

学士特定課題研究論文

全天周ビデオ通信端末を用いた非対称型 コミュニケーションに関する研究

栗岡 保

17B05518

東京工業大学
情報理工学院
情報工学系

指導教員 小池 英樹

2021 年 1 月

概要

学士特定課題研究論文は、シングルカラムでページ数に制限はない。

目次

概要	ii
第 1 章 序論	1
1.1 本章の概要	1
1.2 本研究の背景	1
1.3 本研究の目的	3
1.4 本論文の構成	3
第 2 章 関連研究	4
2.1 360 度カメラに関する研究	4
2.2 全天周球状ディスプレイに関する研究	4
2.3 ビデオ会議中の視線に関する研究	4
2.4 まとめ	4
第 3 章 球体ディスプレイを用いた視線共有システムの提案	5
3.1 従来型ビデオ会議アプリケーションの問題点	5
3.2 提案する視線共有システム	5
3.3 技術的課題	6
第 4 章 OmniEyeBall 対 PC 間通信の実装	7
4.1 ハードウェア実装	7
4.2 ソフトウェア実装	7
第 5 章 実験	8
5.1 実験の目的	8
5.2 実験用アプリケーション	8
5.3 実験手順	8
5.3.1 実験設定	8
5.3.2 実験項目	9
5.3.3 被験者の動き	9
5.4 実験結果	9
5.4.1 定量的評価	9
5.4.2 定性的評価	10
5.5 実験結果から得られる知見と考察	10

第 6 章	考察	12
6.1	本システムの問題点	12
6.2	本システムの応用	12
6.2.1	対話相手表示形式	12
6.2.2	横顔生成形式	12
6.2.3	おかしら会議形式	12
6.3	今後の展望	12
第 7 章	結論	13
	謝辞	14
	参考文献	15
	参考文献	16

図目次

1.1	zoom	1
1.2	全天球パノラマ画像	2
1.3	theta V	2

表目次

5.1	平面ディスプレイ使用時	10
5.2	OEB 使用時	10
5.3	5 段階評価項目の結果	10

第 1 章

序論

1.1 本章の概要

本章では、本研究の背景、目的、本論文の構成について述べる。

1.2 本研究の背景

近年、様々なビデオ会議アプリケーション (図 1.1) が登場している。その例としては、zoom [1] や Google Meet [2] が挙げられる。新型コロナウイルス感染症の流行により、ビデオ会議の需要がさらに高まりつつある。

しかし、ビデオ会議における様々な問題点も指摘されている。例えば Roel [3] は、カメラの視覚外の情報や、人物の情報の不足のために、対面時のようなインタラクションを得られないことを指摘している。



図 1.1 zoom

[1]



図 1.2 全天球パノラマ画像



図 1.3 theta V

[4]

一方で、昨今は誰でも気軽に全天球映像 (図 1.2) を撮影することができるようになっている。その例として、全天球カメラ (360 度カメラ, 全天周カメラ, 全方位カメラなどともいう (図 1.3)) を使用して、全天球映像をヘッドマウントディスプレイを用いて観覧したり、パノラマ映像として動画や静止画を保存することが出来るようになっている。

全天周カメラの使用により、カメラの視野の問題は解決される。実際に、Anthony ら [5] は、全天球の視野の広さによって、リモートユーザーがローカルユーザーの環境をより早く理解できると結論付けている。だが、Johnson ら [6] によって、パノラマ視野によって映像の複雑さが増し、より大きな認知負荷を必要としたことも示されている。

認知負荷を削減するため、パノラマ映像を 3 次元に表示する、球体プロジェクターを用いた方法が

挙げられる。球体プロジェクターの利用としては、Li らの OmniEyeBall

1.3 本研究の目的

1.4 本論文の構成

第 2 章

関連研究

2.1 360 度カメラに関する研究

2.2 全天周球状ディスプレイに関する研究

2.3 ビデオ会議中の視線に関する研究

2.4 まとめ

第 3 章

球体ディスプレイを用いた視線共有システムの提案

3.1 従来型ビデオ会議アプリケーションの問題点

ビデオ会議では、視野の狭さにより、参加者同士の状況が分かりづらいという欠点がある。そのため、コミュニケーションが円滑に行われない場合がある。全天球ビデオカメラを使用することで視野の狭さの問題が解決するが、一方で映像が複雑になり、その影響でやはり快適なコミュニケーションが行われない可能性がある。

一方で、先ほど述べた OmniEyeBall に代表される、全天周ディスプレイを用いたビデオ会議が考えられる。全天周ディスプレイでは、全天球映像が球を用いた立体的かつ自然な形で表現され、全天球映像を視認しやすくなる。

しかし、上記の方法でなお、視線情報の問題が存在する。従来のビデオ会議では、参加者がどの参加者を見ているかなどの視線情報が伝わらない。以下では、PC の使用者 1 人と OmniEyeBall の使用者複数人の非対称な通話状況において、PC 使用者の視線情報を伝える方法を提案する。以下で提案したシステムを用いて、従来のビデオ通話アプリケーションを用いた場合よりも PC の使用者の視線情報が OmniEyeBall の使用者に正確に伝わることを最終的な目標である。

3.2 提案する視線共有システム

PC 側で提案するシステムを以下に述べる。平面ディスプレイ上に OmniEyeBall に取り付けた全天球ビデオカメラのパノラマ映像を表示する。映像上で、参加者の顔をクリックすることで、その参加者が正面に来るようにパノラマ映像を平行移動させる。そのようにして、常に正面を見ていれば見たい人を見ているという状態を作り出す。

(ここに、画像をスライドさせる図を乗せる)

以降は、OmniEyeBall 側で提案するシステムを述べる。PC 側で使用する全天周ビデオカメラの映像を球体ディスプレイ上に表示する。PC 側のクリック操作に合わせクリックされた人物の方向に顔が向くように映像を回転させる。こうすることで、PC 側で見ている人間とみられた人間が向き合う状態を作る。

この時、例えば 3 人が 120 度の間隔で OmniEyeBall の周りに座っている状況を考えると、顔を向

けられていない2人は、ほとんどPC側の参加者の顔を視認できない。このような場合、この2人には画面上部に小さなウィンドウを表示し、そこにPC側参加者の顔部分を表示する。顔部分を表示すると、その顔は大抵正面方向を向いており、自身が見られていると錯覚する恐れがある。それを防ぐため、PC参加者の顔ウィンドウの隣には、見られている参加者の顔ウィンドウを表示させる。

(ここで図の挿入)

また、このままではPC側の使用者が、特定の個人を見ていない場合(例えば、全員に対して語りかける場合)を区別できない。このため、特定の個人を見ていない場合は、全員に対して画面上部にPC参加者のみの顔ウィンドウを表示させる。クリックした人物の方向を向く仕様では、常に前回のクリックの結果が残り、いずれかの人物の顔方向にPC参加者の顔が向くことになる。この方向に対して、あえて顔ウィンドウを顔映像に被せることによって、大きく表示される方の顔映像の視認性を下げ、自身が見られているという錯覚を防ぐことが狙いである。

(ここで図の挿入)

3.3 技術的課題

ハードウェア部分の実装に関しては、全天球ビデオカメラ2つ、PC2台、及びOmniEyeBall1台を用意すればよい。

一方、ソフトウェア部分の実装に関しては、いくつか準備の必要がある要素が存在する。

まず、ビデオストリーミングの方法について考える必要がある。UDP等を用いて映像情報の送受信を行う必要がある。しかし、フレーム落ちの問題など、現在使用されているビデオ会議アプリケーションのような堅牢で高速な通信を実現するためには、高度な技術と時間を要する。そこで、映像送受信については、すでに利用されているビデオ会議アプリケーションを用いることにした。

しかし、ただ映像を送受信しては、PC側の操作に合わせて映像を回転させるといった処理が出来ない。PC側では全天球ビデオカメラの映像を加工した後に送信を行う必要がある。あるいは、PC側で表示するUIにおいて、OmniEyeBall側の全天周パノラマ映像をそのまま表示させるのではなく、やはりクリックした人物の顔を正面に移動させるなどの必要がある。ここで、加工した映像をキャプチャしてビデオ会議アプリケーションに認識させる方法と、一方で、送信されてきた映像をキャプチャーして、以下で説明する画像加工ライブラリで扱えるようにする方法を考えなければならない。

この両方を解決するために、今回は仮想ビデオカメラを用いる。仮想ビデオカメラとは、OBS-virtual-cam [7] や UnityCapture [8] 等のような、特定のアプリケーションの映像をリアルタイムでキャプチャし、その映像をあたかも接続したビデオカメラの映像のように扱えるものである。これを用いることで、ビデオ会議アプリケーションや画像加工ライブラリが、加工済みの映像や送信されてきた映像をビデオカメラの映像として認識し、ストリーミングや加工が出来るようになる。

(以下に、映像送受信 + 加工のフローの図を乗せる)

次に、映像の加工方法について述べる。映像の加工はpythonなどで利用可能な画像加工ライブラリを用いた。同時にPC画面に表示するUIもpythonのライブラリで作成した。

第 4 章

OmniEyeBall 対 PC 間通信の実装

4.1 ハードウェア実装

全天周ビデオカメラ

全天周ビデオカメラとして, PC 側では RICOH THETA S, OmniEyeBall 側では RICOH THETA V を用いた.

球体ディスプレイ

全天球動画の出力装置としては渋谷光学の球体プロジェクターである Glomal350 を使用する.

4.2 ソフトウェア実装

第 5 章

実験

5.1 実験の目的

- ・ OEB と 4 章で紹介した視線情報共有手法を用いて、一人が多人数に対しプレゼンテーションを行う際、被検者がどのように感じたかを定性的に調べる
- ・ 意図した視線が正確に伝わっていたか、定量的な評価も行う
- ・ 定量的データとアンケート結果から、本システムの優れた点と改善すべき点を明らかにし、今後の研究へとつなげる

5.2 実験用アプリケーション

- ・ (4 章と重複? 要修正)
- ・ 画面のレイアウトについて詳しく記述する

5.3 実験手順

5.3.1 実験設定

- ・ プレゼンター 1 人、傍聴者 3 人によるプレゼンテーションを行った
- ・ プレゼンと質問時間合わせて 10 分程度
- ・ 提案アプリケーション + OEB 使用時と、Zoom のみを使用した場合の 2 度行った
- ・ プレゼンター負担削減のため、プレゼン資料は事前に作成した (2 種類)
- ・ また、各傍聴者に、最低一回の質問機会を設けるため、各プレゼンに対し 3 つの質問項目を事前に用意した
- ・ 実験環境は以下の通り
- ・ 共通: THETA V と THETA S を使用、プレゼンターは全天球カメラに対し 30cm ほど離れていた、プレゼンターの PC 詳細 (後で調べる)
- ・ Zoom 使用時: 質問者側は Mac book pro (栗岡机上: 後で詳しく) を使用 (横に 3 人並んで一つの画面を見た)
- ・ OEB 使用時: 質問者側は OmniEyeBall を使用

5.3.2 実験項目

- ・ 定量的な項目と定性的な項目それぞれ設けた
- ・ 以下定量的項目
- ・ プレゼンター側が誰を見ていたのかをリアルタイムで記録
- ・ 同時に、質問者側は、目が合っていると感じる間、キーを押しっぱなしにして時間を記録
- ・ 見ているときに目が合ったと感じた割合 (真陽性率)
- ・ 見ていないのに目が合ったと感じた割合 (偽陽性率) を算出
- ・ 以下定性的項目
- ・ System Usability Scale に基づいた 5 段階評価項目 10 個
- ・ 独自の評価項目 (5 段階評価 2 つ + 記述式項目)
- ・ OEB を使用したケースと PC を使用したケース両方を経験した被検者に対し、どちらの方が好ましいかを問う記述式項目を設けた

5.3.3 被験者の動き

- ・ OEB 使用時と平面ディスプレイ使用時の 2 回の実験を行った (プレゼンター 1 人、質問者 3 人)
- ・ 平面ディスプレイ使用の実験の後、10 分程度の休憩時間を設け、OEB を使用する実験に移った
- ・ 実験の開始前に、使用するアプリケーションの操作説明を行った
- ・ また、5 分程度、事前準備したスライドの内容をある程度覚えてもらった (プレゼン中、出来るだけ映像の方を見てもらうため)
- ・ プレゼンターは、プレゼン中、3 つ存在するプレゼン内容のチェックポイントに合わせ、ランダムな順番で 1 人ずつ、対応する被験者を見てもらった
- ・ また、質問中は質問を受けている被験者の方向を見るように指示した
- ・ OEB 使用時は、見る被験者の映像をクリックし続けることで、視線情報のログを取った
- ・ 平面ディスプレイ使用中は、被検者それぞれに対応したキー (映像に移っている左の人間から A,S,D) を押しっぱなしにしてもらうことで視線情報のログを取った
- ・ 一方、質問者側の被験者 3 人には、プレゼン中はプレゼンを聴くように指示した
- ・ 質問は、事前に準備した質問 3 つをランダムな順番で行った (プレゼンターから質問するように促される)
- ・ 上記の質問が終わったのちは各質問者から自由に質問してもらい、プレゼンと質問合わせて約 10 分使用した

5.4 実験結果

5.4.1 定量的評価

- ・ 定量的評価の結果は以下の通り

	真陽性率	偽陽性率
被検者 1	0.05	0.02
被検者 2	0.57	0.12
被検者 3	0.00	0.00

表 5.1 平面ディスプレイ使用時

	真陽性率	偽陽性率
被検者 1	0.06	0.07
被検者 2	0.87	0.28
被検者 3	0.19	0.00

表 5.2 OEB 使用時

5.4.2 定性的評価

- ・定性的評価の、5 段階評価項目の結果は以下の通り (4 人の平均値を算出)

	平面時	OEB 時
質問 1	2.75	3.5
質問 2	2.5	3.25
質問 3	3.5	3.25
質問 4	1.75	3.75
質問 5	3	3.5
質問 6	3.5	3.25
質問 7	4.75	3.75
質問 8	2.25	1.5
質問 9	3.5	3.5
質問 10	1.5	2.5
質問 11	1.75	3.25
質問 12	4	3.25

表 5.3 5 段階評価項目の結果

- ・記述式項目については 5-5 節にて言及する

5.5 実験結果から得られる知見と考察

・表 5.1,5.2 を見ると OEB 使用時, 平面ディスプレイ使用時共に真陽性率 (プレゼンターが見ていた時間の中で, 目が合ったと感じた時間の割合) の最大値, 最小値の差が非常に大きい (極端な例は平面ディスプレイ使用時の 0.00(一度もキーを押さなかった))

- ・自由記述意見の 1 つ「発表者の顔の画像が暗くて小さいため、何となく正面を向いている時に目

があっけいそうだと感じた。曖昧であり、判断に困った」(平面ディスプレイ使用時)に見られるように、そもそも「目が合っている」という感覚自体があいまいで、人によって見られていると感じた機会が不ぞろいである原因であると考えられる。

- ・真陽性率は3人とも平面ディスプレイ使用時に対し、OEB 使用後の方が増加している
- ・サンプル数の少なさ、及び順序効果により、比較はできないが、本実験で使用したアプリケーションによって視線情報が共有されやすくなったという仮定に対し、一考の余地がある。
- ・一方で、視線情報を小窓の表示状態などからも判断できる OEB 使用時の方が、偽陽性率が下がると予想していたが、2 者は OEB 使用時の方が偽陽性率が増加した。特に、被検者 2 の偽陽性率は 0.16 上昇し 0.28 となっている。

・「正面に発表者がいて、さらに窓に発表者がいることに違和感を感じました。」(OEB 使用時) という意見があった。本アプリケーションは、「特定の個人を見ていない場合、全員に小窓を表示する」という仕様になっている。この仕様では、前回の人物クリックの結果、質問者側でプレゼンターの顔が正面に移動したのにもかかわらず、その上に小窓を表示され 2 つの顔が映ってしまう。そのため、見られていない時も見られていると錯覚した可能性がある。

- ・OEB を使用する場合において、上記以外にも小窓の表示方法に関する意見が多くみられた。
- ・「顔の切り抜きはもう少し下に置くことは可能か？」
- ・「OEB についての説明が少ないので仕様がよく分からなかった。小窓がなくなった時がよくわからなかった。小窓が唐突に消えるのに驚いた。」
- ・よって、小窓の表示方法に関しては再考する必要がある。
- ・また、全天周ビデオカメラにおいて、被写体の顔が小さく表示されるという意見も多くみられた。
- ・「3 人のうち誰か一人を見るということが難しいと感じた (視野に全員が収まっているので誰か一人にフォーカスして話しづらい)」
- ・「顔とカメラが遠い？からなのか、そもそも目が見えなかった」
- ・「発表者の顔の画像が暗くて小さいため、何となく正面を向いている時に目があっけいそうだと感じた。曖昧であり、判断に困った。」(いずれも平面ディスプレイ使用時)
- ・全天周カメラ自体の視野の広さと、撮影領域上下が広がるような歪みによって、相対的に顔が小さくなるのが原因であると考えられる。
- ・よって、全天周カメラを用いて顔を映したコミュニケーションを行う際には、何らかの方法で顔部分をクローズアップする必要がある。

・「ディスプレイのサイズの3人とも常に視野の中に収まっているので、クリックした人が真ん中に来る必要はないかな…と感じた。画像そのものが移動するよりは、クリックしている間はカーソルの周辺に枠がでる…とかのほうが「ひとりにフォーカスしている感」があっけい良いかもしれないと思いました。」(OEB 使用時、プレゼンター側)

・PC を使用するプレゼンター側において、常にカメラの方を向くことで見たい人間を正面に捉えるように設計したが、常に3人とも視界に入っているためかえって混乱を招くことになった。

第 6 章

考察

6.1 本システムの問題点

6.2 本システムの応用

6.2.1 対話相手表示形式

6.2.2 横顔生成形式

6.2.3 おかしら会議形式

6.3 今後の展望

第 7 章

結論

結論は、網羅的にかつ簡潔に。

謝辞

本論文の執筆にあたり、議論して頂いた関係者に感謝する。

参考文献

参考文献

- [1] zoom の説明
- [2] google meet の説明
- [3] The GAZE groupware system: mediating joint attention in multiparty communication and collaboration の文献
- [4] Theta V の説明
- [5] Collaboration in 360° Videochat: Challenges and Opportunities の文献
- [6] Can You See Me Now?: How Field of View Affects Collaboration in Robotic Telepresence の文献
- [7] OBS-virtual-cam の説明
- [8] UnityCapture の説明