

おかしら会議:球体ディスプレイによる遠隔会議での一対一会話支援システム

宮藤 詩緒* 李 正卿* 佐藤 俊樹* 小池 英樹*

概要. 本研究では、離れた場所にいる2つのグループが球体型ディスプレイを用いて隔地会議を行う場合に着目し、複数人会議で発生する一対一会話を支援する新しい遠隔会議支援ディスプレイ「おかしら会議」システムを提案する。「おかしら会議」では、複数に会議中の一対一の会話を行う際に、球体ディスプレイ上に話者の首から上の頭を再現することで遠隔の話者同士のバーチャルな対面会話を再現し、周囲にいる話者以外の参加者にも会話する二人を見守ることのできる状況を実現する。

1 はじめに

会議等で複数人が集まり意見交換を行う場合、参加者は各自自由に発言を行うため、話し手が頻繁に入れ替わる場合が殆どである。しかし、このような会議の中でも、例えば特定の誰かと誰かの意見が対立した場合の当人同士の議論や、個人から個人への説明等、話し手と聞き手が限定された一対一(個人対個人)の持続的な会話が継続する場合がある。このような一対一の会話では、グループ中の特定の2人だけが会話を行うような状態となり、話の当事者同士が「二人で会話を行なっている」ことを意識し、会議室内で対面し、お互いの顔を見ながら会話を行う。また、当事者以外の傍観者は、2人の会話を周囲から見守ることとなる。近年では遠隔地にカメラやマイクを設置し、複数人のグループ間で行う遠隔地会議は Skype 等のビデオ会議支援ツールの登場により一般的となってきている。従来の遠隔地会議では、遠隔地の映像を撮影するカメラや、その映像を参加間で共有する大型のディスプレイが用いられることが主流であった[2]。しかし、従来の遠隔地会議では、視線方向の不一致やカメラ・画面視野の制限により会話相手が誰かを理解しにくいという問題点がある。特に、本研究が目にした複数人同士での遠隔通信を行う際に発生する個人に向けた会話ではこの問題が顕著となり、話者が遠隔にいるグループ内の「誰に話しかけているのか」を判断できないことがある。そのため、ローカルで行う一対一会話で感じられる「二人で会話を行っている」体験が損なわれてしまう問題がある。

これに対し、全天球カメラと球体型ディスプレイを用いる手法もある[3]。全天球カメラと球体型ディスプレイを用いた遠隔通信では、撮影された画像の方向が保たれたまま映像が表示されるため、遠隔地の話者の環境を理解しやすいという利点がある。本研究でも、これまで特に複数人同士で行う遠隔会議・

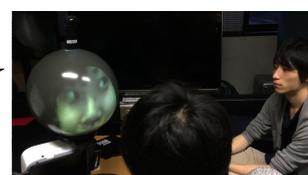
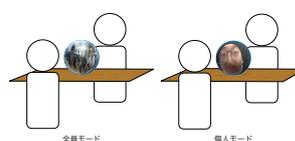


図 1. Concept Image

図 2. 個人モード

作業に特化した球体型ディスプレイシステムの研究を行ってきた[1]。しかし、球体型ディスプレイにそのまま遠隔地の映像を映し出すだけでは、ディスプレイ上に表示された話者が小さく、一対一会話中の相手を遠くに感じてしまうという問題点があった。本論文では、遠隔地会議中に行われる「個人対個人の会話」において、丸い球体ディスプレイそのものを遠隔地にいる「話者の顔」とみため、球上に話者の頭をリアルタイムに再現する表示を行うことで、遠隔にいる者同士が自然な一対一会話を行うことを可能にする「おかしら会議」システムを提案する。

2 提案

本研究が解決すべき課題は以下である。まず、遠隔にいる話者(当事者)同士がディスプレイ(球状ディスプレイ)を介してでも「直接対面しているかのような感覚」を自然に作り出す必要がある。特に本研究ではこの「直接対面している感覚」を、話者同士が顔と顔を向き合わせ、目が合った状態で会話をディスプレイを介して再現することだと考え、これが可能な顔の表示手法を検討する必要がある。これを実現するために、本研究では球状ディスプレイを用い、遠隔にいる話者の首から上の映像をリアルタイムに撮影・球面マッピングする手法を提案する(図1)。同時に、話者の周囲にいる者に対しても、首から上が表示されたディスプレイの側面にも話者の映像を表示すること等を用いて「すぐ近くで対面して会話を行っているかのような感覚」を与える表示を試みる。

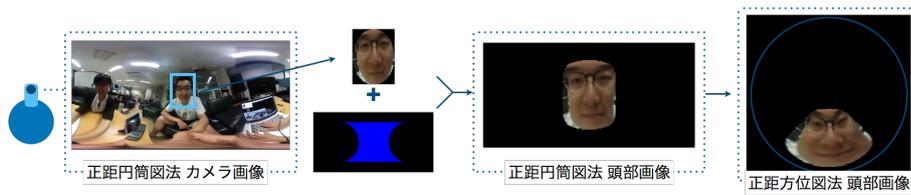


図 3. 頭部画像生成方法

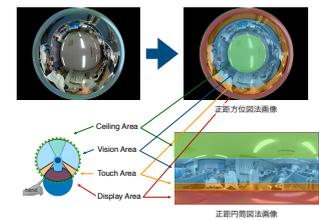


図 4. 画像処理範囲

3 実装

本システムのプロトタイプでは、ハードウェアとしてLiら[1]のOmniEyeballを用いた。このOmniEyeballは上部に全天球カメラが付いた球体型ディスプレイである。なお、カメラのフレームレートは30fps、解像度は1280×640である。

本システムで表示する遠隔地の様子は、意見収集等の複数人での会話時、会話相手部屋の全体像の表示を行う全体モードと、一対一の会話時、会話相手方向に会話話者の頭部の表示を行う個人モードの2種類である。全体モードから個人モードへの切り替え方法として考えられる方法は、画像認識や音声認識を用いて一対一会話の検出を行い、検出された会話話者の頭部を自動で表示する方法、会話を行いたい相手をタッチで選択する方法、等が考えられる。今回のプロトタイプでは、ディスプレイ上に表示された会話相手を直接タッチすると個人モードに切り替わり、カメラを手で覆うと全体モードに切り替わる方法を採用した。

図4に画像処理に用いられる画像範囲を示す。全天球カメラから撮影された画像は図4のように4種類の範囲に分けて処理を行う。Ceiling Areaは天井の映る範囲、Display Areaはカメラ直下の本体ディスプレイ側面が写り込む範囲であり、この2種類の範囲は処理には使われない。ディスプレイ側面に近接するユーザの手が映し出される画像範囲をTouch Areaとし、この範囲でディスプレイの接触判定処理を行う。その他の範囲をVision Areaとし、この範囲で顔検出処理を行う。各範囲の配置はそれぞれ、正距方位図法画像では、円形画像の外側から内側に向かって、正距円筒図法画像では、矩形画像の上側から下側に向かって、Ceiling Area, Vision Area, Touch Area, Display Areaの順である。

顔検出は画像範囲のVision Areaに対して処理を行う。正距方位図法画像では画像全体が歪曲して表示されているため、正距円筒図法画像に対して顔検出処理を行う。正距円筒図法画像においても歪みは存在しているが、Vision Areaでは歪曲が少なく、一般の矩形画像に用いる顔検出を使用することができる。顔検出に用いる分類器はdlibが提供しているfrontal_face_detectorを用いる。図3に頭部画像生成方法を示す。検出された顔画像は球面へのマッピ

ングを行なった際の歪みを考慮して変形を行い、正距円筒図法での画像を生成する。その後、正距方位図法への変換を行なった後、投影を行う。

図2に実際におかしら会議システムを用いて遠隔通信を行なっている様子を示す。遠隔地にいる話者が頭部モードを選択すると、球体ディスプレイに話者の頭部が表示され、目の前に話者の頭があるかのような状態で目を合わせて会話することが可能である。

4 課題と展望

本研究では、遠隔地にいるグループ同士が会議を行う際の一対一会話の問題に着目し、話者同士の対面会話の再現を可能にする球体ディスプレイと全周カメラを用いたディスプレイシステムの提案を行った。今後は顔検出・切り取りの精度向上、音声による自動切り替え、話者の視線のずれや、傍観者の介入、後頭部へ表示する情報の有無などの考慮も含めて実装を行い、システムの完成度を高めていく。またプロジェクタとカメラを内蔵することで、完全な球体型ディスプレイでの実装を行い、球体特有の転がす、投げる、手に取るといった動作を用いたインタラクションなども取り入れたい。

参考文献

- [1] Z. Li, S. Miyafuji, T. Sato, and H. Koike. OmniEyeball: Spherical Display Embedded With Omnidirectional Camera Using Dynamic Spherical Mapping. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16 Adjunct*, pp. 193–194, New York, NY, USA, 2016. ACM.
- [2] P. K. Luff, N. Yamashita, H. Kuzuoka, and C. Heath. Flexible Ecologies And Incongruent Locations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pp. 877–886, New York, NY, USA, 2015. ACM.
- [3] O. Oyekoya, W. Steptoe, and A. Steed. Sphere-Avatar: A Situated Display to Represent a Remote Collaborator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pp. 2551–2560, New York, NY, USA, 2012. ACM.