

## 定位感に基づいて勝ち抜き選択された HRTF による音像定位 Sound Localization with individualized HRTF selected by tournament matches

西塔 宏二<sup>†</sup>  
Koji Saito

岩谷 幸雄<sup>†</sup>  
Yukio Iwaya

鈴木 陽一<sup>†</sup>  
Yôiti Suzuki

### 1. はじめに

オーディオビジュアル技術やバーチャルリアリティシステムには臨場感のあるリアルな音の再現が求められている。この課題を実現するための技術の1つに聴覚ディスプレイがある。聴覚ディスプレイは音源に頭部伝達関数(HRTF:Head-Related Transfer Function)を畳み込みバイノーラル再生することによって、任意の位置に音像を模擬し、リアルな音場を再現する。

聴覚ディスプレイによる提示音像の定位感を最適にするためには、個人ごとに最適な HRTF を与える必要があるといわれている [1]。しかし、本人の HRTF を実際に測定するには、大掛かりな測定環境と時間が必要である。さらに、ヘッドホン受聴の場合には、装着特性を打ち消す必要がある [2]。したがって、個人ごとの HRTF を厳密に合成して優れた定位感を得ることは、一般的な環境での利用を考えると現実的ではない。

一方では多くの聴取者が平均的に定位できる汎用的な HRTF が模索されている [3] が、この場合には HRTF が平均から遠い聴取者への対応ができない。

そこで、これらの問題を解決する手法として定位感に基づく聴覚ディスプレイの個人化手法 DOMISO (Determination of Optimum Impulse-response by Sound Orientation) を提案する。本論文では、DOMISO の詳細と、定位実験によって示唆された DOMISO の有効性について報告する。

### 2. DOMISO について

本手法は、複数人の HRTF を用いて仮想音像を作成し、その中から聴感上最も定位感の良かった音像をトーナメント方式によって選択するものである。手順を以下に示す。

- (1) 聴取者にこれから提示する音像の軌道を教示する
- (2) 複数の HRTF から合成した仮想音像を聴取者に2つずつ提示し、定位感が良かったほうを選択させる
- (3) この選択をトーナメント方式で行い、最も定位感が良かった音像が勝ち残るまで (2) を繰り返す
- (4) (3) によって勝ち残った音像の提示環境を本手法によって個人化された聴覚ディスプレイとする

ここでの定位感とは、「あらかじめ軌道を教示して仮想音像を提示したとき、聴取した音像の軌道が教示のイメージと主観的に合致していた度合い」と定義する。

### 2.1 DOMISO の利点

本手法の利点を以下に述べる。(1) 本手法によって個人化された聴覚ディスプレイは、あらかじめ候補として用意されていた聴覚ディスプレイ群の中で最も定位感の良かった聴覚ディスプレイとなる。すなわち、候補とされる聴覚ディスプレイ群が十分に万人の HRTF 空間を満たすような、バラつきのある集合であれば、万人に対して定位感の良い聴覚ディスプレイを提供することが可能である。(2) 本手法は、聴覚ディスプレイが利用可能な環境であれば個人化を行うことができるため、測定環境の制約がほとんど無い。(3) 定位感が良かったほうの音像を選ぶという簡単なタスクのみで個人化を行っているため、短時間での個人化が可能である。(4) 実際に利用する聴覚ディスプレイの再生系を用いて定位感を最大化しているため、ヘッドホン特性や装着特性を信号処理的に補正する必要が無い。

表1は、従来法と提案法の2つの手法の個人化にかかるコストの対比を示している。表中にあげられている従来法の測定環境は、現在東北大学電気通信研究所で用いられている HRTF の測定環境である。この表からもわかるように、本人の HRTF を測定するには多くの時間が必要である。また、様々な方向から音響信号を提示し、これを耳で録音するという作業が必要のため、大掛かりな測定装置が必要となる。測定中、被測定者は耳栓型マイクロホンを装着しイスに座り、約2時間もの間、頭部や体を動かさないようにじっとしていなければならない。被測定者にかかる負担も大きい。これに比べると、DOMISO は短時間で測定環境の制約もほとんどなく容易に個人化を行うことができると考えられる。

### 3. DOMISO 実施条件

本手法の実施条件を以下に示す。

#### 3.1 候補の HRTF

選択候補となる HRTF は多いほど万人に対応できるが、それだけ選択回数は増え、個人化に時間がかかってしまう。また、トーナメント方式によって選択することを考慮すると、選択候補は  $2^n$  ( $n = 1, 2, 3, \dots$ ) 個が好ましい。なお、本実験で用いた HRTF の選択候補は、東北大学電気通信研究所で測定された120人の HRTF からランダムに32個を選び取っている。

#### 3.2 提示する音像の軌道

これまでの実験では、聴取者の HRTF であっても、提示する音像の軌道によって定位感の評価に差があった [4,5]。よって、どのような音像の定位感を聴覚ディスプレイに求めるかによって、提示する音像の軌道を考慮する必要がある。例えば、聴取者正面方向の定位感が良い聴覚ディスプレイを求めるのなら、提示する音像の軌道

<sup>†</sup>東北大学電気通信研究所 / 大学院情報科学研究科

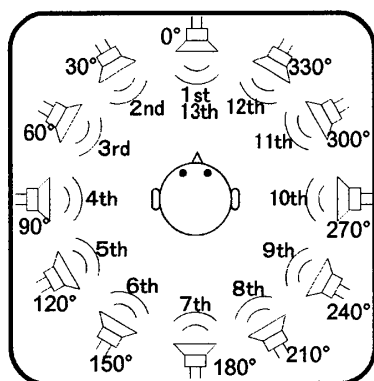


図 1: 定位実験で用いた音像の軌道。(聴取者正面を 0° とし, 1st から 13th の順で提示される。刺激音は 1つの地点につき 1 s 持続する。)

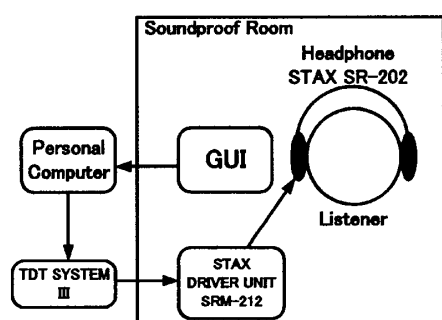


図 2: 提示環境

は正面方向を含んでいなければならない。また、提示する音像が移動音像であるのなら、提示する音像も移動音像にしたほうが良い。本実験では、静止音像を利用して定位精度を検討するため、図 1 に示すような水平面を離散的に回転する音像軌道を使用している。

### 3.3 提示環境

聴取者は実際の提示環境を利用して、定位感の良い音像を選択するため、ヘッドホンの特性や装着特性の補正など、特別な信号処理は必要としないが、正確に定位感の評価が行えるよう雑音などの外部刺激が無い環境が望ましい。なお、本実験で用いた提示環境は、図 2 に示すような環境である。

## 4. DOMISO の有効性の検討

DOMISO によって個人化された聴覚ディスプレイがどの程度の定位精度を持つかを定位実験によって検討した。

### 4.1 実験方法

実験は、第 3.3 節に示す提示環境を用いている。聴取者には、ある方向の HRTF が合成されたピンクノイズ (2 s, 16bit, 48kHz サンプリング) が提示され、聴取者はその提示音がどの方向から聞こえたかを 10° 間隔で回答する。提示する仮想音像の角度は水平角 0° から 330° まで、30° 間隔の 12 方向からである。1 方向あたり 5

回、合計 60 回の提示をランダムに行った。このような定位実験を以下に示す 3 つの条件で行った。

**Own 条件** 本人の HRTF を聴覚ディスプレイに用いた条件

**Individualized 条件** DOMISO によって個人化された聴覚ディスプレイを用いた条件

**Away 条件** DOMISO で用いた HRTF の中で満足な定位感が得られなかった HRTF を聴覚ディスプレイに用いた条件

### 4.2 実験結果

実験参加者は正常な聴力を有する成人男子 4 名である。図 3 は、それぞれの条件での方向定位結果である。図は、提示された仮想音像の角度に対して、それがどの角度に知覚されたかをプロットしており、円の面積は知覚された回数に比例して大きくなっている。この図より、聴取者が若干の前後誤判断をしている傾向がうかがえる。前後誤判断とは、聴取者より前方に提示された音像を聴取者が後方からの音と知覚してしまった状態と、聴取者より後方に提示された音像を聴取者が前方からの音と知覚してしまった状態のことである。このような誤判断を頻繁に起こしてしまう聴覚ディスプレイは定位精度の良いものとは言えない。そこで、各実験条件での前後誤判断の割合を検討した。

図 4 は、それぞれの実験条件での前後誤判断率を示している。図 4 より、全ての聴取者において Individualized 条件は、Own 条件と同程度かそれ以下の前後誤判断率を示している。また、全ての聴取者において Away 条件は前後誤判断が最大になっており、Individualized 条件は Away 条件の前後誤判断率以下を示していることがわかる。

ここで、各実験条件が与える前後誤判断率に差がないかを調べるため、前後誤判断率に対する実験条件 1 要因 3 水準の分散分析を行った。その結果、実験条件による有意傾向 ( $F_{(2,6)} = 4.61, p = .061$ ) がみられたので、どの水準間に差があるのかを調べるため下位検定として多重比較を行った。その結果、Own 条件と Individualized 条件には有意傾向はみられなかったが、Own 条件と Away 条件、Individualized 条件と Away 条件には有意傾向 ( $MS_e = 24.78, p < .1$ ) がみられた。

これより、DOMISO によって個人化された聴覚ディスプレイは、本人の HRTF を聴覚ディスプレイに用いた場合と同程度の定位精度が期待できると考えられる。

## 5. まとめ

定位感に基づく聴覚ディスプレイの個人化手法として DOMISO を提案し、その有効性を検討するため定位実験を行った。その結果、定位感に基づいて勝ち抜き選択された HRTF により、本人のものをを用いた場合と同程度の定位精度が得られる可能性が示唆された。

測定の制約なく短時間で優れた定位感と定位精度を持つ聴覚ディスプレイが得られることから、提案法である DOMISO は非常に有効であると考えられる。

表 1: 個人化にかかるコスト

	HRTF の測定による個人化 (従来法)	DOMISO による個人化 (提案法)
必要時間	約 2 時間	約 15 分
測定装置	球状スピーカアレイ, 耳栓型マイクロホンなど	ヘッドホン, PC など
聴取者課題	頭部や体を動かさないようにしてイスに座る	提示音から定位感の良い音像を選ぶ

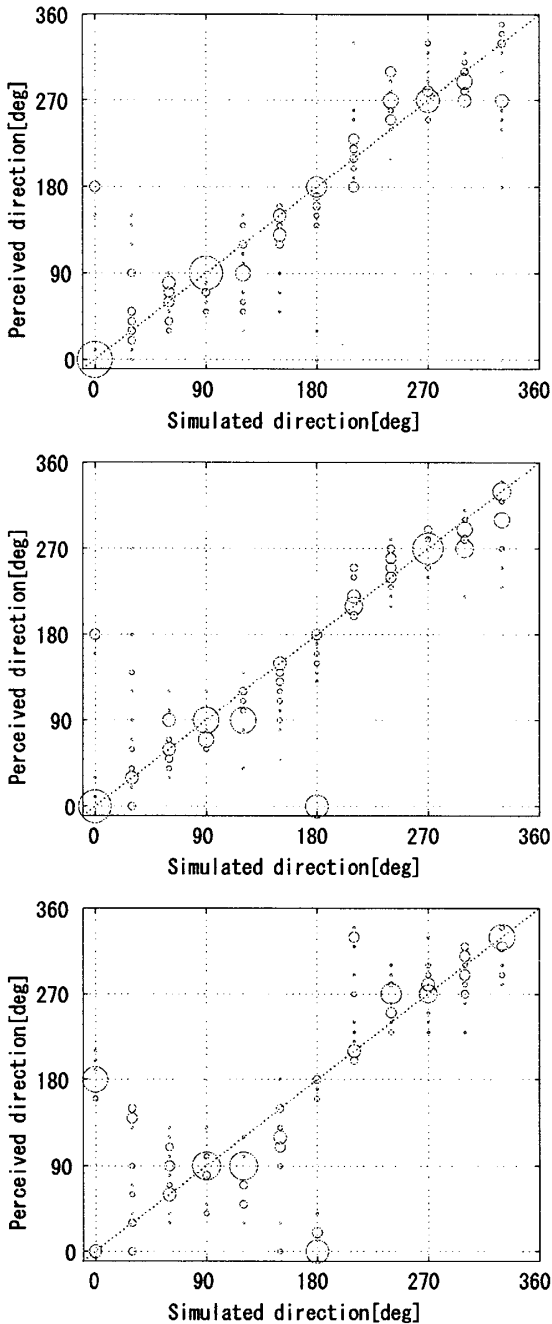


図 3: 定位実験の結果 (上: Own 条件での方向定位, 中: Individualized 条件での方向定位, 下: Away 条件での方向定位)

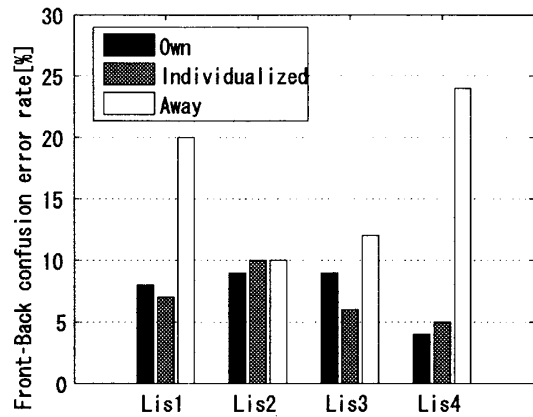


図 4: 定位実験の結果 (各条件での前後誤判断率)

参考文献

- [1] M. Morimoto and Y. Ando, J. of Acoust. Soc. Jpn(E), 29, pp.167-174(1980).
- [2] Henrik Møller, "Fundamentals of Binaural Technology", App. Acoust., 36, pp.171-218 (1992).
- [3] 飯田ら, 音講論 (秋), pp.297-298(2000).
- [4] 西塔宏二, 岩谷幸雄, 鈴木陽一"定位感に基づく個人化頭部伝達関数の勝抜き戦選択". 信学技報, 1-6,(2004).
- [5] 西塔宏二, 岩谷幸雄, 鈴木陽一"移動音像の定位感に基づく聴覚ディスプレイの個人化". 2004 年日本音響学会秋季大会・講演論文集, 659-660(2004).