Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24

遠隔会議を活性化するためのアバター表情演出に関する提案

山田涼介[1]　由井薗隆也[1]

概要：遠隔地間で行うテレビ会議では，顔映像が会議参加者に影響することが知られている．本研究では，アバター表情を人工的に演出することによって，遠隔会議を活性化する手法を提案する．具体的には，アバター表情を利用者の顔表情と関係なく，４種類の表情（喜び，悩み，怒り，悲しみ）を用いた感情ストーリーを提示する．この提示によって，遠隔会議のコミュニケーションや印象に影響を及ぼす可能性を示す．

キーワード：遠隔会議，アバター，顔表情，感情ストーリー

A Proposal of Facial Expression Direction to Activate Remote Meeting

RYOSUKE YAMADA[†1]　TAKAYA YUIZONO[†1]

1. はじめに

近年，テレビ会議が様々な組織で活用されている．その中，顔映像における視線の不一致や不明瞭な表情は会議参加者のモチベーションを低下させることが知られている[1]．この課題を解決するために臨場感通信などの対面環境に近い環境を作り出すアプローチが取り組まれてきている．特に，視線の一致が重要とされ，ゲイズアウェアネスを支援するインタフェース技術が提案されてきた[2]．しかしながら，対面環境と同様な環境を構築するアプローチは現状において，大掛かりな設備が必要となり，設備に高いコストがかかるという課題がある．

一方，電子会議システムは対面会議と異なる心理的作用をもたらすことが知られている．テキスト会話においては匿名性などの機能により，相手の地位に関係のない平等な発言をもたらす会議が実施される可能性も示されてきている．そこで，本研究では人間の心理面に注目した方法について検討する．人間の心理的影響に着目し，人工的に顔映像を演出する機能を検討する．

具体的には従来の映像通信を用いるのでなく，アバターとしての顔画像を用いる．そして，顔の表情による心理的効果を演出するために，第三者が感情ストーリーを用いて顔の表情を変えてゆく手法を検討した．以下，２．では関連知識を述べ，３．は，遠隔会議を活性化するための提案である．４．では提案手法の評価実験について述べ，５．で実験結果を通して，提案の効果を考察する．６．おわりにでは結論を述べる．

2. 関連研究

2.1 テレビ会議と顔映像について

テレビ会議の利点として，遠隔地の参加者同士でも顔を合わせることができることや同じ画面を共有した作業を行える点がある．対面会議を目指す高臨場感を指向したテレビ会議の利用は遠隔地にある専用の会議室に専用のシステムを設置する方法が採られる．一方，Skype に代表されるパソコンやスマートフォンなど多くの人々が所有する機器を用いた利用が一般化している．つまり，無料もしくは安価で利用できる映像通話システムが多くなった近年では，専用設備を用意することは少なく，よりコストを抑えて，遠隔地の参加者とコミュニケーションを図ることができる．

テレビ会議の利用は顔やその周辺を映す場合がほとんどである．つまり，良くも悪くも顔の様子に注目しやすくなる．その中，視線の一致や変化が相手に対しどのような感情を抱かせるか，またどのような影響を及ぼすかを遠隔教育に用いた研究がある [1]. その結果，視線一致型 TV 会議システムは実対面の学習環境に近く，相手に親近感を抱きやすくなるという効果が示されている．また，そのディベート学習において，遠隔学習環境が実対面学習環境より緊張感を感じにくくなるという心理的効果が指摘されており，対面環境と異なる遠隔環境の効果が示唆されている．

また顔映像の表示方法を変えることが，話の内容への意識や集中度，及びコミュニケーションに影響することが示されている[3]. 他者の顔が個別のディスプレイに表示される「個別ディスプレイ環境」と，他者の顔映像が合成されて一つのディスプレイに表示される「共通ディスプレイ環境」が比較されている．その結果，個別ディスプレイ環境のほうが短い発言や割り込みが多くなり，会話が活発にな

1 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

ⓒ 2020 Information Processing Society of Japan

Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24

るという知見が得られている．

テレビ会議だけでなく映像を用いたコミュニケーションにおいて，顔を一つのメディアとして捉えることができる．コミュニケーションは，言語を媒体としたコミュニケーションだけでなく，声質や表情，身振りなどの言語以外の媒体によるノンバーバル・コミュニケーションとに分けて議論されてきている[4]．例えば，話し手が聞き手に何かを話しかけたが聞き手が沈黙したままの状態だとする．音声や文字だけのコミュニケーションの場合，その沈黙の意味や原因を特定できないことが多い．これに対し相手の顔を見ながらのコミュニケーションの場合，顔の表情を通して相手の理解を推察できる[5]．例えば，聞き手が目と眉を上に引き上げて，首を傾げながら話し手を見ている様子を確認することで，話し手が自らの発言内容を分かりにくいと感じていると推察できる．

以上より，顔の表情は，システム利用者の感情やコミュニケーションに影響し，テレビ会議において重要な役割を果たすことがわかる．

2.2 顔映像以外の使用について

テレビ会議システムでは映像通話を利用した共同作業を行うこともあるが，遠隔共同作業を支援するための共同作業空間として電子黒板を提供するアプローチがある．例えば，小幡らは協調的な作業において音声にテレビ電話機能を加えた効果や共通電子黒板を加えた効果を研究室実験で調べている[6]．その研究では，共有黒板がある場合，意見交換を行いながら作業を進める問題解決型の課題において，ビデオ通信を用いる環境が意見を容易に出せるという結果を得ている．一方，電子黒板を用いた場合，議論が長引き，作業時間が長くなるが，結論に対する自信が実験参加者間で一致する傾向が確認されている．従って，ビデオ通信と電子黒板を併用することによって，議論を活性化できる可能性がある．

また映像通信は使用されていないが，遠隔対話者の身体動作が音声コミュニケーションに及ぼす影響が調べられている[7]．身体動作の提示により音声コミュニケーションが円滑化される可能性を示しており，映像を用いたコミュニケーションにおいても適用できる可能性がある．

以上のように顔映像以外に共有黒板や身体動作は会議の活性化に影響を及ぼす可能性がある．一方，本研究では人間の心理面に対する影響に注目した手法を提案しており，それらとは異なる観点から会議の活性化を検討する．

3. 顔画像操作システム EmoFace とストーリー表現

3.1 提案手法について

顔映像はテレビ会議において重要な役割を担う存在である．本研究ではその顔映像をアバターに替え，アバターの顔画像の表情を人工的に操作することでテレビ会議に心理的な効果をもたらす．具体的には，顔画像を利用者の顔表情に関係なく，ポジティブな表情やネガティブな表情を提示する．ポジティブな表情はテレビ会議に対してポジティブな印象を形成し，会議を活性化する効果があると仮定している．一方，ネガティブな表情はネガティブな印象を形成し，会議に不協和感をもたらすと仮定している．従って，人工的にアバターの顔画像を演出することができると想定し，以下に述べる顔画像表示システム EmoFace を開発すると共に，感情ストーリーを考案した．

3.2 顔画像操作システム EmoFace

本研究の提案手法を実行するために，アバターの顔画像を利用者に関係なく第三者が操作できるシステムが必要である．そこで，図 1 に示す顔画像操作システム EmoFace を開発した．EmoFace は Java を用いたサーバ・クライアント型のシステムである．クライアント型プログラムは，起動オプションによって，利用者の役割を会話者または操作者の役割に設定できる．そして，顔画像の表示を変えるためのボタンの表示・非表示を設定できる．

EmoFace を使った会話者以外の操作者のインタフェースについて説明する．図 1 の左側アバター顔画像の左上に利用者 A の名前，右側アバター顔画像の左上に利用者 B の名前があらわれる．また起動時に男性か女性のアバター顔画像のどちらを使うか設定できる．なお，顔画像が示す感情の種類は 4 つであり，喜び，普通，怒り，悲しい，である．図 1 の下側にある感情を示すボタンを押すと，その感情に対応した顔画像がサーバに接続した全クライアントの対応する顔画像に反映される．

つまり，これらボタンを第三者が操作することによって 3.3 で述べる感情ストーリーを演出できる．そして，音声のみのテレビ会議システムと組み合わせることによって，提案するテレビ会議の演出を実現できる．

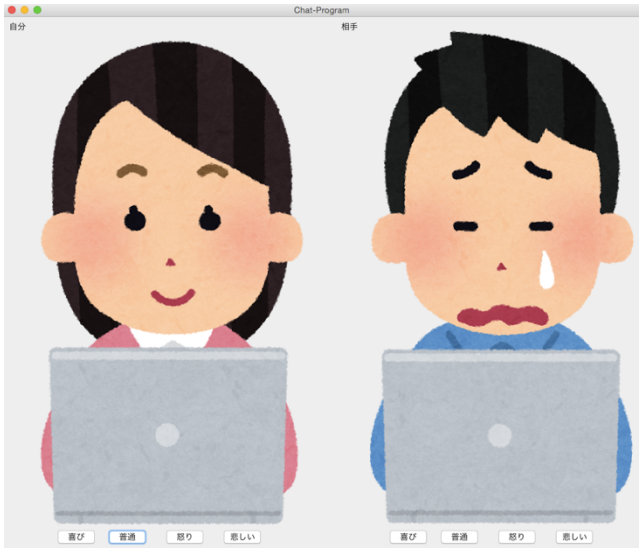


図 1　EmoFace の画面

ⓒ 2020 Information Processing Society of Japan

Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24

3.3 感情ストーリーと効果

アバター表情を演出するための感情ストーリーについて説明する．本研究の提案では，顔表情を利用者の意図と関係なく感情ストーリーに沿って提示する．本研究では，ハッピーエンドパターン(HE 型)とバッドエンドパターン(BE 型)の２通りの感情ストーリーを考えた．

HE 型の感情ストーリーはネガティブな感情を示すアバター表情からポジティブな感情を示すアバター表情に変化させるストーリーである．図２の上部に示すような時系列でアバター表情を変化させ，図２の下部に表示するような感情の起伏効果を狙っている．逆に，BE 型の感情ストーリーはポジティブな感情を示すアバター表情からネガティブな感情を示すアバター表示に変化させるストーリーである．図３の上部に示すような時系列でアバター表情を変化させ，図３の下部に表示するような感情の起伏効果を狙っている．

この感情ストーリーによって，会議参加者に心理的影響を与え，会議を活性化する効果を検討する．基本的に，喜びのアバター表情は会議参加者にポジティブな印象を与え，会議の活性化につながると考える．一方，怒りや悲しみのアバター表情は会議参加者にネガティブな印象を与え，会議に対する不安感や困惑をもたらすと考える．

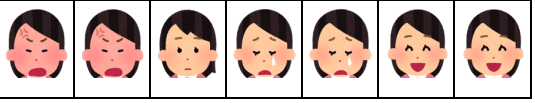
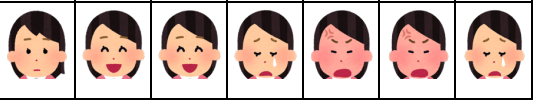
秒	開始 1.00	1.00 2.00	2.00 3.00	3.00 4.00	4.00 5.00	5.00 6.00	6.00 7.00
A さん	悲し	怒り	怒り	悲し	悩み	喜び	喜び
							
B さん	怒り	怒り	悩み	悲し	悲し	喜び	喜び
							



図 2　ハッピーエンドパターンの感情ストーリー

秒	開始 1.00	1.00 2.00	2.00 3.00	3.00 4.00	4.00 5.00	5.00 6.00	6.00 7.00
A さん	喜び	喜び	悩み	怒り	怒り	悲し	悲し
							
B さん	悩み	喜び	喜び	悲し	怒り	怒り	悲し
							

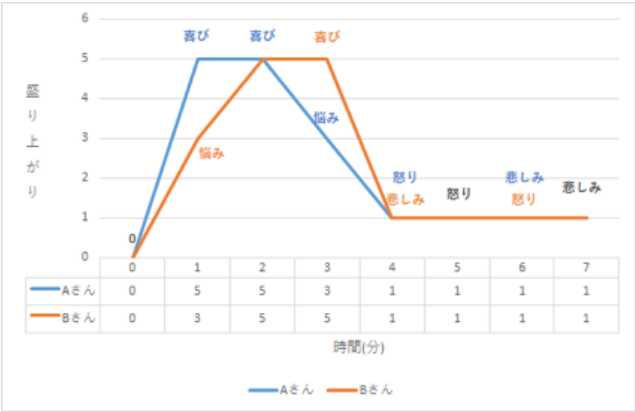


図 3　バッドエンドパターンの感情ストーリー

4. 評価実験

提案手法を評価するために３種類の会話実験を設定した．それらはハッピーエンドパターンを用いた会話実験，バッドエンドパターンを用いた会話実験，及び，ベースラインとなる通常のテレビ会議を用いた会話実験（画像通信型）である．テレビ会議システムとして Skype を用い，提案手法を用いる場合は，画像通信を使わない状態に設定し，音声通信を用いた．実験の様子を図４に示す．離れたブースに座った２人の実験参加者が会話実験を行っている様子がわかる．一方，制御者である実験者は別のブースに座り，EmoFace を操作者設定にして，提案手法を用いる実験において感情ストーリーに合わせて，顔表情を操作した．

実験参加者は北陸先端科学技術大学院大学の大学院生であり，日本人 4 人，中国人留学生 8 人の計 12 名であった．会話実験において実験参加者は２人一組のペアとなり３種類の会話実験を行った．その際，ペアは同一国籍で形成した．すべてのペアは最初にベースラインとなる画像通信型を用いた会話実験を実施し，その後，感情ストーリーを用いた会話実験を行った．その際，ペアごとに HE 型と BE 型の順番を交互に入れ替えて実施した．すべての会話実験の時間は７分であった．

ⓒ 2020 Information Processing Society of Japan

Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24



(a)　実験サイト A の様子



(a)　実験サイト B の様子

図 4　EmoFace を用いた感情ストーリー実験の様子

本実験の会話内容として，以下に示す 3 種類の議題を用いた．

- 議題 1 ：実験参加者の 2 人で食事に行く場合，洋食・和食・中華料理のどれにするか
- 議題 2 : 実験参加者の 2 人で卒業旅行に行く場合どこに行くか(国内)
- 議題 3 : 実験参加者の 2 人で卒業旅行に行く場合どこに行くか(国外)

意思決定型の会議タスクに対してビデオ映像が影響することが知られているため，意思決定に関する議題を設定した．議題 1 のみは通常テレビ会議のみに用い，議題 2 と議題 3 は提案手法を用いる 2 回の実験において，実験ペアごとに交互に用いた．

本実験の会話実験の様子は QuickTime の録画機能を用い，ノート PC 画面および音声会話を記録した．また，各会話実験の終了ごとに，会話実験の印象を尋ねるアンケートを実施した．その質問項目は「設問 A : 実験中，発言を多くできたと感じるか？」，「設問 B : 他人の表情によって発言しやすいと感じたか？」，「設問 C : 会議の結果には満足したか？」，「設問 D : 実際の顔の代わりに，アバターを見るテレビ会議には参加しやすかったか？」であった．それぞれの質問は 4 段階評価（強くそう思う，そう思う，そう思わない，強くそう思わない）で回答してもらった．一方，設問 B～D については，その選択理由を記述させた．そして，最後に自由記述欄を設けた．なお設問 D は提案手法を用いた場合にのみ回答させた．加えて，3 回目の会話実験では，「計 3 回の会議を通して，違いや変化を感じたか」について 4 段階評価をさせると共に，その理由を自由記述させた．

5.　実験結果と考察

5.1　会話分析結果

会話実験中の会話を時系列にテキストデータ化し，それぞれの発話に対して「提案，例示，説明，経験，質問，確認，回答，賛同，否定，その他」に分けてタグづけを行った．会話分析のために 10 種類のタグを用いたが，インタラクション的な会話要素を分析するために「提案・発言」「問合せ」「応答」「その他」という 4 種類の発話にまとめて会話の数や種類を調べた．なお「提案・発言」は「提案，例示，説明，経験」，「問合せ」は「質問，確認」，「応答」は「回答，賛同，否定」をまとめたものであり，それ以外は「その他」となる．また，感情ストーリーの影響を調べるために，ネガティブな表情とポジティブな表情が入れ替わる 3 分 30 秒を基準に，会話データを前半と後半に分けて分析した．以下に，会話データの例を示す．

「提案・発言」に集約した会話の例を次に示す．

●提案…「鹿児島とかいいんじゃないでしょうか」
●例示…「とりあえず，例えば中国行った時だと，焼肉とか」
●説明…「まあ，普通に JAL とか使うと，安い時だと 4500 円とか」
●経験…「私は…東京以外…私も行ったことがあります」

「問合せ」に集約した会話の例を次に示す．

●質問…「海外の方って，何回くらい行ったことありますか？」
●確認…「そうですね．心配ないです．」

「応答」に集約した会話の例を次に示す．

●回答…「だいたいねえ，20 回くらい行ったのか」
●賛同…「まあ確かに，大学で行きたいですね」
●否定…「いや，今のロシアは安いですね」

表 1 に会話分析の結果を示す．ウィルコクソンの符号付順位検定を使用したライアンの方法による多重比較を行った．全会話数において，HE 型より BE 型との間に有意差があり($p=0.028<0.05$)，BE 型の会話数が HE 型のものより多い結果となった．また，個人ごとの会話数において，BE 型が従来方式の画像通信型（$p=0.007<0.01$），HE 型（$p=0.017<0.05$）との間に有意差があり，会話数が多い結果となった．さらに，3 種類の方式を前半と後半に分け，個人ごとの会話数を比較した結果，従来の画像通信型（$p=0.024<0.05$），HE 型（$p=0.008<0.01$）の双方とも前半と後半の会話数に有意差があり，前半より後半の会話数が多くなる結果となった．以上より，ポジティブな顔画像表示

ⓒ 2020 Information Processing Society of Japan

Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24

は，会話数の増加につながるとともに，ネガティブな顔画像表示は，会話数の抑制につながる可能性がわかった．また，意外なことに，BE 型は最初にポジティブな顔画像が表示された影響のためか，全会話数は多くなる可能性がわかった．

インタラクションの種類に分けた個人ごとの会話数に対して，同様な多重比較を実施した．その結果，「提案・発言」において，HE 型は BE 型（$p=0.007<0.01$），画像通信型（$p=0.017<0.05$）との間に有意差があり，会話数が少ない結果となった．また「応答」において HE 型は BE 型（$p=0.028<0.05$）との間に有意差があり，会話数が少ない結果となった．従って，HE 型は最初にネガティブな表示が影響して，これら種類の会話数減少につながった可能性がある．

表 1　会話分析の結果

	画像通信表現	ハッピーエンド (HE 型)	バッドエンド (BE 型)
全会話数	22.3	21.7	**25.3**
前半	10.7	9.8	12.6
後半	**11.7**	**11.8**	12.7
提案・発言	**5.8**	4.6	**5.9**
問合せ	7.5	7.3	8.3
応答	8.8	8.5	**10.4**
その他	0.3	1.3	0.8

下線：有意差がある数値．太字：多い場合

5.2 アンケートおよびインタビュー結果

アンケート結果を質問ごとに図 5,6,7,8 に示す．質問それぞれに対する結果に対して，ウィルコクソンの符号付順位検定を使用したライアンの方法による多重比較を行った．その結果，会議の満足度を尋ねる設問 C に関して，BE 型は画像通信型との間に有意差があり（$p=0.018<0.05$），満足と思わない人が多く， BE 型は不満足感を参加者にもたらすことがわかった．

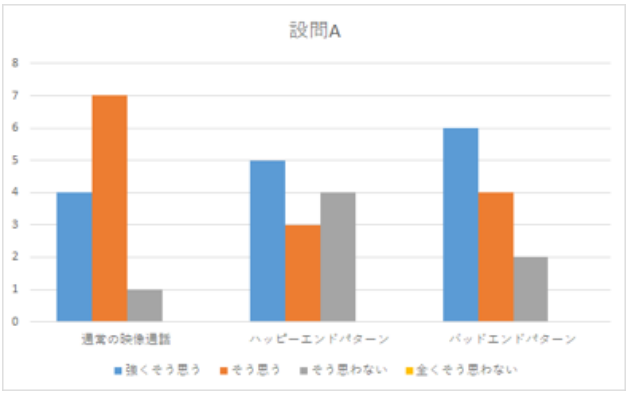


図 5　設問 A：発言を多くできたかの印象

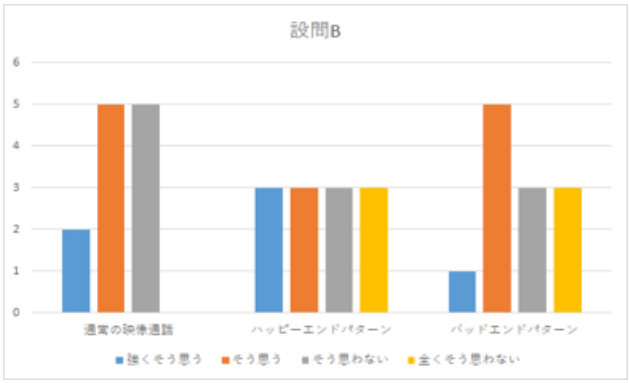


図 6　設問 B：他人の表情による発言への影響

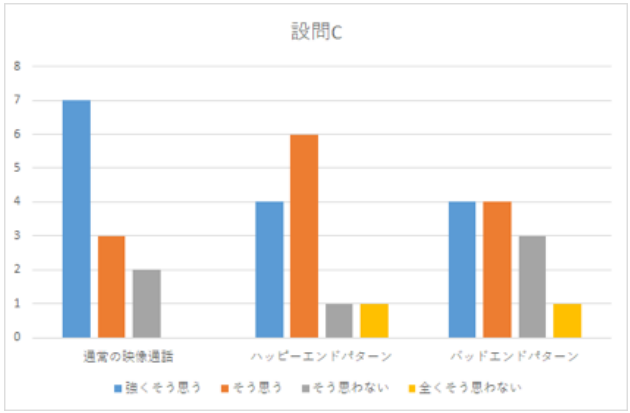


図 7　設問 C：会議の結果には満足したか

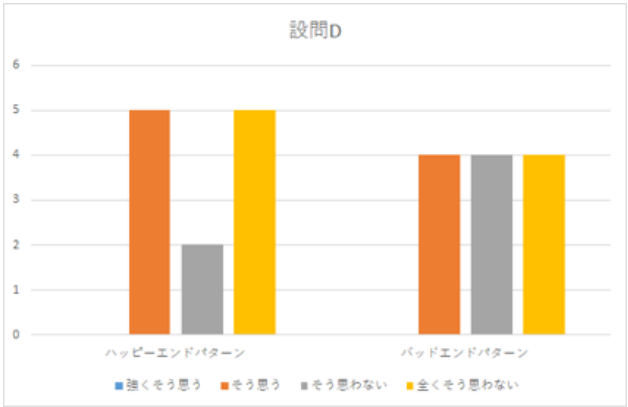


図 8　設問 D:　アバター画像会議の参加しやすさ

最後に行った，３回の実験に対する影響があるかについては，12 人中 1 人が強くそう思う，8 人がそう思う，3 人がそう思わないという結果になった．アンケートの自由記述を調べると，HE 型においては 12 人中 4 名が，BE 型においては 12 人中 2 人が「アバターの存在意義に疑問を感じた」，「アバターにあまり注目できなかった」という意見を挙げていた．しかしながら，それ以外の実験参加者は驚きや戸惑い，興味などを持つ結果であった．総じて，アバター画像は実験会話中に意識せざる負えない影響を与えていたことが伺える．

ⓒ 2020 Information Processing Society of Japan

Vol.2020-GN-109 No.36
Vol.2020-CDS-27 No.36
Vol.2020-DCC-24 No.36
2020/1/24

アバターの変化に対する感想については「感情が端的に表示されるのでそこで会話の方針が立てやすくなった」という肯定的な意見や「自分の感情とは無関係にアバターの表情が変化していくので，ちぐはぐな感じがした．特に表情が変わった瞬間は思考が一瞬止まった」,「今回のアバターはダイナミックに表情が変化するわけではないのでこの方法で会議することでの利点は感じられなかった」という否定的な意見や効果を感じないという意見もあった. 実際, これら意見を書いた参加者の会話数は，３つの会話実験の間において差はほとんどない傾向があった．

5.3 考察

実験結果よりポジティブな表情である喜びの顔画像を用いることによって，会話数や反応が多くなり会議が活性化すると考えられる．また，怒りや悲しみといったネガティブな表情は会話数を停滞させると共に，会議の満足感を下げることがわかる．つまり，アバターを用いた顔画像は利用者の心理に影響を与え遠隔会議に影響しうることがわかる．

今回の実験結果をそのまま踏まえると，ポジティブな表情をそのまま提示すればよい可能性がある．一方，ストーリーに関する心理的効果として，不満足な状態を長く続けて，満足な状態を与えると満足度が高まるということがいわれている．従って，最初に，ポジティブな顔画像を提示し，その次にネガティブな顔画像を提示し，その後，ポジティブな顔画像を提示することによって，満足度を高めるといった演出が考えられる．これら顔画像ストーリーを用いた効果の検証は今後の課題である．

また，総じて，アバター表情の効果はわかったが，アバター表情に心理的親近感を得ないためか，一部の参加者には効果的でなかった場合もあった．この対策としては，参加者の実際の表情を用いた顔画像に変更することや，現在の１分間の画像変化ではなく，よりダイナミックさを感じる演出を検討するなどの工夫が考えられる．

6. おわりに

本報告では，テレビ会議研究において，対面環境と同様な環境の実現を目指す対面環境アプローチと異なり，アバター顔画像の表情を人工的に変える心理環境アプローチを検討した．顔画像にはポジティブな表情やネガティブな表情を用意し，会話者の意図に関係ない感情ストーリーを提示することによって会議の活性化に対する影響を調べた．その結果，得られた知見は次の通りである．

- ポジティブな表情をもつ顔画像を最初に提示すると会話が活性化され，全体的に会議の会話量が増える．
- ネガティブな表情をもつ顔画像は参加者の満足度を下げる傾向がある．

以上のように，アバター顔画像の表情を人工的に制御することによって，会議の活性化を演出できる可能性がわかった．

今後の課題として，２つの知見を融合したストーリー表現の効果を検証すること，利用者の顔画像を用いることにより，ダイナミックさと尤もらしさが感じられる顔画像の演出が考えられる．

参考文献

[1] 谷田貝雅典，永岡慶三，坂井滋和，安田孝美．視線一致及び従来型 TV 会議システムを利用した 2 大学間遠隔交流学習．2011, vol.18, no.1, p.34-41.
[2] 宗森純，由井薗隆也，井上智雄．アイデア発想法と協同作業支援．共立出版, 2014.
[3] 勝山恒男，渡辺理．マルチメディアコンピューティングで触発する．知的触発に向かう情報社会 ~グループウェア維新．共立出版株式会社. 1995, p.146-158.
[4] 長谷川修，森島繁生，金子正秀．「顔」の情報処理．電子情報通信学会論文誌 A. 1997, vol. J80-A, no.8, p.1231-1249.
[5] 伊藤淳子，角所考，美濃導彦．顔メディアに注目したコミュニケーションの分析．情報処理学会第 64 回全国大会講演論文集，2002, vol.2002, no.1, p.555-562.
[6] 小幡明彦．遠隔の共同作業における映像通信，共有電子黒板の効果．情報処理学会論文誌．2000, vol.39, no.10, p.2752-2753.
[7] 玉木秀和，東野豪，小林稔，井原雅行，岡田謙一．遠隔会議における発話衝突低減手法．情報処理学会論文誌．2012, vol.53, no.7, p.1797-1806.
[8] 尾上聡，山本健太，田中一晶，中西英之．遠隔対話者の身体動作の提示による音声コミュニケーションの円滑化．情報処理学会論文誌．2013, vol.54, no.4, p.1462-1469.

ⓒ 2020 Information Processing Society of Japan