



## 2009 年度上期未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

後藤 真孝 PM(産業技術総合研究所 情報技術研究部門  
メディアインタラクション研究グループ長)

### 2. 採択者氏名

チーフクリエイター: 鎌土 記良(奈良先端科学技術大学院大学  
情報科学研究科 音情報処理学講座)  
コクリエイター : 鈴木 翔太(奈良先端科学技術大学院大学  
情報科学研究科 音情報処理学講座)

### 3. プロジェクト管理組織

リトルスタジオインク株式会社

### 4. 委託金支払額

2,971,902 円

### 5. テーマ名

オーディオオブジェクト操作機能を有する圧縮符号化ソフトの開発

### 6. 関連Webサイト

<http://spalab.naist.jp/aocdemo.html>

## 7. テーマ概要

近年、計算機性能向上によって信号のマルチチャンネル化や音質改善、多様なエフェクトなど、多くの音響技術が発展してきている。信号のマルチチャンネル化によって、ホームシアター用 5.1ch サラウンドシステムなどのマルチチャンネルシステムが普及し始めており、7.1ch や 9.1ch とチャンネル数は増加し続けている。DVD や Blu-ray Disc、地上波デジタル放送などがマルチチャンネル信号に対応しているため、今後もマルチチャンネルシステムの普及が進むと考えられる。

また、オーディオのエフェクト技術は、単にトーンを変えるだけでなく、臨場感さえもコントロールすることが可能となっている。単純なトーンコントロールだけでなく、部屋の広さに代表される空間的印象さえも制御できるようになり、これらの技術は制作者のみならず、ユーザー側においても可能になった。そして、近年、オーディオ信号に含まれる各オーディオオブジェクト（例えば、ギター、ピアノ、ボーカルなど）に対してエフェクト処理ができるような、ユーザーがオーディオをより柔軟に制御できるシステムの研究・開発が盛んに行われている。

このような背景の元、Moving Picture Experts Group(MPEG)では、MPEG-Surround を拡張した技術として、ユーザーがオブジェクトの定位感を自由に操作できる圧縮技術 Spatial Audio Object Coding(SAOC)の標準化(MPEG-SAOC)を進めている。しかし、現状の SAOC では、原理上、音質劣化や、現在市場において大半のシェアを占めるミックスダウンが施された CD、DVD、Blu-ray Disk などの混合音源や実環境にて収録された音に対して用いることができない問題点がある。

本提案では、高音質、高圧縮効率で既存混合音源においてもユーザーが各オーディオオブジェクトの操作を行うことができる新たな技術の開発と、この技術を用いユーザーが手元でヴィジュアルに既存音源における空間音情報を操作、好みの音を演出可能なこれまでに類を見ない画期的なオーディオソフトウェアの開発を行うことを目的とする。

## 8. 採択理由

ユーザが自分好みの臨場感を得られるように、オーディオオブジェクトの操作が可能な音響技術を実現する提案である。

要は音源の定位を自由に操作できるようにするインタフェースを提供するということであるが、技術的なポイントは、ミックスダウン前の音源情報を使わずに、音楽 CD のようなステレオ収録された楽曲に対しても適用できることである。

精度と操作性をどこまで上げられるかが大きな挑戦ではあるが、それ以上に、実際に提案しているようなインタフェースを実現するには、各方向の音がどの楽器かを特

定する必要があり、音源分離にも取り組まなければならない。

鎌土君、鈴木君は、面接の質疑でかなり厳しい質問を私が連発したにも関わらず、まったく動じずに、自分達にはやれる、という姿勢で議論ができたのは、素晴らしかった。

本当に使いたいと思えるものに仕上がるかどうかは鍵であり、それには高い信号処理性能と魅力的なインタフェースが不可欠である。どこまで達成してくれるのか、その技術力に裏付けされた意気込みと元気に大いに期待したい。

## 9. 開発目標

マルチチャネル信号において高音質かつ高圧縮効率を実現し、かつユーザが自由にオーディオオブジェクトを操作することができるソフトウェアを構築した。ソフトウェアは次の要素から成り立つ。

- データの圧縮を行うエンコーダ
- データの伸張をリアルタイムで行うデコーダ
- オーディオオブジェクトの操作を行うことのできるコントローラ

## 10. 進捗概要

未踏プロジェクト開始段階では、基本的な信号処理の一部のみができていた状態であった。プロジェクト開始後、前半では、エンコーダ、デコーダに関して、MATLAB でのプロトタイプの実装を C++上で動作するように移植作業を着実に進めて高速化を図り、タッチ式のオブジェクトコントローラの実装に取り掛かり始めていた。

11 月にプロジェクトレビューをした際には、全体をいかにバランスよく進め、トータルのシステムとして魅力的に仕上げるかが重要になることを確認し、主にコントローラをどう仕上げるか、最後にどう見せるかを中心に議論した。その後も開発は順調に進み、操作しているときの適切かつ魅力的な視覚的、聴覚的なフィードバックも検討して、効果的なデモンストレーションシステムを実現した。

## 11. 成果

マルチチャネル信号に含まれるオーディオオブジェクトの操作(定位感操作)を可能とするソフトウェアを開発し、リオーラ (REAULA : REformation of AUditory LAteralization) システムと名付けた。

リオーラシステムは、データの圧縮を行うエンコーダ(リオーラ・エンコーダ)、これらの伸張をリアルタイムで行うデコーダ(リオーラ・デコーダ)、オーディオオブジェクトの

操作を行うことのできるコントローラ(リオーラ・コントローラ)の3つのアプリケーションにより構成される。

コーデックの開発では、音質・圧縮効率共に従来手法より優れるコサイン距離規範 k-means クラスタリングによる圧縮符号化に着目し、この手法において SAOC と同様のオーディオオブジェクト操作が施せないか検討を行った。

コサイン距離規範 k-means クラスタリングによる圧縮符号化では、エンコーダからダウンミックス信号と空間情報が出力される。空間情報はマルチチャンネル信号に含まれるオーディオオブジェクトに対応した空間代表ベクトルと呼ばれる複数本のベクトルから成っており、ベクトルの大きさや角度がオーディオオブジェクトの音量や配置に関連していると考えられる。そのため、空間代表ベクトルの操作を行うことでコサイン距離規範 k-means クラスタリングによる圧縮符号化においても SAOC のようなオーディオオブジェクトの操作が可能になると考えた。

そこで、空間代表ベクトルの大きさはオーディオオブジェクトの距離感、角度はオーディオオブジェクトの方位感に対応していると考え、これら进行操作することにより SAOC のようなオーディオオブジェクトごとの音響的操作が可能と考えた。

図1に提案内容の概要を示す。本システムでは、デコーダ内部にて空間情報进行操作する処理ブロックを追加することにより、コサイン距離規範 k-means クラスタリングによる圧縮符号化の枠組みの中でオーディオオブジェクト操作が可能となっている。

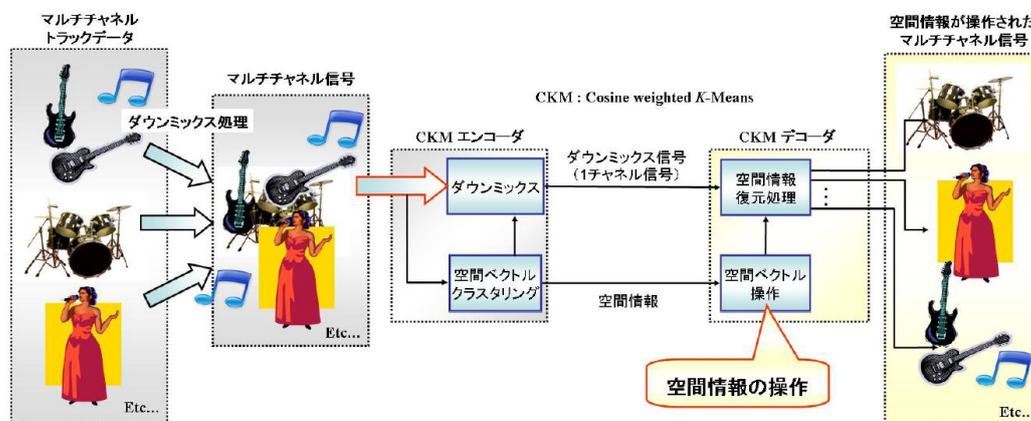


図1 提案システムにおけるマルチチャンネル信号圧縮伸張過程

エンコーダ・デコーダ機能の実装には速度が求められるため、言語には C++ を、マルチメディア操作や科学技術計算には IT++、SDL や PortAudio などの高速・高品質なオープンソースソフトウェアを用い効率的に開発を行った。

また、デコーダにおいてはリアルタイム性が求められるため、計算処理の最適化を行うために Intel C++ コンパイラを用いた。コントローラにおいては、被験者から得た使用感の情報をコントローラの GUI に対し即座に反映できるよう C++ と Python を

用い、効率的な開発を行った。

現存のオーディオ機器に見られるユーザインターフェースではオーディオオブジェクトごとの音響的操作は困難であるため、ユーザが直感的にオーディオオブジェクトを操作可能なコントローラを開発した。

図2にリオーラのコントローラを示す。このコントローラでは、画面中心に受聴者を模したアイコンが表示され、その回りにエンコーダにより識別されたオーディオオブジェクトが配置される。これらのオーディオオブジェクトを、受聴者を中心に回転(方位感操作)・遠近(距離感操作)操作をすることにより好みの位置に音像を定位させることが可能となる。



図2 リオーラコントローラ

実際の圧縮効率、音質、オーディオオブジェクトの操作性の評価を行った結果、圧縮効率、音質は共に従来手法より良いという結果を得た。また、オーディオオブジェクトの操作性に関しては、学習効率が高く、初見で全ての操作方法がわかるという評価を得た。

本プロジェクトの開発の成功により、開発したソフトウェアは競合技術である MPEG-SAOC と比較し圧縮率・音質に関して従来技術より優れているため、従来コンテンツより高音質で臨場感溢れる音の伝送が狭帯域で可能となり、ネットワークを介したよりリッチな音響コンテンツを充実させることが可能となった。

また、従来は不可能であった音源に依存しないオーディオオブジェクトの音響操作が可能となった。これにより、これまで想像もしなかったような音響的体験が可能となった。

## 12. プロジェクト評価

鎌土君、鈴木君は、期間全体を通じて主体的に開発を進めており、完成度の高いシステムを実現する開発力、プロジェクトを推進する実行力、才能を高く評価する。特に、成果のプログラム配布を意識した C++実装にも取り組み、音源種類の自動同定まで

は着手しなかったものの、ユーザがコントローラを操作する際にいかに視認性・操作性の高いコントローラを提供するかを深く検討して実装した点が、優れている。まだシステム自体の配布の目処は立っていないが、成果報告会において、タッチ式のオブジェクトコントローラを体験できる形で、効果的なデモンストレーションを実演したことは、高く評価できる。

### 13. 今後の課題

どこまで本気に社会にこの技術を根付かせようとしているかによって、今後の展開は変わりうるが、まずは、誰もが体験する形で配布する等、Web ページやデモビデオの整備も含め、その良さを一人でも多くの人にわかってもらう工夫がなされることを期待したい。また、産業界と連携した展開も可能になれば素晴らしい。