

分散カメラ・マイク・スピーカアレイによる 対人スポットフォーミング

Sound Spot Forming at Target Person by Camera, Microphone, and Speaker Array

溝口 博¹⁾ 玉井 裕樹¹⁾ 加賀美 聡²⁾ 長嶋 功一³⁾
Hiroshi MIZOGUCHI Yuki TAMAI Satoshi KAGAMI Koichi NAGASHIMA

- 1) 東京理科大学 理工学部 (〒278-8510 千葉県野田市山崎 2641 E-mail: hm@rs.noda.tus.ac.jp)
2) 独立行政法人 産業技術総合研究所 デジタルヒューマン研究センター(〒135-0064 東京都江東区青海 2-41-6 E-mail: s.kagami@aist.go.jp)
3) 有限会社アールラボ(〒166-0000 東京都杉並区阿佐谷南 1-44-15 E-mail: knaga@r-lab.co.jp)

ABSTRACT. This paper describes speaker array system which can form sound beam towards desired direction and microphone array system which can collect only sound from desired direction. By speaker array system which we developed, the authors can form sound beam towards not only one direction but also two directions. Measuring sound pressure by the noise level meter and the gantry crane with 1mm resolution which can be operated by PC, we can evaluate the effect of sound beam forming clearly. By indicating the frequency ingredient of the sound collected by the spectrum using microphone array system that we developed, we can evaluate the effect quantitatively.

1. 背景

来るべき高齢化社会を目前にして、人に優しく、人と共存・協調可能な機械システムへの社会的必要性と期待とが高まっている。これはデジタル・ヒューマン技術の典型的な応用の一つと位置付けられる。そのような人に優しい機械システムの実現には、まず、システムが人をみて判る機能を備えることが不可欠である。

この一環として、人を見守り、人と直接、音声を介して対話したり、身振りを理解したりできるような知覚型ユーザインタフェース(Perceptual User Interface)への期待と要求が高まってきている。しかしながら、音声理解や音声対話以前の問題として、広い範囲で動き回る人に対し、離れたところから雑音無く高品質で音声を授受する手段さえも確立されていないのが現状である。

2. 目的

以上を踏まえ、本開発プロジェクトでは、「人に優しい」知覚型ユーザインタフェースの実現に向け、まず、人の存在を認識してその人に注意を向け、言わば「聞き耳をたてる」ような形で音声を集音し、「耳元で語りかける」ような形で音を聴かせる技術の確立を目的とする。

このために、本プロジェクトでは、対象とする人の頭部周辺に「スポット状」の高感度・高音圧分布を作り出して S/N 比の高い集音と音声伝送とを実現し、たとえその人が動いても、それに対応してスポット状高感度・高音圧分布を追従させることが可能な対象人物追従型集音・伝声インタフェースの実現を行う。これにより、言わば「口元で耳をすます」効果と「耳元で語りかける」効果の実現を図る。さらに、人の存在の認識結果を、音

声伝送だけでなく、他の感覚への働きかけや物理的な動きとしても提示するような、データ取得からモデル利用、提示までを一貫したシステムとしての開発を目指す。すなわち、人とシステムとのインタラクションの質をより豊かなものとする。

具体的に、本開発プロジェクトは、大規模マイクアレイ・スピーカアレイによる「スポットフォーミング」技術の実現に取り組んだ。このために以下の項目を実施した。

- 1)スピーカアレイの数学モデル開発
- 2)スピーカアレイ・シミュレータの開発
- 3)同シミュレータを用いた各種条件下でのシミュレーション並びに構築システムの設計
- 4)実験およびソフト開発用小規模(16台)スピーカアレイシステムの構築と動作実験
- 5)大規模(64～128台規模)システム用多チャンネル同時サンプル AD/DA 変換ボード開発
- 6)同変換ボード用ドライバソフトウェアの開発
- 7)小規模(8台)マイクアレイシステムの構築と動作実験
- 8)大規模システムの構築と動作実験

3. スピーカアレイ・マイクアレイの原理

スピーカアレイでは、焦点から各スピーカまでの距離差による音声到達遅延時間の計算を行う。各スピーカから出力する値にスピーカ毎の遅延を付加する。その値を、実時間 OS と同時出力 D/A ボードを用いて全チャンネル同時等周期出力を行う。結果、各スピーカから送出された音声の位相が焦点でのみ揃い、増幅する。マイクアレイでは、同時入力 A/D ボードを用いて全チャンネル同時等周期入力を行う。入力してきた値に遅延を付加し、音声の合成を行う。結果、焦点で生じた音声のみ増幅し、

選択的に収拾することが可能となる。

Fig.1 に著者らが構築したスピーカアレイを、Fig.2 に構築したマイクアレイを示す。

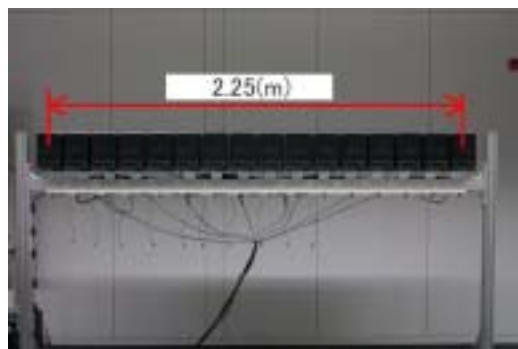


Fig.1 Speaker Array



Fig.2 Microphone Array

4．単一方向への選択的音響ビーム形成

著者らはこれまで既に、一直線上に配置したスピーカアレイによって単一方向の音圧を選択的に高める音響ビーム形成に成功している。しかし、手動で測定していたため、思い通りの結果を得ることはできなかった。著者らは、1(mm)の分解能を持ち PC での操作が可能なガントリークレーンを用いて計測を行うことで、音響ビーム形成効果をよりはっきりとした形で図示することに成功した。

本プロジェクトで構築したスピーカアレイシステムのスピーカ個数は 16 個、間隔は 0.15(m)、サンプリング周波数は 5512.5(Hz)である。格子点状に 0.25(m)間隔・11×11=121 箇所 で音圧を測定し、測定結果を元に Fig.3 に図示した。

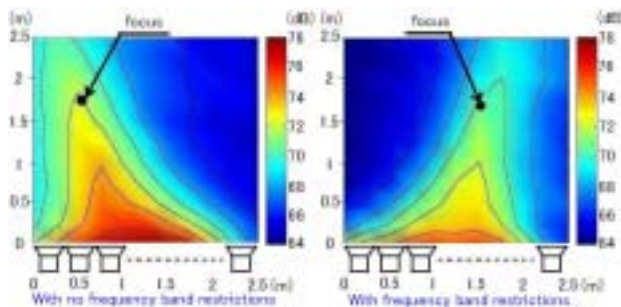


Fig.3 One Direction Sound Beam

Fig.3 の左側は、ニュース(男声)を単一の焦点へ向け

てビームフォーミングを行った際に形成された音場の音圧分布を図示したものがある。右側は、同じニュースの内容を、人の声の主成分である 200 ~ 2000(Hz)の範囲で周波数帯域制限を行ったものを焦点へ向けて送出した際に形成された音場の音圧分布を図示したものである。

どちらも焦点へ向けて帯状の音響ビームが形成されているのを確認できる。また、右図の方が形成された音響ビームが細いことが分かる。以上のことから、所望する単一方向への音声送出、及び音声送出によって形成された音響ビームの効果を、周波数帯域制限を行うことで上げることが、視覚で確認することに成功したと言える。

5．二方向への選択的音声送出

次に、同じスピーカアレイシステムで、各々別の位置にある焦点へ向けて音声送出を行った。Fig.3 と同様に 0.25(m)間隔・121 箇所 で音圧測定を行った。Fig.4 が測定結果である。左図は、2 つの焦点方向へ同一の男声を音声送出したもの、右図は、左の焦点方向へ男声を、右の焦点方向へ女声を音声送出したものである。

どちらも 2 つの焦点方向へ帯状の音響ビームを形成しているのが分かる。右図では、左右で高音圧領域の大きさに差があるが、これは用いた音声の相対的な音量の差から生じたものであると考えられる。また 2 つの焦点の間に、ある程度的高音圧領域が広がっているが、この領域ではホワイトノイズとなっており、聴覚上どちらの音声も情報として認識することは不可能である。

以上のことから、所望する 2 方向への選択的な音声送出による効果を視覚で確認することに成功したと言える。また、理論上 3 方向以上への選択的音響ビーム形成も可能であると考えられるが、今回構築したシステムでは、焦点位置でのホワイトノイズの影響が大きくなり、音声情報の認識が困難になった。しかしながら、スピーカ個数を増やすことでこの問題は解決できると考えられる。

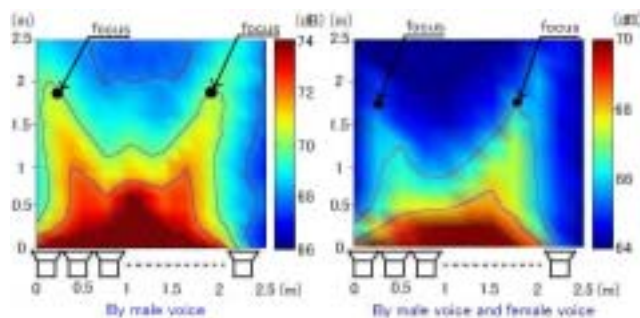


Fig.4 Two Direction Sound Beams

6．マイクアレイによる方向選択的音声収拾

本プロジェクトで構築したマイクアレイシステムのマイク個数は 16 個、間隔は 0.15(m)サンプリング周波数は 44100(Hz)である。

焦点位置に 1000(Hz)のサイン波が流れるスピーカを配置し、焦点位置から x 軸方向へ 1.5(m)ずれた位置に 500(Hz)のサイン波が流れる別のスピーカを配置する。マイクアレイを用いて焦点で生じる音のみを収拾した。結果が Fig.3 である。

Fig.5 の上図は 1 個のマイクで収拾した WAVE、中図は 8 個のマイクで拾った音をそのまま合成した WAVE、下図はマイクアレイシステムを用いて焦点で生じる音のみを収拾した WAVE、それぞれをスペクトル表示したものである。上図、中図、下図の順で 1000(Hz)と 500(Hz)の

周波数成分の音圧差は、0(dB)、11(dB)、20(dB)となっている。

以上のことから、マイクアレイシステムを用いることで、焦点で生じる音のみを増幅させて収拾することが可能であると言える。また、マイク個数を増やす、マイク間隔を狭くする、以上のことを行うことで収音効果の向上を狙えると考えられる。

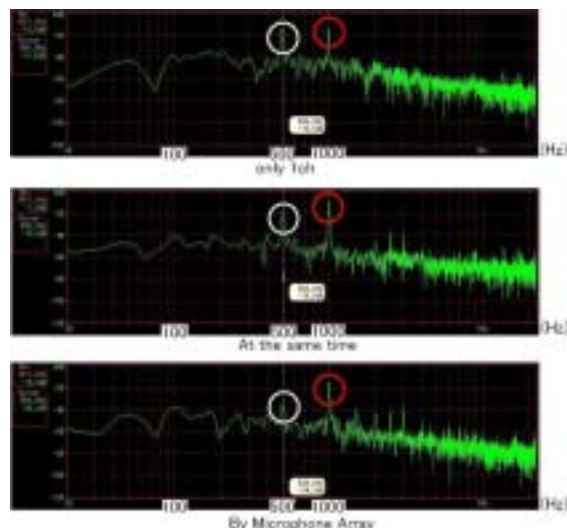


Fig.5 Spectrum of Microphone Array Output

7. おわりに

以上述べてきたように、小規模な一直線上スピーカアレイの構築は完成し、動作実験により当初予定の機能が実現できていることを確認した。これはいわば「耳元で語りかける」効果に相当するもので、本プロジェクトの成果として、その効果の実現に成功したことを意味する。

ただし、小規模アレイの音圧分布特性は、本文中で述べてきたように、まだビームフォーミングの段階である。本プロジェクトが最終的に目指すスポットフォーミングは、複数ビームの重ね合わせにより実現できることが、シミュレーションにより確かめられている。したがって、構築中の大規模スピーカアレイを、複数のビーム重ね合わせるような形で配置することにより、スポットフォーミング実現に成功することが期待できる。大規模スピーカアレイを稼働させ、スポットフォーミングを実現することが当面の課題である。

スピーカアレイと同様、マイクアレイに関しても、大規模化してビームを重ね合わせるような配置をとることにより、スポットフォーミングの実現が期待できる。既に小規模マイクアレイは構築し、実験により効果は確認できている。スピーカアレイ同様、これを大規模化し、スポットフォーミングを実現することも今後の課題である。

本プロジェクトの申請段階では想定していなかったこととして、スピーカアレイとマイクアレイの連携がある。これはプロジェクトの実施過程で考案したもので、一部についてはシミュレーションを実施し、可能性を確認している。両者を連携させることにより、局所的な消音効果が期待できることから、任意の位置に配置された、任意種類の音源に対する能動消音が、事前の環境設定や設計を行うことなく、その場ですぐに実現できることが期待できる。これを実際に試し、実環境における効果を確

認してみることも、今後の展開の一つである。

本プロジェクトの副産物としては、大規模アレイ用に開発した多チャンネル同時サンプル AD/DA 変換ボードの、他用途への応用がある。数十チャンネル規模で、十数マイクロ秒オーダの高速なデータのやり取りを、全チャンネル同期してできるような変換ボードは、従来全く存在していない。このため、数十チャンネル規模で、十数マイクロ秒のサンプリングを必要とするような、多計測点の時間的変化データを高速に採取可能な分布型のセンサ群や、高速動作する多自由度メカニズムなどは、潜在的な必要性があるにもかかわらず、これまで実現されてこなかった。開発者らは、本プロジェクトで開発した多チャンネル高速同時サンプリング変換ボードを活用し、これらの用途への展開も図ってゆく予定である。

8. 参加企業及び機関

東京理科大学

独立行政法人産業技術総合研究所

有限会社アールラボ

