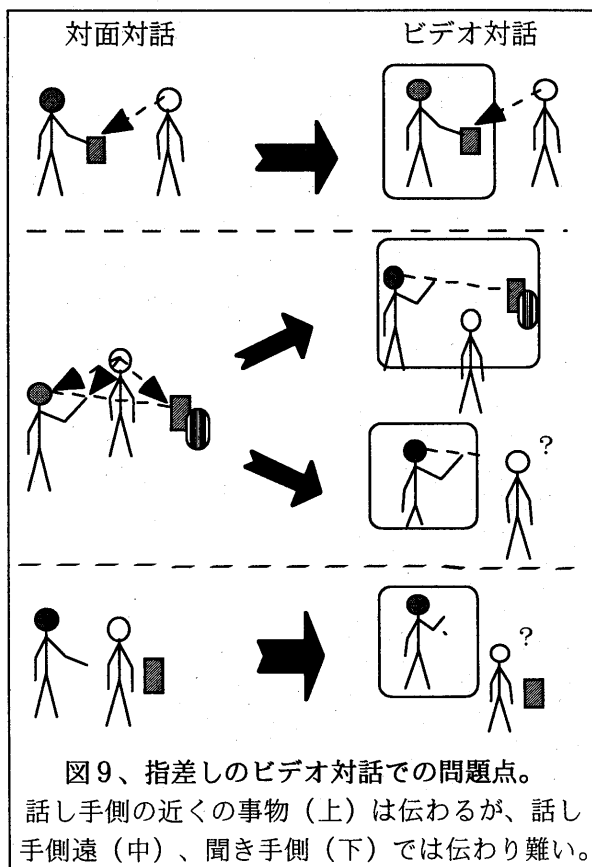


図 9、指差しのビデオ対話での問題点。

話し手側の近くの事物（上）は伝わるが、話し手側遠（中）、聞き手側（下）では伝わり難い。

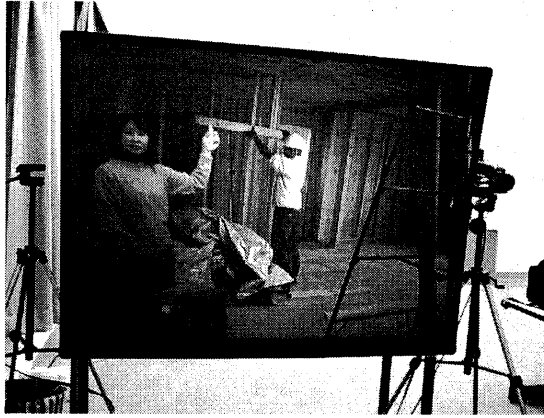
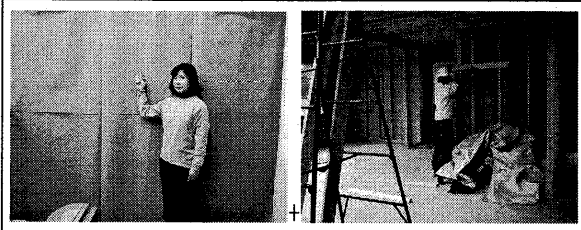


図10、超鏡指差し例。

置によらず、対話者の誰でもが指差すことができる。さらに遠くの山のように、実世界では手に届かない事物であっても、画面上に表示されていれば、指差すことができる。画面

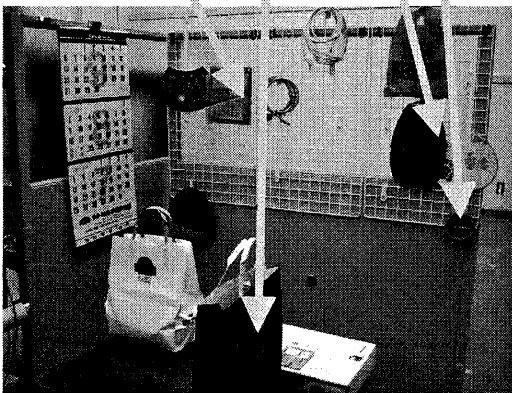
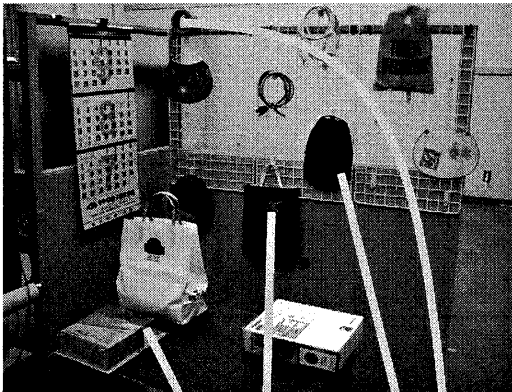


図11、課題例

上で指差しが成功すれば、WISIWYS のため、全員が誤解なく、指差しの事物を直接的に同定できる。ここでは、話し手の目の位置の理解も、目の位置と指先の作る直線の検索も必要ない。

4、実験

超鏡指差しを用いることにより、形式は対面対話と異なるが、対面対話と同等の積極的なジェスチャーを含んだ対話が可能になる。しかし、この超鏡指差しが利用者に学習しやすく、分かりやすく、使いやすいものでなければ有用とは言えない。そこで、超鏡指差しが利用者に受け入れられる事を確認する実験を行った。なお本実験の被験者には、他の被験者の指示行為を観測したり、超鏡対話で対話を楽しんでいる人の様子を観測することを許した。

超鏡システムを使って3次元事物の移動の遠隔指示課題を被験者に課した。実験に使用した超鏡システムは、2地点、クロマキーによる合成法のシステムである(図7)。実験は2名一組で行った。ブルーバック側の被験者は、紙に書かれた配置換えの指示書を見る(図11)。その内容に従った配置換えの指示を超鏡画面越しに、実作業役の相手に伝える。制限時間は特に設けないが、被験者には、正確に、できるだけ速く課題を終了するように求めた。一回の指示課題には4つの事物の移動を課した。指示方法については、言葉だけでも良く、ジェスチャーは使っても使わなくても良く、指示は一方通行ではなく、実作業役の被験者からの質問なども構わない事を教示した。どのような方法で指示するかは、被験者に任せた。被験者の年齢は10才代から50才代までの28組56名である。

5、観察結果

多くの人が、超鏡指差しを使うと予想した。実際、ほとんどの被験者が超鏡指差しを使って課題遂行をした。しかし当然の事ながら、被験者が初めから超鏡指差しを使った訳ではない。

5.1 観測された指示行為の種類

観測された指示行為は、以下の5種類に分類できた(図12)。

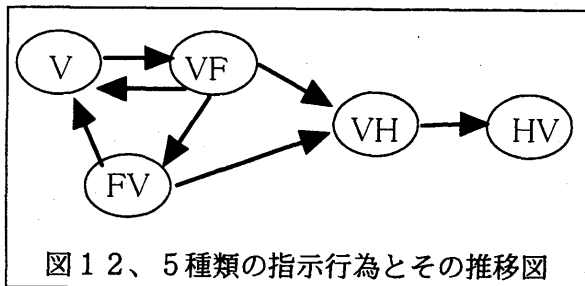


図 1 2、5 種類の指示行為とその推移図

- V: 言葉だけで指示する
 - VF: 言葉が主で対面対話用の指差しを伴う
 - VH: 言葉が主で、超鏡指差しを伴う
 - HV: 超鏡指差しが主で、言葉が伴う
 - FV: 対面対話用の指差しが主で言葉が伴う
- これらの指示行為間には推移関係があり、1 回の指示行為の中にも推移が見られた。

言葉だけの指示 V からは、指示しているうちに、思わず手が動いてしまい、対面対話用の指差し VF に移行する。そして、うまく指示が伝われば、再び手は下ろされて言葉だけの指示 V に戻る。このときにうまく指示が伝わらないと、ある人は、(対面対話用の) 指差しが主になって、説明を続ける(FV)。しかし、FV では指示がうまく伝わらずに、再び V に移行する。

また、ある人は、指示 VF、FV において、指差しが超鏡対話で正しく伝わらない事に気が付くと、超鏡指差しに変更する(VH)。超鏡指差しが成功すると、次第に言葉が「あれ」、「これ」といった指示代名詞が主になってきて超鏡指差しが主になってくる (HV)。

4 つの課題により、指示行為がどのように

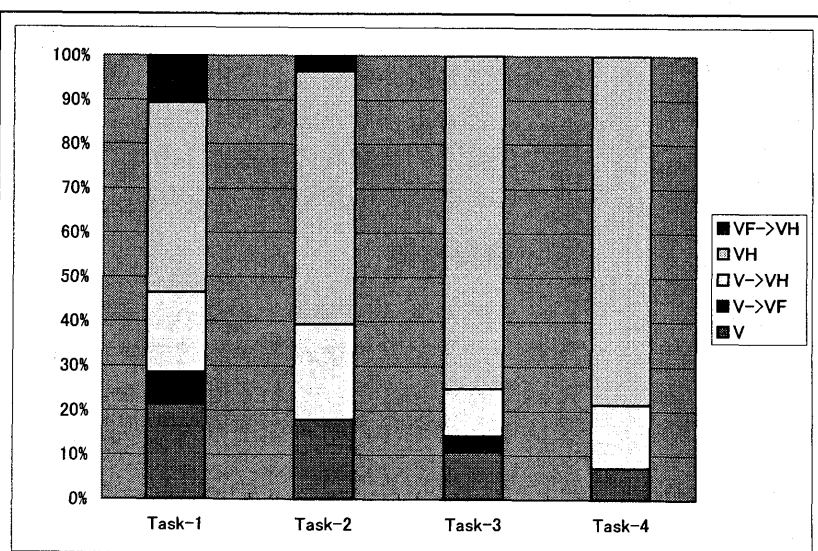


図 1 3、課題遂行と指示行為の推移

変化したかを図 1 3 に示す。1 回目の課題では約 30% の被験者が、言葉による指示を行っている。4 回目の課題では、93% の被験者が超鏡指差しを習得し、それにより指示を行っているが、7% の被験者は言葉による指示に留まっている。

5.2 「あれ」から「これ」へ

興味深いことに、超鏡指差しが成功すると、事物を指すときの言葉が、近くのものでも、遠くのものでも、すべて「これ」に変化した。この事は、画面上に見えている物はすべて手で触れることができる身近な存在に(少なくとも心理的には) 変化したことを意味する。

5.3 「言葉による指示」から「指差し」へ

言葉で説明した方が、誤解が少なく、簡便な場合と、直接指す方が簡便な場合があり、これらの境目は、利用者によって異なる。一般に、大人と子供を比較すると、大人の方が事物を指定する場合、言葉による指示が多く使われた。また、事物の特徴が曖昧になり捕らえにくくなるに従って、言葉より超鏡画面上での指示行為に移行する傾向があった。

個別には例外は有るものの、概して 10 才台の被験者は、超鏡指差しがメインで「これ」「ここ」「違う」「そうそう」という単語レベルでの発話で課題が遂行されていた。20 才台の利用者も、超鏡指差しがメインで、「この赤いもの」「この帽子」という様に、指差しを補助する言葉を使うことが多かった。30 才、40 才台と年齢が上がるに従って、「この帽子を移動させて下さい」といながら超鏡指差しをするように、言葉での情報量の割合が増えてきた。

6、考察

多くの利用者が、超鏡指差しを自主的に習得し、利用することが観測された。そこで、超鏡指差しの習得過程について考察する。

超鏡指差しの習得は他人の対話を観測することによっても引き起こされる。実際、40% の被験者は、課題開始直後から超鏡指差し

だけで課題を遂行した。

残りの被験者の観察から、53%の被験者は対面の指差しを伴う指示(VF、FV)において、指差しが失敗することにより、超鏡指差しを学習することが分かった。失敗は新しい方法を学ぶ強い動機付け状況であるが、この状況でどのようにして超鏡指差しが、被験者に「発見」され学習されたかは興味ある現象である。

前述の通り超鏡は、自己像を表示した対話システムであり、対話者全員が同じ映像を見て(WISIWYS)対話を行う。失敗した時点で超鏡のWISIWYSという性質により、失敗の原因、すなわち、なぜ相手が自分の指差しが理解できないのかを知ることができる。そして、観客の視点に立った指差し、すなわち超鏡指差しを試行錯誤しながら習得する。図14にその様子を示す。観測によれば、この学習には2、3秒しか費やしていない。すなわち、学習の容易性が示されたと言える。

一方、苦勞せずに簡単に学習した新しい知識は簡単に忘れ去られて使われないことも多い。実際、ある課題で超鏡指差しが学習・習得されたはずの被験者が、その後の課題で再度、言葉だけの指示や対面対話用の指差しを伴った指示で課題を遂行しようとする場面が多く観測された。しかし、そのような場合

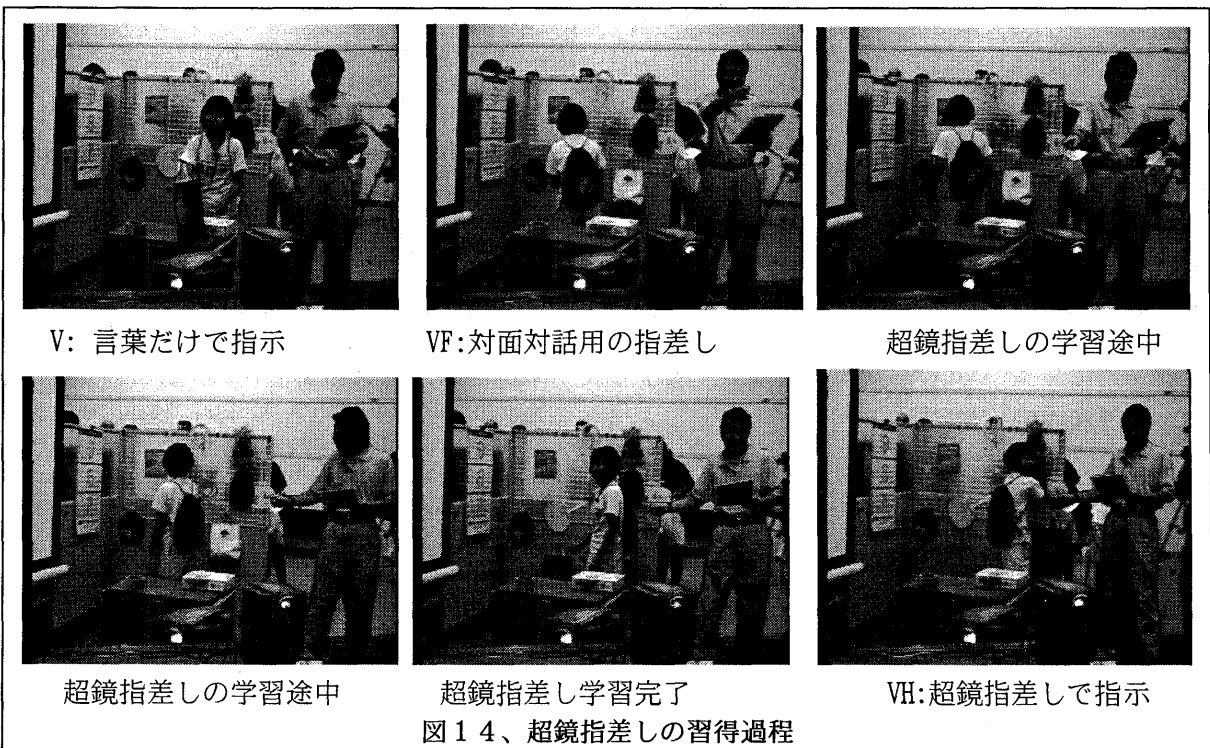
でも、不都合が表面化するとたんに、WISIWYSにより、超鏡指差しが再学習され、課題遂行されていた。すなわち、超鏡対話環境には、利用者を超鏡指差しを自主的に学習させる力を持っていると言える。

なお前述の通り、7%の被験者は、VおよびFVで失敗をしなかった為、最後まで言葉だけの指示で課題を遂行した。彼らは言葉による説明がうまく、言葉だけでも問題無く課題を遂行する事ができた。そこで実験終了後、彼らに対し、超鏡指差しを教示し、超鏡指差しの困難さについて質問をした。その結果、全員が問題なく超鏡指差しを使うことができ、「言葉で指示するより簡単そうだ」という感想も得た。

7、関連研究

指示行為を有効にするには、指示対象に双方がアクセスできることが必要である。

Tangらは、描画キャンパスを使った共同描画に限定し、テレビ電話の映像とは別に、描画するキャンパスと描いている両者の手のビデオ映像を加えたVideoDraw[8]システムを開発した。描画キャンパスは双方が話し手の視点となれるWISIWYSであるが、相手の身体は上半身と手が分離し、相手の手だけが自分



の側から生えている構図になっている。

葛岡らの Agora[9]では、机の上の3次元の事物を対話の共有資源とし、机の上には相手方の机上の映像が上面より投影される。これにより、遠隔地の机上の書類や、それを操作する対話相手の手振りが伝わる。人物の上半身と手の映像の整合性は良いが、対面対話と同様に対話者ごとに視点の差異が存在する。

Tang らの VideoWhireBoard[10]では、顔画像の表示をあきらめる代わりに、WhiteBoard とその WhiteBoard に書き込みをしている人の姿のシルエットをそこにダブらせて表示する。これにより、人物と描画内容の整合性をとり、1つのディスプレイだけで対話することができ **WISIWYS** を確保し、表示のシームレス化を実現している。また、対話相手の注目点などをシルエットにより知る事ができる。さらに、石井の ClearBoard [11]は、この WhiteBoard を透明にして、書き手の表情なども読みとれるように工夫してある。

しかしいずれの対話方式でも、対話相手は、両者を物理的に分割するモニタ、スクリーン、WhiteBoard や ClearBoard の向こうに存在する。お互いの位置は依然、固定されたままであり、作業内容もシステムの近傍に限られる。共有できる空間は、作業対象の2次元のキャンバス、あるいは両者間にある机に限られる。

8、おわりに

本論文では、自己像を表示する新しいビデオ対話システム「超鏡」における指差し行為について述べた。

実験により、超鏡対話における指示行為が利用者に実際に受け入れられる事を確かめた。実験では4つの指示課題を与えたところ、課題ごとに超鏡指示の割合が増加し、4番目の課題では93%の利用者が、超鏡指差しを使用して課題遂行した。

しかし、最後まで、超鏡指差しを学習できなかった利用者も7%いた。学習できなかった利用者は、言葉の説明がうまく、そのため、言葉だけでも問題なく、作業の指示が伝わった。すなわち、超鏡指差しを学習するには、言葉だけの指示が失敗に終わる事が必要であったと言える。彼らは、今回の実験では自主的な学習が行われなかったが、教示によっ

て超鏡指差しの容易性は確かめられた。

この結果は、超鏡の応用として、機器操作（例えばビデオデッキ）のわからない人に対し、家電メーカーの専門家が、消費者の所有している実際の対象機器を指差しながら、具体的な操作方法を教える（図3）といった新サービスの可能性を示唆している。また、工事現場（図10）、宇宙ロケット、事故現場など時間的・物理的に同席が困難な場所での、事物を対象とした対話にも応用できると考える。

参考文献

- [1] Birdwhistell R.L.; Kinesics and Context: Essays on Body Motion Communication, Univ. of Pennsylvania Press (1970)
- [2] Morikawa, O. and Maesako, T.: HyperMirror: a Video-Mediated communication system, CHI'97 extended abstracts, 317-318 (1997)
- [3] 森川、前迫：「超鏡」：自己像を表示するビデオ対話方式、情報処理学会ヒューマンインタフェース研究会資料 HI72-5,25-30 (1997)
- [4] Morikawa, O. and Maesako, T.: HyperMirror: Toward Pleasant-to-use Video Mediated Communication System, CSCW'98, 149-158 (1998)
- [5] 森川治：ビデオ対話における自己像の表示による対話相手の存在感への影響、ヒューマンインタフェース学会誌、Vol.1-1,61-68, (1999)
- [6] Krueger, M., Artificial Reality II, Addison-Wesley Pub. (1991)
- [7] Sommerer, C. Mignonneau, L. MIC Exploration Space, ACM SIGGRAPH'96 Visual Proceedings, 17(1996)
- [8] Tang, J. and Minneman, S., VideoDraw: A Video Interface for Collaborative Drawing, ACM Transactions on information Systems (TOIS), ACM, 9-2, 170-184 (1991)
- [9] Kuzuoka, H., Yamashita, J., Yamazaki, K., Yamazaki, A.; Agora: A Remote Collaboration System that Enables Mutual Monitoring, in CHI'99 Extended Abstracts, 190-191 (1999)
- [10] Tang, J. and Minneman, S., VideoWhireBoard Video Shadow to Support Remote Collaboration, Proc. CHI'91, 315-322 (1991)
- [11] Ishii, H., Kobayashi, M. and Grudin, J., Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments, ACM Transactions on information Systems (TOIS), ACM, 11-4, 349-375 (1993)