

平成 22 年 6 月 22 日現在

研究種目：基盤研究（C）

研究期間：2007～2009

課題番号：19500178

研究課題名（和文）没入感のある空間共有のための 3 次元音響技術に関する研究

研究課題名（英文）A study for 3D audio technique to share immersive space

研究代表者

柳川 博文（YANAGAWA HIROFUMI）

千葉工業大学 情報科学部 教授

研究者番号：70296309

研究成果の概要（和文）：

前方 2 つのスピーカの入力信号間レベル差 ICLD と同時時間差 ICTD により真横音像定位を可能にする音像定位制御方式 AVS に関して、（1）側方反射音があってもその状況で ICLD と ICTD の最適値が得られ、真横音像定位できること、（2）ICTD にこれまでの音像定位理論では説明できない許容幅が存在すること、（3）複数話者同時発話においても AVS 方式であれば、それらの音声の話者ごとに個別のスピーカで再生したのと同等に聞き取りやすいことなどを明らかにした。

研究成果の概要（英文）：

AVS is a method of localizing a sound image aside even by two loudspeakers in front of a listener. In this study, it was found that (1) by searching ICLD and ICTD, the sound image localized aside even in strong lateral reflections, (2) an allowance of ICTD was found, which cannot be explained by conventional theory of sound image localization, (3) three speeches reproduced by AVS shows same easiness of hearing as reproduced those by three loudspeakers separately.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	1,300,000	390,000	1,690,000
2008年度	1,000,000	300,000	1,300,000
2009年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：総合領域

科研費の分科・細目：情報学，感性情報学，ソフトコンピューティング

キーワード：音像定位，信号間レベル差，信号間時間差，複数話者，同時発話

## 1. 研究開始当初の背景

従来、聴取者前方 2 つのステレオスピーカで真横方向まで音像定位させるバーチャル

サラウンド技術はスピーカから両耳への伝達関数に基づいて開発された。そのためスピーカからの直接音以外の再生音場の反射音

などにより効果が阻害される。加えて、信号処理上頭部伝達関数の逆関数が必要であり、そのことが音像定位を制御するフィルタの収束性悪化などの問題を引き起こしている。本研究代表者の柳川は2つのスピーカの入力信号を帯域分割し、最も真横に定位する信号間レベル差 ICLD (Inter-Channel Level Difference) および時間差 (Inter-Channel Time Difference) を聴覚的に探索することにより、音像定位制御を行うという手法 AVS (Adjustable Virtual Surround) を提案した。その成果はこれまで、オーセンティック社により AUTHENSURROUND (オーセンサラウンド) ⑧の商標で商品化され、NEC 製パーソナルコンピュータに搭載された。これまで AVS は環境の反射音に影響されにくく、誰が聞いても同様な効果があることは経験的に承知しているが、客観的かつ定量的な確認はまだ行われていない。また、単なるバーチャルサラウンド以外に、例えば複数話者間の音声コミュニケーションに応用できるのか、またその効果はどの程度か明らかにすることは、ネットワークを介した対話システムが普及しつつある今日大いに有意義である。

## 2. 研究の目的

これまで開発した、聴覚に基づく音像定位制御の信号処理の理論的解析・検証を行い、その原理を解明する。次に、実際の聴取環境を音響反射の面などから調査し、聴取環境モデルを構築し、環境にロバストな音像定位制御信号処理アルゴリズムを開発する。さらに、本研究成果を応用したコミュニケーションシステムのユーザーとなる聴取者が、ユーザー自身の聴取環境を変えることなく、聴覚実験によって3次元の音像定位制御のための信号処理パラメータを効率的に探索するための手法を開発する。

## 3. 研究の方法

### (1) 側方反射音の影響

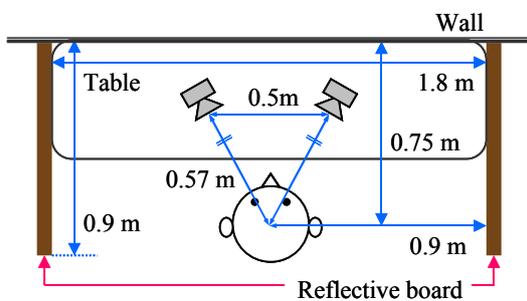


図1. 反射壁, スピーカ, 被験者の配置

いくつかの実験により、環境の間接音の中で、側方反射音がもっとも影響することが明

らかとなったので、実環境で考えられる側方反射音でも影響が大きい対向する側壁からの反射音がある状況を、最も本方式に不利な環境のモデルとして選び、本手法を試みる。反射壁は91 cm 183 cmのラワン合板2枚を用いた。スピーカ、被験者の位置関係を図1. に示す。

また、同時に本手法の狭帯域 (1/4 オクターブ) バンドパルス OBP と同バンドノイズ OBN を比較し、バンドパルスの有効性を検証する。これは信号の性質と音像定位の関係を調べるためでもある。

### (2) 個人差の影響

ICLD, ICTD の探索結果の個人差の程度を被験者 6 人による探索実験により明らかにする。またこのとき真横音像定位する ICTD の許容幅を求める。なお、実験は反射音の無い無響室で行った。被験者とスピーカの配置を図2. に示す。

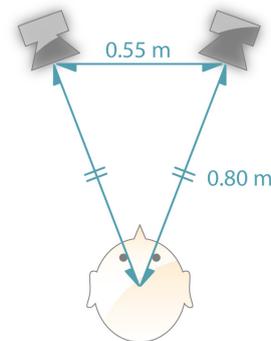


図2. 被験者とスピーカの位置関係

### (3) 複数話者同時発話の聞き取りやすさ

3人の話者 A, B, C の約10秒間の音声を同時に、図3. に示す (i) ステレオ再生, (ii) 3つのスピーカによる独立3ch再生, (iii) AVSによる仮想3chの3通りの方法で再生し、分離感に着目した聞き取りやすさを比較した。

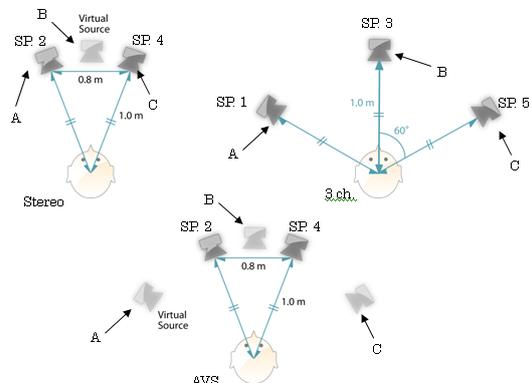


図3. 話者3人の音声の再生方法

## 4. 研究成果

(1) 側方反射音の影響

図4. に狭帯域バンドパルスの結果を、図5. に狭帯域バンドノイズの結果を示す. 図の楕円状の領域は、被験者が定位した音像の位置、大きさを表し、中心周波数 250 Hz から 8 kHz の 21 個の結果をまとめて表示している.

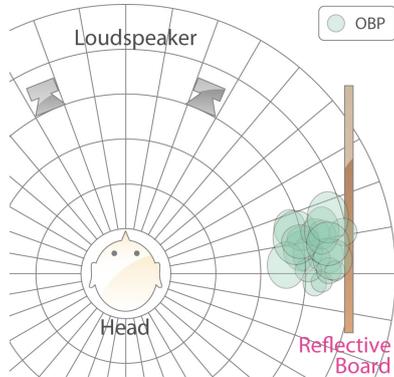


図4. 反射音の影響 (OBP)

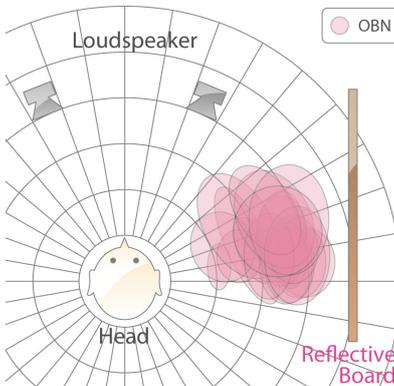


図5. 反射音の影響 (OBN)

いずれの信号でも 250 Hz を除き、横方向に定位した. しかし、OBP の方が音像までの距離は遠く、音像の大きさも小さい. さらに被験者の内観報告によると ICLD, ICTD の探索も OBP の方が容易であるとのことであった. 以上のことから、対向壁による側方反射音があっても 1 つの帯域を除いて横方向に音像定位し、さらに探索に OBP を用いる方が音像の質も良いことが示された. 一般に、音声や映画の効果音などは概ね過渡的であり、OBP の結果に基づいた音像定位制御が有効である. また、250 Hz で横方向定位しなかったが、音声など通常の音源はスペクトルの帯域幅がある程度あるので、殆ど効果に影響しない.

OBP による音像が OBN より鮮明であることは、音像定位に関わる聴覚が、スペクトルが同一でも、定常的で振幅変化のない音よりも過渡的で振幅変化のある音に、より性能を発揮していると言え、信号のオンセットが定位に寄与していることを示唆している.

(2) 個人差の影響

図6. に ICLD, 図7. に ICTD の平均値と標準偏差を示す. なお、ICTD はサンプル数で表しており、サンプリング周波数は 48 kHz である.

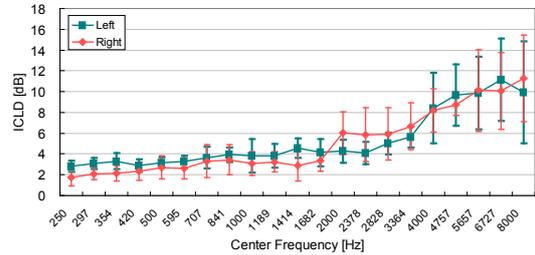


図6. ICLD の平均値と標準偏差

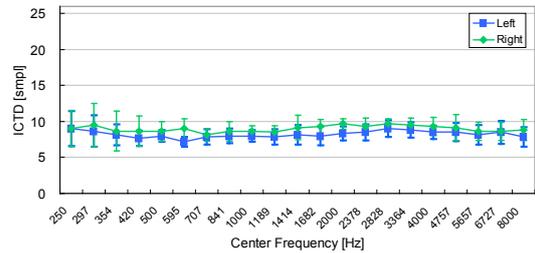


図7. ICTD の平均値と標準偏差

図6. より ICLD の平均値は周波数とともにゆるやかに上昇している. これは単一音源に対する両耳間レベル差が、周波数が高くなるほど増加する傾向と一致している. 標準偏差も特に 2 kHz 以上で増加している. このことは 2 kHz 以上で両耳間レベル差が音像定位に優位であることと合わせ、個人差が無視できないことを示しているが、平均値の ICLD で設計した AVS の音像定位が許容できないほどであるかは検証する必要がある. 図7. より ICTD のばらつきは殆ど無く、平均値が万人に通用することが分かる. なお、両図には色分けして、左右差が分かるようにしているが、この程度の差であればさほど問題ないと思われる.

次に、6 人の被験者の ICTD の許容幅を累積した頻度分布を図8. に示す. これを見ると 1 kHz 以下で低域に向かって顕著に増加している. ただし、ICLD はその都度探索しているので、ICTD そのものがこの範囲のどこでも良いわけではない. しかし、ここに見られる傾向は従来の頭部伝達関数による音像定位理論、すなわち両耳間レベル差や両耳間時間差で一意に定まる音像定位では説明困難である. このことは本方式の音像定位メカニズムが先行音効果や time intensity trade のような非線形な聴覚の作用に関係していることを伺わせる. 今後さらに検証する価値のある興味深い結果である.

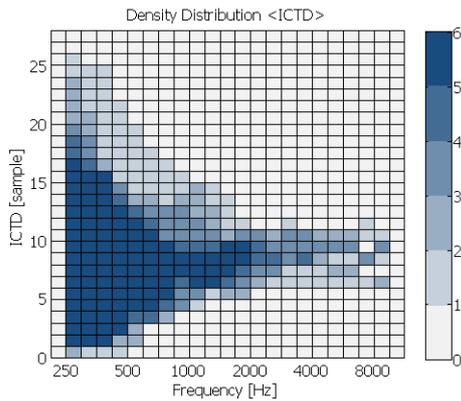


図8. ICTD 許容幅の頻度分布

(4) 複数話者同時発話音声の分離感

3つの再生方式を対比較し、被験者11名の5段階評価による評点を図9. に示す. なお、3人の音声すべてが男性の場合とすべて女性の場合の2つの結果を描いている.

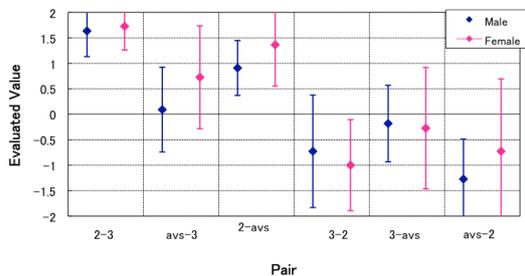


図9. 聞き取りやすさの評点

図の横軸、例えば2-3はステレオと独立3ch再生の比較を示し、この例では3ch再生の方が約1.7だけ聞き取りやすいことを示す. これから、(ii)3つのスピーカによる独立3ch再生がもっとも分離感があり、(iii) AVSによる仮想3chがこれに続き、(i)ステレオ再生がもっとも悪い結果となった. このことは音像定位分布の分離度と対応し、定位が離れているほど分離して聞き取れることが明らかになった. さらに、シェッフエの方法で分散分析したところ、(i)と(iii)は1%および5%危険率でも有意差があり、(iii)と(ii)は男声では1%および5%危険率でも有意差が無く、女声では1%危険率では有意差が無く、5%危険率では有意差ありとなった. よって本方式 AVS は3スピーカ独立再生とほぼ同等の分離感あるいは聞き取りやすさをもたらすことが分かった.

以上の結果を通し、本研究が当初の目的に沿って遂行され、聴取環境のモデル化と本方式の環境ロバスト性、探索を容易にする信号の特性、聴覚に基づく音像定位メカニズム解明の糸口となる ICTD 特性、さらに音声コミュニケーションへの応用可能性とその利点など

を明らかにすることができた. また、2007年と2009年の2度の国際会議 AES Tokyo Convention においてエキジビション会場のブースを借り、デモを行ったところ海外からの参加者に好評であった.

5. 主な発表論文等

[学会発表] (計6件)

① 榎田美奈子, 柳川博文 (千葉工業大学), 末廣一美, 高山泰典, 福島学 (日本文理大学), 「低周波数帯域における両耳間相関係数の弁別」, 日本音響学会講演論文集, 3-P-11, pp. 781- 782, 2008年9月12日 (福岡)

② 久保田薫, 柳川博文 (千葉工業大学), 末廣一美, 高山泰典, 福島学 (日本文理大学), 「マルチチャンネル音楽コンテンツのチャンネル間相関と空間的印象」, 日本音響学会講演論文集, 1-P-6, pp. 727- 728, 2008年9月10日 (福岡)

③ 西村一行, 岩上知広, 柳川博文 (千葉工業大学), 波々伯部龍人, 福島学 (日本文理大学), 「信号間レベル差及び時間差を用いた音像定位制御における帯域幅の検討」, 日本音響学会講演論文集, 1-8-6, pp. 595- 596, 2008年9月10日 (福岡)

④ 高岡創, 棚原憲照 (日本文理大学), 西村一行, 白石健, 柳川博文 (千葉工業大学), 福島学, 岡本壽夫 (日本文理大学), 「反射のある室における狭帯域 ICTD および狭帯域 ICLD による音像定位領域拡大手法の実験的検討」, 日本音響学会講演論文集, 3-P-23, pp. 789 - 790, 2007年9月21日 (甲府)

⑤ 西村一行, 白石健, 柳川博文 (千葉工業大学), 高岡創, 棚原憲照, 福島学 (日本文理大学), 「音像定位領域拡大手法における試験音の検討 ～狭帯域パルスと狭帯域ノイズの比較検討～」, 日本音響学会講演論文集, 3-P-24, pp. 791- 792, 2007年9月21日 (甲府)

⑥ Takeshi Shiraishi, Kazuyuki Nishimura, Hirofumi Yanagawa, and Manabu Fukushima, Expanding sound image localization by using narrowband inter-channel time and level differences Audio Engineering Society Convention Paper, 5 pages, 2007, 21 July (Tokyo)

[産業財産権]

○出願状況 (計2件)

名称: 「音像定位フィルタ及びこれを用いた音響信号処理装置並びに音響信号処理方法」

発明者: 柳川博文

権利者: 千葉工業大学

種類：特許  
番号：特願 2008-145500  
出願年月日：平成 20 年 6 月 3 日  
国内外の別：国内

名称：「反射音生成装置」  
発明者：柳川博文  
権利者：千葉工業大学  
種類：特許  
番号：特願 2008-145501  
出願年月日：平成 20 年 6 月 3 日  
国内外の別：国内

## 6. 研究組織

### (1) 研究代表者

柳川 博文 (YANAGAWA HIROFUMI)  
千葉工業大学・情報科学部・教授  
研究者番号：70296309