

平成29年度 卒業論文

テレビ会議システムにおける遠隔単独参加者の  
瞳拡大表示による存在感の増強

指導教員  
舟橋 健司 准教授

名古屋工業大学 工学部 情報工学科  
平成23年度入学 23115148番

名前 南 純一

# 目次

第1章	はじめに	1
第2章	テレビ会議システムにおける瞳の拡大表示	3
2.1	瞳の拡大表示手法	3
2.2	瞳を拡大表示する条件	6
2.2.1	体を前傾させたとき	6
2.2.2	口を開いたとき	8
第3章	評価実験	10
3.1	実験手続き	10
3.2	実験環境	12
3.3	実験の結果と考察	15
3.3.1	実験後アンケートによる評価と考察	15
3.3.2	会議の効率による評価と考察	17
第4章	むすび	18
	謝辞	19
	参考文献	20

## 第1章 はじめに

近年、雇用形態の多様化やオフィスの多地点化に伴い分散環境における会議の重要性が増している。分散環境において電話のような音声のみの会議では相手のソーシャルプレゼンス（遠隔地にいる相手と対面している感覚の度合）が希薄になるという問題がある。ソーシャルプレゼンスを高める方法としてテレビ会議が知られているが [1], 対面会議との比較では十分ではない。これは遠隔地にいる会議参加者の会話中に発生する表情や視線、身振り手振りなどの非言語情報 [2] が、十分に伝達されていないからだと考えられる。他者とのコミュニケーションにおいては言語情報だけでなく、非言語情報も重要になっており、Albert の実験によると、人が他人から受け取る情報の割合は非言語情報が 93 % を占めている [3]。そのため分散環境における会議で失われるしまう非言語情報を補い、会議をより効率的に行うための支援システムの研究が行われている。例えば分散環境で行われるテレビ会議では遠隔地側の参加者の視線情報が失われてしまうのでその視線情報を弱い光によって気づかせる方法 [4] や、参加者の発話要求、発話終了に関するジェスチャーを検出しそれを音で知らせることで話者交替を円滑に行う方法 [5] などがある。また、遠隔地側の参加者の身体動作をロボットと連動させることでソーシャルプレゼンスを強化する方法 [6][7] もある。しかしこれらの手法は、支援システムによって発生する光や音、ロボットの動作が会議参加者の注意を引き過ぎると考えられ、特にこれらの手法を導入してから慣れるまでは、会議の円滑な進行を阻害してしまう恐れがある。

ところで非言語情報のなかでも視線は、その送り手の意図がどこに注がれているのかを指し示すものであり、人と人が双方向的に情報を伝え受け取るための重要なコミュニケーション・ツールとして機能している。人には他者の視線を自動的に察知する視線検出器が備わっており、この視線検出器は他者の瞳の黒い部分を自動的に検出する性質を持っていることが分かっている [8]。また小田ら [9] らが行った実験

では、実験参加者は与えられた700円を自身と他の参加者との間で分け合うように教示された。瞳の絵を置いた場合では、そうでない場合と比べ参加者の分配額が増えることが確認されている。瞳の絵には規範意識を高め社会性行動を促進させる効果があると考えられる。これらのことから人は無意識のうちに他者の瞳の存在を感じていることが分かる。

そこで本研究では、多人数と遠隔地の一人に分かれて行われるテレビ会議の効率を上げるために、ディスプレイに映る遠隔地側の参加者の瞳を一時的に拡大する方法を提案する。瞳を拡大する条件と拡大される時間を設定することで、多人数側の参加者に違和感を与えることなく遠隔地側の参加者の存在感を増強でき、遠隔テレビ会議の効率が上がると考える。また提案する手法の効果を検証するために提案する手法を使わないテレビ会議との比較実験を行う。

本論文では、第二章で本提案を実現するための方法について説明し、第三章で本手法の妥当性を検証するための実験について述べる。第四章では、本研究のまとめ、今後の課題について述べる。

## 第2章 テレビ会議システムにおける瞳の拡大表示

本章では、多人数と一人で二拠点に分かれて行われる遠隔テレビ会議において、ディスプレイに映る遠隔側の参加者の瞳を一時的に拡大表示する手法について説明する。多人数対少人数で行われる分散環境におけるテレビ会議では、少人数側の参加者が議論に取り残されやすい。これは少人数側の参加者が「話したい」と感じているときに起こす表情や視線、身振り手振りなどの身体動作に多人数側の参加者が気づかないためだと考えられる。

そこで本手法では、少人数側の参加者の存在感を増強するために少人数側の参加者が話したいと感じているときに起こす身体動作を認識し、それに応じてディスプレイに映る少人数側の参加者の瞳を一時的に拡大表示する。以下では、遠隔テレビ会議が行われる二拠点のうち、人数が多い側の会議室を「現地会議室」と呼び、一人だけの側を「遠隔会議室」と呼ぶ。またそれぞれの拠点での会議参加者を「現地参加者」、「遠隔参加者」と呼ぶ。

### 2.1 瞳の拡大表示手法

本手法は、両会議室の参加者の様子をカメラで映しながら行う遠隔テレビ会議を想定しており、そのカメラによって映し出される遠隔参加者の瞳を認識し、一時的に拡大表示する。まず、OpenCVの標準ライブラリを利用して瞳の認識を行い、瞳の周りを囲んだ矩形(図 2.1)の座標を得る。(2.2.1項, 2.2.2項で顔の認識, 口の認識を行う場合も同様にOpenCVの標準ライブラリを利用する。)

次に、瞳を拡大表示するため矩形に囲われた画像を、魚眼レンズを通してその画像を見たような画像に変換する。

認識した矩形の横幅を  $w_e$ , 縦幅を  $h_e$ , 矩形内のある画素の座標を  $(x, y)$  とする。画

像変換における歪みの大きさを *weight* として,  $r, rg$  を式 2.1, 式 2.2 のように定める.

$$r = \sqrt{\frac{\left(x - \frac{w_e}{2}\right)^2 + \left(y - \frac{h_e}{2}\right)^2}{\left(\frac{w_e}{2}\right)^2 + \left(\frac{h_e}{2}\right)^2}} \quad (2.1)$$

$$rg = r^{weight} \quad (2.2)$$

ここで  $r$  は, 矩形の中心から矩形内の画素  $(x, y)$  への距離と, 矩形の対角線の半分との比の平方根である. また  $rg$  は  $r$  を歪みの大きさの程度を決定する重みだけ乗じたものである. この  $r, rg$  を用いて変換後の座標  $(vx, vy)$  を式 2.3, 式 2.4 のように求める.

$$vx = \frac{(x - w_e) \times rg}{r} + w_e + 0.5 \quad (2.3)$$

$$vy = \frac{(y - h_e) \times rg}{r} + h_e + 0.5 \quad (2.4)$$

元の画像と瞳の拡大処理をした後の画像を図 2.2, 図 2.3 に示す.

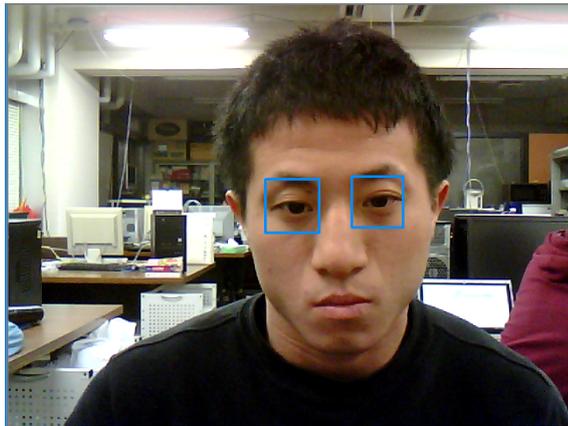


図 2.1: 両目の認識



図 2.2: 元の画像



図 2.3: 瞳の拡大

## 2.2 瞳を拡大表示する条件

複数人の会話において、誰かが話者になると他の人が聞き手になり、話者が話を終わると自然に他の人が次の話者になる。このような話者の自然な順番交替は、質問応答などの言語により行われるだけでない。表情、身体動作などの非言語により行われることもあり議論進行には重要である。身体は会話の順番交替において、次に話し出したい、まだ聞き続けたい態度を表現していると考えられる。

人が「話したい」と感じたときに、どのような身体動作をするだろうか。話しを聞いている人が会話中に徐々に身を乗り出していけば、その人は次に「話したい」態度を表現していると解釈できる。また徳永らの実験 [10] では人は「話したい」と感じたときには発話の開始を連想させる口を少し開く仕草をすることがわかった。

本手法は遠隔参加者の存在感の増強により遠隔テレビ会議の会話を円滑にし、会議の効率を高めるのが目的である。したがって、遠隔参加者が「話したい」と感じたときに起こす身体動作として考えられる「体を前傾させたとき」、「口を開いたとき」に瞳を一時的に拡大表示する。

### 2.2.1 体を前傾させたとき

人は体を前傾させるときに顔も同時に前に動く。ここで顔の横幅の変化率により「体を前傾させた」かどうかの判定を行う。

ある時刻  $t$  において認識した顔の横幅を  $w_f(t)$  とすると時刻  $t - D_1$  における顔の横幅は  $w_f(t - D_1)$  と表せる。ここで  $D_1$  は変化率を判断するための時間間隔である。時刻  $t - D_1$  の顔の横幅から、時刻  $t$  の顔の横幅の変化率を求め、閾値  $c_1$  を超えたときに「体を前傾させた」と判定する (式 2.5, 図 2.4)。

$$c_1 < \frac{w_f(t)}{w_f(t - D_1)} \quad (2.5)$$

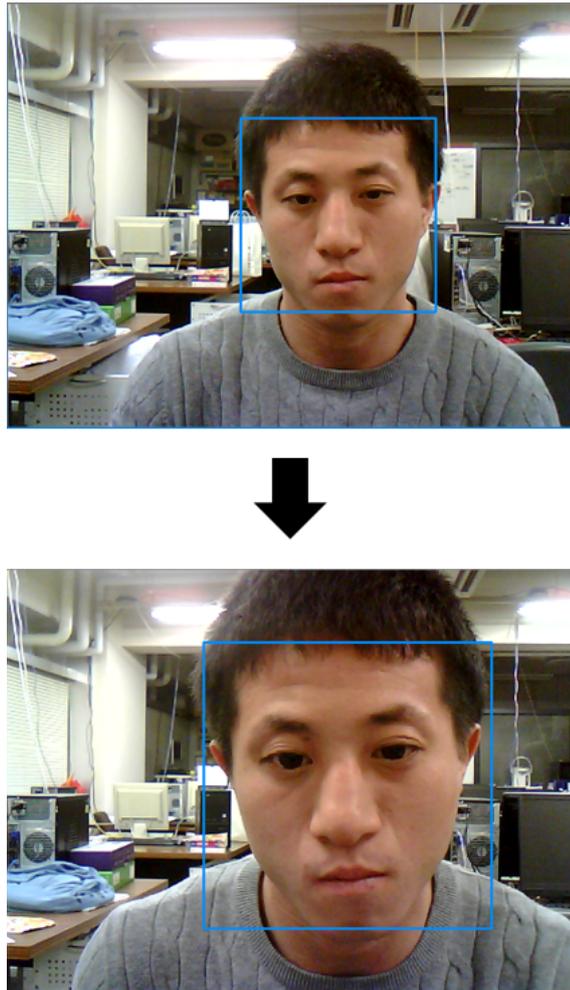


図 2.4: 体の前傾による瞳の拡大表示

### 2.2.2 口を開いたとき

画像から口の認識を行った場合、特に OpenCV の標準ライブラリを使用した場合には、口を開く動作をしたときに口の認識領域の横幅が増加することが分かっている。ここで顔の横幅に対する口の横幅の割合を  $P(t)$  とする。また、時間間隔  $D_2$  において  $P(t)$  の平均値  $P_{ave}(t)$  を求める。これらの比により「口を開いた」かどうかを判定する。口の横幅の変化率でなく、顔の横幅に対する口の横幅の割合によって判定するのは、口を動かさずに顔を近づけたり遠ざけたりする動作をするだけで、口の横幅は変化するためである。また  $P(t)$  の平均値を用いたのは、認識誤差による影響を避けるためである。

ある時刻  $t$  において認識した顔の横幅  $w_f(t)$  における口の横幅  $w_m(t)$  の割合  $P(t) = w_m(t)/w_f(t)$  を求める。時間間隔  $D_2$  における  $P(t)$  の平均値  $P_{ave}(t)$  は式 2.6 で表すことができる。

$$P_{ave}(t) = \frac{\sum_{k=t-D_2}^t P(k)}{D_2} \quad (2.6)$$

その上で、 $P_{ave}(t)$  と  $P(t)$  の比が閾値  $c_2$  を超えたとき、「口を開いた」と判定する (式 2.7, 図 2.5)。

$$c_2 < \frac{P(t)}{P_{ave}} \quad (2.7)$$

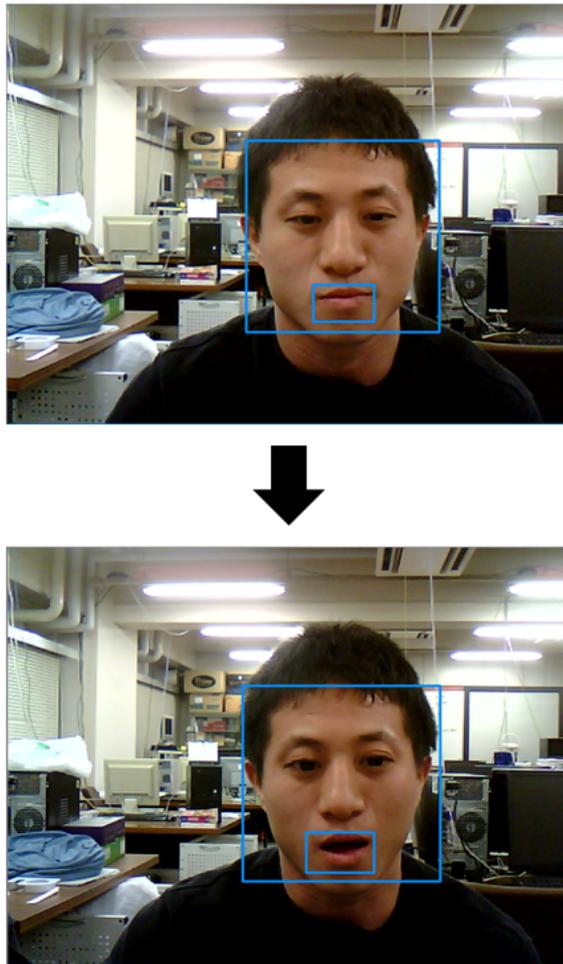


図 2.5: 口を開くことによる瞳の拡大表示

## 第3章 評価実験

本章では、本手法を用いることで、遠隔参加者の存在感が増強され、それによって遠隔テレビ会議の効率が高まるかどうか評価する。そのために本手法を用いたテレビ会議と本手法を用いていないテレビ会議のそれぞれを行い、実験後に回答してもらったアンケートの結果と会議の効率を比較し考察する。

### 3.1 実験手続き

被験者となる大学生8人(男性7人, 女性1人)を4人ずつの2グループに分けた。それぞれのグループで現地会議室3人, 遠隔会議室1人に分かれて行う遠隔テレビ会議を4回行った。この4回のうち本手法を用いたテレビ会議, 本手法を用いないテレビ会議をそれぞれ2回ずつである。また本手法は, 会議参加者に違和感を与えずに遠隔参加者の存在感を増強することを目的としている。本手法を用いたテレビ会議であることを会議参加者に知らせてしまうと, それに気を取られてしまうことが考えられるため, 本手法を用いているか用いていないかを伝えずに会議を行ってもらった。また一回の会議の時間をそれぞれ5分, 会議の間の休憩を3分とした。

本実験で行った会議の順序を表3.1に示す(以下では, 本手法を用いたテレビ会議を“Exp.Eye”, 本手法を用いないテレビ会議を“Video”と表記する)。第一グループと第二グループでは本手法を用いたテレビ会議, 本手法を用いないテレビ会議の順序を入れ替えており, 各条件の会議で遠隔会議への慣れや, 会議をすることによる疲れなどの条件に差が出ないようにした。会議が2回終わるごとに現地参加者, 遠隔参加者それぞれにアンケートを取った。またアンケートを取る前にどちらが本手法を用いた会議か説明した。

会議の方式としてブレインストーミングを採用した。ブレインストーミングは他

の会議方式に比べて発話数や発話時間がテーマに影響されにくく、均一な条件で実験を繰り返せるためである。ブレインストーミングのテーマは「食堂の改善法」のように被験者である学生に身近な話題とし、個人の興味によって発話数に極端な差が出ないように工夫した。

会議の順序	第一グループ	第二グループ
1	Exp_Eye	Video
2	Video	Exp_Eye
-	(アンケートに回答、遠隔参加者の交替)	
3	Exp_Eye	Video
4	Video	Exp_Eye
-	(アンケートに回答)	

表 3.1: 会議の順序

### 3.2 実験環境

会議室間の音声の通信は Skype[11] を用いて行った。現地会議室の音声の入出力装置には YAMAHA の会議用マイクスピーカー “PJP-10UR” [12] を用いた。また遠隔会議室の音声の入出力装置には BUFFALO のヘッドセット “BSHSUH13” [13] を用いた。現地会議室の映像を PC 内蔵カメラにより撮影し, Skype により遠隔会議室側の PC (CPU: Inter Core i5-6500 3.20GHz RAM:8.0GB) へ送信した。また遠隔会議室の映像を BUFFALO の USB カメラ “BSW20K04” により撮影し, 現地会議室側の PC (CPU: Inter Core i5-7200U 2.50GHz RAM:4.0G) へ有線接続し送信した。

現地会議室側の PC に 2 章で説明した本手法を実現するためのシステムを C++ で実装した。瞳を拡大表示する時間, 指定時間  $D_1, D_2$ , 閾値  $c_1, c_2$  の値を表 3.2 のように設定した。また瞳の拡大表示処理による歪みの大きさを決定する weight は 1.15 に設定した (図 3.2)。

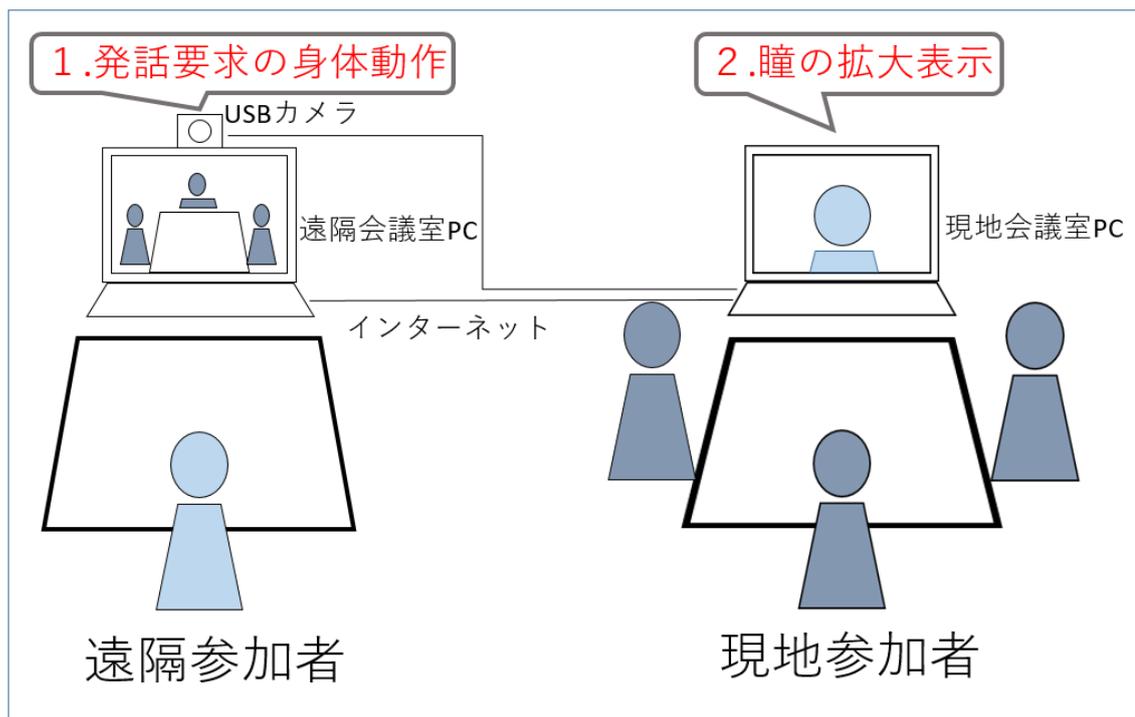


図 3.1: 本手法を用いたシステムの全体構成

	値
瞳を拡大表示する時間	1 秒間
指定時間 $D_1$	0.27 秒
指定時間 $D_2$	0.53 秒
閾値 $c_1$	1.13
閾値 $c_2$	1.08

表 3.2: 各パラメータの設定



図 3.2: 拡大処理後に表示される瞳の大きさ



図 3.3: 現地会議室の様子

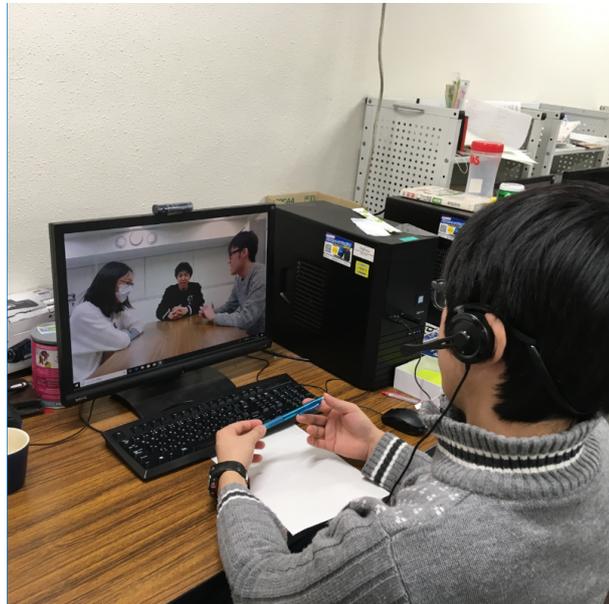


図 3.4: 遠隔会議室の様子

### 3.3 実験の結果と考察

#### 3.3.1 実験後アンケートによる評価と考察

実験後に現地参加者、遠隔参加者それぞれに回答してもらったアンケート結果を表 3.3, 表 3.4 に示す。現地参加者、遠隔参加者への  $n$  個目の質問をそれぞれ“現  $n$ ”、“遠  $n$ ”と表現する。また回答は1から5の5段階評価であり、数値は平均値を示している。

表 3.3 の質問“現 1”、“現 2”を見ると本手法を用いていない会議では、それぞれ 1.17, 1.58 と低いのに対し、本手法を用いた会議では 3.25, 2.75 と大きな改善が見られる。これは、遠隔参加者が「話したい」と感じているときに瞳を拡大表示することで現地参加者が、遠隔参加者の視線を強く感じられたことを示している。またそれにより遠隔参加者が話し出そうとしていることに気づいたと考えられる。また質問“現 3”を見ると本手法を用いた条件の方が低い数値になっており、画面に違和感を覚えた現地参加者が本手法を用いていない会議より多かったことが分かる。

表 3.4 の質問“遠 1”、“遠 2”を見ると、本手法を用いた会議の方が高い数値を示している。しかし、それぞれの質問に対して  $t$  検定を用いたところ  $p$  値が 0.18, 0.76 となり、有意差は見られなかった。質問“現 2”に対して  $t$  検定を用いたところ  $p$  値が 0.0004 となり、有意差が見られた。これらのことから、提案手法を用いた会議の方が遠隔参加者が話し出そうとしていることに現地参加者はより多く気づいたと考えられる。しかし現地参加者は遠隔参加者の話を聞こうとしなかった、または、現地参加者は、遠隔参加者の話を聞こうとしたが、遠隔参加者はそれに気づかなかったとも考えられる。

記号	質問事項	Video	Exp.Eye
現 1	あなたが現地参加者と話しているとき、遠隔参加者の視線を感じたか。	1.17	3.25
現 2	あなたが現地参加者と話しているとき、遠隔参加者が話しだそうとしていることに気づいたか。	1.58	2.75
現 3	遠隔参加者が映る画面は自然だったか。	4.67	3.83

表 3.3: 現地参加者へのアンケート結果

記号	質問事項	Video	Exp_Eye
遠1	あなたは現地会議室での会話中に、注意を払われていると感じたか.	2.5	3
遠2	あなたは現地会議室での会話中に、話し始めることは容易だと感じたか.	2.75	3

表 3.4: 遠隔参加者へのアンケート結果

### 3.3.2 会議の効率による評価と考察

各条件での会議の効率を評価するため、実験の様子を撮影したビデオを分析した。ここで「会議の効率」とは、会議の目的に沿った発言を会議の時間内にどれだけできたかを表現している。この実験で行った会議は、3.1節で示したようにブレインストーミングを採用している。ブレインストーミングは、お互いに意見を出し合うことで沢山のアイデアを生産し問題の解決に結びつける集団発想法である。そこで、「会議の効率」が高いとは新規のアイデアの提案回数が多いことであると考え、新規のアイデアの提案回数を調査した。

本手法を用いた会議と用いていない会議の提案回数と遠隔参加者の提案回数の合計を表3.5に示す。本手法を用いた会議の方が、新規アイデアの提案回数が多く会議の効率が高くなっている。また遠隔参加者の新規アイデアの提案回数の割合も本手法を用いた会議の方が高くなっている。よって本手法を用いることにより遠隔参加者が「話したい」と感じたときに話し出すことが出来たと考えられる。そしてそれにより会議の効率も高くなったと考える。

	Video	Exp_Eye
会議全体	40回	50回
遠隔参加者	9回	13回
(遠隔参加者の提案回数)/(会議全体の提案回数)	0.23	0.26

表 3.5: 新規アイデアの提案回数と遠隔参加者が提案した割合

## 第4章 むすび

本研究では、多人数と遠隔地の一人に分かれて行われるテレビ会議の効率を高めるために、遠隔参加者が「話したい」と感じたときに起こす身体動作に応じてディスプレイに映る遠隔参加者の瞳を一時的に拡大表示することで、遠隔参加者の存在感を増大させる手法を提案した。作成したシステムでは、人が「話したい」と感じたときに起こす身体動作として考えられる「体を前傾させたとき」、「口を開いたとき」に瞳を一時的に拡大表示させた。提案した手法の妥当性を検証するために、本手法を用いたテレビ会議と本手法を用いていないテレビ会議のそれぞれを行い、会議が終了した後に会議参加者に評価してもらった。その結果、本手法を用いた会議の方が、現地参加者が遠隔参加者の視線を強く感じられ、それにより遠隔参加者が話し出そうとしていることに容易に気づくことを確認できた。また会議の効率についてそれぞれ分析すると、本手法を用いることにより遠隔参加者が「話したい」と感じたときに話し出すことができたと考えられ、それにより会議の効率も高くなることが確認出来た。

本研究の課題として、経験的に設定した値(瞳を拡大表示する時間, 指定時間  $D_1, D_2$  閾値  $c_1, c_2$ , 瞳の拡大表示処理による歪みの大きさを決定する *weight*) を実験により適切に調整することが挙げられる。また「話したい」と感じたときに起こす身体動作は本手法で取り上げた二つだけでない。認識する身体動作を増やすことで、遠隔テレビ会議の効率が高まると考えられる。また今回の実験方法では、実施回数が限られていた。より多くの実験を行うことで、最適な条件を調査することも今後の課題である。

## 謝辞

本研究を進めるにあたって、日頃から多大なご尽力を頂き、ご指導を賜りました名古屋工業大学、舟橋健司 准教授、伊藤宏隆 助教に心から感謝致します。

また、本研究に多大なご協力頂きました舟橋研究室諸氏に心から感謝致します。

## 参考文献

- [1] Ellen A.Isaacs, John C.Tang, “What Video Can and Can’t Do for Collaboration : A case study”, MULTIMEDIA '93 Proceedings of the first ACM international conference on Multimedia, 199-206, 1993.
- [2] W・フォン・ラフラー＝エンゲル, “ノンバーバルコミュニケーション:ことばによらない伝達”, 本名信行 (編), (社) 大修館書店, 1981.
- [3] 松下 温, “コラボレーションとコミュニケーション”, 岡田謙一 (編), (社) 共立出版, 1995.
- [4] 藪田 幹文, アルニーラティカン, “人数が不均衡な遠隔テレビ会議における弱い光を用いた視線アウェアネス”, 情報処理学会論文誌, Vol.58, No.1, 166-75, 2017.
- [5] 藪田 幹文, 増田 雄亮, “分散環境における話者交替のアウェアネス支援”, 情報処理学会論文誌, Vol.56, No.1, 126-136, 2015.
- [6] 田中 一唱, 中西 英之, 石黒 浩, “ロボット会議:物理的実体を介した身体動作の提示によるソーシャルテレプレゼンスの強化”, 情報処理学会論文誌, Vol.57, No.1, 209-217, 2016.
- [7] 芝 智偉, 今井 順一, “人物動画像とロボットの連動により話者の存在感を表出する遠隔対話システム”, 日本感性工学会論文誌, Vol.14, No.4, 497-504, 2015.
- [8] 吉崎 一人, 瀧澤 志保, 安田 恭子, 亀井 宗, 中村 純一, 佐々木 洋, 杉本 助男, “他者の視線が視覚的注意に与える影響”, 第五回認知神経心理学研究会, 2002.
- [9] Ryo Oda, Ryota Ichihashi, “The Watching Eyes Effect on Charitable Donation is Boosted by Fewer People in the Vicinity”, Letters on Evolutionary Behavioral Science, Vol.7, No.2, 2016.

- [10] 徳永 弘子, 武川 直樹, 寺井 仁, 湯浅 将栄, 大和 淳司, “会話における参与役割志向態度に基づく順番交替分析” 知能と情報 (日本知能情報フレンジィ学会誌), Vol25, No.5, 889-900, 2013.
- [11] Microsoft Skype, “<https://www.skype.com/ja/>”.
- [12] YAMAHA PJP-10UR, “<https://sound-solution.yamaha.com/products/uc/pjp-10ur/index>”.
- [13] BUFFALO BSHSUH13, “<http://buffalo.jp/product/multimedia/headset/bshsuh13/>”.
- [14] BUFFALO BSW20k04, “<http://buffalo.jp/products/catalog/supply/multimedia/webcamera/200/bsw20k04/>”.