

遠隔講義を支援する音響情報ズームング

Acoustic Zooming that Supports Remote Classroom

檜山 敦 泉 雅彦 廣瀬 通孝
Atsushi Hiyama Masahiko Izumi Michitaka Hirose

東京大学大学院 情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo

In this paper, in order to enable a user to easily understand the features of particular sounds under a noisy circumstance, we propose a concept “acoustic zooming” which derived from a metaphor of visual zooming effects. We developed a map-based interface that enables the user to listen to sounds in a specified area that the user can adjust. In order to discuss integrated interface designs, we conducted two exploratory experiments to evaluate the performance of the systems.

1. はじめに

空間を隔てたコミュニケーションにおいてできることが増えると、テレワークという働き方が社会においてより一層浸透していくと考えられる。特に、活動に様々な制約や要望がある高齢者がテレワークという形で社会参加や就労を行う機会を飛躍的に増やすことができれば、超高齢社会における高齢者層の活力の維持と労働力不足の両面に関する課題解決につながる技術的な貢献となる。

遠隔コミュニケーションの基本である、音声、映像を伝達する情報技術の発展は遠隔で共有できる情報を増大させ、遠隔 TV 会議などは広く使われるようになった。ここに遠隔の参加者における臨場感・存在感を増強させるテレプレゼンス技術を適用すれば、遠隔コミュニケーションをより対面会話に近づけられると期待されている。しかし操作インタフェースの煩雑さ、明瞭性の低い音声など、カジュアルに様々な場へ導入するには、未だ多くの課題が残されている [Izumi 2014]。

対人コミュニケーションにおいて、音声は様々な役割を果たす [Kitawaki 1996]。遠隔コミュニケーションにおいても同様で、むしろ伝達される情報が制限される分、音声はより重要である。遠隔地間の音声通信では音声の周波数成分や位置に関する情報が削られ、遠隔ユーザの声が聞き分けにくい。そこで様々な音空間可視化技術が研究されているが、これらの研究は音源の可視化に比重を置き、ユーザが能動的に音源選択できるインタフェースに関する研究は少ない [Yasumoto 2005]。

そこで本研究では、遠隔音空間を可視化した上で、ユーザが自由に聞きたい音を選択できるシステムを構築することを目的とする。音源選択のため、これまでに提案した音響ズームングという概念を導入したインタフェースを構築し [Izumi 2013]、実験室・実地での評価を行った。

2. 音響情報ズームング

「ズーム」は通常は視覚における概念で、日本国語大辞典によれば、「ズームレンズを用いて被写体を拡大、縮小すること」とある。本研究ではこの概念を捉え直し、「空間中のある一点における情報をピンポイントに取得すること」をズームと定義する。これを聴覚に適用し、「空間中に多数の音源がある状況において、任意音源の音のみを聞くこと」を音響情報ズームング(以下、音

響ズームング)と定義する。

音響ズームングの実現にあたり、聴体積という概念を導入する。聴体積はコンピュータグラフィックスの分野の視体積という概念を音空間に転用したもので、内包される音源の音のみをユーザに提示するための領域である (Figure 1)。聴体積を操作することで、音声を選択的に体験できる。

これまでの研究で、音響ズームングを用いて複数音源下で特定音を探す操作をモデル化した [Izumi 2013]。この結果、全体を俯瞰的に聞くフェーズと部分を集中的に聞くフェーズがあり、二者をシームレスに繋ぐことの重要性が示された。

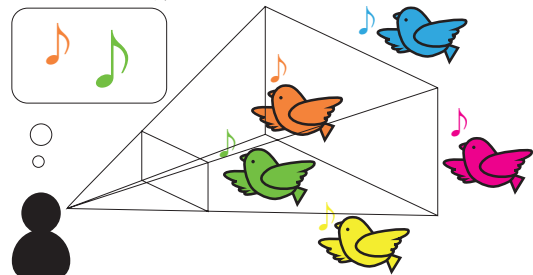


Figure 1: 聴体積コンセプト

3. ZooMap:Map ベース音響ズームング

3.1 遠隔コミュニケーションと音響ズームング

遠隔コミュニケーションには様々な形態があり、人数、目的、手段・媒体と、利用シーンにより重視される要素は異なる。そのため簡便なモジュールとして既存システムに統合できるよう遠隔コミュニケーションの支援システムを構築すると、柔軟性、汎用性の面で有効と考えられる。

本研究では特に、不特定多数の参加者が自由に会話を楽しむような場への遠隔参加をターゲットケースとした。この場は一般にカクテルパーティと呼ばれ、ここに遠隔参加する場合、様々な音声がやり取りされるため個々の会話の聞き分けが難しい。ここに音響ズームングを導入することで、遠隔参加者の音声聞き分けを支援することとした。

3.2 Map(地図)ベース音響ズームング

先行研究の知見をふまえ、遠隔コミュニケーション支援の音響ズームングの形として、Map ベース音響ズームングを提案する (Figure 2)。本手法では、遠隔地の音源位置・音量を Map 上で可視化し、この上で聴体積を直接ドラッグ、ピンチして聴取音を選択する体験を提供する。これにより状況把握、およびシームレ

すな全体・部分の切替が可能となり、かつ操作負荷の低いインタフェースにできると考えられる。

Figure 3 に Map ベース音響ズームング ZooMap のシステム全体図を示す。ZooMap は、会話現地で入力音声 mixes する Zooming Mixer と、遠隔ユーザが可視化された音源位置・音量を調整する Zooming Map から構成される。Mixer は OSC メッセージで操作可能であり、Map の様な外部操作インタフェースの構築が容易である。また Mixer への入力音声を変更するだけで、様々な既存システムに統合できる。

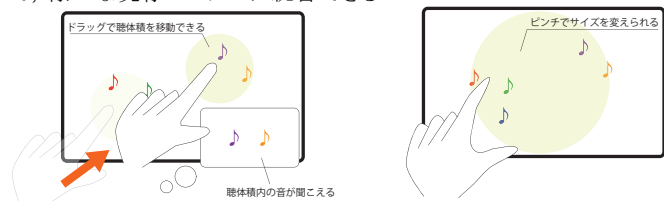


Figure 2: Map ベース音響ズームング

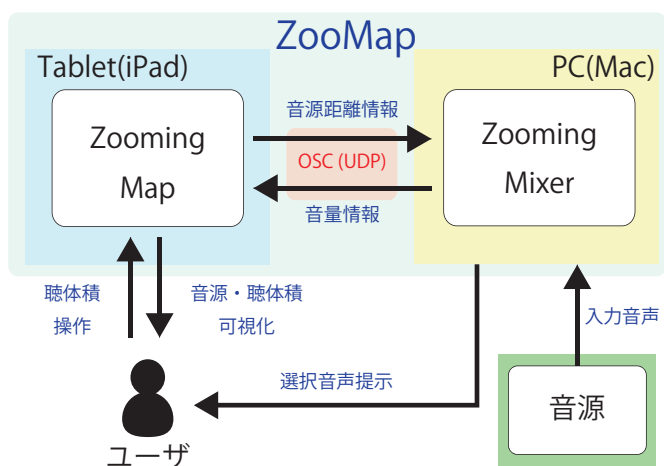


Figure 3: ZooMap システム

4. ZooMap 評価実験

4.1 実験概要

Map ベース音響ズームング(以下 map 条件)を用いた音源聞き分けの容易さを評価するため、一般的なスライダー型インタフェース(以下 slider 条件)との比較実験を行った。各インタフェースの概観を Figure 4 に示す。実験参加者は 20 代男女 8 名である。彼らには、仮想空間中に配置された複数の会話音源から、対象のトピックに関する音源を探すタスクを課した。参加者には、150 秒間の中で出来る限り多くの音源を探すよう教示し、各インタフェースで 3 回ずつ、計 6 タスクを行わせた。実験後、ターゲットの探し易さ、音の絞り込み易さ、タスクの難しさを 5 段階で評価させた。

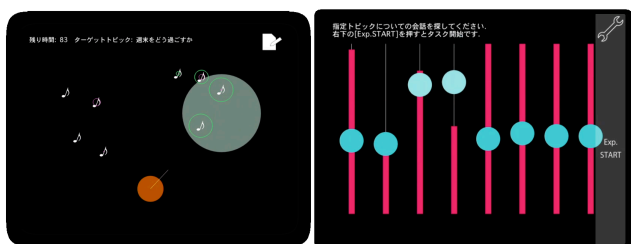


Figure 4: Map 操作(左)とスライダー操作(右)

4.2 結果・考察

探索タスクにおいて、map 条件が slider 条件に比べ回答数が多い傾向があった(対応ある t 検定, $p=0.0578$)。タスク後のアンケートの結果を Figure 5 にまとめる。map 条件の方が、ターゲットを探し易い傾向にあり($p=0.0514$)、音を有意に絞りやすく($p=0.0492$)、タスクの難易度が有意に低いことが分かった($p=0.008$)。以上より、Map ベース音響ズームングは従来のスライダー型インタフェースに比べ複数音源から特定音源を探しやすい方法と言える。また自由回答に「map 条件は 1 操作で音量増減両方が出来て良い」という記述があり、slider 条件において注目音源への絞り込みに 2 段階要する所を、1 段階に簡略化できたことが分かった。

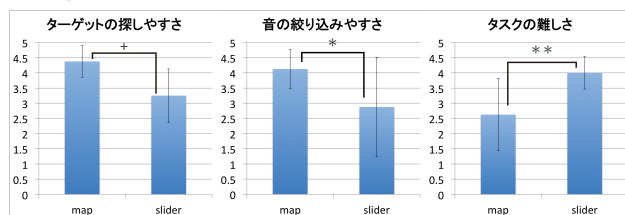


Figure 5: アンケート結果

5. 遠隔講義フィールドにおける投入実験

5.1 実験概要

Map ベース音響ズームングの作業負荷を評価するため、実際の遠隔講義の場に ZooMap を導入した。仙台シニアネットワークの 3 名の講師に本システムを使わせ、兵庫県西宮市清瀬台地区の高齢者に向けた遠隔タブレット講習会を行わせた。構築した音声環境を Figure 6 に示す。講師は全体音声を会場スピーカー、ZooMap による選択音声をヘッドセットで聞いた。結果は操作ログとアンケートより評価した。

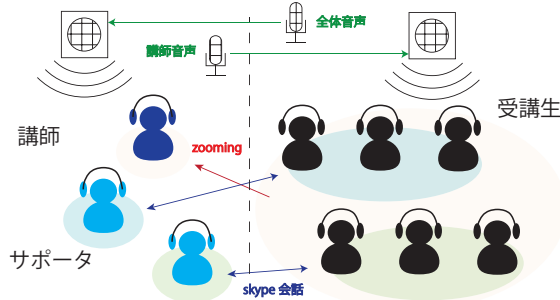


Figure 6: 遠隔講義への展開

5.2 結果・考察

アンケートにより「個人の声を聞き易かったか」を聞いた結果、ZooMap 導入により有意に聞きやすいという結果が得られた(対応のある t 検定, $p=0.0198$)。本システムが遠隔地の音声やり取りの把握を支援可能である事を示している。

一方で操作負荷に関しては「操作は直観的で良いが、会話中に操作するには慣れが必要」との回答を得た。実際のインタラクションの様子を考察するため、操作ログの一例を Figure 7 に示す。講義時の講師の動きと照合すると、彼らは生徒への説明中には操作していない事が分かった。一方で生徒に実習させている間、操作することが度々見られた(図中オレンジ網掛部分)。これらの操作は、遠隔側での会話内容を聞き取るため、ZooMap を使っていたと考えられる。また講義中、現地の会話を聞いた講師が「音が見えるし誰が何を喋っているか分かる」と伝

えるシーンがあり、音源可視化による状況把握支援効果が確認できた。

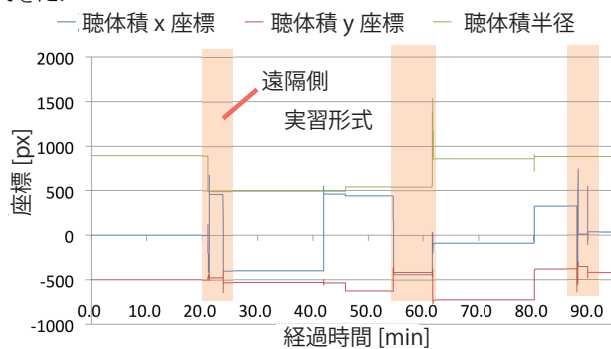


Figure 7: ZooMap 操作ログ

6. 結論と今後の展望

本研究では、遠隔コミュニケーションにおける音環境の把握を支援するため、Map ベース音響ズームリングシステム, ZooMap を実装し、実験室内での性能評価と実際の遠隔講義での評価を行った。結果として、ZooMap が高い音源探索能力を提供し、高齢の講義者でも簡単に使用できることが明らかになった。今後は視線検出を用いるなど、ユーザの意図を汲んだ音源選択をできるインタフェースを構築し、操作負荷の更なる軽減を目指す。

参考文献

- [Izumi 2014] M. Izumi, et al. “Practical Use of a Remote Movable Avatar Robot with an Immersive interfaces for Seniors”, Proceedings Universal Access in Human and Computer Interaction, (2014), pp648-659.
- [Kitawaki 1996] 北脇 信彦. “音のコミュニケーション工学 - マルチメディア時代の音声・音響技術 (音響テクノロジーシリーズ)”, コロナ社, 東京, (1996), pp.50-61.
- [Yasumoto 2005] K. Yasumoto, K. Nahrstedt. “Ravitas: Realistic Voice Chat Framework for Cooperative Virtual Spaces.”, ICME, (2005), pp.1046-1049.
- [Izumi 2013] M. Izumi, et al. “Implementation on a tablet computer and evaluation of acoustic information zooming effects”, Proceedings Asiagraph, (2013), pp91-95.