

令和 4 年 6 月 9 日現在

機関番号：32503

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K12068

研究課題名(和文)音の高精度バーチャルリアリティを実現するための個人化頭部伝達関数の生成

研究課題名(英文)Generation of head-related transfer functions of an individual listener for acoustic virtual reality

研究代表者

飯田 一博 (Iida, Kazuhiro)

千葉工業大学・先進工学部・教授

研究者番号：60458627

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：音のVR(3次元音響再生)を実現するために、個人の頭部伝達関数を生成する方法を検討した。その結果、耳介形状が得られれば、正面方向の音像定位が可能な個人化頭部伝達関数を生成できることが示された。また、より高精度な音像定位を実現するために着目すべき新たな耳介部位に関する知見が得られた。ただし、一般消費者がスマートフォンなどで撮影可能な画像から耳介形状データを得るためには、機械学習ベースのアルゴリズムを構築する必要があることが示唆された。またコロナ禍による行動制限を受けて、当初計画を一部変更して、聴取者が試聴しながら全天空の個人化頭部伝達関数を自身で生成できるツールキットを開発した。

研究成果の学術的意義や社会的意義

従来、耳介形状から個人化頭部伝達関数を推定する研究は多くあるが、生成した頭部伝達関数を用いた音像定位実験により、目標方向は正面だけではあるものの、その有効性を確認したのは世界で最初である。また実用面においては、本研究の成果を用いれば、従来の研究では必要とされていた頭部伝達関数データベースが不要となるため、一般の聴取者が本人に適した頭部伝達関数を自ら生成することが可能になり、音のVRの普及が期待できる。

研究成果の概要(英文)：In order to realize VR in sound reproduction (3D sound reproduction), we examined a method to generate the individual head-related transfer functions (HRTFs). As a result, it was shown that the individualized HRTFs, for which an individual listener localized a sound image in the front direction, can be generated using the data of pinnae shape. In addition, we obtained knowledge on the pinna region that should be used in the HRTF individualization in order to improve the accuracy of sound image localization. However, it was suggested that a machine learning-based algorithm should be considered for general consumers to obtain the pinna shape data from the pinna images that can be taken with a smartphone. In addition, due to the behavioral restrictions caused by COVID-19, we have partially modified the initial research plan and developed a toolkit, by which listeners can generate the individual HRTFs in the upper hemisphere by themselves.

研究分野：空間音響学

キーワード：バーチャルリアリティ 頭部伝達関数 HRTF 個人化 3次元音響

1. 研究開始当初の背景

バーチャルリアリティ (VR) の研究開発が世界中で進められている。音の VR (3次元音響再生) は、聴取者本人の頭部伝達関数 (HRTF) を用いれば実現できるが、他人の HRTF では前後方向を誤って知覚するなどの問題が発生することが知られている。この問題を解決するため、各聴取者に適合する個別の HRTF を提供する方法、すなわち HRTF の個人化方法が研究されている。これまで研究代表者は、HRTF に含まれる本質的な情報、すなわち方向知覚の手掛かりを解明する音像定位実験を行い4つのファクタ (4kHz以上で最も周波数の低い2つのノッチ N1、N2 と2つのピーク P1、P2) が3次元音響再生において特に重要であることを示している [1、2]。加えて研究代表者は各聴取者の HRTF の振幅スペクトルを耳介形状から重回帰式により推定して生成することにより、4ファクタを含む振幅スペクトルを生成できる可能性を示している [3]。しかし、音の VR を実用化し、社会に普及するためには、HRTF の生成精度は不十分である。また、聴取者の耳介形状をスマートホンのカメラなどから自動計測するシステムの開発も必要である。

2. 研究の目的

本研究では、以下の4点を目的とした。1) 聴取者の耳介形状から HRTF のノッチ・ピーク周波数を推定する方法を確立し、デジタルフィルタで個人化 HRTF を生成する。2) HRTF 生成を高精度化する耳介形状計測箇所を探索する。3) 耳介形状パラメータを自動計測することを試みる。4) 生成した個人の HRTF の有効性を音像定位実験により確認する。

3. 研究の方法

(1) テーマ1：耳介形状パラメータによる個人の HRTF 生成

2019年度は個人化 HRTF の生成方法の検討に先立って、まず HRTF のノッチとピークを精度よく抽出するために HRTF のモデル化方法を検討し、モデルのパラメータの最適化を試みた。さらに、このモデルと判別分析により、個人の HRTF のノッチ周波数が存在する周波数帯域を推定する方法を検討した。

2020年度は N1、N2 周波数帯域の推定精度の向上を検討した。さらに、推定した個人の N1、N2 周波数 bin を用いて個人の HRTF を生成し、正面方向を対象とした音像定位実験を実施した。

2021年度はコロナ禍において蜜を避けるため、耳介形状データ入手に大幅な行動制約が生じ、前年度までに確立した個人の HRTF の生成方法を多数の被験者を用いて検証することができなかった。そこで代替案として、聴取者が試聴しながら適合する HRTF を自身で生成できる方法を検討した。

(2) テーマ2：HRTF 生成を高精度化する耳介形状計測箇所の探索

2019年度は音響計測用マネキン (KEMAR) の形状データを用いて、初期 HRTF に胴体からの反射波の影響がどのように現れるかシミュレーションにより検討した。また、耳介を左右・上下・前後および全方向に伸長したときの上半球正中面における N1、N2、P1、P2 の変動を MRI で計測した4名の頭頸部から切り出した耳介形状と KEMAR の耳介形状をシミュレーションにより検討した。

2020年度は耳介形状から初期 HRTF の振幅スペクトルを求める際により精度が高くなる可能性がある計測点を探索するために、MRI で計測した12名の頭頸部から切り出した耳介形状でトポロジカルに同じ点を薄板スプラインで求めて微小変形を加え、点ごとに伝達関数の変動量を計算した (図1)。また、耳介の側面観で陰になる空間が伝達関数に与える影響を検討した。

2021年度は2020年度までにコンピューターシミュレーションによって特徴点となる可能性が高い箇所を見だし、それらを機械学習によって自動検出することを試みた。耳介の側面画像から耳介各部の輪郭抽出を機械学習で実現できるか否かを検討するために計算機によって学習データを増やしたが、十分な効果が得られず、今後の発展が期待できる性能は得られなかった。そのため、従来法の輪郭抽出方法による特徴点の検討へ方針転換し、テーマ3と融合して検討することとなった。

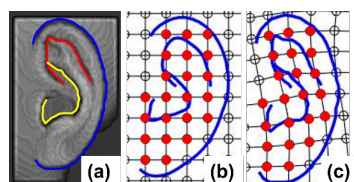


図1 (a) 耳介のトレース画像、(b) 平均耳介と直交格子 (赤は微小変形を加える点)、(c) ある耳介に薄板スプラインを用いて写像した格子点。

(3) テーマ3：耳介形状パラメータの自動計測

耳介形状の特徴量と頭部伝達関数の関連を見いだす実験を実施するための自動計測手法およびそのシステム開発を目標として研究を進めた。2019年度には、ユーザー自身で簡便に耳介形状を計測できることを目指し、ポータブルデバイスを利用して自ら耳介形状を撮影する方法を検討した。設計では、耳介輪郭を自動検出して画角内に計測対象が収まることを音などで通知することが想定した。特殊なセンサーを有する情報機器に限らずに利用できることを目指し、可能な限り画像から耳介特徴量を抽出する方針とした。具体的な手法案としては、焦点を連続的に変化させ、得られた画像群からボケの情報を利用して奥行を推定する手法の開発に取り組んだ。

2020年度には、耳介形状パラメータの自動抽出アルゴリズムの開発をしつつ、自動計測システムを開発にも着手した。システムは、初期段階として、ある程度の実験者の手作業によるパラメータの補正もできるように設計した。複数の実験者が多くの耳介画像を同時に処理するためにwebベースで作業できることをシステムの設計指針とした。フロントエンドは、ブラウザとjavascriptで画像のアップロード、計測ポイントの入力、計測値のリアルタイム表示、計測したすべてのデータを所定の書式で出力およびダウンロードできることとした。

2021年度は、実験計画での評価実験の実施を目指し、1) 耳介形状パラメータの自動抽出の検討、2) システムへの統合に取り組んだ。

1) 耳介形状パラメータの自動抽出

飯田らは耳珠から耳介の輪郭の各部までの距離を特徴量とすることを提案しており[3]、その特徴量を抽出することを検討した。機械学習を利用することにより直接的に耳珠および輪郭を抽出する方法が考えられたために、適切な機械学習のアルゴリズムの検討を行った。また、機械学習ではなく画像処理分野で多くの実績がある手法をベースとしたアルゴリズムも並行して検討した。

2) 耳介形状パラメータの自動計測システム

2020年度に設計したシステムのフロントエンドおよびデータの入出力などの耳珠位置の推定および輪郭抽出ブロック以外の実装を行った。

(4) テーマ4：耳介画像から自動計測した個人のHRTFの有効性実証

コロナ禍において蜜を避けるため、耳介形状データ入手に大幅な行動制約が生じ、2020年度までに確立した個人のHRTFの生成方法を多数の被験者を用いて検証することができなかった。そこで代替案として、(1)に記載したように聴取者が試聴しながら適合するHRTFを自身で生成できる方法を検討することとした。

4. 研究成果

(1) テーマ1：耳介形状パラメータによる個人のHRTF生成

2019年度の研究の結果、上昇角知覚において重要な手掛かりとなるN1、N2、P1、P2を適切に抽出するためのHRTFモデルの分割帯域幅は1/12 oct. もしくは1/6 oct. 以下にする必要があることが明らかになった。さらに、このモデルを用いて判別分析によりN1、N2が存在する周波数帯域(1/12 oct. 間隔)を推定するアルゴリズムを構築した。112耳のHRTFに対して検証を行ったところ、教師データに対する弁別閾内的中率はN1、N2それぞれ76.0%、80.0%であった。未知の耳介では65.0%、62.1%であった。

2020年度の研究の結果、未知の耳介に対する正面方向のN1周波数binカテゴリー的中率は83.1%、N2周波数binカテゴリー的中率は77.8%に向上した。後方のN1周波数binカテゴリー的中率は74.3%、N2周波数binカテゴリー的中率は77.1%であった。さらに、推定した個人のN1・N2周波数binカテゴリーを用いて個人のHRTFを生成し、正面方向を対象とした音像定位実験を実施した結果、精度の高い音像定位が得られる可能性を示した(図2)。

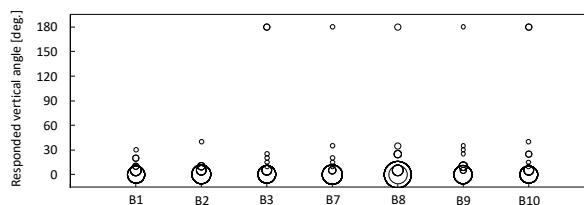


図2 正面方向(0°)を目標方向とした音像定位実験結果。横軸のB1からB10は刺激の種類を表す。縦軸は被験者の回答方向で、0°、90°、180°はそれぞれ、正面、天頂、後方を表す。○の大きさは回答の頻度を表す。

2021年度の研究の結果、正面、天頂、後方の3方向のPNP HRTFモデルを構築した。PNPモデルとは、HRTFを2つのノッチと2つのピークで構成し、それらをIIRフィルタを用いてパラメトリックに表現するモデルである。このHRTFをある聴取者に適合させようとする、3パラメータ×4ピークノッチ=12次元での最適化問題となるため解決は容易ではない。そこで、他のパラメータからの回帰、もしくは定数化することにより、実質的に1パラメータで個人最適化が可能

なモデルを構築した。

さらに、3 方向から正中面の任意の方向の個人化 HRTF を推定する回帰式を求めた。これに両耳間差情報を付加することにより、全天空の任意の方向の個人化 HRTF が生成できるようになった。生成した個人化 HRTF の音像定位実験を行った結果、水平面内の任意の方向への精度よい音像定位が可能であることを示した。しかし、上方での定位精度においては改善の余地を残している (図 3)。

また、PNP モデルを用いた全天空個人化 HRTF の生成と 3D レンダリングを具現化するツールキット (アプリケーションソフト) を開発した (図 4)。

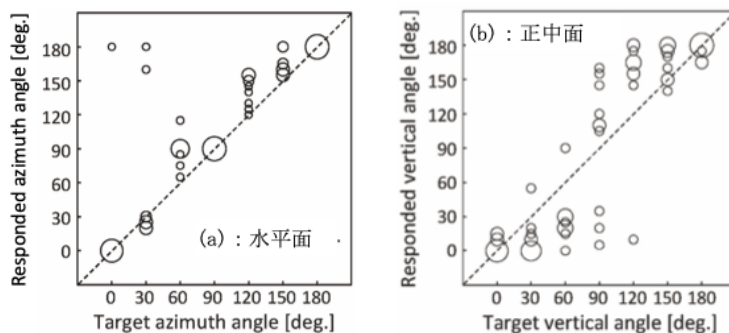


図 3 生成した個人化 HRTF による音像定位実験結果。(a) : 水平面、(b) : 正中面。横軸は目標方向、縦軸は被験者の回答方向を示す。○の大きさは回答頻度を表す。

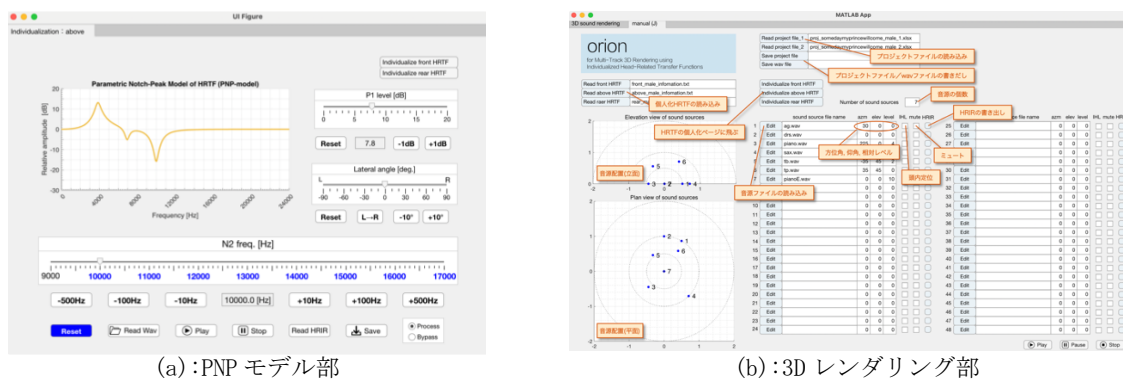


図 4 開発したツールキット (アプリケーション) の GUI。

(2) テーマ 2 : HRTF 生成を高精度化する耳介形状計測箇所探索

2019 年度の研究の結果、胴体からの反射波の初期 HRTF への影響は真上方向に近づくほど小さくなり、上半球正中面ではほとんど現れないことが明らかになった。しかし、下半球正中面では 2.5 kHz の帯域に影響が現れることも明らかになった。また、耳介を全方向に伸長すると N1、N2、P1、P2 周波数は伸長率の逆数倍で低下したが、どれか一方に伸長すると低下の度合いが弱まった。一方、伸長方向と N1、N2、P1、P2 レベルの変動には一定の傾向はみられないことが明らかになった。

2020 年度の研究の結果、耳介切痕と耳甲介舟の微小変形が伝達関数に与える変動量が最も大きいことが明らかになった (図 5)。また、耳介の側面観で陰になる空間の部分埋めた場合、耳珠の内側など耳甲介腔外縁部を埋めたときに伝達関数への影響が最も大きいことが明らかになった。すなわち、これらの部分の形状が耳介形状から HRTF の振幅スペクトルを重回帰で求める際に重要である可能性が高いことが明らかになった。

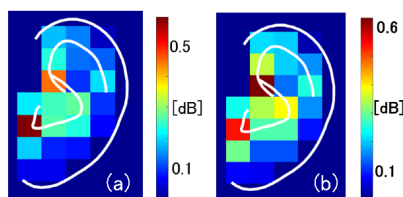


図 5 耳介各部に微小変形を与えたときのスペクトルの平均変動量。(a) 膨張時 (b) 収縮時。

(3) テーマ3：耳介形状パラメータの自動計測

奥行情報を特殊なセンサーを利用せずに画像から抽出することを試みたが、一般消費者が所有する情報端末機器では、被写体深度が浅い場合は十分な奥行の解像度が得られないことがわかり、画像処理を施しても目的とする精度は得られなかった。機械学習をベースとしたアルゴリズムでは、コロナ禍のために教師用データが十分に確保できない状況となり、期待する精度は得られなかった。そのため、被写体深度を固定して得られた画像を用いて、既存の輪郭抽出手法をベースにしてアルゴリズムを開発することが望ましい方針が得られた。

飯田らは耳珠から耳介の輪郭の各部までの距離を特徴量とすることを提案しており[3]、起点を耳珠からマーカーに置き換えてマーカーと耳介主要部の輪郭の交点を特徴量とすることを検討した。その結果、従来の輪郭抽出手法ベースとしたアルゴリズムを用いると、耳介撮影時の光源や顔の向きによる影響により輪郭が安定して得られないことが明らかになった。飯田らが提案している耳珠を起点とする特徴量と音像定位の個人性の関係の知見を有効的に活用するために、耳珠の自動検出の検討を行った。その結果、従来法の輪郭抽出方法をベースに開発した手法では、光源位置などの画像撮影時の条件をある程度整えることにより、耳穴付近の点を検出する可能性が得られたが、耳珠を安定して検出することはできなかった。

以上の検討により、今後は耳珠および輪郭に拘らずに画像処理によって得られる特徴点を基準とした頭部伝達関数の個人性の検討することが効率的であるという知見が得られた。

自動計測システムの実現においては、耳介形状パラメータの自動抽出アルゴリズムのブロック以外は完成しており、今後は、耳介画像と頭部伝達関数が紐付いたデータベースを充実させて、機械学習ベースのアルゴリズムによる手法を開発する、もしくはテーマ2で得られた特徴量に關係する陰影をベースとした特徴量を定めて頭部伝達関数との關係を導くアプローチが一案であることが知見として得られた。

(4) テーマ4：耳介画像から自動計測した個人のHRTFの有効性実証

コロナ禍において蜜を避けるため、耳介形状データ入手に大幅な行動制約が生じ、2020年度までに確立した個人のHRTFの生成方法を多数の被験者を用いて検証することができなかった。そこで代替案として、(1)に記載したように聴取者が試聴しながら適合するHRTFを自身で生成できる方法を検討し、水平面内の任意の方向への精度よい音像定位が可能であることを示した。

<引用文献>

[1] Iida *et al.*, “Median plane localization using parametric model of the head-related transfer function based on spectral cues,” *Applied Acoustics*, 68, 835-850 (2007).

[2] Iida and Ishii, “Effects of adding a spectral peak generated by the second pinna resonance to a parametric model of head-related transfer functions on upper median plane sound localization,” *Applied Acoustics*, 129, 239-247 (2018).

[3] Iida *et al.*, “Generation of the individual head-related transfer functions in the upper median plane based on the anthropometry of the listener's pinnae,” *Applied Acoustics*, 155, 280-285 (2019).

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Kazuhiro Iida, Tsubasa Aizaki, Takeshige Kikuchi	4. 巻 189
2. 論文標題 Toolkit for individualization of head-related transfer functions using parametric notch-peak model	5. 発行年 2022年
3. 雑誌名 Applied Acoustics	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.apacoust.2021.108610	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kazuhiro Iida, Oriie Nishiyama, and Tsubasa Aizaki	4. 巻 177
2. 論文標題 Estimation of the category of notch frequency bins of the individual head-related transfer functions using the anthropometry of the listener's pinnae	5. 発行年 2021年
3. 雑誌名 Applied Acoustics	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.apacoust.2021.107929	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計18件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 西山織絵, 飯田一博
2. 発表標題 正面方向の個人化パラメトリック頭部伝達関数の生成 - ノッチとピークのレベル・尖鋭度設定における固定値の適用性検証 -
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2021年

1. 発表者名 相崎翼, 菊地勇成, 飯田一博
2. 発表標題 パラメトリックノッチピークモデルを用いた頭部伝達関数の個人化ツールキット
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2021年

1. 発表者名 相崎翼, 菊地勇成, 飯田一博
2. 発表標題 パラメトリックノッチピークモデルを用いた頭部伝達関数の個人化ソフトウェア
3. 学会等名 日本音響学会2021年秋季研究発表会
4. 発表年 2021年

1. 発表者名 相崎翼, 飯田一博
2. 発表標題 頭部伝達関数のノッチとピークの振幅レベルが正中面の音像定位精度に及ぼす影響
3. 学会等名 日本音響学会2021年秋季研究発表会
4. 発表年 2021年

1. 発表者名 菊地勇成, 飯田一博
2. 発表標題 頭部伝達関数のノッチ周波数の変移が音像の前後誤判定に及ぼす影響
3. 学会等名 日本音響学会2021年秋季研究発表会
4. 発表年 2021年

1. 発表者名 中村風香, 飯田一博
2. 発表標題 全天空個人化頭部伝達関数による3Dレンダリングツールキット
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2022年

1. 発表者名 中村風香, 飯田一博
2. 発表標題 全天空個人化頭部伝達関数による3Dレンダリングツールキット
3. 学会等名 日本音響学会2022年春季研究発表会
4. 発表年 2022年

1. 発表者名 Kazuhiro Iida, Orié Nishiyama, and Tsubasa Aizaki
2. 発表標題 Estimation of the category of notch frequency bins of the individual head-related transfer functions using the anthropometry of the listener's pinnae
3. 学会等名 179th Joint Meeting Acoustical Society of America (国際学会)
4. 発表年 2020年

1. 発表者名 西山織絵, 飯田一博
2. 発表標題 耳介形状による個人の頭部伝達関数のノッチ周波数のカテゴリ推定
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 飯田一博, 西山織絵, 相崎翼
2. 発表標題 個人の頭部伝達関数のノッチ周波数 binのカテゴリ推定
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2020年

1. 発表者名 飯田一博, 西山織絵, 相崎翼
2. 発表標題 個人の頭部伝達関数のノッチ周波数 binの推定
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2020年

1. 発表者名 川村海斗, 竹本浩典
2. 発表標題 耳介各部の微小変形が伝達関数に与える影響の検討
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 金子祐太郎, 竹本浩典
2. 発表標題 耳介の側面観で陰になる部分が伝達関数に及ぼす影響
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 川村海斗, 竹本浩典, 菅木禎史, 飯田一博
2. 発表標題 頭部インパルス応答の初期部分に含まれる胴体からの反射波に関する検討
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 金子祐太郎, 竹本浩典, 菅木禎史, 飯田一博
2. 発表標題 耳介形状の伸長が伝達関数に及ぼす影響の検討
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 西山織絵, 相崎翼, 飯田一博
2. 発表標題 頭部伝達関数の帯域分割ノッチ・ピークモデル - 分割帯域幅と音像定位精度の関係 -
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 飯田一博, 西山織絵, 相崎 翼
2. 発表標題 個人の頭部伝達関数のノッチ周波数binの推定
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2020年

1. 発表者名 Kazuhiro IIDA and Fuka NAKAMURA
2. 発表標題 TOOLKIT FOR 3D AUDIO RENDERING USING INDIVIDUALIZED HEAD-RELATED TRANSFER FUNCTIONS IN THE UPPER HEMISPHERE
3. 学会等名 The 27th International Conference on Auditory Display (ICAD 2022) (国際学会)
4. 発表年 2022年

〔図書〕 計0件

〔出願〕 計2件

産業財産権の名称 頭部伝達関数生成装置、プログラム及び頭部伝達関数生成方法	発明者 飯田一博，中村風香	権利者 学校法人千葉工業大学
産業財産権の種類、番号 特許、特願2022-024205	出願年 2022年	国内・外国の別 国内

産業財産権の名称 頭部伝達関数生成装置、頭部伝達関数生成プログラム及び頭部伝達関数生成方法	発明者 飯田一博	権利者 学校法人千葉工業大学
産業財産権の種類、番号 特許、特願2020-90035	出願年 2020年	国内・外国の別 国内

〔取得〕 計1件

産業財産権の名称 頭部伝達関数生成装置，頭部伝達関数生成プログラム及び頭部伝達関数生成方法	発明者 飯田一博	権利者 学校法人千葉工業大学
産業財産権の種類、番号 特許、特許第6986778	取得年 2021年	国内・外国の別 国内

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	竹本 浩典 (Takemoto Hironori) (40374102)	千葉工業大学・先進工学部・教授 (32503)	
研究分担者	菅木 禎史 (Chisaki Yoshifumi) (50284740)	千葉工業大学・先進工学部・教授 (32503)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------