遮音カバー付きスピーカユニットを持つディスプレイのためのメディア伝送サーバの開発

田中　健太郎†　　片桐　滋†　　大崎　美穂†
†同志社大

1. はじめに

遠隔コラボレーション支援に必要な臨場感の確保を阻害する要因の一つに，視覚情報と聴覚情報の不一致が挙げられる．本稿では，先に提案された遮音カバー付きスピーカユニット[1]を用いる音像生成法と音源映像（発話者等）の再生法を組み合わせて視聴覚オブジェクトを再生するためのメディア伝送サーバの開発を報告する．

2. 遮音カバー付きスピーカユニット

本方式は，大型ディスプレイ上の音像生成において，比較的広い受聴位置で正確に音像方向を定位できるようにすることを目指す（図1）．遮音カバーで囲まれた 3 基のスピーカから成るスピーカユニットを，ディスプレイ左右端に一対向かい合わせに設置する．スピーカからの直接音が抑制され，ディスプレイ面からの反射音が利用されることで，ディスプレイ上の正確な音像定位の実現が期待される．

3. メディア伝送サーバ

本サーバは，遠隔地で撮影・収音されたメディア信号を受信し，ディスプレイ上に映像・音像として生成する．その実現のため，音声伝送機能については，先に開発された多チャンネル出力機構を持つ音声伝送サーバ[2]を基盤とした．また，映像伝送機能については，受信した映像から顔画像検出を行い，音像生成に必要な音源位置座標を算出する機能を，既存の映像伝送サーバに組み入れた．

4. 評価実験

4.1 実験手法

開発結果を評価するため，再生音声の正確性及び通信の安定性を調査した．サーバ起動時から 1 時間の観測を 5 回行い，映像データと音声データの間のズレやそれぞれにおける遅延の大きさを調べた．1 台のコンピュータ上でストップウォッチを動作させ，5 分毎にアラームを鳴らすように設定した．そのコンピュータのディスプレイとアラーム音をそれぞれカメラとマイクで収録し，そのメディアデータを開発したサーバ経由で送信先の遮音カバー付きスピーカユニットを搭載したディスプレイに出力した．そうして得られた出力結果をカメラで録画し，動画と音声を解析した．こうした評価法を採ることで，本サーバの利用者が実際に長時間の遠隔コラボレーションを行う際に得られると思われる「見え」と「聞こえ」の評価を試みた．

4.2 実験結果

図 2 に，5 回の観測で得られた音声波形を示す．各段の派遣がそれぞれの観測回に対応する．図から，5 分おきに繰り返されるアラーム音が正確に再生されている様子を読み取ることができる．なお，不規則に混入している雑音は，観測における環境音である．

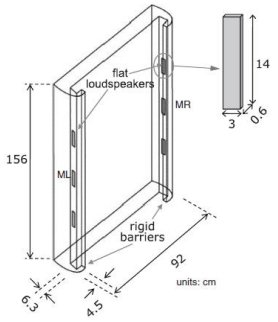

図1．遮音カバー付きスピーカユニットの構成図

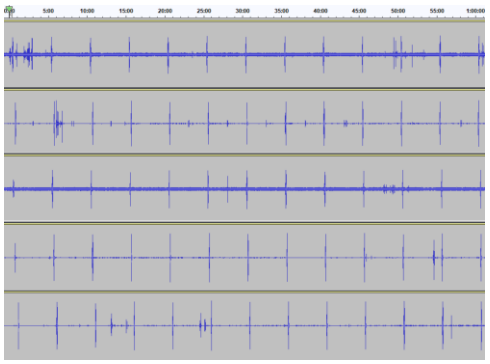

図2．収録した音声波形

それぞれ 1 時間の観測を 5 回繰り返した全体における遅延量は，音声で平均 134.4ms，標準偏差 0.01685 であり，映像で平均 574ms，標準偏差 0.40198 であった．映像と比べ音声の遅延が小さかったのは，前者の伝送に UDP を用いたのに対し後者には TCP を用いたことによるものと考えられる．

5. まとめと今後の課題

遅延の原因としては，メディア伝送サーバ内およびネットワーキングにおけるデータバッファリングが考えられる．今後，それらの見直しが必要と考えている．

参考文献

[1] G.Pablo ほか，“A loudspeaker design for sound image localization on large flat screens”, Acoustical Science and Technology Vol.31, No.4, pp.278-287, 2010.
[2] 松尾，“大型ディスプレイ画面上に音像を生成する機能を持つ音声伝送サーバの開発”，情報処理学会研究報告会，2017.

Copyright © 2018 IEICE