

# 現実の音イベントの方向・位置の推定と その複合現実空間への反映

Direction and Position Estimation of Sound Events in Real World and its Reflection to Mixed Reality Space

大槻麻衣, 西浦敬信, 木村朝子, 柴田史久, 田村秀行

Mai OTSUKI, Takanobu NISHIURA, Asako KIMURA, Fumihisa SHIBATA, and Hideyuki TAMURA

立命館大学情報理工学部

(〒525-8577 滋賀県 草津市 野路東 1-1-1)

**Abstract:** We are developing the Mixed Reality (MR) system which merges real and virtual worlds in both audio and visual senses. Our new approaches using sound for MR can be applied to (1) mixing sound of real and virtual world and (2) detecting a purposely generated sound in real world to be used as an input event to MR space. In this paper, we describe the second approach. Firstly, we have developed the sound detection system which can localize the sound source by using single and multiple microphone arrays. In this system, two types of microphone arrays, wearable type and fixed type are implemented. Secondly, we have evaluated the accuracy of this system, and propose practical usages of sound event such as interactive devices in MR attractions.

**Key Words:** Mixed Reality, Sound Input, Microphone Array, Sound Direction and Position, Interactive Device.

## 1. はじめに

我々は、視覚・聴覚の両方で現実空間と仮想空間を融合する複合現実感 (Mixed Reality; MR) システムの開発を進めている。従来の視覚のみの MR システムに比べ、聴覚の MR も扱えるシステムでは、現実空間の音と仮想空間の音を混在できるだけでなく、現実空間で意図的に発した音イベントを MR 空間に反映させる新たな用法が考えられる。

本研究では、後者を実現するシステムとして、小規模直線型マイクロホンアレーを用いて、現実空間で意図的に発した音イベントの方向・位置推定を行い、その結果を MR 空間に反映させるシステムを開発した。従来、マイクロホンアレーは位置を固定した据置型 (図 1 左) が利用されて来たが、本研究では、マイクロホンアレーを HMD に取り付けたウェアラブル型 (図 1 右) も用意し、構成を自由に変更可能とした。

本稿では、上記システムを用いて、音源方向推定結果を MR 空間への入力とするインタフェースシステムを試作し、その性能を実験によって評価した結果を報告する。また、各種 MR システムに音イベント入力を導入することを想定し、ウェアラブル型と据置型マイクロホンアレーを併用した位置推定を行い、その精度を従来の据置型マイクロホンアレーを用いた場合と比較し、その結果を元に音イベントを MR 空間に導入する際の指針について考察する。

## 2. システム構成

システム構成を図 2 に示す。MR 空間の管理と提示には、Canon MR Platform System を用いている。体験者は、ビデ



図1 マイクロホンアレー外観  
(左:据置型, 右:ウェアラブル型)

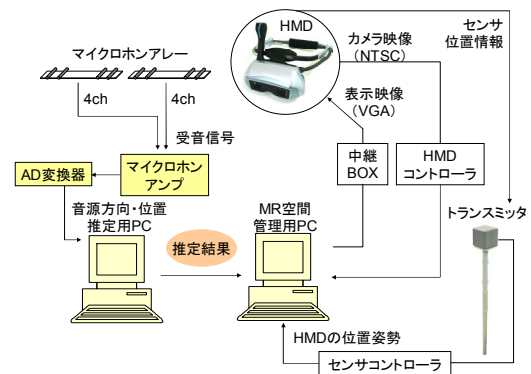


図2 システム構成

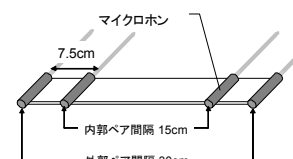


図3 マイクロホンアレーの構成

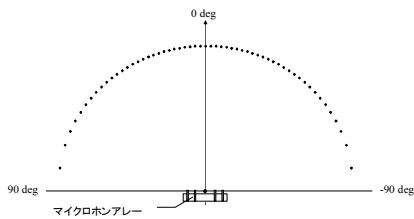


図4 音源方向推定に関する分解能

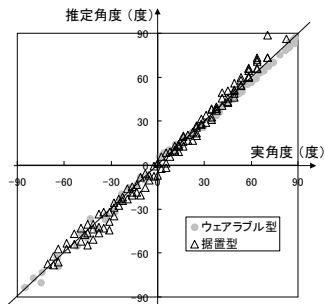


図5 据置型とウェアラブル型の音源方向推定精度の比較

オシスルー型 HMD (Canon VH-2002) を装着し、MR 空間を観察する。HMD の位置姿勢検出には、Polhemus 社製の磁気センサ 3SPACE FASTRAK を使用している。

現実空間における音の検出と方向・位置推定には、複数のマイクロホン配置した小規模直線型マイクロホンアレーを利用している (図 3)。マイクロホンアレーによって受信した信号はマイクロホンアンプ (Thinknet MA-2016C) でラインレベルに増幅し、AD 変換器 (Thinknet DF-2X16-2) を経由した後、音源方向・位置推定用 PC に入力される。

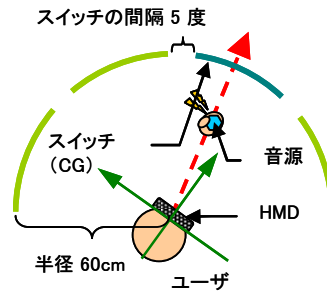
### 3. 音イベントの方向推定と MR 空間での利用

#### 3.1 音源方向の推定方法

音源方向は、マイクロホンアレー1 台を利用して推定することができる。音源方向推定アルゴリズムには CSP 法 (Cross-power Spectrum Phase analysis) [2]を用いた。これにより、水平面での音源方向が得られる。また、システムの利用環境として、暗騒音のある環境を想定しており、このような環境下では雑音のために推定結果に誤りが生じることが予想される。そこで、西浦らが提案する、複数個のマイクロホンペアを用いて CSP 係数を加算するという、より雑音に強い推定方法を採用する[3]。

#### 3.2 音源方向の推定精度

本研究で使用する据置型マイクロホンアレー1 台で推定可能な方向とその精度を図 4 に示す。正面を 0 度とすると、 $\pm 90$  度に近づくほど分解能が低下していることが分かる。一方、ウェアラブル型マイクロホンアレーを利用する場合は、体験者が音源方向を向くと、音源を常にマイクロホンアレーの正面で捉えることが可能となるため、精度が向上する。図 5 は据置型、ウェアラブル型において推定角度を x 軸、実角度を y 軸にとり、各データをプロットしたものである。グラフ中の右肩上がりの線は推定角度と実角度が一致する理想的な結果を示したものである。これより、ウェアラブル型の方が据置型よりも理想的な結果に近いことがわかる。



(a) スイッチ数 4 個の場合 (b) スイッチ操作の様子

図6 音入力によるスイッチの選択

### 3.3 方向推定を利用した入力インタフェースの試作

本研究では、MR 空間において現実空間に重畳描画されたスイッチを模した CG オブジェクトを体験者に提示し、音を入力することによって選択するインタフェースを試作した。本システムでは、体験者に提示する CG のスイッチを、体験者の目の高さ、体験者の頭部位置を中心に距離 60cm の位置に、前面 180 度の半円状に並べ (図 6)、体験者が発した音源方向の推定結果が、あるスイッチの領域に含まれるとき、そのスイッチを「選択した」と判定する。

### 3.4 性能評価実験

#### 3.4.1 目的

3.3 節で構築したインタフェースについて、被験者が意図したとおり入力することができるか、実験を通して確認する。実験では、スイッチ数と入力成功率の関係、スイッチの位置と入力成功率の関係を調査した。

#### 3.4.2 実験手順

スイッチ数 2, 4, 6 個の場合について、以下の手順で実験を行った。ここではユーザの前方 180 度をスイッチ数で分割したものを 1 つのスイッチとし、5 度おきに配置した。

- (1) 被験者に全スイッチの中から 1 つをランダムに提示
- (2) 被験者は提示されたスイッチ上で音を発生
- (3) 音の入力があれば入力成否に関わらずスイッチを消去

以上の試行を各スイッチに対して計 5 回ずつ行い、これを 1 セッションとする。これを、スイッチ数が 2, 4, 6 個の各場合について 2 セッションずつ連続して行った。音源にはカスタネットを利用した。被験者は 22-23 歳の男性 4 名、女性 2 名の計 6 名、実験は 46dB A 程度のオフィス環境で行った。

#### 3.4.3 結果と考察

実験の結果、スイッチ数 2, 4, 6 個全ての場合で、入力成功率は約 99% であり、また、スイッチの位置による入力成功率の変動は見られなかった。このことから、音イベントは安定した入力デバイスとして利用できるといえる。但し、今回は雑音による誤入力を防ぐため、被験者には実験中可能な限り物音を立てないよう、発話をしないよう指示を行った。しかし、実際に MR アトラクションへの応用を考えると、雑音に対してロバストなシステムの構築が必要だと考えられる。

### 4. 音イベントの位置推定と MR 空間での利用

従来、音響分野における位置推定には、複数台の据置型

マイクロホンアレーが用いられていた。ここでは、据置型のみを利用する方法、据置型とウェアラブル型を併用する方法について各種MRシステムに音イベント入力を導入する際の指針について検討する。

#### 4.1 音源位置の推定方法

音源位置の推定は4chマイクロホンアレーを2台用い、各マイクロホンアレーから得られた方向より、三角測量の要領で求める。ここで、据置型マイクロホンアレーの位置姿勢はあらかじめ実測で測定しておき、また、ウェアラブル型は磁気センサによってその位置姿勢を取得するものとする。なお、今回実装したマイクロホンアレーによる位置推定は水平面での音源位置の推定のみを行い、高さの推定は行わないこととする。

#### 4.2 音源位置推定精度の仮説

3.2節に示したマイクロホンアレーの特性より、位置推定を行なう際、両マイクロホンアレーがともに正面方向で音源を捉えたときに精度が高くなると考えられる。また、音源とマイクロホンアレーの距離が近ければ、方向推定段階における数度の誤差は位置推定にはほとんど影響しないが、距離が遠くなればなるほどその誤差が位置推定精度に影響するようになる。三角形の性質より、マイクロホンアレーから音源までの距離が2倍になれば、位置推定の誤差も2倍になるはずである。よって、音源とマイクロホンアレーの距離に近いほうが、位置推定精度が高くなると考えられる。以上のことを踏まえると、各マイクロホンアレーの正面方向が交差する領域が広く、かつマイクロホンアレーから近い距離にある場合に、精度よく推定できる範囲がもっとも広くなると考えられる。また、ウェアラブル型を導入すると、磁気センサの影響により現実空間と仮想空間とでマイクロホンアレー位置にずれが生じるため、その精度が全体的に低下すると考えられる。

#### 4.3 実験

##### 4.3.1 目的

4.2節の仮説が正しいかどうか確認するために、据置型マイクロホンアレー2台を利用する場合について、2台のマイクロホンアレーの配置と精度の関係を調査する。また、据置型2台を用いたときの結果が最も良かったアレー配置において、2台のマイクロホンアレーのうち1台に磁気センサを取り付け、位置姿勢を取得することで、据置型からウェアラブル型に置き換えた場合についても調査する。その結果、据置型2台の場合と比較し、精度がどの程度劣化するのかを調べる。

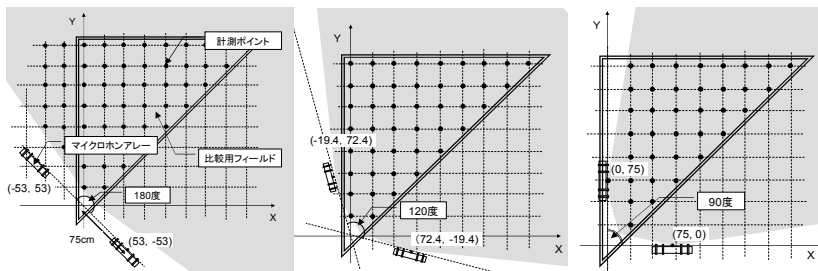


図7 マイクロホンアレーの配置 (■は計測可能範囲を示す)  
(左:180度, 中央:120度, 右:90度)

##### 4.3.2 実験手順

まず、据置型のマイクロホンアレー2台を90, 120, 180度に配置した場合について実験を行う(図7)。実験では、床に25cm間隔のグリッドを組み、各交点で音を発生させ、音源位置推定を行う。計測ポイントは、各マイクロホンアレー配置での計測可能範囲内とする。各点につき5回計測を行い、その誤差平均を求める。ここでの誤差は音源の実測位置と推定位置のユークリッド距離とする。

次に、ウェアラブル型と据置型を1台ずつ利用する場合は、上記の90, 120, 180度の中で最も推定精度の高かった配置を採用し、200cm×200cmの正方形のフィールド上にあるポイント全てについて計測を行う。音源位置の計算に利用するウェアラブル型マイクロホンアレーの位置姿勢は、実測値ではなく、マイクロホンアレーに取り付けた磁気センサの値を使用する。実験時のトランスミッタからレーザーまでの距離は73.1cmとする。

##### 4.3.3 結果

まず、据置型2台を用いた場合の各マイクロホンアレー配置について、比較用フィールド(1辺を200cmとする直角三角形領域)内の全計測ポイントの誤差平均を比較した結果(図8)、アレー配置が120度の場合において、誤差平均が最も小さかったため、据置型とウェアラブル型を併用した実験では、アレー配置を120度とした。

次に、全計測ポイントに関して誤差マップを作成したものを図9, 図10に示す。各バブルの直径は誤差の大きさを示している。また、3.2節の図4より、各マイクロホンアレーの方向推定精度が2.5度より小さい範囲(-34.5~+34.5度)を正面方向(推定精度の高い方向)として点線で示す。図10で、Aは現実空間でのマイクロホンアレー位置を、Bは磁気センサが取得したマイクロホンアレー位置を平均した位置を示している。磁気センサで取得した値を実際の位置姿勢と比較すると、距離は10cm程度、角度は7度程度のずれが生じていた。図10には、マイクロホンアレーBと据置型マイクロホンアレーの正面方向、トランスミッタ位置も併せて示す。

図8の比較用フィールド内全計測ポイントの誤差平均を見ると120度, 90度, 180度の順に大きくなっており、ウェアラブル型を導入した場合は据置型のみの場合の約2.2倍誤差が増加している。また、図9より、マイクロホンアレーの正面方向が交差する部分で推定精度が高いこと、正面方向であっても、音源とマイクロホンアレーの距離が離れるほど誤差が大きくなっていることがわかる。

ウェアラブル型を導入した場合は、マイクロホンアレー

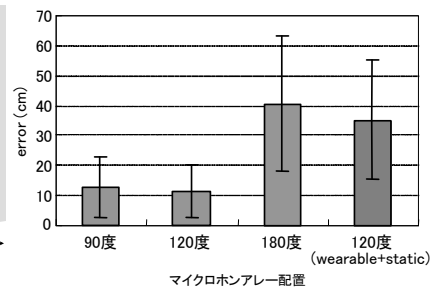


図8 位置推定の誤差平均

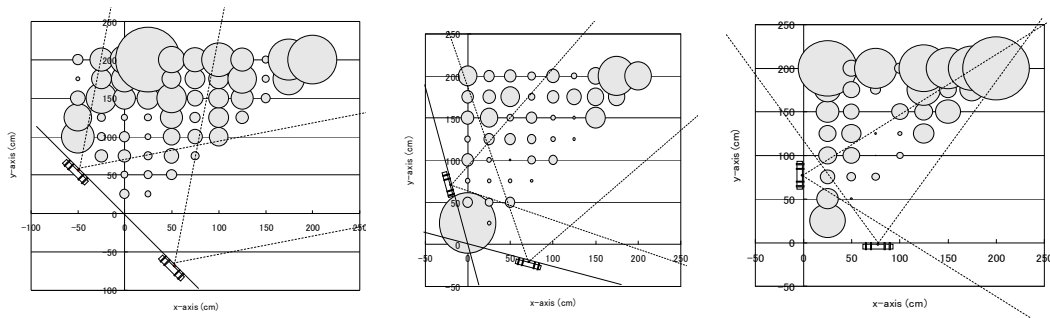


図9 誤差マップ

(左:180度, 中央:120度, 右:90度)

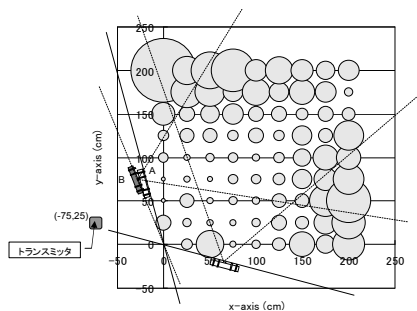


図10 誤差マップ (据置型・ウェアラブル型併用)

の正面方向が交差する部分について、誤差マップの傾向に変化は見られなかったが、実際の位置姿勢と、磁気センサで取得した位置姿勢に誤差が生じているため、正面方向にずれが発生する、音源位置の推定結果自体にも誤差が増加するという傾向が見られた。

以上の結果より、4.2節の仮説が正しいことを確認した。

#### 4.4 考察

実験結果を踏まえ、MR システムへの応用を考える。具体的には、複数台のマイクロホンアレーを用いて位置推定を行い、結果をMR 空間へ反映させるシステムを想定する。

実験結果から、まず考慮しなければならないのは、各マイクロホンアレーの正面方向が交差する部分とMR 空間とが重なる領域をできるだけ大きくするという点である。今回の実験の配置ではマイクロホンアレー配置が120度の場合が最も推定精度が高かったが、90度の場合でも正面方向が交差する部分と音源位置を推定したい領域が重なる部分をできるだけ広くするような配置にすることで、精度が向上すると考えられる。

次に、ウェアラブル型マイクロホンアレーを導入した場合について考えると、推定精度に大きく影響するのはHMDに取り付けられた磁気センサの精度である。磁気センサはレシーバがトランスミッタから遠ざかるほど誤差が大きくなる。ゆえに、単純に考えるとウェアラブル型を導入した場合には、据置型2台を用いて位置推定を行うよりも、精度が低下することになる。しかし、ウェアラブル型の導入により、磁気センサの有効範囲が許す限りで、体験者が音源方向を向くことができるため、常に音源をマイクロホンアレーの正面方向で捉えられるようになる。また、体験者が音源に近づくことによってマイクロホンアレーと音源の距離を短くすることができるという2つの利点が挙げられる。

一方、複数の体験者が参加するMR アトラクションを想定すると、それぞれの体験者がウェアラブル型マイクロホンアレーを装着するだけでは、体験者の向きによって位置推定可能領域が変化するという問題が発生する。また、実験結果より、ウェアラブル型マイクロホンアレーのみを利用して位置推定を行うと、磁気センサの誤差により位置推定精度の低下が生じてしまう。これらの問題を回避するために、2台のウェアラブル型マイクロホンアレーのみで位置推定を行うのではなく、据置型を含めた多数のマイクロホンアレーを組み合わせ、位置推定を行い、明らかな推定誤りを棄却した上で、各マイクロホンアレーペアから算出された位置推定結果の重心を最終的な推定結果とするといった手法が考えられる。

#### 5. むすび

本稿では、小規模直線型マイクロホンアレーを用いて、現実空間で意図的に発した音イベントをMR 空間に反映させるシステムを開発し、音イベントをスイッチ入力として利用する手法とその性能評価について述べた。その結果、音イベントはMR 空間への入力として利用できることが分かった。さらに、複数台のマイクロホンアレーを用いて、音源位置推定を行うシステムを構築した。このシステムを用いて、従来の据置型のマイクロホンアレー2台を用いた場合と、1台をウェアラブル型に置き換えた場合の位置推定実験を行い、実際にMR システムへの応用を行う際の指針について考察した。

今後は音源の3次元位置の推定や、音源の違いの判別などを実装することでMR コンテンツとしての応用の幅を広げる必要があると考えられる。その上で、それらを利用したMR コンテンツを作成することを予定している。

**謝辞** 本研究の一部は、科研費・基盤研究(A)「三感融合型複合現実空間の構成法に関する研究」による。

#### 参考文献

- [1] 大槻他：“現実空間での音検出とその複合現実空間へのイベント入力利用”，信学総大, A-16-30, p.292 (2006.3)
- [2] M. Omologo et al: “Acoustic Event Location Using a Crosspower-Spectrum Phase Based Technique,” IEEE ISASSP 1994 (1994)
- [3] 西浦他：“マイクロホンアレーを用いたCSP法に基づく複数音源位置推定”，信学論, Vol. 83-D-II, No. 8, pp. 1713-1721 (2000)