



研究ノート

組込み向け空間音響処理とその応用

尾上 孝雄*

Fundamentals and Applications of 3D Acoustics for Embedded Systems

Key Words : Embedded System, 3D Sound, Portable System

1. はじめに

『空間音響処理』というあまり馴染みがない言葉を書くと、何か特別な効果を狙った処理を彷彿とさせますが、よく考えると我々が日常生活で聞いている「音」は立体的な空間を通して聞こえており、その音は我々を取り囲んでいる世界から実に多くの影響を受けています。

身近な音の聴取環境として、オーディオ機器がありますが、このオーディオ機器は、モノラルからステレオへ、ステレオからサラウンド(立体音響)へ、技術の進歩とともに変遷し、それにとまってソフトコンテンツも、求められるエンタテインメント性などの付加価値を高めながら、着実に変わってきています。この潮流はめまぐるしい速度で進み、将来においては、立体音響が今のステレオの地位を脅かす位置を占める可能性が十分にあります。

では、立体的な音のごく普通に利用される時が来たとすれば、いったいどのような音響空間が求められるものなのでしょうか。特殊効果としてではなく、普段の生活の中で接しているような自然な空間生成が基本要件として求められると考えられます。

このような観点から、「音」に求められるエンタテインメント性とは何なのか、立体音響の実現手段はツールとしてコンテンツの魅力を最大限に引き出

すことができるだろうか、ユーザ・フレンドリになるだろうか、など克服すべきテーマを見つけ出し一つ一つ解決していくべきではないでしょうか？我々は、産学連携の枠組みのもと、HRTF(頭部伝達関数)を用いたモノラル音の音像定位付加アルゴリズムの研究開発を進めてきました^[1]。本稿では、このアルゴリズムを応用して、さまざまな立体音響の世界を作ろうと進めている研究開発についてご紹介します。

2. 音像定位アルゴリズムの概要

「見られるものの領域は表層である。聞かれるものの領域は深層である。目は表面をなでる。だが耳は、侵入してくるものでなければ知覚できない。——聞く人間はだから見る人間よりも、深層にはいりこめる機会が多いのである。新しい人間は聴く人間となる、ということだ。」^[2]

このような概念のもと、従来の「音を立体化する」というテーマを、空気中の音そのものを再現する「空間としての音」と捉えた音響処理について考えてみます。人間が音の空間的位置を感じ取る、いわゆる音像定位を人間の聴覚を通して知覚するために必要となる基本的な演算だけに注目すれば、その処



* Takao ONOYE
1968年5月生
1993年大阪大学・大学院工学研究科・
電子工学専攻・博士前期課程修了
現在、大阪大学・大学院情報科学研究科、教授、博士(工学)、メディア
処理、組込みシステム、SoCアーキ
テクチャ
TEL 06-6879-4525
FAX 06-6879-4529
E-Mail onoye@ist.osaka-u.ac.jp

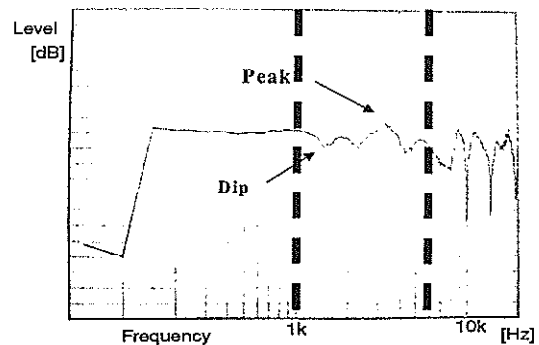


図1 頭部伝達関数の周波数特性

理の複雑さを低減し、演算量を減らすことができると思われます。図1は、頭部伝達関数の周波数特性を示していますが、この周波数特性は、点線で示すように大きく3つの周波数帯域に分割することができます。

低周波数帯域では、平坦な特性を示し、中間周波数帯域では、ピークやディップの複雑な特性を示し、さらに、高周波数帯域では、楕状の形状を示しています。従って、我々のアルゴリズムは、音源からのオーディオ信号をこれらの周波数帯域に分割し、各帯域のオーディオ信号に最適な音像定位処理を施します。結果として、各周波数帯域に必要な最低限の演算だけを実行することにより、音像定位の実行演算量を大幅に減らすことができます。

〔低周波数帯域〕

まず、人間の頭部による回折に注目します。人間の頭部は、個人差はあるが、概ね直径が150~200mm程度の球体とみなすことができます。この直径を半波長とする周波数以下の音に対しては、半波長がこの球体の直径よりも大きくなるため、人間の頭部による音の反射及び回折による影響が少なく、無視できるものと考えられます。この場合、音源から両耳に入る時間差及び音量差のみをパラメータとして入力オーディオ信号を処理することにより、音像定位が実現できます。この境界周波数(f_1)は音速(v)を340m/s、直径(d_1)を150~200mmとすると、 $f_1 = v / (d_1/2)$ より850Hz~1.1kHz程度となります。

〔高周波数帯域〕

次に、人間の耳介による音の回折に注目します。人間の耳について、その耳介を円錐形とみなせば、その底面の直径は概ね35mm~55mm程度であるので、この底面の直径を半波長とする周波数以上の音に対しては、耳介を物理的要因とする影響が大きいと考えられます。即ち、この周波数帯域においては、コムフィルタで近似できます。したがって、この処理及び時間差と音量差をパラメータとして制御することにより、入力信号を処理できます。この耳介による影響を考慮する境界周波数(f_2)は、直径(d_2)が35mm~55mmであることから、3kHz~5kHz程度となります。

この境界周波数(f_2)とコムフィルタの特性は、音源の方向により変化します。これは、外耳道入口に到達する音源からの音が、耳介により音源方向別に

異なる影響を受けるためです。直接外耳道に到達する音は、耳介による反射や回折した音が遅延して外耳道入口に到達し、それを重畳したものであり、前述の通りこれはコムフィルタで近似することができ、この込むフィルタの特性と境界周波数(f_2)を変化させることにより方向感を創出できます。

また、上下方向の音像定位に関しては、音源と同じ側にある耳の伝達関数は、肩による反射音を考慮に入れなければならない、これらについてもコムフィルタ構成を変更することにより実現できます。

〔中間周波数帯域〕

上記2帯域の間の中間帯域では、頭部や耳介による回折の影響を考慮する必要があります。このような頭部伝達関数の複雑な周波数特性はパラメトリックイコライザ(PEQ)を用い、入力信号に対する音像定位処理を実行します。この中間周波数帯域においては、処理内容が最も複雑となります。しかしながら、2帯域を取り除いた後の周波数特性は、ピーク及びディップ数が3程度となり、簡潔なフィルタ構成で実現できます。

これらの考察の結果、低演算量で実現する我々の音像定位アルゴリズムは、図2のような構成で実現できます。

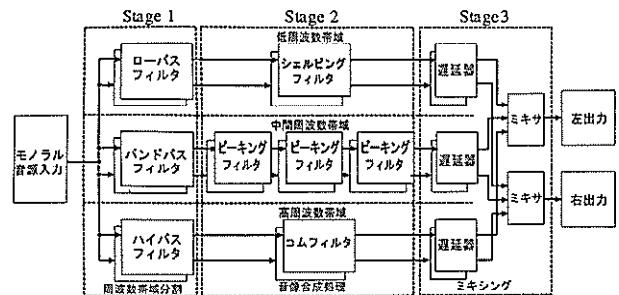


図2 音像定位アルゴリズムの構成

3. ステレオ音源頭外音場拡大への応用

上記アルゴリズムの一応用として、携帯型プレーヤでの聴取を前提としたアプローチをご紹介します。携帯型プレーヤの場合、システムとしてマルチチャンネル化(5.1chなどの出力を持たせること)は困難です。また、音源となる音楽などのソフトも一般のステレオで作られたものになります。さらに小型化、省電力化なども要求されます。このような条件下で、ステレオ音源の音場を頭外へ拡大する手法を実装し

ました。

一般に、ステレオ音をヘッドホンで再生する際、左右のヘッドホンから出る音はそれぞれ逆側の耳には届かないため、左右の音の時間差や音圧差など、音像の定位位置を判断する手掛かりが失われてしまいます。そのため、絶えず音像が頭中に定位する違和感から、疲労感を伴うこととなります(図3,4参照)。



図3 スピーカでの音像定位イメージ



図4 ヘッドホンでの音像定位イメージ

このような問題を解決するため、リスニングルーム壁面での反射や頭部による回折を付加することで、スピーカからの出力音を仮想的に再現し、図5のように音像の定位位置を頭外へ拡大することによって、音場の違和感をなくし、疲労を抑制するものです。

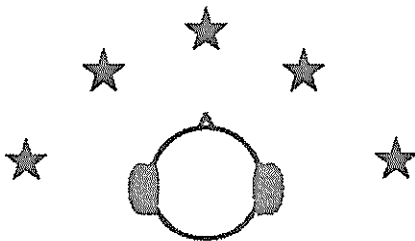


図5 音場頭外拡大定位イメージ

具体的には、図6に示すようにスピーカから聴取者の耳に直接到達する音(直接音)、壁面等に反射し耳に到達する音(反射音)、ならびに、頭部により回折し、他方の耳に到達するクロストーク音を考慮し、

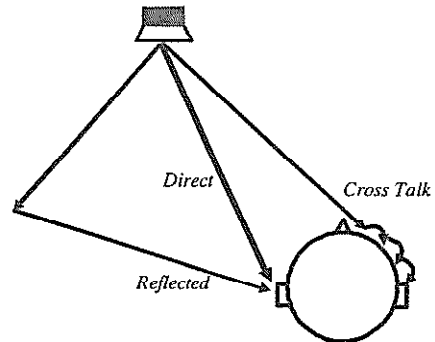


図6 音場拡大のための間接音の付加

音場拡大効果を実現する処理を行います。クロストーク音は周波数によって位相が変化するため、先のアルゴリズムに従い、3つの周波数帯域に分割し、それぞれ異なる位相の変化率を持たせます。

さらに、選択的にIIRフィルタを位相器として採用することにより、より正確なクロストーク音の再生を行うこととします。

上記の構成でDSP実装を行いました。実装には、一般に携帯型アプリケーションに用いられているTexas Instruments社製TMS320C5409を利用しました。入出力のサンプリング周波数は44.1kHzで、このサンプリング周期内で音場拡大処理と効果の入/切などオーバーヘッド処理を完了させることとしました。

使用メモリ量としては、ROMにプログラムと4種類の頭外拡大設定パラメータを格納して1.9kword、RAMは計算中間データや伝達時間差制御用のバッファなど含めて3.8kwordに収めています。

演算量は、分割後の処理の差により違いはあるものの35MIPS程度で処理が可能です。これは、位相器などの処理に最適化を施すことで、さらに削ることが可能と考えています。

消費電力は、DSPのデータシートによると、これらの処理を行った際のDSPコアの平均電流値は0.9mA/MIPSであるため、電源電圧1.8Vに対して56mW程度と見積もることができます。

4. ま と め

個別化された聴取環境が今後ますます普及していくのではないかと予測の元に、ヘッドホンでの立体音響聴取と、現存するインフラおよび今後予想されるユビキタス・コンピューティングのためのイ

ソフラ環境へ立体音響を適用することを目標に研究開発を行っております。このように研究・開発・実験・実証成果の出た分野もある一方、立体音響の研究対象は際限がないこともまた事実であり、一つ一つ着実にこなして行く必要があります。

参考文献

- [1] W.Kobayashi, N.Sakamoto, T.Onoye, and I.Shirakawa: "3Dacoustic image localization algorithm by embedded DSP", *IEICE Trans. Fundamentals*, vol.E84-A, no.6, pp.1423-1430(June 2001).
- [2] J.E. ベーレント『世界は音』
- [3] N.Iwanaga, W.Kobayashi, K.Furuya, T.Onoye, and I.Shirakawa: "Embedded implementation of acoustic field enhancement for stereo sound sources", *IEEE Trans. Consumer Electronics*, vol.49, no.3, pp.737-741 (Aug. 2003).

この記事をお読みになり、著者の研究室の訪問見学をご希望の方は、当協会事務局へご連絡ください。事務局で著者と日程を調整して、お知らせいたします。

申し込み期限：本誌発行から2か月後の月末日

申し込み先：生産技術振興協会 tel 06-6395-4895 E-mail seisan@maple.ocn.ne.jp

必要事項：お名前、ご所属、希望日時(選択の幅をもたせてください)、複数人の場合はそれぞれのお名前、ご所属、代表者の連絡先

著者の都合でご希望に沿えない場合もありますので、予めご了承ください。