LM-002

複合現実感による分散会議における映像音声配置の影響

Effects of Audio and Visual Placement in Mixed Reality Meeting

野口　康人[†]　　井上　智雄[†,‡]
Yasuhito Noguchi,　　Tomoo Inoue

１．はじめに

会議は企業における情報伝達，創造，調整，決定を行うための重要なプロセスであり[1]～[3]，分散環境においてもその需要は高い．会議支援研究では，遠隔会議の臨場感を高めるための研究が多くなされてきた[4][5]．高臨場感とは現実との差異が人間にとって小さいことを意味し，高い臨場感は人間の感性を励起するなど，スムーズなコミュニケーションにおいて重要な要素である[6]．臨場感を対面会議に近づけるための方法の一つとして，複合現実感を用いた手法が研究されている[7][8]．これらは複合現実空間内に遠隔地点の参加者のアバタを表示することで分散会議のコミュニケーションを支援する．

一方，視聴覚から得られる情報が相互に影響を及ぼしあっていることは以前から知られており[9]，複合現実空間内での視聴覚効果について明らかにする必要がある．我々はこの点に着目し，複合現実空間内での視覚刺激の効果と聴覚刺激の効果の関係性について実験的に検討した．

以下，2 章では関連研究について述べ，3 章で実施した実験について述べる．最後に 4 章においてまとめを述べる．

２．関連研究

我々は五感を利用して情報を取得することを通じて，身の回りの状況を把握している．その中でも視覚と聴覚は大きな役割を担っており，これらの感覚から得られた情報は，単独ではなく相互に補完するように利用されると考えられている[10]．これは複合現実空間内においても例外ではないと考えられる．

視覚刺激にテレビを用いた場合の，視覚刺激と聴覚刺激のずれについて検討している研究がある．この研究では 20 インチのテレビとステレオ音像を用いて，ニュースのアナウンス映像を注視した場合の，映像と音像定位のずれの許容範囲に関する検討を行っている．その結果では，提示した映像と音像のずれが 10°以内であれば，音像は映像の方向に強く引き寄せられ，20 度以上になると，知覚される音像の位置は映像の影響をほとんど受けなくなると報告されている[11]．同じくテレビ環境での映像と音像定位のずれについて検討している研究がもう一つ別にある．その結果では，ニュースのアナウンス映像とスピーカの位置が被験者から見て 8°以上ずれると，被験者の半数以上が違和感を覚えると報告されている[12]．これらのようにテレビ環境においては，視覚刺激と聴覚刺激，両者間の方向のずれに対する知覚が研究されているが，複合現実空間内においてもこれらの知見が有効であるかどうかはまだ定かではない．

一方，視覚と聴覚を併用した仮想現実感の実現例[13][14]や，複合現実空間内での視覚と聴覚の融合を目的としたシステムも研究されている[15]．しかしながら，仮想現実空間内や複合現実空間内での視覚刺激の効果と聴覚刺激の効果の関係性について明らかにしたものは少ない．そこで我々は複合現実分散会議条件において，アバタとスピーカの位置をずらした環境で会議実験を行い，評価を行った．

３．視聴覚効果評価実験

3.1. 実験目的

関連研究により，視覚刺激にテレビを用いた条件では，視覚刺激と聴覚刺激の位置が 8°ないし 20°以上ずれると視聴者は視覚と聴覚に違和感を覚えることが分かっている．複合現実空間内でも視聴者に対して同じような効果があることが仮説として考えられる．

複合現実感を用いた分散会議システムにおいて，複合現実空間内での視覚刺激の効果として，座席配置の効果が示されている[16]．このシステムでは遠隔地にいる会議参加者をアバタとして表示しており，会議参加者の音声はスピーカによって本人のアバタの前から出力されている．この研究では視覚刺激と聴覚刺激の位置が一致している場合について会議実験を行っている．これに対し本研究では，視覚刺激と聴覚刺激の位置が一致していない場合について会議実験を行う．視覚刺激と聴覚刺激の位置が一致している場合としていない場合の両者の間で，会議システムの評価にどのような違いがあるのかについて検討する．

3.2. 実験環境

3.2.1. 使用機器

システムの実装にはキヤノン社製の MR Platform システム[17]を用いている．HMD は同じくキヤノン社製の VH-2002 である．表示素子数は 640×480×GBA で 92 万画素，表示画角は H51°× V37°（2m 先で 92 インチ相当）である．位置センサは Polhemus 社の FASTRAK である．開発言語は C++，グラフィック開発環境には Open Inventor（SGI 社）と OpenGL を用いた．

3.2.2. システム構成

システムの構成図を図 1，システム使用者の外観を図 2 に示す．システム使用者は HMD を付けて会議に参加する．それ以外の会議参加者は，遠隔地として想定された遮音ブースにて会議に参加する．遮音ブースは会議システムのある部屋内にあるが，直接その中の様子が見えたり，音声が聞こえたりすることはない．音声の伝送にはコンピュータネットワークは利用していない．

[†]筑波大学大学院 図書館情報メディア研究科
Graduate School of Library, Information and Media Studies, University of Tsukuba
[‡]国立情報学研究所
National Institute of Informatics

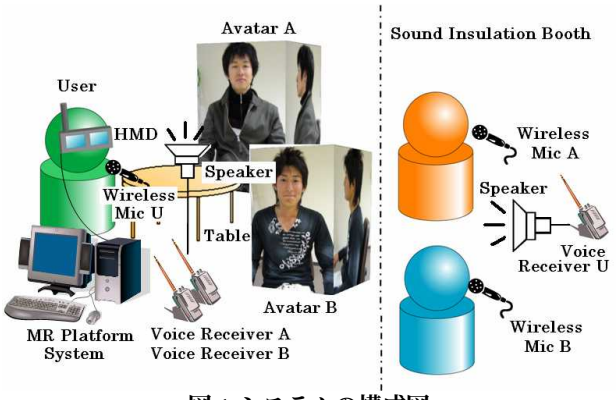

図 1 システムの構成図

図 2 システム使用者の外観

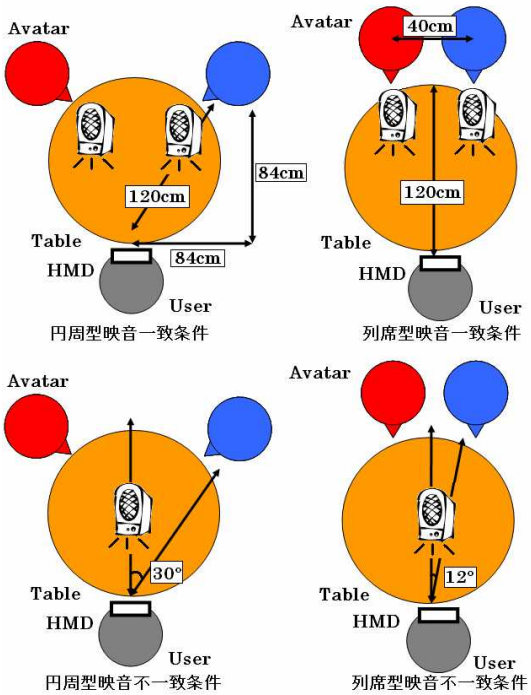

図 3 アバタとスピーカの配置

音声処理 ユーザの音声はワイヤレスマイクで取得され，遮音ブース内の音声レシーバへと送られる．遮音ブースにいる各参加者の音声はそれぞれのワイヤレスマイクにより取得され，机中央に設置されたスピーカから出力される．

映像処理 システム使用者は，現実世界の映像上に表示されたアバタを相手に会議に参加する．遮音ブース内の参加者の静止画を用いたアバタを事前に用意しておく．アバタは図 1 のように，全体が四角柱で構成され，正面には正面の静止画像，側面には側面の静止画像が貼り付けられている．これらのアバタは会議での利用を前提としているため上半身のみで構成されている．これらは，アバタの空間的配置の影響を実験するためのものであり，いずれも，ある一定の位置に存在するほかに動くことはない．つまり，参加者の視線や表情，動きを反映することはない．

3.3. 実験方法

アバタとスピーカの配置 机を囲むようにしてアバタを配置し，各アバタの前にスピーカを配置した条件（図 3 の円周型映音一致条件）と机の真ん中にスピーカを配置した条件（図 3 の円周型映音不一致条件），従来の大型 2D ディスプレイを用いた会議システムのようにアバタを横一列に並べて配置し，各アバタの前にスピーカを配置した条件（図 3 の列席型映音一致条件）と机の真ん中にスピーカを配置した条件（図 3 の列席型映音不一致条件）の 4 条件を比較する．

被験者の視界 円周型映音不一致条件の被験者の視界を図 4，列席型映音不一致条件の被験者の視界を図 5 に示す．

会議実験 円周型映音不一致条件と列席型映音不一致条件における実験について述べる．被験者は大学生，大学院生で構成された 20 人である．被験者は HMD を被って会議に参加する．被験者とは別の 2 名が実験協力者として遮音ブースから会議に参加した．実験協力者は，図 1 のように被験者との音声のやりとりのみで会議に参加した．会議は各参加者が同じ立場として行った．会議のテーマは，実験者が提示した 6 つのテーマから被験者に選んでもらった．

図 4 円周型映音不一致条件の被験者の視界

図 5 列席型映音不一致条件の被験者の視界

最初に約 3 分間で会議テーマを決めてもらい，1 つの座席配置について，約 5 分間の議論を行った．被験者は，議論の直後に，体験した配置条件について質問紙に記入した．以上の手順を円周型映音不一致条件，列席型映音不一致条件に対し繰り返した．順序効果の打ち消しは，先に円周型映音不一致条件で会議を行う者と先に列席型映音不一致条件で会議を行う者の人数を等しくすることで行なった．円周型映音一致条件と列席型映音一致条件における実験について述べる．被験者は大学生で構成された 6 人である．実験の手順は円周型映音不一致条件，列席型映音不一致条件に対して行った実験と変わりないが，遠隔地にいる実験協力者は被験者の音声のみではなく，被験者の映像も見ながら会議に参加した．また，アバタは図 5 で示したデザインのアバタとは若干異なるデザインのものを使用した．実験条件の詳細については先行研究[16]で詳しく記述している．

アバタの大きさと音声 アバタの大きさは，等身大の86.3%大を採用した．これは先行研究[16]で被験者が等身大だと感じた大きさの平均値である．音声はアナログ伝送で遅延を考慮する必要はない．また，マイクの音質は会議に支障のないものであった．

質問紙 被験者に質問紙に記入してもらうことで評価を行った．質問紙は 13 の評価項目と座席配置に対する自由記述欄で構成される．評価項目は円周型映音不一致条件，列席型映音不一致条件で共通である．円周型映音一致条件，列席型映音一致条件に対して行った実験ではこの 13 の質問項目を全て含んでいる．評価項目には対面式の会議に近いと感じたか，会議の臨場感の高さ，座席配置による参加者間の位置関係の快適さを評価する質問を設けた．また，会議空間を考慮している先行研究[16]の評価項目を参考に，(1) 会議目的の達成度の評価，(2) 参加者自身にかかわる評価，(3) 相手参加者の認識にかかわる評価のいずれも欠けることのないようにした．

3.4. 実験結果と考察

各質問は「全く当てはまらない」「当てはまらない」「あまり当てはまらない」「どちらともいえない」「やや当てはまる」「当てはまる」「非常に当てはまる」の 7 段階で評価してもらった．これら 7 段階に順に 1 点から 7 点を与え，配置ごとに各質問に対する平均得点を求めた．円周型映音一致条件と列席型映音不一致条件間の比較，円周型映音不一致条件と列席型映音不一致条件間の比較については Wilcoxon の符号付順位検定を行った．また，円周型映音一致条件と円周型映音不一致条件間の比較，列席型映音一致条件と列席型映音不一致条件間の比較については Mann-Whitney の U 検定を行った．評価結果のうち，平均得点に統計的有意差があった項目を表 1，表 2，表 3 に示す．

円周型映音一致条件と列席型映音一致条件間の比較 円周型映音一致条件と列席型映音一致条件間の質問紙による評価結果を表 1 に示す．評価項目「発言内容を聞き取れた」において有意差があった．また，「音声と映像に違和感があった」にも有意傾向にあり，音声の認識しやすさにおいて両者の間に差が見られた．両条件ともアバタの位置から本人の音声が出力されているにもかかわらず，円周型条件の方が列席型条件よりも音声の聞き取りやすさにおいて良い結果となった．これは，列席型条件よりも円周型条件の方がアバタ間での物理的距離があるため，相手を個々に認識しやすかったためだと考えられる．

また，評価項目「このシステムは会議空間に適している」において有意差があった．また，「相手との位置関係が快適だった」，「会議に参加していると感じた」について有意傾向が見られた．このように，アバタの空間的配置が参加者の受ける印象に影響することが分かった．

円周型映音一致条件と円周型映音不一致条件間の比較 円周型映音一致条件と円周型映音不一致条件間の質問紙による評価結果を表 2 に示す．評価項目「臨場感の高い会議が行えた」,「相手の存在感がある」,「十分なコミュニケーションがとれると感じた」,「相手に話しかけやすい」,「相手との位置関係が快適だった」において円周型映音一致条件の方が，有意に平均得点が高かった．アバタが同じ配置である場合，映像と音声の位置が一致していることが会議の臨場感，相手の存在感，十分なコミュニケーション，相手に対する話しかけやすさ，位置関係の快適さを高める上で重要な役割を果たしていると言える．

列席型映音一致条件と列席型映音不一致条件間の比較 列席型映音一致条件と列席型映音不一致条件間の質問紙による評価結果を表 3 に示す．ひとつの評価項目に対して有意傾向があったが，その他の評価項目については有意差がなかった．スピーカがアバタと同じ位置にある場合と，机の中央に位置している場合では，結果的にどちらもアバタと音声がある方向がほぼ一致することになったため，被験者に与えた影響は小さかったものと考えられる．映音不一致条件のスピーカの方が映音一致条件よりも被験者に近い場所に位置している．評価項目「発言内容を聞き取れた」に関しては，列席型映音不一致条件の方が平均得点において上回っていた．これは，列席型映音不一致条件におけるスピーカの位置の方が，より被験者に近く，被験者にとって大きい音声で聞こえたためだと考えられる．

円周型映音不一致条件と列席型映音不一致条件間の比較 円周型映音一致条件と列席型映音一致条件間で有意差があった評価項目についても，映像と音声の位置が一致していない条件では有意差がなかった．円周型映音一致条件と列席型映音一致条件ではアバタのある方向から音声が出力されるため，システム使用者は頻繁にアバタが表示されている方向に顔を向けていた．しかしながら，円周型映音不一致条件と列席型映音不一致条件では常に前方から会議相手の声が出力されるため，特に円周型映音不一致条件において，アバタが表示されている方向に顔を向けない被験者が目立った．そのため両配置の間に差がなくなってしまったものと考えられる．特に今回使用したシステムでは，表示されているアバタの映像に変化がないため，実験の前半では頻繁にアバタに顔を向けるものの，後半ではまったくアバタが表示されている方向に顔を向けない被験者も見られた．

4 条件間の比較 4 条件間の比較から，アバタとスピーカの方向が一致していることが，会議に対する評価項目の結果に影響をおよぼすことが分かった．また，視覚刺激と聴覚刺激の位置が一致している条件間であれば，アバタの座席配置の効果が表れることが分かった．

表 1 円周型映音一致条件，列席型映音一致条件間の質問紙による評価結果

評価項目	円周型映音一致条件 平均得点（N = 6）	列席型映音一致条件 平均得点（N = 6）	Willcoxon 符号付 順位検定 p 値
発言内容を聞き取れた	5.50	4.33	**0.034
このシステムは会議空間に適している	5.67	4.50	**0.038
音声と映像に違和感があった	3.00	4.67	**0.072
相手との位置関係が快適だった	5.83	5.33	**0.083
会議に参加していると感じた	5.67	5.17	**0.083

(** : $p < 0.05$, * : $p < 0.10$)

表 2 円周型映音一致条件，円周型映音不一致条件間の質問紙による評価結果

評価項目	円周型映音一致条件 平均得点（N = 6）	円周型映音不一致条件 平均得点（N = 20）	Mann-Whitney U 検定 p 値
臨場感の高い会議が行えた	5.66	3.95	**0.015
相手の存在感がある	6.00	4.35	**0.021
十分なコミュニケーションがとれると感じた	5.50	3.95	**0.024
相手に話しかけやすい	5.66	4.20	**0.034
相手との位置関係が快適だった	5.83	4.70	**0.039

(** : $p < 0.05$, * : $p < 0.10$)

表 3 列席型映音一致条件，列席型映音不一致条件間の質問紙による評価結果

評価項目	列席型映音一致条件 平均得点（N = 6）	列席型映音不一致条件 平均得点（N = 20）	Mann-Whitney U 検定 p 値
発言内容を聞き取れた	4.33	5.25	*0.098

(** : $p < 0.05$, * : $p < 0.10$)

４．まとめ

本稿では，複合現実空間内におけるアバタとスピーカの位置をずらした条件において会議実験を行い，アバタとスピーカの位置が一致している場合の結果と比較検討した．その結果，アバタとスピーカの位置が一致していることが，評価項目「臨場感の高い会議が行えた」,「相手の存在感がある」,「十分なコミュニケーションがとれると感じた」,「相手に話しかけやすい」,「相手との位置関係が快適だった」に影響をおよぼすことが分かった．また, アバタとスピーカの位置が一致している条件間であれば,アバタの座席配置が会議の評価に影響することが分かった．今後は，スピーカを固定する位置を変化させた場合についての検討と，アバタとスピーカのずれがどの程度会議効率に影響を与えるかについて検討するべきであると考える．

参考文献

[1] 高橋誠：会議の進め方，日本経済新聞社，1992.
[2] 野口音光：会議の技術，ダイヤモンド社，1967.
[3] 石川弘義： 会議の心理学，ちくま文庫，1986.
[4] 並木育夫，目黒義隆，青木茂明，入江一成，野村知義，斎藤一重：高臨場感マルチメディア通信会議システムの構築と評価，電子情報通信学会論文誌 B-I，Vol.J80-B-I, No.6, pp.338- 347, 1997.
[5] 伴野明，岸野文郎：臨場感通信会議におけるヒューマンインタフェース技術，人工知能学会誌，Vol.6, No.3, pp. 358 - 369, 1991.
[6] 長原一，八木康史，谷内田正彦：超臨場感ビジュアライゼーションシステム，情報処理学会研究報告, Vol.CVIM-140, No.13, pp.95-102, 2003.
[7] S. Gibbs, C. Arapis, and C.J. Breiteneder: TELEPORT - Towards Immersive Copresence , ACM Multimedia Systems, Vol.7, pp.214-221, May. 1999.
[8] H. Kato, M. Billinghurst, K. Morinaga, K. Tachibana: THE EFFECT OF SPATIAL CUES IN AUGMENTED REALITY VIDEO CONFERENCING, Proc. of the 9th HCI International, August 5-10th, 2001.
[9] C.E. Jack, W.R. Thurlow: Effects of degree of visual association and angle of displacement on the “ventriloquism” effect, Percept. Motor Skills, Vol.37, pp.967-979, 1973.
[10] 吉田拓矢，安倍幸治，高根昭一，曽根敏夫：音源定位に対する視覚的注意の影響，日本音響学会聴覚研究会資料 H-2004-80, Vol.34, 2004.
[11] 中林克己：ステレオ音像とテレビ映像の相乗効果，テレビジョン学会誌，Vol.37, pp.984-991, 1983.
[12] 八木信忠：音声多重放送とその問題点，映画テレビ技術，Vol.315, pp.24-28, 1978.
[13] 田村祐一，佐藤哲也，陰山聡，藤原進，中村浩章：数値シミュレーションデータ表現のための音情報機能を付加したバーチャルリアリティシステムの開発, 日本バーチャルリアリティ学会論文誌, Vol.5, No.3, pp.943-948, 2000.
[14] 大久保洋幸，小野一穂，小宮山摂，湯山一郎：CG 画像と同期したインタラクティブ音場再生システムについて，日本バーチャルリアリティ学会論文誌, Vol.5, 6 No.3, pp.943-948, 2000.
[15] 比嘉恭太, 西浦敬信, 木村朝子, 柴田史久, 田村秀行：視覚・聴覚を併用した複合現実感システムの開発(1) ―視覚的 MR と聴覚的 MR の同時提示の実現―, 日本バーチャルリアリティ学会第 11 回大会, 2A2-3, 2006.
[16] 野口康人，井上智雄：複合現実感を用いた分散会議における複数アバタの配置と表現，情報処理学会論文誌, Vol.48, No.1, pp.54-62, 2007.
[17] S. Uchiyama, K. Takemoto, K. Satoh, H. Yamamoto and H. Tamura: MR Platform: A basic body on which mixed reality applications are built, Proc. IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2002), pp.246-253, 2002.