

仮想空間を使った多地点遠隔会議システム：e-MulCS

福井 健太郎[†] 喜多野 美鈴[†] 岡田 謙 一[†]

現在、対面環境での会議よりも、多地点遠隔会議が見劣りすることは否めない事実である。その要因の1つとして、「だれがだれを見ているか」認識できないことが考えられる。そこで我々は多地点遠隔会議システムに仮想空間を導入し、その空間中にユーザの分身であるアバタを置いた。また、会議を行ううえで入力インタフェースを“操作”することは、会話の流れを止めかねない。そこで、我々はモーションプロセッサを利用した直感的な入力インタフェースを導入し、「だれがだれを見ているか」を自動的に認識できるような多地点遠隔会議システムを構築した。さらに、ユーザ自身が向いている方向のフィードバックを得るための画面表示の仕方も開発した。そして、これらを盛り込んだシステムとして、e-MulCSを実装した。評価を行った結果、従来の多地点遠隔会議システムに比べ、「だれがだれの方向を向いているか」がよく分かり、それにともない会話がよりスムーズに進んだことが分かった。また、我々の導入した入力インタフェースと表示法が適切であったという結果を得ることができた。

Multiparty Conference System enhanced by Virtual Space: e-MulCS

KENTARO FUKUI,[†] MISUZU KITANO[†] and KEN-ICHI OKADA[†]

Currently, multi-party video conference does not provide equivalent quality in comparison to face-to-face conference. One assumed reason is that participants can not be aware of “who is focusing on whom”. We introduce virtual space to the multi-party conference system, allocating avatars in the space. We also introduce intuitive input interface using motion processor in order to construct a multi-party conference system where participants can be aware of “who is focusing on whom”. A new displaying method is essential for this system, therefore, we introduce a way by which an user can obtain the feedback of which user he/she is focusing on. We introduce e-MulCS as the system which fulfils those proposals. By comparing this system with the video conference system, we were able to obtain the result that our system is better in the sense that participants were aware of “who is focusing on whom”.

1. はじめに

近年、コンピュータを使いネットワークを利用することにより、遠隔地の人同士の共同作業を支援するシステムが目立っている。その中でも、多地点遠隔地間でのコミュニケーションを実現した、多地点遠隔会議システムに注目が集まりつつある。

多地点遠隔会議システムで最も一般的だと考えられているのが、図1のように、1つの画面に、各拠点から送られてきた動画を分割して表示することにより、遠隔地の人同士と一緒に会議を行うことができるシステム（以後、ビデオ会議システム）である。このようなシステムは、すでにさまざまなところで実用化されつつある¹⁾。しかし、これらのシステムを使用してみ

ると、画面が区切られて表示されていることにより、だれがだれに注目しているか分かりにくい。その結果、会話がスムーズに進まなくなる、参加者が同じ空間にいる気分になれない（一体感が得られない）といった問題が発生してしまう。また、動画を多地点にリアルタイム配信するためには、通信コストがかかる。

多地点遠隔会議において、視線一致を実現したシステムとして、MAJIC²⁾があげられる。MAJICでは、等身大の相手画像との視線一致が可能となる。しかし、このシステムは、カメラ位置と対話参加者との物理的位置合わせが困難であり、また、各参加者ごとに大がかりな装置が必要になるなど、実用化には解決すべき問題点が多数残っている。PAW³⁾、Valentine⁴⁾、InterSpace⁵⁾のように仮想空間を構築し、その中を自由に動き回れるアバタを介してのコミュニケーションも実用化されつつある。しかし、このようなシステムではキーボードやマウスなどを利用して操作をする必

[†] 慶應義塾大学理工学部

Faculty of Science and Technology, Keio University



図1 ビデオ画像を取り入れた多地点遠隔会議システム
Fig. 1 Window of the general video conference system.

要があり、会議に適さない。The GAZE Groupware⁶⁾のように、会議室のような仮想空間内に、ユーザの動画画像を貼り付けた板を配置するシステムも存在する。しかし、このようなシステムでは、板に顔が貼り付けられているため、臨場感を得られない。また、板をどの方向に動かしても、板の上の顔画像がつねにユーザ自身の方向を向いているように見えてしまうモノリザ効果⁷⁾が生じ、だれがだれを向いているかを正確に把握しにくくなる。

このような背景のもと、本研究では「シームレスなコミュニケーション空間」をどのような表示方法を用いれば実現できるかに着目した。そして、「だれがだれの方向を向いているかを認識できること」を PFA (Participant Focusing Awareness) と定義し、シームレスなコミュニケーション空間を実現した多地点遠隔会議での PFA の支援に焦点をあてた。また、ユーザに負担を与えることなく、ユーザの向いている方向を認識できるデバイスについても考慮した。さらに、ユーザ自身に、向いている方向のフィードバックを与えるための画面についても検討を加えた。そして、これらの検討を組み込んだ、多地点遠隔会議システム “e-MulCS” (enhanced Multiparty Conference System) を実装した。

以下本稿では、2章で多地点遠隔会議に求められる要件を、3章で e-MulCS のデザインについて述べ、4章で e-MulCS の実装について説明する。5章では評価とその結果について考察し、6章ではまとめと今後の課題について述べる。

2. 多地点遠隔会議に求められる要件

本章では、多地点遠隔会議に求められる要件について、「だれがだれの方向を向いているかを認識できること」の重要性と、自然なインタラクションに焦点を絞って述べる。

2.1 PFA (Participant Focusing Awareness) の重要性

対面会議では、参加者についての詳細な情報を手に入れられる。参加者は話をする際、話しかけたい相手の方向に視線を向けるか、体全体を向ける。そして、

その行動をもとに、その人がだれに注目しているかを他の参加者は認識できる。このように、対面会議では自動的に PFA を支援している。

それに対して、ビデオ会議システムでは、PFA を支援しない。PFA が支援されていないため、「だれに対して話しているのですか?」「あなたに話しているではありません。」といった対面会議では起こらない会話の食い違いがしばしば発生する。つまり、PFA が支援されないシステムを利用することにより、言葉によるコミュニケーション(パーバルコミュニケーション)がうまくとれなくなるという問題が発生する。

言葉以外のコミュニケーション(ノンバーバルコミュニケーション)においても、PFA の欠如による弊害は出てくる。人間はコミュニケーションを行う際、自身の話に参加者がどのくらい興味を持っているのかをつねに感じている⁸⁾。そして、興味を持っている人が少なければ、話を早く切り上げたり、話の方向性を変えたりするなどの工夫をする。PFA を支援しないビデオ会議システムでは、皆がユーザ自身の方向を向いており、会話にだれが興味を持っているのかわからない。したがって、会話に対するフィードバックが少なくなり、会話をしながらも不安な気分になりやすい⁹⁾。

このように、PFA は、パーバルコミュニケーション、ノンバーバルコミュニケーションの双方に関して、重要な役割を果たしていることが分かる。

2.2 自然なインタラクションの実現

多地点遠隔会議において自然なインタラクションを行うためには、以下の2つが重要である。

直感的な操作

コンピュータと人間とのインタラクションには、主にキーボードとマウスを利用することが多い¹⁰⁾。しかし、会議の最中にこれらの装置を利用することは、現実世界とかけ離れた操作をユーザに強要することになる¹¹⁾。ユーザが会議に集中できるようにするためには、対面会議となんら変わりのない操作で会議に臨めるようにする必要がある。

ユーザに負担がなく、無意識な操作

情報機器の技術革新により、HMD や立体視ディスプレイという表示装置を利用することができるようになった。また、入力装置として、PHANToM やデータグローブなど、触覚を利用したインタフェースも実用化の段階に入っている^{12),13)}。

しかし、これらの装置はすべて、ユーザに装着、または触れることを強要している。これらを装着することにより、ユーザは負担を感じ、会議に集中できなくなる可能性もある。また、負担をあまり感じない人で

も、意識的に装置を操作する必要があり、これが会議に影響を与える可能性もある。よって、ユーザに負荷のかかる装置を会議システムにて使用するべきではない。

3. 多地点遠隔会議システム“e-MulCS”のデザイン

本研究では、多地点遠隔会議システムでの PFA 支援について考え、また、より直感的にシステムとインタラクションをできる入力インタフェースに着目した。そして、これらを可能にしたシステムとして、“e-MulCS”を実装する^{14),15)}。

3.1 シームレスなコミュニケーション空間における PFA の支援

我々は PFA を支援するためにはシームレスなコミュニケーション空間が不可欠であると考え、そこで、これを実現する手法として、仮想空間を利用した。

3.1.1 シームレスなコミュニケーション空間の必要性

画面を分割したビデオ会議システムでも、参加者の意識次第で、PFA を支援することはある程度可能である。たとえば、図 2 のような構成のビデオ会議システムでは、各ユーザが意識的に自身の注目する人の方向に首を大げさに動かすと、その動画を見て、他のユーザはどのユーザに話そうとしているかを、ある程度認識することが可能である。

このように、意識的に PFA をビデオ会議システムで実現する方法はあるが、この手法だと、直感的な PFA を支援しているとはいえない。つまり、あるユーザが自分の方向を向いたとしても、だれの方向を向いているかを認識するのに、ユーザの配置と顔の向きとの 2 つの情報から計算する必要がある。また、全員正面を向いているため、ユーザ自身、話しかけている相手がだれであるのかあいまいになりがちである。

このような問題は、「個々のユーザが独立した空間

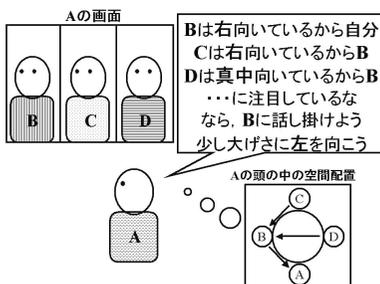


図 2 ビデオ会議での PFA 支援
Fig.2 PFA in video conference system.

に表示されている」ことに起因する。参加者が存在する空間が共有されれば、直感的な PFA を支援できる。

3.1.2 仮想空間の利用

直感的に PFA を支援するには、ある空間の中に、参加者全員を配置することが望まれる。そこで我々は、参加者全員を仮想空間の中に配置する手法を考案した。そして、仮想空間の中にユーザの分身である“アバター”を配置し、向きを変えられるようにした。

このように仮想空間を構築する手法をとることにより、以下のような利点が考えられる。

- PFA を直感的に支援する。
- 1 つの空間の中に皆が存在するため、「一緒にいる」という意識を得られる。

3.2 自然な PFA の支援

PFA を支援した会議システムでは、どのユーザがどのユーザに注目しているかを、システム側で認識する必要がある。そのための入力インタフェースには、以下のような条件が求められる。

- ユーザに負担をかけず、無意識に利用できる。
- 直感的に使える。
- リアルタイムにユーザの注目方向を認識できる。
- 精度が高い。

このような条件を満足する入力インタフェースとして、モーションプロセッサを導入した。

モーションプロセッサ(図 3 参照)⁶⁾とは、搭載されている LED から近赤外光を発生し、その光の反射をとらえる装置である。遠方からの反射光は急速に弱くなるので、近くの被写体からの反射光のみが撮像される。背景が自動的に除去されているため、背景処理の必要がなくなり、高精度な物体認識をリアルタイムに行える。モーションプロセッサを利用し、あらかじめユーザの各方向を向いたときの画像を登録しておくことにより、ユーザが顔の向きを変えたら、それがどちらの方向かをリアルタイムに認識できる。

3.3 ユーザの“注意”に基づいた空間表示法

多地点会議システムなどで、自身の向いた方向に対

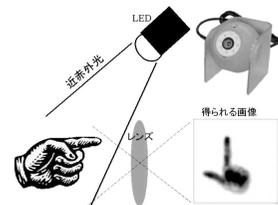


図 3 モーションプロセッサとモーションプロセッサの動作概要
Fig.3 Picture of motion processor and overview of how it works.

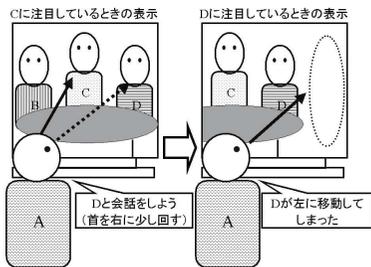


図 4 従来の仮想空間表示法の問題

Fig. 4 Defect of common displaying method.

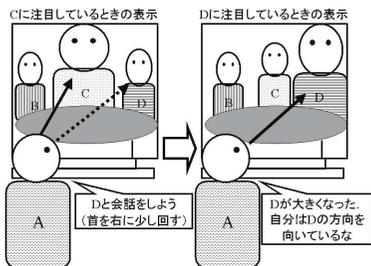


図 5 新しい仮想空間表示法

Fig. 5 New displaying method.

するフィードバックがないと、次第に人間は顔を動かさなくなる¹⁷⁾。そこで、ユーザに、“向いた方向”に対するフィードバックを与える必要がある。

一般的に、仮想空間の中でユーザの向きに対するフィードバックを与える手法として、ユーザ自身の見たいものを画面の中心に持ってくる表示法を使うことが多い。しかし、顔の向きで方向を判定するようなシステムでは、この表示法を適用すると、ユーザに違和感を与える。すなわち、図 4 のようにユーザがアバタ D に注目しようとして、画面の右側を向くと、アバタ D が画面の中央に来てしまう。さらに中央に来たアバタ D の方向を見ようと画面の中央に首を戻すと、また、アバタ D は画面の右側に表示される。このように、いつまでたっても目的のアバタを注目できなくなる。

人間は、相手に注目するとき、その人に意識が集中し、その相手が鮮明に見える代わりに、周りの背景などがぼやけてあまり見えなくなる⁴⁾。本システムでは、このことに着目し、注目した相手をより鮮明に表示するために、図 5 のようにユーザが向いた方向のアバタが“大きく”なる表示方法を開発した。この表示方法では、ユーザが右側のアバタ D に注目したいと考え、画面の右側を向くと、アバタ D が大きくなる。こうすることにより、ユーザは違和感なく、向いた方向へのフィードバックを得られる。



図 6 システム画面

Fig. 6 Implemented system.

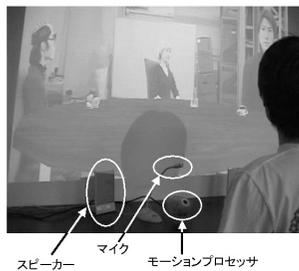


図 7 システム利用風景

Fig. 7 e-MulCS in use.

4. “e-MulCS” の実装

4.1 システム概略

本システムの基盤は、Java を利用して実装した。また、仮想空間の構築に関しては、Java の標準 3 次元グラフィック API である Java3D を利用した。モーションプロセッサの制御プログラムは VisualBasic で作成し、音声は JavaSound を用いた。全体的には、約 12,000 行程度（うち、5,000 行程度は評価実験用）のソースコードで実現されている。

図 6 にシステム画面、図 7 にシステム利用風景を表示する。なお、図 7 では、没入感を出すために、プロジェクトによる大画面を利用しているが、普通の PC の画面でも、十分本システムを利用できる。

4.2 仮想空間の構築

本システムは、3~4 人の会議を支援する。各ユーザはアバタとして仮想空間内に存在し、テーブルの周りを同じ間隔（角度）で座っているように設定した。

仮想空間内でのアバタは、以下のようなことができるように設計されている。

- 向きは、360 度回転できる。しかし、4 人の会議では左右 45 度回転できれば十分であり、それ以上は回転しない。図 6 で、左と中央のアバタは、右のアバタの方向を向いており、右のアバタは左のアバタの方向を向いていることになる。

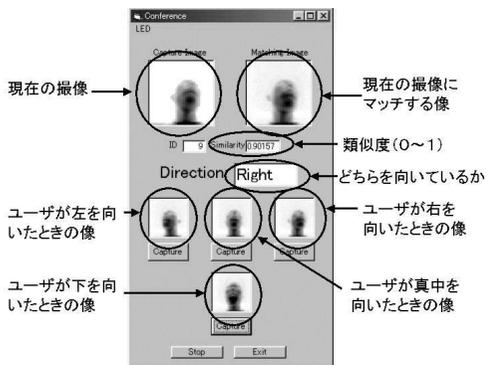


図8 モーションプロセッサプログラムの実装図

Fig.8 Implementation of the Motion Processor Program.

- 会議の最中、机の上に置いてある書類などに目を通すことはよくある。そこで、ユーザが下を向いているときは、アバタにも下を向いたしぐさをさせることにより、書類などに目を通してしている状態も、他のユーザが認識できるようにする。
- ユーザが話したとき、マイクからの音声入力レベルを検出し、それに応じてアバタの口を動かす。
- より人間らしさを出すために、適当なタイミングでまばたきをさせる。

図6から、各々のアバタが、どちらの方向を向いているかが分かる。この場合、全体的に右側の参加者が話の進行役をしているか、話の中心になっている可能性が高いことが一目で読み取れる。これは、図1や図2のようなシステムではできなかったことである。

本システムでは、3Dアバタを用いて実装しているためビデオ画像を利用できないが、3Dアバタにビデオ画像をマッピングすることにより、理論的にはビデオ画像を用いた会議システムも実現可能である。

4.3 モーションプロセッサの実装

モーションプロセッサは、撮像した映像をフレームデータの中に、256ピクセル(16×16)の形式で格納する。撮像された生画像の各ピクセルは、0~255の値を持つ。この値が大きければ大きいほど、その物体はモーションプロセッサの近くに存在することを示す。

e-MulCSで会議を行う前に、モーションプロセッサプログラム上で、ユーザの、左・中央・右・下を向いた状態を撮像し、各画像をフレームバッファ上に格納する。会議中には現画像と格納していた画像のパターンマッチングを実行することにより、ユーザがどちらを向いたかをリアルタイムに認識する(図8参照)。

このプログラムを利用することにより、通常のPCディスプレイ(液晶15inch)を利用した環境で、87%の精度で向いた方向を認識できた。



図9 e-MulCSでの空間表示

Fig.9 Displaying method of e-MulCS.

4.4 本システムに適した空間表示の実装

ユーザ自身の見たいものを画面の中央に持ってくる従来の表示法は、本システムには適用しにくいということは前章で述べた。そこで本システムでは、向いた方向のアバタが大きくなる表示方法を実装した。この表示法では、ユーザが左側を向いたときは、図9の左側のように左のアバタが大きくなり、逆にユーザが右側を向いたときは、右のアバタが大きくなる。

5. 実験と評価

e-MulCSで実装したデザイン要件が適切であったかを評価するために、PFAの実現性評価、e-MulCSの実用面の評価、入力装置と画面表示の適性に関して評価実験を行った。

5.1 PFAの実現性評価

コミュニケーションを行ううえで、PFAの支援は重要であることはすでに述べた。そして、我々は仮想空間を用いることにより、直感的にPFAを支援できると主張した。そこで、実際にe-MulCSがビデオ会議システムに比べ、直感的にPFAを支援できていることを実証する実験を行った。

5.1.1 実験内容

本実験では、ビデオ会議システムにおけるPFAと、e-MulCSにおけるPFAの比較を行った。「画面が区切られているか否か」以外をまったく同じ条件にしないと、PFAの支援度合いを正確に測れないため、図10のように、ビデオ会議システムに似たシステム(以下、ビデオ会議風システム)とe-MulCS(アバタはビデオ会議システムと同じもの)を比較した。

ビデオ会議風システムは、図2を元に設計している。したがって、図10の上の図で、左側のアバタは右側のアバタに、中央のアバタは右側のアバタに、右側のアバタは被験者に注目していることになる。

実験の流れは以下のとおりである。

- (1) スタートボタンを押す。
- (2) 自動的にあるアバタが「はい!」と声を出し、他のアバタ、またはユーザの方向を向く。

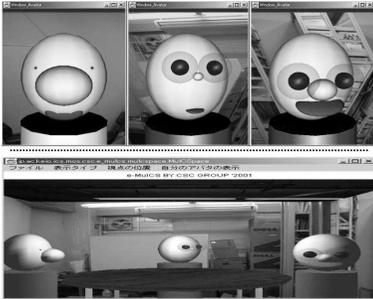


図 10 実験で使ったビデオ会議システム(上)と e-MulCS(下)
Fig. 10 Video conference system (above) and e-MulCS (below) for this experiment.

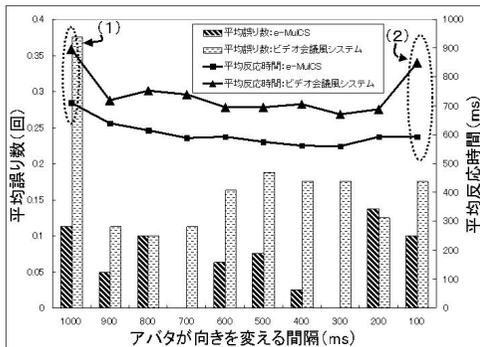


図 11 アバタが向きを変える間隔と平均反応速度・平均誤り数の関係
Fig. 11 Relation between avatar's focusing interval, the reaction time and the accuracy.

- (3) アバタが向かれた場合、そのアバタは「はい!!」と言って他のアバタ、または被験者の方向を向く。
- (4) 被験者が向かれた場合、被験者は自身の方向を向いたアバタを示すキーをできるだけ早く押す。
- (5) 初期条件として、アバタは 1000 [ms] ごとに向きを変える。そして、ユーザの方向を 5 回向いた後、向きを変える間隔を 100 [ms] 短くする。この実験を 10 回繰り返す、最終的には 100 [ms] ごとにアバタが向きを変える。

両方のシステムで、被験者 48 人にこのタスクを行ってもらい、反応速度、誤り数を測定した。24 人はビデオ会議システムを先に、残り 24 人は e-MulCS を先にやってもらうことにより、順序効果がないようにした。実験後、アンケートに記入してもらった。

5.1.2 実験結果と考察

アバタが向きを変える時間間隔と被験者の平均反応速度、平均誤り数の関係を示したのが図 11 である。なお、それぞれのシステムにおいて、各アバタが向きを変える間隔を基準に Wilcoxon 符号付き順位検定を

表 1 被験者の反応速度と正確性

Table 1 Results: the reaction time and the accuracy.

	ビデオ会議システム	
	システム	e-MulCS
平均反応速度 (msec)	741	603
平均誤り数 (回)	3.40	1.33

Wilcoxon 符号付順位検定を行った結果、すべての項目が危険率 1% 以下で有意であった

行った結果、両システムでの差は危険率 1% 以下で有意である ($p = 0.00195$) ということが分かった。この図と検定の結果から、アバタの向きを変える間隔に関係なく、反応速度は e-MulCS の方がつねに早いことが分かる。アバタが向きを変える間隔が長い (1000 [ms]) とき (この図の (1) 参照) は、反応速度が遅いが、これは慣れが影響していると考えられる。また、アバタの向きを変える間隔が早くなったとき、この図中の (2) で示した部分に大きな差が出てくる。これは、ビデオ会議システムでは、向きを変える間隔が長い間は、頭の中で考えながら操作してもあまり反応が遅れないが、間隔が短くなったとき、頭で考えている分反応が遅れたと考えられる。つまり、ビデオ会議システムで支障なく PFA を得るための、発言者の移り変わり速度の閾値は 100 [ms] から 200 [ms] の間にあることが考えられる。被験者の反応速度と正確性を示した表 1 から、e-MulCS の方が全体的に早く反応できたことが分かる。

アバタの向きを変える時間間隔と誤り数は、あまり関係がないことを、図 11 は示している。これは被験者が反応速度を上げることに重点を置いたため、正確性に差が出なかったと考える。全体的には、PFA を支援している e-MulCS の方が正確に反応できたことが、表 1 より分かった。

表 2 のアンケート結果から、e-MulCS のシームレスなコミュニケーション空間を有しているという特徴が、被験者にとって自然であり、早い会話の移り変わりにも対応することができ、だれがだれの方向を向いているのか直感的に分かる環境を実現したことが分かる。

これらの結果から、直感的に PFA を支援する手法として、仮想空間を用い、その空間の中のアバタの向きを変える手法の有用性を確認できた。

5.2 e-MulCS の実用面評価

システム全体として、実用的側面からの評価を得るため、スケジュール調整会議を行った。

5.2.1 実験方法

本実験では、e-MulCS の比較対象として iVisit (<http://www.िवisit.com>) を用いた (図 12 参照)。そ

表 2 PFA 実現性評価のアンケート結果 (1: まったく思わない, 7: 強く思う)

Table 2 Evaluation of how well PFA is supported: results of the questionnaire (1: strongly disagree, 7: strongly agree).

質問	ビデオ会議システム	
	ビデオ会議システム	e-MulCS
違和感なく入力できたか (直感的に正解を出せたか)	1.21	5.50
速さに対応できたか	1.43	5.11
現実空間と近かったか	1.11	5.36
だれがどちらの方向を向いているか分かりやすかったか	1.21	5.50
「向かれた」気になったか	1.71	4.89
楽だったか	2.00	5.43

Wilcoxon 符号付順位検定を行った結果, すべての質問が危険率 1%以下で有意であった



図 12 実験に利用したビデオ会議システム：iVisit

Fig. 12 iVisit: The video conference system which we used for the comparative experiment.

して, 作業をより早く完了することができ, 参加者に余計な操作や発言をさせることなく, また, つねに「だれがだれの方向を向いているか」が分かることを「実用面からみて有用」とであると定義した。

被験者には, それぞれのスケジュールを合わせ, 都合の良い時間にディスカッションを入れる, というタスクを, 両方のシステムで行ってもらった。タスクは, 以下のような制約のもとに実施された。

- なるべく早く予定を合わせる。
- 相手の名前を極力呼ばない。
- 自分のスケジュールを順に読み上げたりせず, 話し合いながら決めていく。

実験は, 48 人を 16 グループに分け (一グループあたり 3 人), 8 グループはビデオ会議システムを, 残り 8 グループは e-MulCS を先にやってもらうことにより, 順序効果が出ないようにした。そして以下を測定し, 最後にアンケートに答えてもらった。なお, 実験の様子はビデオで撮影した。

- タスク完了時間
- 名前使用回数
- 1 分あたりの発言時間
- 1 分あたりの発言回数

5.2.2 実験結果と考察

タスク完了時間, 名前を使用した回数, 1 分あたりの発言時間, 1 分あたりの発言回数を記したのが, 表 3 である。この表から, e-MulCS の方が早くタスクを完了できたことが分かる。また, 本実験では, 相手の名前を呼ばないという制約があるにもかかわらず, ビデオ会議システムでは 3.12 回名前を呼んでいるが, e-MulCS では 0.20 回にとどまっている。これは, e-MulCS が, ビデオ会議システムより直感的に PFA を支援しており, 相手の名前を呼んで, 「話しかけようとしている」ことを伝える必要がほとんどなかったことを示していると考えられる。本システムでのモーションプロセッサには, 87% の認識率があると 4.3 節で述べたが, e-MulCS で名前を呼ぶ回数が大変少ないというこの結果から, モーションプロセッサによる誤認識が起きて, 自身が向いた方向に対するフィードバックが画面にあるため, 被験者はすぐに向きを修正することができたということが読み取れる。撮影したビデオを解析すると, ビデオ会議システムの方では, 会話の最中に「え?」や「だれに言ってるの?」といった発言が頻繁に出てきているが, e-MulCS ではそのような発言がほとんどなかった。この影響は, 表 3 にも出てきている。この表で, 1 分あたりの発言時間と発言回数を見ると, 発言時間はほぼ同じだが, 発言回数は e-MulCS の方が少なくなっている。つまり, e-MulCS の方が, 1 回の発言時間が長かったことを示しており, これは「え?」などの聞き返しが e-MulCS では少なかったため, このような結果になったと考える。

表 4 のアンケート結果を見ると, ビデオ会議システムに比べて, e-MulCS の方が, 空間を共有した気になれたという結果を得られたが, 有意差は出ていない。これは, どちらのシステムでも画面が平面であり, 画面以外の周りの様子が視野に入ってきたため表示法に大きな差が出なかったためと想像される。表示デバイスとして, 外界の情報が視野に入らない HMD などを利用した場合は, もっと大きな差が出るのではないかと想像する。他の質問では, 「だれがだれの方向を向いているか」が分かり, 「見られた・話された」気になれたという結果を得ることができ, PFA を支援したことによる利点がこの結果から分かる。「プライバシーは守られたか」という質問は, 各システムを利用することにより, 各被験者の部屋の雰囲気など, 他の人に知られたいくない情報が伝わる度合いを測定するために設問した。その結果, e-MulCS の方が良い結果を出しており, プライバシーを守るという観点から, 仮想空間を使うことは有用であると考えられる。また,

表3 ビデオ会議システムと e-MulCS の実用面評価結果

Table 3 Results of the comparative about the practicability experiment between video conference system and e-MulCS.

	ビデオ会議 システム	e-MulCS	Wilcoxon 符号付 順位検定 p 値
タスク完了時間(秒)	272	246	*0.01152
名前使用回数(回)	3.12	0.20	**0.00143
発言時間(秒/1分)	18.5	18.4	0.05875
発言回数(回/1分)	10.2	9.31	0.08033

(N = 48; **: $p < 0.01$, *: $p < 0.05$)

表4 実用面評価のアンケート(1:まったく思わない, 5:強く思う)

Table 4 System practicability: Results of the questionnaire (1: strongly disagree, 5: strongly agree).

質問			Wilcoxon
	ビデオ会議 システム	e-MulCS	符号付順位 検定 p 値
空間を共有している気になったか	3.52	3.91	0.12406
空間的距離感を感じられたか	2.44	3.69	**0.00015
だれがだれを向いているか分かったか	1.75	4.56	**0.00000
見られた・話された気になったか	2.29	4.23	**0.00000
固有名詞を使わず平気だったか	2.52	3.80	**0.00005
プライバシーは守られたか	2.98	4.33	**0.00000
画面を「向く」気になったか	3.17	4.23	**0.00011

(N = 48; **: $p < 0.01$, *: $p < 0.05$)

e-MulCS では、顔を動かすと、自身の注目した方向へフィードバックを得ることができる。そのため「画面を向く気になったか」という質問に対して、我々は大幅に e-MulCS の方が良い結果を出すことを予想していた。しかしながら、実際には予想していたほどの差は出なかった。これは、本システムでのモーションプロセッサプログラムの認識率が 87% であり、ときどき誤認識を起こすため、それが被験者にとって負担となり、このような結果になったと想像する。

e-MulCS は、タスクをより早く完了でき、相手の名前を意識的に呼ぶ必要もあまりなく、つねに「だれがだれの方向を向いているか」を認識することができるということが分かった。つまり、実用面からみて有用であるということが確認できた。

5.3 入力装置と画面表示の適性評価

本システムでは、モーションプロセッサと新しい表示方法を使うことにより、対面会議と同じように、注目したい人の方向をつねに直感的に向くことができ、PFA がサポートされているコミュニケーションを実現できる。そこで、このシステムにおいて利用した入力装置と、その装置に適した画面表示方法が適切であったかを調べるための実験を行うため、スケジュール調整会議を行った。比較システムとして、キーボードと従来の空間表示方法(図13参照)を適用した e-MulCS を実装した。なお、比較システムの表示法は、図4を



図13 キーボードを利用したときの表示法

Fig. 13 Displaying method for the keyboard version of e-MulCS.

もとに設計されているが、向きを変えても画面上からアバタが消えないように、少し離れたところから会議を眺めるように設計した。

5.3.1 実験内容

本実験は、5.2.1 項と同じ方法でスケジュール調整会議を行い、さらに被験者の「向きを変えた回数」も測定した。

5.3.2 実験結果と考察

モーションプロセッサと新しい表示法が、本システムにおいて直感的、かつ操作に負担がかからないインタフェースなら、被験者は、システムを操作する時間が減り、その分会話の時間が長くなることが予想される。また、負担がかからないと、向きを変える労力もないため、向きを変える回数が増えることが予想される。タスク完了時間、発言時間、発言回数、向き変更

表5 入力インタフェースと画面表示法の比較結果
Table 5 Comparison of the input interface.

	キーボード と一般表示	モーションプロ セッサと新表示法	Wilcoxon 符号付き 順位検定 p 値
タスク完了時間 (秒)	248	246	0.53844
発言時間 (秒/1分)	17.4	18.4	0.28399
発言回数 (回/1分)	9.55	9.31	0.36132
向き変更数 (回/1分)	9.89	36.23	**0.00000

($N = 48$; **: $p < 0.01$, *: $p < 0.05$)

表6 入力インタフェースと画面表示法の比較のアンケート (1:まったく思わない, 5:強く思う)

Table 6 Input interface and display method: results of the questionnaire (1: strongly disagree, 5: strongly agree).

質問	キーボード と一般表示	モーション プロセッサと 新表示法	Wilcoxon 符号付順位 検定 p 値
画面を向く気になったか	3.04	4.27	**0.00001
直感的に向きを伝えられたか	2.95	3.94	**0.00356
相手を向いた気になったか	3.04	4.29	**0.00000
目は疲れなかったか	3.73	3.46	0.17619

($N = 48$; **: $p < 0.01$, *: $p < 0.05$)

回数の平均を記したのが、表5であるが、この表を見ると、モーションプロセッサを使ったシステムの方が、1分あたりの発言時間が長くなり、向き変更数が多くなっているため、より直感的にシステムを操作することができたと推測する。

表6のアンケートより、「画面を向く気になったか」「直感的に向きを伝えられたか」という質問に対してモーションプロセッサの方が良い結果を示した。これも、モーションプロセッサの直感的に操作できる点と、被験者に負担がかからないことが評価されたと推測される。また、「相手を向いた気になったか」という質問に対して、新しい表示法の方が良い結果を出し、新しい表示法が被験者にとって、直感的であったことがこの結果から分かる。「目は疲れなかったか」という質問に対しても、新しい表示法の方が良い結果を出すと我々は推測していた。しかし、従来の仮想空間表示法が良い結果を出した。これは、従来の表示法の方が被験者にとって慣れていないためこのような結果になったと考えられる。だが、Wilcoxon 符号付順位検定の結果から、有意な差ではないことが分かる。新しい表示法も慣れさえすれば問題ないと我々は考える。

これらの結果から、多地点遠隔会議システムにおいて、モーションプロセッサとモーションプロセッサに適している表示方法を適用した方が、キーボードと従来の表示方法を適用するより、適切であることが分かった。本実験では3人での会議を行ったが、参加人数が増えれば増えるほどモーションプロセッサの誤認識は大きくなると考えられる。よって、本システムは3~

4人程度の多地点遠隔会議を行うのに適していると考える。

6. まとめ

本稿では、ビデオ会議システムの問題点として、だれがだれに注目しているのかが分からなくなってしまうこと (PFAの欠如) をあげた。そこで、PFAを支援することの重要性について述べ、直感的にPFAを支援させる手法として、仮想空間を用いることの有用性について述べた。

また、PFAを支援させるためには、だれがだれに注目しているかをシステムが認識する必要がある。そのための入力インタフェースとして、モーションプロセッサを利用することにより、ユーザが直感的に、なおかつ負担を感じずにシステムを利用できることについて述べた。さらに、モーションプロセッサを利用した場合、従来のように見たいものを画面の中央に持ってくるという表示方法では、不具合が発生する。そこで、ユーザの向いている方向のアバタが拡大する手法を開発し、これを用いることにより、ユーザ自身が自然に向いている方向のフィードバックを得ることができるとして述べた。

これらを支援したシステムとして、e-MulCSを開発し、その有用性を評価実験を通して確認した。今後の多地点遠隔会議システムにおいて、このような手法を用いることにより、より対面に近い会議を実現できると思われる。今後の課題としては、アバタの表情をより豊かにする、対面会議と同じように「場の雰囲気」

などが伝わるようにすることなどがあると考える。

参考文献

- 1) 松下 温, 岡田謙一, 勝山恒男, 西村 孝, 山上俊彦(編): 知的触発に向かう情報社会—グループウェア維新, 共立出版(1995).
- 2) Okada, K., Maeda, F., Ichikawa, Y. and Matsushita, Y.: Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design, *Proc. CSCW '94*, pp.385-393 (1994).
- 3) 松田晃一, 三宅真浩: パーソナルエージェント指向仮想社会 PAW(第2版)の構築と評価, 情報処理学会論文誌, Vol.41, No.10, pp.2698-2707 (2000).
- 4) 本田新九郎, 富岡展也, 木村尚亮, 大澤隆治, 岡田謙一, 松下 温: アウェアネススペースを実現し気配を考慮した大部屋仮想オフィス, 情報処理学会研究会資料: マルチメディア通信と分散処理 80-26・グループウェア 21-26, pp.151-156 (1997).
- 5) 井上雅之, 宇佐美潔志, 清末梯之, 石橋 聡, 長谷雅彦: 3次元仮想社会 InterSpaceにおけるコミュニティ形成過程とコミュニケーションメディア利用推移に関する考察, 情報処理学会論文誌, Vol.41, No.10, pp.2670-2678 (2000).
- 6) Vertegaal, R.: The GAZE groupware system: mediating joint attention in multiparty communication and collaboration, *Proc. CHI '99*, pp.294-301 (1999).
- 7) Zitnick, C.L., Gemmell, J. and Toyama, K.: Manipulation of Video Eye Gaze and Head Orientation for Video Teleconferencing, Microsoft Research Technical Report, MSR-TR-99-46 (1999).
- 8) 松下 温, 岡田謙一(編著): コラボレーションとコミュニケーション, 共立出版(1995).
- 9) Vertegaal, R., Slagter, R., van der Veer, G. and Nijholt, A.: Eye Gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes, *Proc. CHI 2001*, pp.301-308 (2001).
- 10) 村田厚生: ヒューマンインタフェースの基礎と応用, 日本出版サービス(1998).
- 11) 重野 寛, 本田新九郎, 大澤隆治, 永野 豊, 岡田謙一, 松下 温: 仮想空間における風と香りの表現手法—仮想空間システム Friend Park, 情報処理学会論文誌, Vol.42, No.7, pp.1922-1932 (2001).
- 12) 藤田欣也, 河内卓朗: グローブ型簡易仮想物体硬さ感覚呈示装置の開発, 信学技報 MVE97-20, 電子情報通信学会(1997).
- 13) 蒲原新一, 竹田 仰: 分類システムと遺伝アルゴリズムによる腕相撲対戦の試み, 信学技報 MVE95-21, 電子情報通信学会(1995).
- 14) 福井健太郎, 島中美穂子, 岡田謙一: 仮想空間を使った多地点コミュニケーションを促進させる技法, 日本バーチャルリアリティ学会研究報告, CSVC2001-11, pp.27-32 (2001).
- 15) 喜多野美鈴, 福井健太郎, 岡田謙一: 仮想空間を使った多地点コミュニケーションシステム e-MulCSの評価, 日本バーチャルリアリティ学会研究報告 CSVC2002-03, pp.15-20 (2002).
- 16) 沼崎俊一, 森下 明, 梅木直子, 土井美和子: ジェスチャ入力に適した画像入力装置の提案とその3次元情報検出性能の検討, 情報処理学会論文誌, Vol.41, No.5, pp.1267-1275 (2000).
- 17) Sellen, A.J.: Speech Patterns in Video Mediated Conversations, *Proc. CHI '92*, pp.49-59 (1992).

(平成14年3月18日受付)

(平成14年9月5日採録)



福井健太郎(学生会員)

2001年慶應義塾大学理工学部情報工学科卒業。現在、同大学大学院理工学研究科開放環境科学専攻情報通信メディア工学専修修士課程に在学中。グループウェアの研究に従事。



喜多野美鈴

2002年慶應義塾大学理工学部情報工学科卒業。現在、アサヒビール株式会社勤務。



岡田 謙一(正会員)

慶應義塾大学理工学部情報工学科教授, 工学博士。専門は, グループウェア, コンピュータ・ヒューマン・インタラクション。「コンピュータヒューマンインタラクション」(オーム社), 「コラボレーションとコミュニケーション」(共立出版)をはじめ著書多数。情報学会誌編集主査, 論文誌編集主査, GW研究会主査等を歴任。現在, GN研究会運営委員, BCC研究グループ幹事, 日本VR学会仮想都市研究会副委員長。IEEE, ACM, 電子情報通信学会, 人工知能学会各会員。1995年度情報処理学会論文賞, 情報処理学会40周年記念論文賞, 2000年度情報処理学会論文賞受賞。