

仮想空間におけるミーティングコンテンツの 効果的な表示とその利用

横山 勇斗[†] 長尾 確[‡]

名古屋大学 工学部電気電子・情報工学科[†] 名古屋大学 大学院情報学研究科[‡]

1. はじめに

近年, 仮想現実に関わる様々な技術やコンテンツが増大傾向にあるといえるが, その利用の多くはゲームへの利用に留まっている. その一方で, Microsoft 社などは仮想現実技術をオフィスや生活空間に適応しようとする動きを見せている[1]. そこで, 本研究ではゲーム以外への利用として, この技術を会議議事録に利用する手法を提案する.

一般的な議事録の問題として, テキストから読み取れる実際の会議の雰囲気には違いがあり, また動画の議事録でも聴衆の反応がわかりづらいことなどが挙げられる. ディスカッションマイニング[2]などの, 会議を構造的に記録して振り返る技術が開発されているが, それらでは実際の会議の臨場感や参加者の反応を知るのには不十分であり, 会議の情報を利用者に伝えきれているとは言えない. そこで, 本研究では仮想現実を用いて参加者の反応や議論の様子を再現し, 利用者に会議に参加しているのと近い経験を与えることを目的とする.

2. 議事録データ収集

本研究では, 360° カメラであるリコー社の RICOH THETA (以下 THETA とする) を用いて撮影した動画と, Apple 社の iPad を用いて音声認識の結果を書き起こしたものを利用して議事録を作成し, ルームスケール VR が可能な Head Mounted Display (以下 HMD) である HTC Vive (以下 Vive) を用いてそれを視聴し, より臨場感を持った仮想現実における議事録コンテンツの提供を行う.

本システムではまず, THETA と iPad を用いて会議を撮影・記録する. その際, 参加者は iPad を用いて資料を閲覧し, そこにメモなどを書き込む. それら資料の切り替わった時刻や書き込みの内

容も同様に記録する. 音声は参加者それぞれの iPad 上で録音し, 同時に発言に対して音声認識を行い, 発言内容とその時刻を保存する. 会議終了後, 音声認識のテキストに含まれる誤認識を手動で修正する. また 360° 動画から顔の位置, 向きを推定し, 話者の方向を向いているか, 頷いているか等の条件から会議において参加者が集中して会議に取り組んでいた時刻を推定する.

収録時の参加者と各機器の様子を図1に示す.



図1 収録風景

3. VR コンテンツの制作

3.1 開発と構成

本研究では, Unity Technologies 社が開発を行っているゲームエンジン Unity を使用して VR コンテンツを制作した. 本システムでは HMD に加え, 2つのコントローラを使用する.

図2に利用者からの実際の視点を示す.

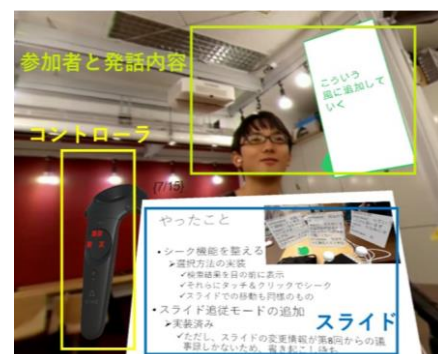


図2 利用者からの視点

基本構成として, 利用者を中心に 360° 動画を再生し, 利用者は辺りを見回すことで全体を見る

Effective Visualization of Meeting Contents in Virtual Space and its Utilization

[†]YOKOYAMA, Yuto (yokoyama@nagao.nuie.nagoya-u.ac.jp)

[‡]NAGAO, Katashi (nagao@nuie.nagoya-u.ac.jp)

[†]Department of Information Engineering, Nagoya University

[‡]Graduate School of Informatics, Nagoya University

ことができる。動画とともに各参加者の音声を再生し、発言者の横に発言内容を表示する。

片方のコントローラには会議で使用された資料や発言内容のログが表示される。これらは、コントローラのボタンを押すことで両者が切り替えられ、コントローラに付属したタッチパッドを操作することで閲覧する内容を切り替えることができる。また、これらは空間に固定して閲覧することや動画と同期した表示も可能である。

また、もう片方のコントローラを用いて動画操作や検索などを行うが、その際、図 3 に示すように利用者の目の前に操作内容を表すアイコンを出現させ、それらに実際に触れることで操作を行うといった視覚的な操作方法とした。



図 3 操作の様子

3.2 操作と機能

仮想現実を用いたコンテンツ、特に HMD を用いたものは臨場感や雰囲気を利用者に提供することに適しているが、マウスやキーボード等の入力機器が使えないため、円滑な操作が困難である。そこで本システムでは視覚的な情報を提示し、コントローラを用いて触るなど体感的な操作を行う。

3.2.1 動画操作

より短い時間での閲覧を促すため、基本的な再生、停止といった動画操作に加え、倍速での閲覧や、発言・資料が切り替わったタイミングに動画を切り替えることができる。また、画像処理によって推定した、特に会議に集中していた時刻を利用者にハイライトとして提示し、それらの時刻に動画を切り替えることも可能である。

3.2.2 検索

VR では一般にキーボードの利用が困難なため、単語検索には音声認識を用いる。利用者が発話した単語を検索して提示し、その発話時刻に動画を切り替えることができる。また、修正されたテキストから TF-IDF 値を算出し、重要単語を発言者ごとのトピックとして提示する。利用者はそれらトピックを選択することでそのトピックに関する発言時刻へと切り替えることが可能である。図 4 に示すように、コントローラからポイ

ンタを表示し、ボタンを押して選択することで動画を切り替える。また発話者がトピックに関して発言する前後においてトピックを強調表示する。

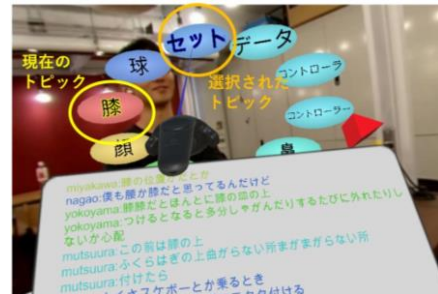


図 4 トピックの提示と選択

3.2.3 参加者の表示

周辺を見回すことで参加者の様子をを知ることができるが、視界が制限されるため、全ての参加者を同時に見ることができない。そのため、参加者全員の顔を一度にすべて表示する状態に切り替えられる。また、図 3 に示すように次の発言者とその方向を提示することで、利用者の視線を誘導することができる。

4. 被験者実験

学生を対象にした、システムユーザビリティスケール(SUS)とシミュレータ酔いアンケート(SSQ)を用いたユーザビリティと映像酔いに関する被験者実験を行う予定である。

5. まとめ・今後の課題

本研究では、HMD による没入型のコンテンツと体感的な操作を用いることで利用者に対し会議の臨場感を再現し、仮想空間での効果的な会議の表示を可能にした。今後の課題として閲覧速度や情報の参照し易さなどの改善が挙げられる。また、本システムは基本的に過去の議事録を閲覧するものだが、リアルタイムかつ遠隔で会議に参加し、仮想空間でその様子を視聴することができれば、より臨場感を伝える意義を持つことができる。そのためには、持ち運び可能な携帯端末を用いた仮想現実での閲覧も視野に入れて開発を行う。

参考文献

- [1] Ben Cutler, Spencer Fowers, Wayne Chang. "Holoportation". Microsoft Research. <https://www.microsoft.com/en-us/research/project/holoportation-3/>, (参照 2018-01-09)
- [2] K. Nagao, K. Kaji, D. Yamamoto, and H. Tomobe. Discussion Mining: Annotation-Based Knowledge Discovery from Real World Activities. Advances in Multimedia Information Processing - PCM 2004, LNCS, Vol.3331, pp. 522-531. Springer, 2005.