

機関番号：15501

研究種目：若手研究(B)

研究期間：2007～2010

課題番号：19760176

研究課題名(和文)音フィードバック手動制御系のための適応型音響ディスプレイの開発

研究課題名(英文) Development of adaptive auditory display for use in manual auditory feedback control system

研究代表者

藤井 文武 (FUJII FUMITAKE)

山口大学・大学院理工学研究科・准教授

研究者番号：30274179

研究成果の概要(和文)：

ステレオヘッドホンにより提示される音信号により聞き手に空間内の任意の位置に仮想的に音源が存在するような知覚を与える「仮想音像定位」技術は、その原理から聞き手により高い定位感を感じる場合とそうでない場合があることが知られている。本研究では、聞き手が誰でもあっても次第に高い定位感を感じる事が出来るような提示音を適応的に構成することのできる音響ディスプレイシステムの一構成法を提案した。

研究成果の概要(英文)：

Virtual sound localization is a technology which gives the listener the sensation of existence of a virtual sound source at a specified spatial location in the binaural listening environment. It is well known that the effect of sound localization varies listener to listener, for the technology depends on some transfer functions of the sound transmission which suffers large personal difference. This research proposes an auditory display system of the binaural listening environment which adaptively absorbs the personal difference and improves sound localization precision.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,200,000	0	1,200,000
2008年度	600,000	180,000	780,000
2009年度	400,000	120,000	520,000
2010年度	800,000	240,000	1,040,000
年度			
総計	3,000,000	540,000	3,540,000

研究分野：工学

科研費の分科・細目：機械工学・知能機械学・機械システム

キーワード：マンマシンインターフェース、頭外音像定位、頭部伝達関数(HRTF)、頭外音像定位伝達関数(SLTF)、外耳道伝達関数(ECTF)、ニューラルネットワーク、手動制御系

## 1. 研究開始当初の背景

人間がフィードバック制御系におけるコントローラの役割を担う手動フィードバック制御系に関する研究は、制御を行う人間の制御行為および特性のモデル化や、それを踏まえた人間の操作支援と系全体としての制御性能向上に関する研究が、1960年代を端

緒として活発に行われてきた。普遍的に有効なモデルが示されるとともに、数多くの支援手法が提案されているが、これら先行研究のほとんどは、制御量と目標値の偏差の検出手段として視覚を想定している。

一方、手動制御系におけるセンサデバイスとして聴覚のみを用いる系については、視覚

による手動フィードバック制御系の古典的解析結果である，人間の閉ループ系の制御調整能力を示すモデルであるクロスオーバーモデルが聴覚フィードバック手動制御系においても成立するかという観点で 70 年代にいくつかの研究が行われ，クロスオーバーモデルは聴覚フィードバック手動制御系においても成立するものの，そのクロスオーバー周波数は視覚フィードバックの場合と比較して低くなる，つまり制御性能は視覚利用の場合に比して劣るという結果が報告されている．その後聴覚フィードバック手動制御系の研究は一時途絶えているように見えるが，近年の音響工学の進展により人間への音情報提示技術は進歩しており，かつ計算機の処理能力の飛躍的向上によって，人間に知覚される音情報を提示側の意図どおり知覚させるためにある程度複雑な演算処理を行っても実時間で再生することも可能になっていると考えられる．そこで，今回の研究では，改めて「聴覚フィードバック手動制御系の性能向上と操作支援」に取り組むことにより，将来的には視覚障害を有する方にも利用していただけるような，新たな人間機械系のインターフェースを開発することを目指す．1970 年代の時点ではなし得なかった性能を有する，実用に足るシステムの構築を目標とする．

## 2. 研究の目的

聴覚フィードバック手動制御系においては，制御のために必要な信号の値を，音のパラメータと関連付けて提示し，その量を人間に可能な限り正確に知覚させる必要がある．信号値の音パラメータへの関連付けには，70 年代に試されたものも含め様々な方法が考えられる．研究代表者が本研究に先立って行った，制御に関する信号値の音パラメータへの関連付け方法が，外性目標信号への追従を課した手動制御系の制御性能に与える影響を検証する実験の結果として，「信号の音情報への変換には，信号値の極性と大きさを空間内を仮想的に移動する仮想音像の位置と対応付けて提示する方法が最も性能が高い」，「人間を 2 自由度制御系のコントローラとして 2 つの音像を提示すると逆に混乱し制御性能が悪化する」の 2 つの知見が得られた．これより，提示音を空間内に仮想的に定位させる頭外音像定位技術の定位精度を向上させることが結果として音フィードバック手動制御系の制御性能向上につながると期待される．そこで本研究では，この仮想音像の頭外音像定位の精度を制御工学的手法を用いて適応的に向上させていく音提示手段である，適応型音響ディスプレイの開発を目指すこととした．

## 3. 研究の方法

頭外音像定位伝達関数 (SLTF) とは，空間内のある位置に存在する音源から発せられた音信号を入力，対応して外耳道の入口で観測される音信号を出力としたときに入出力間の伝達特性を表す関数として定義される．音源の位置が異なれば，外耳道入口までの伝達経路および経路中に存在する障害物も異なるため，SLTF は音源位置 (方位角および仰角) により異なる．仮想音像定位は，人間に聞かせたい音源信号を，音像を定位させたい方位角と仰角に対応する SLTF に畳み込み，その結果をバイノーラル受聴と呼ばれるステレオヘッドホンを利用した再生系で人間に提示することで実現できる．しかしながら，この仮想音像定位の原理に仮想音像定位の精度を劣化させる要因が含まれている．

頭外音像定位では，頭部伝達関数 (HRTF) が良く参照される．HRTF は空間内に存在するラウドスピーカー出口で收音される信号を入力，鼓膜で観測される信号を出力とした時の伝達関数である．外耳道入口で提示する信号を入力，鼓膜直前で観測される信号を出力とした時の，外耳道内における音響信号の伝播特性を表す外耳道伝達関数 (ECTF) を用い

れば，  

$$\text{HRTF} = \text{SLTF} * \text{ECTF}$$
 なる関係式が成り立つ．

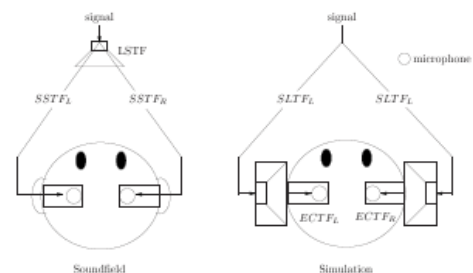


図 1: SLTF と ECTF

従来の研究結果から，HRTF として見た場合には個人差はそれほど顕著ではないが，ECTF により顕著に現れることが知られている．そこで今回は，個人差が顕著である ECTF を個人に適合するように適応的に更新していく機能を持った音響ディスプレイを開発することとした．この目的のため，以下のサブタスクに分割して研究を実施した．

### (1) HRTF の測定実験

仮想音像定位技術の理解を進めるため，無響室を借用し HRTF の測定実験を行う．

### (2) ECTF の同定実験

個人差が顕著であるとされる ECTF を測定するための測定系を構築し，測定実験を行う．

### (3) ECTF 補正による定位精度の変化検証

HRTF として MIT により公開されているダミーヘッドで測定したデータベースを用い，受

聴者個人の ECTF を利用して SLTF を算出した場合と、別人の ECTF を利用して SLTF を算出した場合の 2 とおりについて、仮想音像定位実験を行い、ECTF の個人差補正が定位精度にどのような影響を与えるかを評価・確認する。

(4) 音フィードバック手動制御系の実験系開発

変化する目標値への追従制御を行う場合、時間に連れて変化する誤差信号を仮想音像の方位角の変化として表現し再生する系を構築する必要がある。この実験系の開発を行う。

(5) ECTF の個人差を定位誤差の利用により適応的に吸収する ECTF 適応調節アルゴリズムの開発と評価

頭外音像定位伝達関数の個人差の適応的吸収に関する調整アルゴリズムを開発する。また、その効果を検証する実験を行う。

4. 研究成果

(1) HRTF の測定実験

民間企業の無響室を借用し、ダミーヘッドおよび人間の被験者 3 名について、次ページの写真に示すように HRTF の実測実験を行った。測定は、仰角方向が 0, -15, -30[度]の 3 方向、方位角は正面を基準として ±90 度の範囲を 5 度刻みで測定した。測定結果データは膨大であるため、下図に示す一例を除いて省略する。得られたデータは事後の評価検討に利用した。

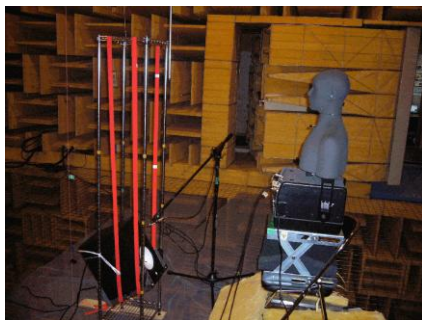


図 2: ダミーヘッドを利用した測定の様子

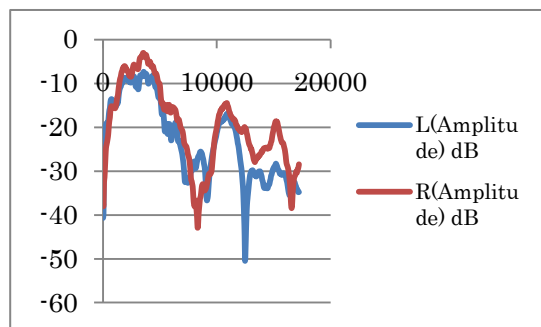


図 3: 人間に対する測定結果例 (仰角・方位角とも 0 度)

(2) ECTF の同定実験

ECTF の同定実験は、TSP (Time Stretched Pulse) 信号を同定用信号として用いた。TSP はその周波数スペクトルが全周波数において 1 となる白色性の信号である。インパルス

信号の位相を周波数の二乗に比例するだけシフトさせ時間軸上にエネルギーを分散させた chirp 信号であり、高い S/N 比でのインパルス応答測定を可能とする。これを図 4 のような測定系のヘッドホンより再生、図 5 に示すマイクロホンで観測される信号をサンプリング周波数 44.1[kHz]にて収録、その信号に逆 TSP スペクトルを畳み込むことで ECTF を測定した。測定結果の例を図 6 に示す。被験者によりゲインスペクトルのピークとノッチの位置が異なっていることが分かる。

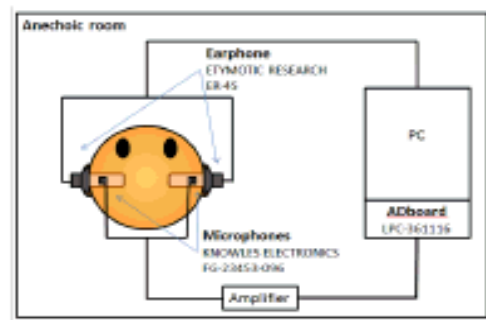


図 4: 外耳道伝達関数の測定系

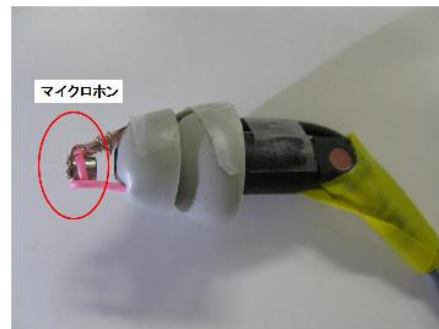


図 5: ECTF 測定に用いるマイクつきイヤホン

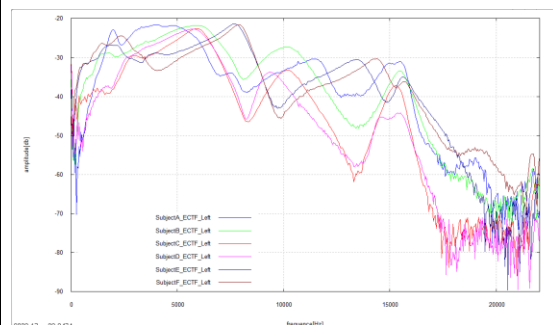


図 6: 被験者 6 名の左耳の ECTF の測定結果

(3) ECTF 補正による定位精度の変化検証

(2) で測定した ECTF を利用し、HRTF を被験者個人の ECTF で除した場合とそうでない場合 (図 6 の 6 名の被験者のうち 1 名のデータを利用し、頭外データに対応する被験者は実験から除外) の 2 とおりについて、計算した SLTF を 1 秒間の音源信号に畳み込んで作

成した音像を被験者にステレオヘッドホンを通じて提示し、定位したと感じた方向を答えてもらう試験を行った。SLTF は水平面内の正面から±90 度の範囲について 5 度刻みで 37 方向のものを用意して用いている。受聴者には仮想音像が 5 度刻みとなっていることは知らされておらず、定位した方位角は 1 度刻みで答えてもらった。音源信号は 37 方向からランダムに選択して提示され、それが N=200 セット提示された後の平均定位誤差を

$$E_N = \frac{1}{N} \sum_{i=1}^N |\theta[i] - \theta_r[i]|$$

式で計算した結果を下表に示す。

	other's ECTF [°]	individual ECTF [°]
subjectA	44.45	56.215
subjectB	29.005	26.94
subjectC	27.27	21.045
subjectD	22.08	21.905
subjectE	23.33	20.32

表 1: 仮想音像定位実験の結果

これを見ると、1 名の被験者を除き、自身の ECTF を利用して計算した SLTF を畳み込んだ場合の方が精度が向上していることが分かる。

#### (4) 音フィードバック手動制御系の実験系開発

音フィードバック手動制御系とは、以下の図で示されるようなシステムで、コントローラである人間へのフィードバック信号を音として人間に提示するものである。制御系の構成には 2 自由度制御構造も考えられるが、外性の目標信号への追従が求められる手動フィードバック制御系は数少ないと考えられる。今回は以下の 1 自由度系の構成について実験システムを構築した。

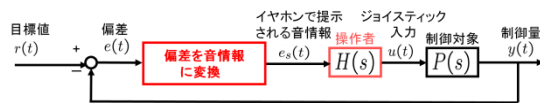


図 7: 1 自由度音フィードバック手動制御系

この系で、 $e(t)$  を仮想音像定位の方位角情報として操作者に提示するためには、連続的に変化していく  $e(t)$  の値に対応して音源信号に畳み込む SLTF の方位角を変更していく必要がある。このような状況で利用できる畳み込み演算の手法として、overlap add 法および overlap save 法が知られているが、今回は overlap save 法を利用してプログラミングを行った。これらの方法の詳細については成書を参照されたい。実験環境は windows を OS として搭載する PC 上に構築し

ている。音信号  $es(t)$  はステレオヘッドホンを通じて操作者に提示され、操作者は  $es(t)$  を通じて認知した  $e(t)$  を 0 にするべく、ジョイスティックを操作する。制御対象の伝達関数  $P(s)$  はプログラムにより変更可能である。

#### (5) ECTF の個人差を定位誤差信号の利用により適応的に吸収する適応アルゴリズムの開発と評価

ある方位角  $\theta_r(n)$  ( $n$  は時刻を表す整数) の方位に定位すると期待される音信号は、受聴者に提示したい音源信号時系列を  $r(n)$  とすると、 $\theta_r(n)$  方向に対応する左右の HRTF をそれぞれ左右の ECTF で除した伝達関数 SLTF と  $r(n)$  を畳み込んで作成することができる。これをステレオヘッドホンを通じて左右の耳に提示することで受聴者には  $\theta_r$  方向に仮想音源の存在を感じる仮想音像定位が発生すると期待されるが、様々な要因によりこの信号を聞いた受聴者が音像を認知する方位角  $\theta(n)$  は  $\theta_r(n)$  とは異なり、誤差が生じる。

そこで、この誤差を吸収するためのシステムを考案した。その考え方は図 8 のとおりである。より個人差が顕著であるとされている左右の外耳道伝達関数 ECTF をいくつかの帯域に分割する。図中の文字  $k$  は周波数帯域に付した番号を表す。その帯域のゲインを補正する係数  $\alpha_R(k, n)$  および  $\alpha_L(k, n)$  を置き、この 2 つの係数を  $e(n)$  が減少する方向に適応的に調節することで、定位精度を改善することを狙う。

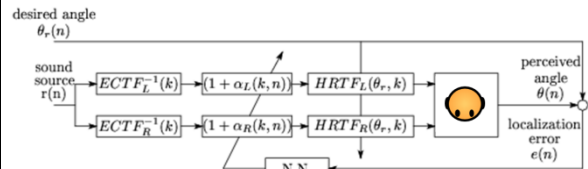


図 8: 仮想音像定位の定位誤差を学習・修正するシステム

しかし、上述の 2 つの係数の調節方向を  $e(n)$  の減少に寄与する方向に正しく設定するためには、係数を増加/減少させた際に  $e(n)$  が増加/減少のどちら向きに変化するのかわ知る必要がある。図 8 からそれは、定位誤差の評価関数を

$$E(n) = \frac{1}{2} e^2(n)$$

として、この関数の係数に対する勾配  $\partial E(n)/\partial \alpha_R$  および  $\partial E(n)/\partial \alpha_L$  を知る必要がある。しかしながら、この勾配を知ろうとすると、人間が左右の耳に提示された音響信号をどのように処理して角度を知覚しているかを定式化する必要がある。そこで今回は、

この部分をニューラルネットワークでモデリングすることとした。その仕組みを図9に示す。

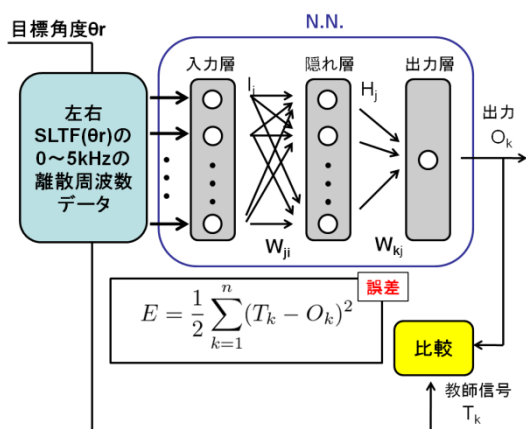


図9：ニューラルネットワークによる人間の音像方位角認知機構のモデル化

図9では、ネットワークの入力として図のとおり左右 SLTF の限られた帯域のデータを与え (512 入力), 出力が与えた SLTF の方位角となるように教師信号を設定して学習を行わせた。5 度刻みの方位角のそれぞれについて学習の結果 0.003 度以内の出力誤差に収まるようにした。この仕組みを導入することで、先に必要であった勾配の推定値を得ることができ、個人差を吸収して定位精度を向上させる仕組みを構築することができた。

以上構築したシステムを利用し、改めて (3) と同一の評価試験を実施したところ、下表のような結果を得た。

	other ECTF <sup>[9]</sup> [°]	modified ECTF <sup>[9]</sup> [°]
Subject A	44.45	<b>75.39</b>
Subject B	29.01	<b>24.53</b>
Subject C	27.27	<b>29.79</b>
Subject D	22.08	<b>22.23</b>
Subject E	23.33	<b>22.41</b>

表2：個人差を適応的に吸収する音響ディスプレイを用いた仮想音像定位実験結果

#### (6) 結果のまとめ、今後の展開

本研究の中心となる結果は、(5) で提示したシステムである。表2に示した結果ではその有効性が明白に明らかであるということとはできないが、これには改善の余地がある。図9で示したニューラルネットワークでは、教師データとして、SLTFの元になったMITのHRTFデータベースにおけるHRTFの方位角情報をそのまま正解として与えている。これは、事前に各人の音像定位特性を測定する実験を課すことが実用化への障害となることを懸念しての選択であるが、そのように作られたSLTFを被験者各人がどの方位の音像として知覚したかのデータを、ネットワーク学習の際の正解として用いれば、学習が終了した

ニューラルネットワークは個人の仮想音像の方位角知覚特性を踏まえたモデルとなっており、 $\alpha_R, \alpha_L$ の修正もより適切な方向に進むことが期待される。

このような設定の下での実験は、時間の関係で1名分しか行うことが出来なかったが、表2中の被験者Cについてこれを行ったところ、平均定位誤差が18.51度まで改善したという結果が得られており、かなり改善効果が高いことが確認できて( )いる。したがって、本研究課題終了後も引き続いてこの実験に取り組み、個人差の適応的の吸収に関する有効性の検証を行うとともに、出来上がった音響ディスプレイを図7における音情報変換機構として仮想音像定位を用いる音フィードバック手動制御系に組み込み、手動制御系の制御性能向上に向けた調整を継続する予定である。

なお、(5)で示した仮想音像知覚の個人差を吸収する音響ディスプレイシステムの構築に関しては、国内の学会における口頭発表は完了しているが、今後は前段落で述べた実験・検討を継続して行い、それがまとめ次第、国内学会の査読付雑誌への投稿と国際会議での発表申し込みを行い、今後鋭意成果公表の作業を進めていく。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計0件)

〔学会発表〕(計4件)

①河原直哉、藤井文武、井上友倫、穂垣暢久、「外耳道伝達関数の適応的調節による聴覚フィードバック手動制御系の制御性能向上」、日本音響学会2011年春季研究発表会、2011年3月11日、早稲田大学

②小田雄介、河原直哉、和田憲造、井上友倫、藤井文武、「外耳道伝達関数の逐次修正による聴覚フィードバック手動制御系の制御性能向上」、日本音響学会2010年春季研究発表会、2010年3月10日、電気通信大学

③小田雄介、河原直哉、井上友倫、藤井文武、和田憲造、「音像移動音の定位感の向上に関する研究」、第18回計測自動制御学会中国支部学術講演会論文集、2009年11月29日、鳥取大学

④松下正裕、藤井文武、和田憲造、「HRTFの適応的更新に基づく音情報を用いた手動制御系に関する研究」、第16回計測自動制御学会中国支部学術講演会論文集、2007年11月10日、山口大学

〔図書〕(計0件)

〔産業財産権〕

○出願状況(計0件)

○取得状況(計0件)

[その他]  
なし

6. 研究組織

(1) 研究代表者

藤井 文武 (FUJII FUMITAKE)  
山口大学・大学院理工学研究科・准教授  
研究者番号：30274179

(2) 研究分担者

なし

(3) 連携研究者

なし