

# オフィスコミュニケーションを円滑化する エリア收音システムの開発

片桐 一浩 矢頭 隆

今日、ワーク・ライフ・バランスの実現や非常災害時の事業継続のために、ICT（情報通信技術）を活用したテレワークの需要が高まっている。『超臨場感テレワークシステム』は、離れた場所に存在する複数のオフィス間を映像や音情報などで相互に接続して、お互いが、あたかも同じオフィスにいるかのように体感させることを目指している<sup>1)</sup>。このテレワークシステムでは、遠隔地に設置した複数のカメラにより、ユーザーが自由な視点で遠隔地の映像を見ることができ、さらに画面上の任意のエリアをズームする機能を備えている。遠隔地と円滑にコミュニケーションを取るためには、画面のズームに伴い音もズームし、画面に映っているエリアの音のみを聞くことができる機能が求められる。しかし、オフィス内では様々な場所で同時に音が発生しているため、マイクロホンエリアの中心や近くに配置するだけでは、エリア外の音も收音してしまうという問題があった。

本稿では、複数のマイクロホンアレイを用い、様々な音が存在する環境下においても、任意のエリアの音のみを收音できるエリア收音技術と、この技術を実装した実時間動作するエリア收音システムを紹介する。



図1 映像と音のズームアップ

## 特定方向の收音技術

特定方向に存在する音を強調し、それ以外の音を抑圧して收音する技術として、複数のマイクロホンから構成されるマイクロホンアレイを用いたビームフォーマ（BF：Beam Former）がある<sup>2)</sup>。BFとは、各マイクロホンに到達する音の時間差を利用して指向性を形成する

技術である。BFは、加算型と減算型の大きく2つの種類に分けられ、特に減算型BFは、加算型BFに比べ少ないマイクロホン数で鋭い指向性を形成できるという利点がある<sup>3)</sup>。図2は、マイクロホン数が2個の場合の減算型BFのブロック図である。減算型BFは、まず減算器により、一方のマイクロホンの入力信号 $X_1$ からもう一方の入力信号 $X_2$ を減算する。すると正面方向に存在する音は、2つのマイクロホンの入力信号に含まれる位相が同じであるため抑圧される。このとき形成される指向性は、図3に示すような正面方向に死角を向ける8の字型の双指向性となる。つまり收音したい音（目的音）が正面にある場合、目的音以外の左右方向に存在する音（雑音）が抽出されることになる。その後、入力信号 $X_1$ のパワーから、抽出した雑音Aのパワーを周波数領域で減算するスペクトル減算（SS：Spectral Subtraction）を行う。これにより、入力信号 $X_1$ に含まれる雑音が抑圧され、目的音を強調することができる。減算型BFは、最終的に図3の死角方向（正面方向）に鋭い指向性を形成することになる。

このようにマイクロホンアレイを用いたBFによって指向性を形成すれば、特定の方向に存在する音を收音することができる。一方で、ある特定のエリア内に存在する音（目的エリア音）だけを收音したい場合、実環境ではそのエリアの周囲に多数の雑音が存在する状況が考えられる。通常BFは、前後左右へ直線的にしか指向性を形成することができない。それ故、図4に示すような目的エリアと同方向に雑音源が存在する場合、目的エリア音だけでなく雑音まで強調してしまうという問題がある。

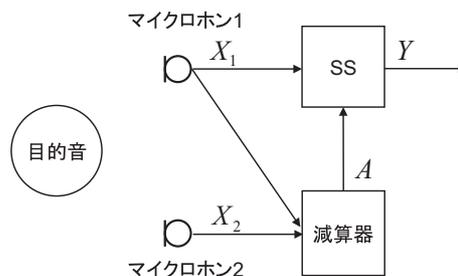


図2 減算型BFのブロック図

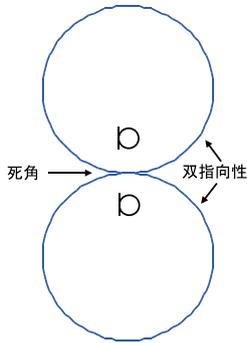


図3 双指向性

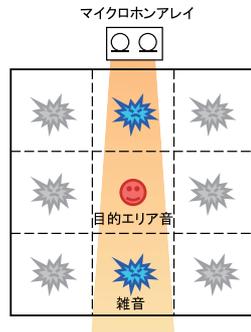


図4 減算型BFの指向性

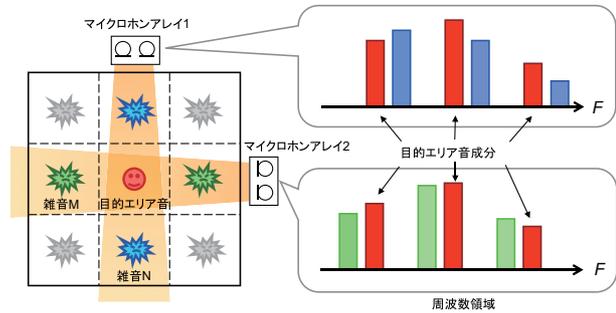


図5 各マイクロホンアレイのBF出力の比較

### エリア收音技術

エリア收音は、目的エリアを含む空間内に複数のマイクロホンアレイを任意に配置し、それぞれBFにより目的エリア方向へ指向性を形成する。この状態では、依然として各マイクロホンアレイのBFの指向性に目的エリア音だけでなく、目的エリア音方向の雑音も含まれている（図5左）。一般的に音声は、周波数領域に成分がまばらにしか存在しないため、複数の音声と同時に存在しても、周波数領域で音声成分が重なる確率は低くなるという特性がある。従って、各マイクロホンアレイのBF出力を周波数領域で比較すると、目的エリア音成分はどちらの出力にも含まれるが、雑音成分は各マイクロホンアレイで異なることになる（図5右）。エリア收音は、この特性を利用して目的エリア音のみを收音する。

図6に2つの2チャンネル・マイクロホンアレイを用いたときのエリア收音のブロック図を示す。また図7は、エリア收音における入力信号中の目的エリア音と雑音成分の変化を示した図である。まずマイクロホンアレイ毎に減算型BFにより指向性を形成し、目的エリア方向に存在する音 $Y_{MA1}$ と $Y_{MA2}$ を抽出する。 $Y_{MA1}$ と $Y_{MA2}$ の周波数成分を比較すると、どちらにも目的エリア音の成分が含まれている。その一方、目的エリア外に存在する雑音成分は、 $Y_{MA1}$ 、 $Y_{MA2}$ で異なる。このため、 $Y_{MA1}$ から $Y_{MA2}$ をSSすると、共通に含まれる目的エリア音成分が消去され、 $Y_{MA1}$ に含まれる目的エリア音以外の雑音 $N$ が抽出される（図7①）。この際、マイクロホンアレイ2のBF出力に含まれる雑音 $M$ は、もともと $Y_{MA1}$ には含まれていないため影響しない。その後、 $Y_{MA1}$ からさらに $N$ をSSすることにより、入力信号に含まれる全ての雑音を抑圧し、目的エリア音 $Z$ が抽出される（図7②）。

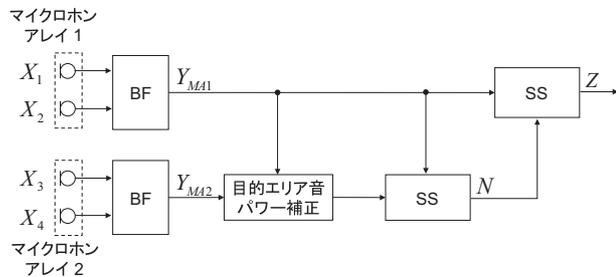


図6 エリア收音ブロック図

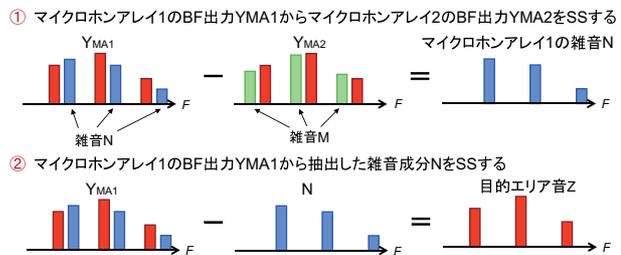


図7 周波数成分の変化

マイクロホンアレイ1と2のBF出力に含まれる目的エリア音のパワーは、目的エリアと各マイクロホンアレイの距離の違いや、エリア内の話者の向きにより変動する。各BF出力中の目的エリア音のパワーが同じでないと、SSにより $N$ を抽出することができない。そのため、目的エリア音のパワーを同じにする補正係数 $\alpha$ が必要となるが、通常目的エリア音の成分は未知であるので、 $\alpha$ を以下のように算出する。まず(1)式に従い、各マイクロホンアレイのBF出力間で周波数毎にパワーの比を求める。ここで $k$ は周波数を表している。次に全ての $\alpha_k$ から最頻値を算出し、この値を $\alpha$ とする。

$$\alpha_k = \frac{Y_{MA1k}}{Y_{MA2k}} \quad (1)$$

目的エリア音成分は、各BF出力に同じ割合、分布で含まれているため、比率は同じになる。逆に目的エリア音以外の成分は、BF出力毎に異なるので比率にはばらつきがある。この特性から、各周波数の比率に対して最頻値を求めれば、その値がそのまま各BF出力間で目的エリア音のパワーを等しくする補正係数となる。

図8(a)は、目的エリアに対し、マイクロホンアレイ1と2がほぼ同じ距離に配置されている場合の各BF出力間のパワー比のヒストグラムである。BF出力中の目的エリア音のパワーはマイクロホンアレイ1と2でほぼ同じため、 $\alpha$ は1に近い値となる。また図8(b)は、目的エリアに対し、マイクロホンアレイ1よりも2のほうが近い距離に配置されている場合のヒストグラムである。目的エリア音のパワーはマイクロホンアレイ2の方が大きいいため、 $\alpha$ が1より小さい値になることが分かる。

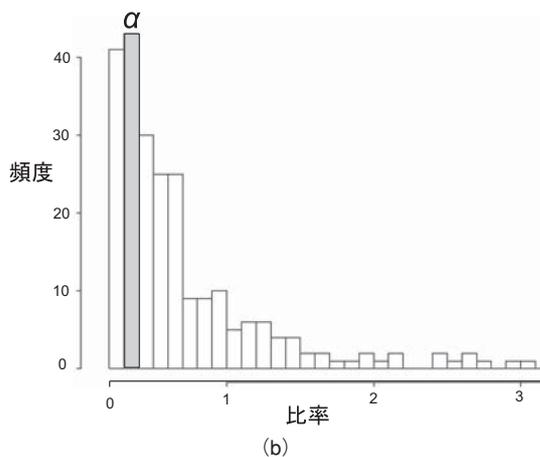
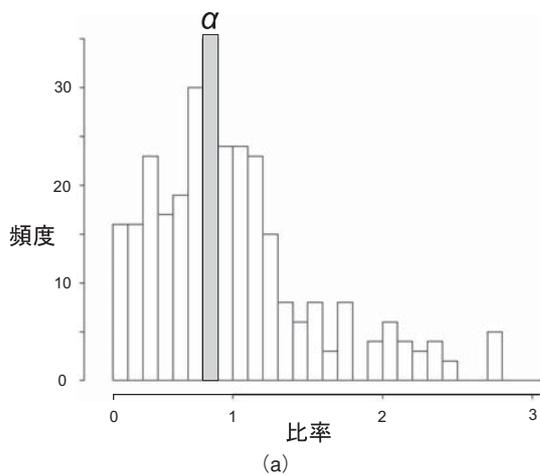


図8 パワー比のヒストグラム

## エリア收音システム

本システムは、Windows PC (Core™i7-2600 / RAM 16GB)、多チャンネルオーディオインターフェース (ASIO)、全指向性のピンマイクから構成される<sup>4)</sup>。図9(a)はマイクロホンアレイの配置の一例である。マイクロホンアレイは、図9(b)のように3チャンネル・マイクロホンアレイとしている。3チャンネルにすることで指向性の方向を変えることができ、図10に示すように2つのマイクロホンアレイを用いて2つのエリアを收音することが可能となる。本システムでは、リアルタイムに2つ以上のエリアを切替えて、もしくは同時に收音することができる。

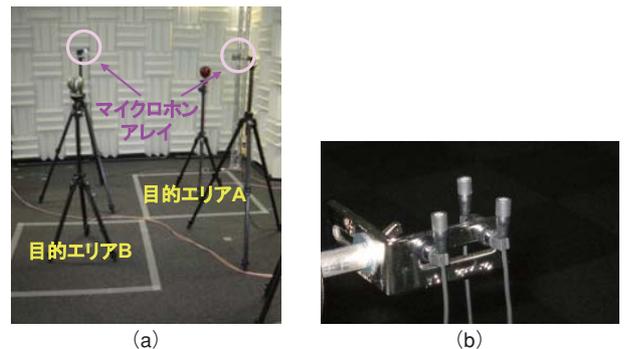


図9 エリア收音システムのマイクロホンアレイ

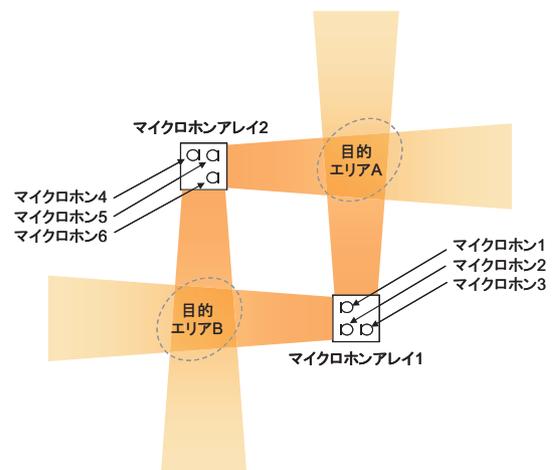


図10 3チャンネル・マイクロホンアレイによるエリア收音

## 性能評価実験

エリア收音システムが、目的エリアの周囲に存在する雑音源を実際にどの程度抑圧できるのかを調べるために性能評価実験を行った。目的エリアとマイクロ

ホンアレイ、雑音源の配置を図 11に示す。2つのマイクロホンアレイの指向性をBFにより別々の方向から目的エリアに向け、同時に3つの雑音（音声）を再生したときの目的エリア内の雑音抑圧量を調べた。評価の指標としてNoise Reduction Rate(NRR)を用いる。またマイクロホンアレイ単独のBFを比較対象とした。

エリア收音とマイクロホンアレイ1単独での目的エリア内の雑音の抑圧量を図 12に示す。図より、マイクロホンアレイ1単独ではNRR が5dBであるのに対し、エリア收音は15dBと約10dB抑圧量が優れていることがわかる。これはエリア收音が、マイクロホンアレイ1単独では抑圧できない雑音1、2をほぼ完全に抑圧できていることを示している。

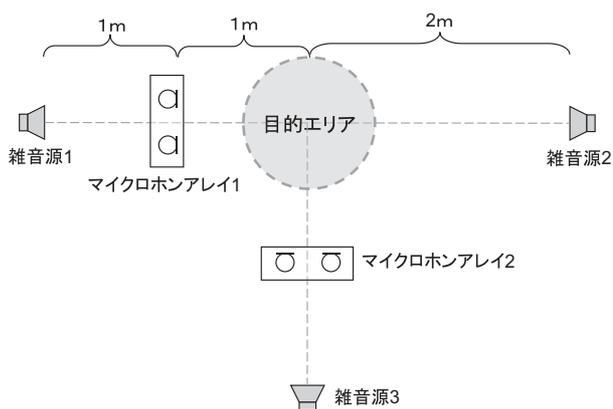


図 11 評価実験レイアウト

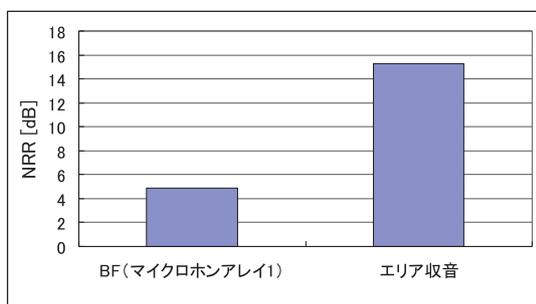


図 12 目的エリア内の NRR

## まとめ

本稿では、複数のマイクロホンアレイを用い、目的エリアの周囲に雑音が存在していても目的音のみを收音するエリア收音技術と、実時間動作するエリア收音システムを紹介した。また評価実験を行い、エリア收音システムが目的エリア外に存在する音を抑圧できることを示した。

今後は、エリア收音システムを超臨場感テレワークシステムに組み込み、実証実験を行う予定である。

## 謝辞

本研究の一部は、独立行政法人情報通信研究機構からの研究委託「革新的な三次元映像技術による超臨場感コミュニケーション技術の研究開発」により実施したものである。関係各位に感謝致します。◆◆

## 参考文献

- 1) 徳満 他：超臨場感テレワークシステム、OKIテクニカルレビュー、No.215、pp.58-61、2009
- 2) 浅野：音のアレイ信号処理、2011、コロナ社
- 3) 矢頭 他：臨場感テレワークにおける音処理技術、OKIテクニカルレビュー、No.213、pp.16-21、2008
- 4) 片桐 他：複数の3チャンネル・マイクロホンアレイを用いたエリア收音システムの開発、第18回日本バーチャルリアリティ学会大会論文集、pp.458-461、2013

## 筆者紹介

片桐一浩：Kazuhiro Katagiri. 研究開発センタ メディア処理技術研究開発部

矢頭隆：Takashi Yazu. 研究開発センタ メディア処理技術研究開発部