基礎論文

複合現実空間との新しい マルチモーダル・インタラクション方法の提案と実現

大槻 麻衣[*1], 木村 朝子[*2], 西浦 敬信[*1], 柴田 史久[*1], 田村 秀行[*1]

Design and Implementation of a Novel Method of Interacting with Mixed Reality Space

Mai Ohtsuki[*1], Asako Kimura[*2], Takanobu Nishiura[*1], Fumihisa Shibata[*1], and Hideyuki Tamura[*1]

Abstract – In mixed reality (MR) space, there is the characteristic that the users can see their hands and the interactive device in their hand. In this study, we propose the novel interaction method with MR space using this characteristic. This multi-model interface was developed based on the traditional pointing device that can manipulate 2D plane or 3D space directly and its feedback uses sound event in the real world, and the visual feedback (and tactile feedback in some cases) in the MR space. More specifically, we attached the small liner microphone array onto a head mounted display (HMD), used the direction and position of sound source for input into MR space, and extended it from 2D to 3D pointing. This paper describes some implementation using our method such as the menu selection and the MR attraction.

Keywords: Mixed reality, Multimodal, Microphone array, Sound source localization, Input device, Interactive device, and 3D space.

1. はじめに

人工現実感 (Virtual Reality; VR) という概念の登場以来，仮想環境体験のリアリティを高め，対話のモダリティを増す努力が続けられてきた．即ち，没入感のある立体映像空間，3 次元音場環境，力覚フィードバック・デバイスを導入し，それぞれの表現能力を増すことで，体験者に臨場感や没入感を与えてきた．また，仮想空間の入力では，身体を動かす行為の他は，スイッチを押す，棒状のデバイスを動かす，手袋状デバイスで掴む，ペダルを漕ぐ等の操作で仮想環境への働きかけを行ってきた．

VR の発展形である「複合現実感」（Mixed Reality; MR）[1][2]では，体験者の目の前にある現実世界が操作対象であり，自らの手足や共同体験者の姿も視認できる．このため，実世界指向インタフェースを導入して，より効果的に複合現実空間を操作できる方法が期待されている．

こうした背景を踏まえ，本研究では対話操作における入力面での機能拡張を図り，視覚・聴覚・触覚がバランスよく配置された MR 空間とのマルチモーダル・インタフェースの実現をめざす．従来の MR 研究は視覚的 MR の実現に集中していたが，我々は既に出力面で視覚・聴覚の併用・共存する MR システム[3]を実現したので，次に音の検出機能を利用して，それを VR/MR 空間の入力として活用することを試みる．具体的には，マイクロホンアレーを用いて現実空間で発した音刺激（音イベント）の方向と位置を検出し，それを MR 空間に反映する方法を採用する．

本研究は，

- 2 次元位置検出に（超）音波が使われているならば 3 次元位置入力にも音を利用できないか
- 意図的な聴覚フィードバックが煩わしいならば，耳に快い音を入力にし，視覚・触覚をフィードバック側に使う逆転の発想もあるのではないか

という 2 つの着想に基づいている．音イベントとして我々が想定したのは，「手をたたく」「簡単な楽器・道具を操作する」などの様々な楽音，効果音を活用することであり，

- 実際の楽器・道具などをそのまま入力デバイスとして利用可能になる
- 可聴音を採用することで音そのものがフィードバックとなり，入力行為をその場で確認できる

という特長を有している．

*1 立命館大学大学院 理工学研究科
*1 Graduate School of Science and Engineering, Ritsumeikan University
*2 科学技術振興機構 さきがけ
*2 PRESTO, Japan Science and Technology Agency

以下，第 2 章では関連研究との比較を行い，第 3 章では本手法の要素技術と開発したシステムの構成を，第 4 章では本システムを用いて作成したインタラクションの実現例をいくつか紹介し，提案する入力手法と従来の入力手法との比較実験結果について述べる．

2. 関連研究との比較

VR システムで音入力を採用した例として, Irawati ら[4]は音声コマンドで VR 空間内の CG オブジェクトを操作した．また，Mihara ら[5]は，非言語（例えば,「あー」といった声）または簡単な音声コマンドを併用し，2 次元ディスプレイ上のカーソル操作を行うシステムを構築した．しかし，音声コマンドはメニュー選択には利用できても表現力が不足していることや，言語に依存するという問題点がある．また，非言語は日常会話で発しない不自然な発話であるため，抵抗を感じる利用者も少なくない．本研究では，音イベントに拍手や楽器といった非音声を採用しているため，言語に関する制約がなく，利用者が扱いやすく親しみやすいインタフェースとなることが期待される．

音イベントの検出に関しては，音響分野においてマイクロホンを複数個並べたマイクロホンアレーが活発に利用されている．マイクロホンアレーを利用した例としては Nagai ら[6]の研究が挙げられる．この研究では，マイクロホンアレーを用いて音源（話者）方向を推定し, 指向性を形成することによって，車内という雑音環境下でも話者の音声コマンドを正確に認識するシステムを実装している．我々はこうした先行研究に依拠し，音イベントの検出にマイクロホンアレーを用いる．これにより，単一のマイクロホンによって音の ON/OFF だけを取得する従来の VR 分野の実現例とは異なり，その方向や位置も取得し入力とすることが可能となる．

音の位置情報を利用したシステムでは, Patel ら[7]が，単一のマイクでディスプレイに吹きかけられた呼気の位置情報を取得し，その位置に応じて表示オブジェクトを変化させている．伊賀ら[8]は，頭部の磁気センサと 1 つのマイクを併用し，スクリーンに吹きかけられた呼気の位置情報を取得，ペイントツール，GUI 操作に利用している．いずれも対象作業領域は 2 次元ディスプレイやスクリーンであり，機能的には GUI の操作にとどまっている．本手法は 3 次元空間を対象とする点に新規性がある．

音イベントの位置検出とそれに伴う視覚的な応答を効果的に利用したアトラクションに，Ishii らの「PingPongPlus」[9]がある．これは複数のマイクロホンアレーを用いて実際にピンポン球が卓球台にぶつかる衝突音を検出して，上部から CG 映像を投影し，卓球台上に波形を描くという演出を行なっている．音の位置を入力として利用する点と視覚によるフィードバックと応答という組み合わせは本提案手法と同じであるが，卓球という特定の目的に特化したシステムになっている．換言すれば，本論文で提案する手法は 3 次元空間に対するポインティング操作やメニュー選択など，さまざまなインタラクションに利用できる汎用性を有している．

3. MR システム構成とマイクロホンアレー

3.1 システム概要

上述の目的で開発したシステムの構成を**図 1** に示す．MR 空間の管理と提示には，キヤノン製の MR プラットフォーム・システムを用いる．体験者は，ビデオシースルー型 HMD (Canon VH-2002) を装着し，現実空間に CG 映像が重畳描画された 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 を使用する．

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を系統的に配置したマイクロホンアレーを利用する．マイクロホンアレーによって受音され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され，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．1 基のマイクロホンアレーを用いることで音源方向の，複数基のマイクロホンアレーを用いることで音源位置の推定が可能となる．

従来，マイクロホンアレーはその位置を固定した据置型（**図 2 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 に取り付けた装着型（**図 2 右**）という新たな方式を導入し，用途に応じて構成を変更可能とした．

なお，類似の用例として，HRP-2[10]や PaPeRo[11]

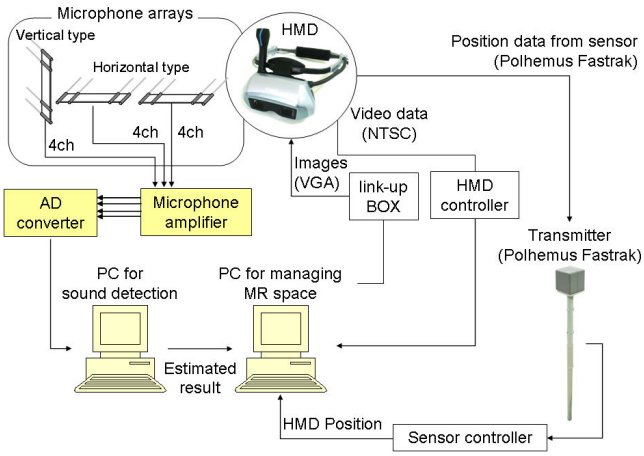


図 1　システム構成
Fig.1 System configuration

など，マイクロホンアレーをロボットに搭載して音像定位能力の向上を図ったものがある．本研究では，体験者，つまり人間の頭部に HMD とともに装着し，体験者の意志でその向きや位置を制御できる点に特徴がある．

3.2 マイクロホンアレーの設計

本研究で用いるマイクロホンアレーは，騒音レベル 50 dBA（多少のざわめきがある）程度の環境で，利用者がおよそ 1.5 m 四方の体験フィールドにおいて MR 空間を体験することを想定し，以下のように設計した．

これまでに開発されたマイクロホンアレーの形状には円形にマイクロホンを配置したもの[12]，渦巻状に配置したもの[13]などがある．しかしながら，本研究では MR 空間での利用を想定している，つまり実時間での処理が要求されるため，計算が簡便な小規模直線型マイクロホンアレーを採用した．

また，マイクロホンアレーの大きさは装着型として用いることを想定し，HMD の上に取り付けても装着者の負担とならないよう 20 cm とした．

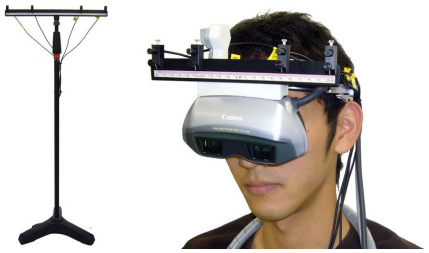


図 2　マイクロホンアレー外観
（左:据置型，右:装着型）
Fig.2 Microphone arrays
(left: fixed type, right: wearable type)

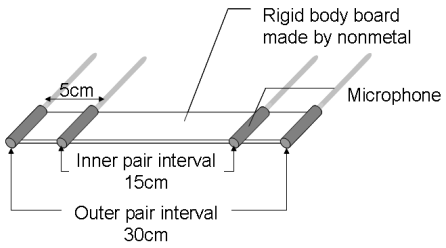


図 3　マイクロホンアレー構成
Fig.3 Microphone array configuration

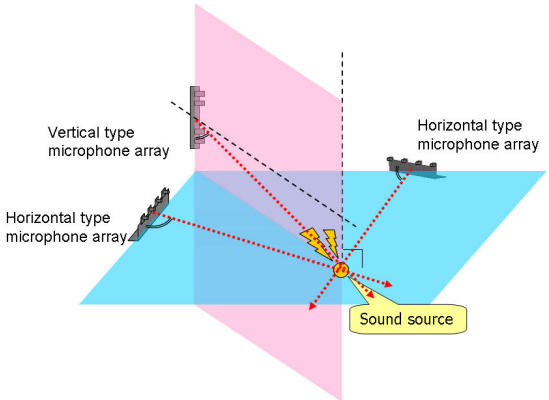


図 4　3 次元音源位置推定
Fig.4 3D localization estimation method

音源方向推定は最低 1 組のマイクロホンペア，つまり 2 ch あれば可能となるが，マイクロホンペアの数を増やすことによって雑音に対して頑健な推定が可能となる．しかし，マイクロホン数が増えるに従ってケーブルが煩雑になり，体験者の動きを制限してしまう恐れがあることから，本研究では 2 組のマイクロホンペア（計 4 ch）を用いた（**図 3**）．

3.3 音源方向の推定

音源方向は，マイクロホンアレー1 基を利用して推定できる．本研究では，雑音に頑健な推定を実現するために，音源方向推定アルゴリズムに西浦ら[14]の提案する CSP 法 (Cross-power Spectrum Phase analysis)[15]を拡張した方法を採用した．なお，本システムで採用したマイクロホンアレーの方向推定精度に関しては，付録 A に記した．

従来，小規模直線型マイクロホンアレーを用いた CSP 法による音源方向推定では，

- マイクロホンアレーの限られた方向でしか推定できない
- 側面方向の精度が正面方向に比べ，低くなる

といった欠点があった．しかし，本研究で初めて導入した装着型マイクロホンアレーでは，ユーザが頭部に装着して頭を動かすことで，音源を正面方向で捉えることができる．このため，装着型ではユーザ周り全方向において精度の高い推定が可能となった．

3.4 音源位置の推定

2 基以上の横型マイクロホンアレーを用い，各マイクロホンアレーから得られた方向と各マイクロホンアレーの位置姿勢から，現実空間の水平面における音源の 2 次元位置が算出できる．それには，マイクロホンアレー自体の MR 空間内での絶対位置が既知でなければならない．本研究では，据置型マイクロホンアレーの位置姿勢は予め実測で測定しておき，装着型マイクロホンアレーは磁気センサによってその位置姿勢を取得する方式を採用した．この方式での推定精度の評価は，付録 B に記した．

さらに，縦型マイクロホンアレーを 1 基追加することによって音源の 3 次元位置が取得できる（**図 4**）．

4. インタラクション実現例

4.1 メニュー選択

まず，体験者の周囲に配した CG のメニューを音イベントにより選択するインタラクションの実現を試みた．ここでは，装着型マイクロホンアレー1 基による音源方向推定結果を利用した．

メニュー項目は現実空間に CG オブジェクトとして重畳描画され，体験者はそのメニュー内容を視認

し，選択したい項目に向けて音イベントを発生させる．メニュー項目は，頭部位置を中心として，距離 60 cm（手を伸ばして選択しやすい距離）の位置に円筒状の離散的帯領域として配置した（図 5）．

体験者が手を伸ばして発した音の音源方向推定結果が含まれる領域を選択項目と見なし（図 6），その領域の色を変化させるという視覚フィードバックを採用した（図 7）．

音イベントの種類に多様性を持たせることもまた，本研究の狙いである．小規模直線型マイクロホンアレーを用いた CSP 法による音源方向推定では，正確な方向推定を行うために，面音源ではなく，点音源であること，全周波数を含む衝突音（インパルス音）を音源に用いることが望ましい．

そこで，この両特性を備えた音源として，まず両手をあわせて打つ「拍手」を採用し，所望の良好な結果を得た．音の種類を増やして，楽器やその他の道具で音を発生する場合には，上記の条件に加え，安定性と使いやすさの観点から，

- 持ちやすい
- 鳴らしやすい
- 操作者が意図している時だけ音が出せる

を満たしていることが望ましい．この条件を満たす物として，カスタネットとブザーを採用した．

音イベントを利用したメニュー選択の操作性を確認するため成人男女 6 名を被験者として実験を行った．被験者の前方 180 度にメニュー項目を 2，4，6 個配置し，カスタネットを使ってメニュー選択を行わせた．上記 3 種類について，1 つのメニュー項目をランダムに 5 回選択する試行を 2 回（計メニュー項目数×5×2 回）繰り返した結果，全ての場合で入力成功率は約 99%であり，さらにメニュー項目の位置による入力成功率の違いは見られなかった．

この機能は，視聴覚併用型 MR アトラクション「Watch the Birdie!」[16]に導入された．このアトラクションの中で，HMD を装着した体験者は音デバイスとしてバードコール[1]を用い，眼前に提示された CG の鳥メニューの中から 1 羽を選択し，（図 8），選択した鳥（CG）が体験スペース内を飛ぶ様を観察することができる．

4.2 方向指示

前述の離散的なメニュー選択が成功したので，次に任意方向での無段階の（すなわち，離散的でない）方向指示のインタラクションを実現した．マイクロホンアレーの構成は，4.1 と同じである．

まず，ユーザが任意方向に向けて音イベントを発生させると，その方向に CG オブジェクトを出現させるといった利用方法を想定し，これを実現した．

[1] バードコール：鳥のさえずりに似た音を発生させる道具．バードウォッチングで使用される．

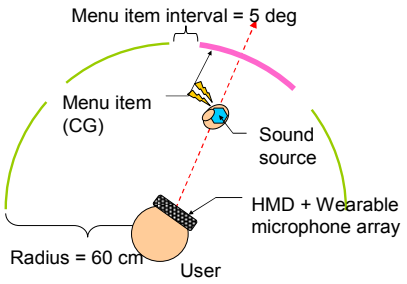


図 5　メニュー項目の配置（項目数 4 の場合）
Fig. 5 Layout of menu items (in the case of four items)

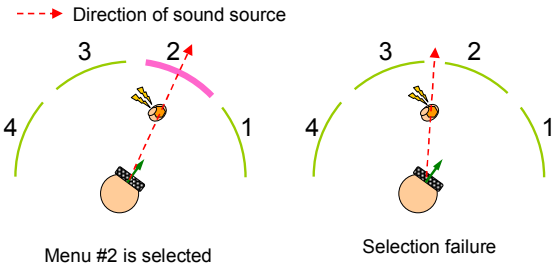


図 6　メニュー選択の成否
Fig. 6 Success and failure in menu selection

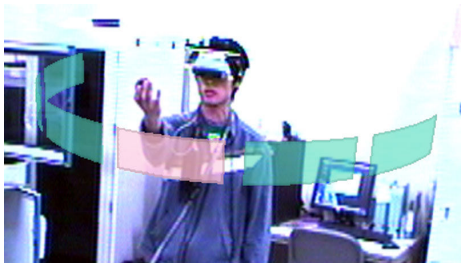


図 7　音入力によるメニュー選択
Fig.7 Selecting a menu item using the sound device



図 8　Watch the birdie!におけるメニュー選択
Fig.8 Menu selection in “Watch the Birdie!”

図 9 に音源（携帯電話のスピーカー）の方向推定を行い，体験者を中心として距離 50 cm，体験者の目線の高さの位置に円錐形の CG オブジェクトを描画している様子を示す．結果は概ね良好で，十分対話デバイスとして利用に耐えるという感触を得た．

この方向選択機能をもっと積極的に用いることは，先述の MR アトラクション「Watch the Birdie!」で実現した．体験者はバードコールを用いて所望の方向にいる鳥を呼び寄せることができる．具体的には，体験者が現実空間で鳴らしたバードコールの方向推定を行い，その方向から鳥が体験者に向かって飛来するというインタラクションを実現した（図 10）．



図 9　音源方向に CG オブジェクトが重畳描画されている様子
Fig.9 The virtual object is superimposed on to the direction of the mobile phone speaker.

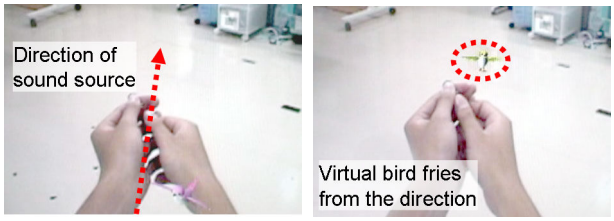


図 10　音デバイスによる方向指示
Fig.10 Pointing a direction using the birdcall

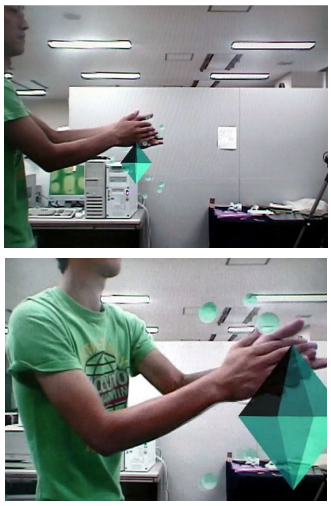


図 11　音源位置に CG オブジェクトが重畳描画されている様子
Fig.11 The virtual object is superimposed on the location of handclap.

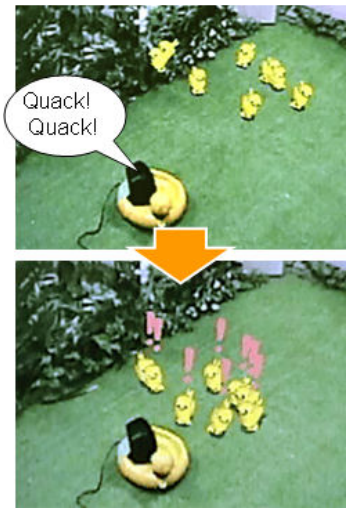


図 12　音源（スピーカー）に集まる子アヒル (CG)
Fig.12 The Virtual ducklings gather toward the position of the real mother toy duck where the squawking sound is played.

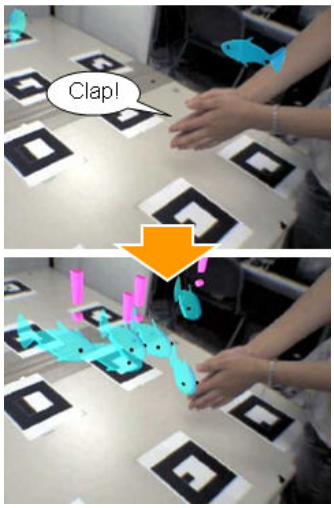


図 13　音源（拍手）に集まる魚 (CG)
Fig.13 Virtual fish could be gathered by handclap.

4.3 2 次元平面のポインティング

3 つ目のインタラクションとして，床や机上など平面上をポインティングするインタラクションを実現した．ここでは，装着型マイクロホンアレーと据置型マイクロホンアレーを各 1 基ずつ用いて，音源位置の推定を行う．

まず，ユーザが任意の位置で音イベントを発生させると，その位置（高さは固定）に CG オブジェクトを出現させるといった利用方法を想定し，これを実現した．**図 11** に音源（拍手）の位置推定を行い，推定された水平位置（目の高さ）に CG オブジェクトを重畳描画している様子を示す（図 9 では，体験者から見た方向に CG オブジェクトが重なって見えているに過ぎない）．

この機能を「Watch the Birdie!」に導入し，CG で描かれた子アヒルが体験スペース内を自由に歩き回っているところに，現実空間で親アヒルが鳴き声を発すると，子アヒルが集まるというインタラクションを実装した．

親アヒルは実在する玩具で，これにスピーカを取り付け，このスピーカから発するアヒルの鳴き声の位置に，子アヒルが集まっていく様子を体験者に提示する（**図 12**）．また，体験者自身がバードコールを鳴らすことで，自分の周りに子アヒルを呼ぶことも可能とした．これは，2 次元平面状の位置のポインティングに音イベントの位置推定を用いているのであって，前述の指定した方向の鳥を呼び寄せるのとは異なる音イベントの利用方法である．

4.4 3 次元空間のポインティング

音イベントによる 2 次元平面のポインティングが実現できたので，続いてこれを 3 次元に拡張した．マイクロホンアレーの構成は，4.3 の構成に縦型のマイクロホンアレーを追加した．

音を使った 3 次元位置ポインティングの事例として，空間内を自由に泳ぐ仮想の魚を，拍手によって任意の場所に呼び寄せるというインタラクションを実装した（**図 13**）．

この他にも，連続して音を発生している音源の位置を常に推定し，その位置に CG オブジェクトを追従させることで，CG オブジェクトを音によって誘導するという利用方法や，一度描画した CG オブジェクトを消去せず，その場に残すことで，連続音を用いて空間に軌跡を描くことのできる空間型ペイントツールなど，新しい表現手法としても利用できると考えられる．

表 1　質問項目
Table 1 Questionnaire items for evaluation

上段：質問内容，下段：回答項目（カッコ内は点数）
質問 1 聴覚フィードバックが加わることで入力行為の確認しやすさがどう変わったか (1) 低下，(2) やや低下，(3) 変わらない，(4) やや向上，(5) 向上
質問 2 音による入力はペン型デバイスと比較して入力しやすいか (1) 困難，(2) やや困難，(3) 変わらない，(4) やや容易，(5) 容易
質問 3 本アトラクションにはどちらのデバイスが相応しいと思ったか (1) ペン型，(2) どちらかといえばペン型，(3) どちらでもよい，(4) どちらかといえば音，(5) 音

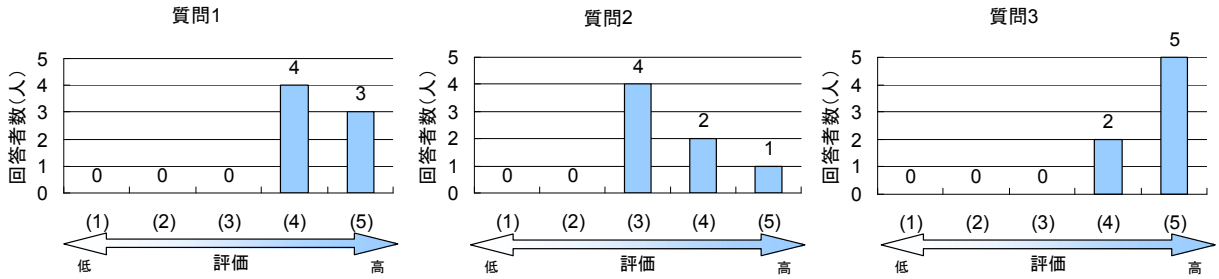


図 14　評価結果
Fig.14 Evaluation result

4.5 従来法との比較実験

提案手法の有用性を確認するため，既存の入力デバイスと本研究で提案する音を利用したインタラクションを比較する実験を行った．

【実験内容】

既存の入力デバイスは，MR 空間とのインタラクションデバイスとしてよく利用されるボタン付きペン型デバイス（Polhemus 社製ペン型デジタイザー ST-8）で，その位置姿勢は磁気センサで検出される．一方，音入力には前述のバードコールを採用した．

実験では 7 名の被験者に，「Watch the Birdie!」における 3 つのインタラクション（方向推定：メニューから鳥を選択する，鳥の飛来方向を指示する，位置推定：子アヒルを呼び寄せる）を，上記 2 つの方法を用いて行なわせた．

実験は，ペン型デバイス，音による入力の順に行い，最後に**表 1** の 3 つの質問に回答させた．

【結果と考察】

実験結果を**図 14** に示す．質問 1 では，すべての被験者が，聴覚フィードバックが加わることで入力行為の確認しやすさが向上，またはやや向上したと回答した．ペン型デバイスでは，ボタン操作による触感と映像変化の視認により入力行為の確認を行なう．これに対して提案手法では，音を鳴らすという行為による触感と音，そして映像変化の 3 つを体感することで確認を行なう．このようにモダリティが増えることで入力確認が容易になったと考えられる．

質問 2 では，音による入力はペン型デバイスによる入力よりもやや容易であるという回答が多く見られた．音による入力はペン型デバイスと比べ精度が低いが，この結果から，今回のように高い精度を求めない用途に対しては提案手法でも十分入力しやすいということがわかる．

質問 3 は，ほとんどの被験者が「Watch the Birdie!」にはバードコールを用いた入力の方が適していると回答した．

以上の回答の他に，被験者からは「バードコールを使って鳥を呼ぶのは，とても自然だ」というコメントも得られた．この評価実験から，方向・位置推定能力では劣っていても，提案手法はその性質に適した使い方をすることで，既存の入力デバイスと同等か，それ以上の優れた働きをすることが確認されたといえる．

5. むすび

本論文では，複合現実 (MR) 空間とのマルチモーダル・インタラクションの新しい手法として，現実空間で発した音イベントの検出結果を 3 次元ポインティング・デバイスとして利用する方法を提案し，複数の実現例を示した．音源の方向・位置推定手法としては，複数のマイクロホンを系統的に配置したマイクロホンアレーを利用する方法を採用したが，本研究は音響分野での新規性を主張するものではなく，MR 空間との実時間対話に新しいモダリティを付加することを主目的としている．

音源の方向・位置推定は，限界性能を追求するのではなく，本利用目的に十分妥当と考えられる精度を確認した上で，各種インタラクション事例を開発した．本手法の発展経過と利用法を段階的に示すために，離散的な方向推定，連続的な方向推定，平面

上の位置推定，3 次元空間内の位置推定の順に実現例を報告したが，この例からもわかるように，本手法は MR 型映像表示との併用，操作時に触感を直接感じられる音発生デバイスの利用で威力を発揮する．また，今回の事例では示さなかったが，振動や反力を返す触力覚フィードバックを組み合せることも容易である．

この種のヒューマンインタフェース研究は，方式の提案だけではなく，実現例を公開し，体験者の意見を求めることも大切である．前述の 4 つの事例の一部は，日本バーチャルリアリティ学会第 11 回大会（2006 年 9 月，於仙台），インタラクション 2007（07 年 3 月，於東京），UIST 2007（07 年 10 月，於米国ロードアイランド州）で技術展示し，以下のような知見や評価を得た．

- 体験者の大半が，磁気センサ等の付加物を身に付けることなく，操作できることを好ましく感じた．
- 拍手，カスタネット，ブザーから発する音は，入力動作を確認しやすく，操作時の触感が操作感を向上させるという意見が得られた．
- 体験者の多くは，特に指示がなくても音源を視野に入れて操作していた．
- 展示会場にはかなりの騒音があったが，若干の吸音材を配置するだけで，（予想以上に）方向・位置検出能力は安定していた．
- 現在，音の種類（音色）は情報として用いていないが，楽器の使い分けや，馴染みのない物でも心地よい音を出すデバイスがあると面白いという意見があった．
- HMD をつけていない体験者も客観視点ディスプレイを眺めながら拍手で音入力を行っていた．

音刺激の検出精度に関して，その物理的性質上，高精度は望んでいなかったが，体験者のほとんどはポインティングの精度向上よりも，新しいモダリティの登場を素直に楽しむ好意的な意見が多かった．

また，装着型マイクロホンアレーについては，体験者から「重い」「わずらわしい」といったコメントが無かったことから，本研究で作成したマイクロホンアレーの設計は妥当であり，体験者は従来の HMD を装着するのと変わらない感覚で，マイクロホンアレーを利用することができていたといえる．

本研究の当初の目的はほぼ達成できたと考えられるので，今後は，上述の意見にもあるように，音の種類に符号情報をもたせることや，音や光を発する新しい対話デバイスの開発に進む計画である．

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] Y. Ohta and H. Tamura (eds.): *Mixed Reality—Merging Real and Virtual Worlds*, Ohm-sha & Springer-Verlag, 1999.

[2] “複合現実感特集 1～3”, 日本バーチャルリアリティ学会論文誌，1999，2002，2005.

[3] K. Higa, T. Nishiura, A. Kimura, F. Shibata, and H. Tamura: “A two-by-two mixed reality system that merges real and virtual worlds in both audio and visual senses,” Proc. of ISMAR 2007, 2007.

[4] S. Irawati, D. Calderón, and H. Ko: “Spatial ontology for semantic integration in 3D multimodal interaction framework,” Proc. of VRCIA 06, pp. 129 - 135, 2006.

[5] Y. Mihara, E. Shibayama, and S. Takahashi: “The migratory cursor: Accurate speech-based cursor movement by moving multiple ghost cursors using non-verbal vocalization,” Proc. of ASSETS 05, pp. 76 - 83, 2005.

[6] T. Nagai, K. Kondo, M. Kaneko and A. Kurematsu: “Estimation of source location based on 2-D MUSIC and its application to speech recognition in cars,” IEEE Proc. of ICASSP 2001, vol. 5, pp. 3041 -3044, 2001.

[7] S. N. Patel and G. D. Abowd: “BLUI: Low-cost localized blowable user interfaces,” Proc. of UIST 2007, pp.217 - 220, 2007.

[8] 伊賀聡一郎, 樋口文人: “Kirifuki: 呼気・吸気によるエンターテイメントシステム”, 日本バーチャルリアリティ学会論文誌, Vol. 7, No. 4, pp. 445 - 452, 2002.

[9] H. Ishii, C. Wisneski, J. Orbanes, B. Chun, and J. Paradiso, “PingPongPlus: Design of an athletic-tangible interface for computer-supported cooperative play,” Proc. of ACM CHI 1999, pp. 394 - 401, 1999.

[10] I. Hara, F. Asano, H. Asoh, J. Ogata, N. Ichimura, Y. Kawai, F. Kanehitro, H.Hirukawa and K. Yamakamoto: “Robust speech interface based on audio and video information fusion for humanoid HRP-2,” IEEE/RSJ Proc. of IROS 2004, pp. 2404 - 2410, Vol. 3, 2004

[11] M. Sato, A. Sugiyama and S. Ohnaka: "Auditory system in a personal robot PaPeRo," IEEE Proc. of ICCE 2006, pp. 19 - 20, 2006.

[12] 中村雅也, 西浦敬信, 李晃伸, 猿渡洋, 鹿野清宏: “マイクロホンアレーを用いた自律移動型ロボットにおける話者位置推定”, 電子情報通信学会 電気音響研究会（音声研究会）, EA2001-4, pp. 25 - 32, 2001.

[13] 東日本旅客鉄道株式会社 エコロジー推進委員会: “社会環境報告書 2004”, p. 9, 2004.

[14] 西浦敬信, 山田武志, 中村哲, 鹿野清宏: “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”, 電子情報通信学会論文誌, Vol. 83-D-II, No. 8, pp. 1713 - 1721, 2000.

[15] M. Omologo and P. Svaizer: “Acoustic event location using a crosspower-spectrum phase based technique,” IEEE Proc. of ICASSP 1994, vol. 2, pp. 273 - 276, 1994.

[16] 石黒祥生, 大槻麻衣, 比嘉恭太, 木村朝子, 柴田史久, 田村秀行: “Watch the Birdie!－三感融合型複合現実感アトラクション”, 日本バーチャルリアリティ学会論文誌, Vol. 12, No. 3, pp. 259 - 268, 2007.

（2007 年 12 月 10 日 受付）

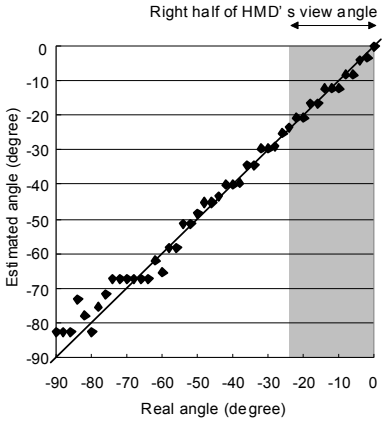


図 15　音源方向推定精度
Fig.15 The result of direction estimation

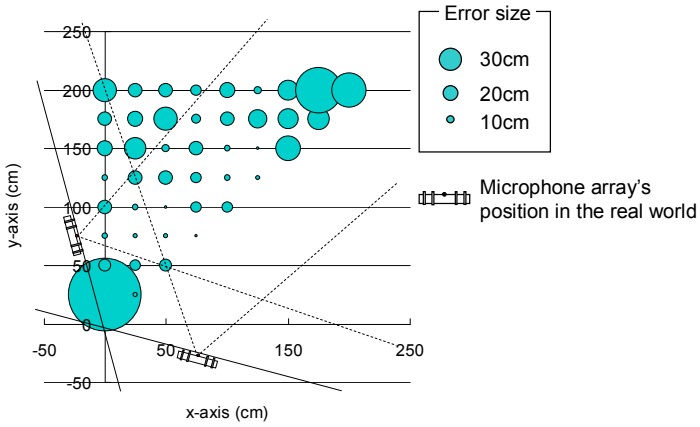


図 16　誤差マップ
Fig.16 Error map

付録

A. 音源方向推定精度の評価

本研究で設計・実装したマイクロホンアレーの方向推定精度評価実験を行なった．

【実験方法】マイクロホンアレーと音源間の距離を 1 m に固定し，正面方向を 0 度として，音源方向を 0 度から 90 度まで 2 度ずつ移動させながら音を発生させ，音源方向と推定結果を比較した．実験は騒音レベル 50 dBA 程度のオフィス環境で行い，音源にはカスタネットを用いた．

【実験結果】(1) 90 度（側面）に近づくほど，推定された方向と音源方向のずれが大きい，(2) HMD の水平視野角内（0 ～ 25.5 度）では精度が高いことがわかった（**図 15**）．以上の結果から，装着型マイクロホンアレーでは，利用者が音源方向を向くことで，利用者の周囲 360 度全方向の音源を精度よく捉え得ることがわかる．

B. 音源位置推定精度の評価

マイクロホンアレー2 基による音源位置推定精度の評価実験を行なった．

【実験方法】据置型マイクロホンアレーは床面から 1 m の高さに，2 基のなす角度が 120 度となるように配置し，音源を高さ 1 m に固定した状態で，水平方向に前後 25 cm ずつ移動させながら音を発生させ，音源位置と推定結果を比較した．実験環境は付録 A と同じである．

【実験結果】実験結果を**図 16** に示す．この図から，計測点がマイクロホンアレーから近く，かつ正面方向であるほど精度が高いことがわかる．

装着型マイクロホンアレーは，利用者が音源方向を向くことでマイクロホンアレーの正面方向で音源を捉えることが可能であり，利用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が短くなるため，上記結果から安定した位置推定精度が期待できる．

ただし，装着型の場合マイクロホンアレーの位置姿勢を常に検出する必要があるため，音源位置推定精度が，利用する位置姿勢センサ自体の有効範囲や精度に左右されるという欠点もある．論文中のシステムでは，比較的手軽に利用可能な磁気センサを頭部位置姿勢センサとして用いているが，もっと高精度な光学式センサ，超音波式センサとのハイブリッド方式を採用することで，装着型の利点はより大きくなる．

また，我々はマイクロホンアレー3 基による音源の 3 次元位置推定についても精度評価実験を行い，同様の結果を確認している．

［著者紹介］

大槻 麻衣　（学生会員）



2006 年立命館大学理工学部情報学科卒．2008 年同 大学院理工学研究科博士前期課程修了．現在，同 研究科博士後期課程に在学中．2008 年日本学術振興会特別研究員．複合現実感，ポスト WIMP 型インタフェースの研究に従事．情報処理学会学生会員．

木村 朝子　（正会員）



1996 年大阪大学基礎工学部卒．1998 年同 大学院基礎工学研究科博士前期課程修了．同 大学助手，2003 年立命館大学理工学部助教授，2004 年同情報理工学部助教授を経て，2007 年 4 月より科学技術振興機構さきがけ研究員，立命館大学総合理工学研究機構客員教授．博士（工学）．実世界指向インタフェース，複合現実感，ハプテックインタフェースの研究に従事．2001 年より 2002 年まで Mayo Clinic にて Special Project Associate．電子情報通信学会，情報処理学会，ヒューマンインタフェース学会，ACM，IEEE 各会員．2006 年本学会学術奨励賞等受賞．

西浦 敬信　（正会員）



1999 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了．2001 年同 研究科博士後期課程修了．和歌山大学システム工学部助手を経て，2004 年 4 月より立命館大学情報理工学部メディア情報学科助教授．現在，同 准教授．博士（工学）．音響信号処理，主として音環境の理解・生成に関する研究に従事．電子情報通信学会，日本音響学会，ISCA の会員．

柴田 史久　（正会員）



1996 年大阪大学大学院基礎工学研究科博士前期課程修了．1999 年同 研究科博士後期課程修了．大阪大学産業科学研究所助手を経て，2003 年 4 月より立命館大学理工学部助教授．現在，同 情報理工学部情報コミュニケーション学科准教授．博士（工学）．モバイルコンピューティング，複合現実感等の研究に従事．現在，University of Central Florida, Media Convergence Lab. 客員研究員．IEEE，電子情報通信学会，日本ロボット学会，情報処理学会等の会員．2005 年本学会学術奨励賞受賞．

田村 秀行　（正会員）



1970 年京都大学工学部電気工学科卒．工業技術院電子技術総合研究所，キヤノン（株）等を経て，2003 年 4 月より立命館大学理工学部教授．現在，同 情報理工学部メディア情報学科教授．工学博士．1997 年より 2001 年まで，MR システム研究所にて「複合現実感研究プロジェクト」を率いた．本学会元理事，現在，評議員，複合現実感研究委員会顧問．編著書「Mixed Reality」(Ohmsha & Springer)「コンピュータ画像処理」（オーム社）など．電子情報通信学会フェロー，IEEE，ACM，情報処理学会，人工知能学会，映像情報メディア学会等の会員．情報処理学会論文賞，人工知能学会功労賞等を受賞．