

## 楽音半開示システムの研究

西原 祐一\* 岩上 直樹+ 藤井 寛\* 串間 和彦\*

\* NTT情報通信研究所

+ NTTヒューマンインタフェース研究所

楽音に対する高品質高圧縮技術の出現と、コンピュータのCPU処理能力の向上に伴い、CD並みに品質の高いデジタル化された楽音情報を、インターネット等を介してリアルタイムに再生することが可能になってきており、今後ネットワーク上で、楽音情報が大量に流通する可能性が増えている。しかし、デジタル情報は、質的劣化のないコピーが可能であるから、著作権者の権利を保護する仕組みの開発は急務である。一方、購入者の立場からは、支払い前に商品の中身を確認する機会も必要である。本研究では、デジタル化された楽音ファイルに適度なスクランブルをかけることで、内容確認が可能で、かつオリジナルの違法コピーを防ぐ楽音半開示方式を考案し、実際の楽音に適用した。

## A Research on Semi-Disclosure Method of Audio-Music Contents

Yuichi Nishihara\*, Naoki Iwakami+, Hiroshi Fujii\*,  
and Kazuhiko Kushima\*

\*NTT Information and Communication Systems Laboratories

+NTT Human Interface Laboratories

Recently, it has become possible to play a music file at the same time as downloading it from the Internet with a quality as high as that of a Compact Disk. Therefore, it is likely that a huge number of music files would be traded on the Internet from now on. However, digital data can be copied exactly, so it is strongly desired to develop a system to protect the rights of the author. On the other hand, the purchaser of the music file would like to have a chance to check the contents before paying for it. In this research, a method is introduced, and actually applied, to both prevent illegal copying of the original music file and make it possible to check the contents by scrambling the digitized music file to a controlled extent.

### 1. はじめに

マルチメディアコンテンツの中でも、音の情報は早くからデジタルの形で蓄積され、広く流通されてきた。CDの広範な普及は、デジタル化された楽音が、非常に大きな市場であることを明示している。

現在、TwinVQ方式にみられるような、圧縮率の高い楽音符号化方式の出現と、コンピュータのCPU速度の向上によって、CD並みの品質の高い楽音を、電話線のような狭い帯域で送り、リアルタイムで再生することが可能となってきた。インターネットで、オンタイムに高品質の楽音を取り出せるようになれば、インターネット上で巨大市場が誕生する可能性は大きい。

しかし、デジタル情報では、質的劣化のないコピーが可能であり、ネットワーク上での楽音女王の安全な流通を行うためには、違法コピーを防ぎ、著作権者の権利を保護する仕組みを実現する必要がある。著作権保護に関しては、デジタル情報への著作権情報の刷り込みを行う電子透かしが有望な手段として注目を浴びているが、直接的なコピー防止という観点では、デジタル情報をスクランブルし、意味のない情報に変えてしまうことの方がより有効な手段となる。デジタル情報においては、暗号化された情報を元に戻すことができなければ、情報をコピーしたことはない。

一方で、楽音情報を購入しようという者の立場からは、料金を支払う前に商品の中身を確認する必要がある。商品の確認という意味では、その商品の見本となるようなプロモーション用のサンプルを作成、配布し、料金を払った者に対して本体の方を送付するという方法が考えられる。しかし、後述する半開示システムが実現すれば、スクランブルのかけられた商品自身が、商品見本の役割を兼ねることができる。このことにより、手間

をかけてプロモーション用のサンプルを作成する手間が省け、また、商品見本と、実際の商品が違うものだったという誤りも防ぐことが可能である。

本研究において提案する楽音半開示のシステムは、オリジナルの違法コピー防止のためにスクランブルがかけられたファイルに、試聴版としての役割も持たせ、本体より圧倒的に情報量の少ない鍵情報によって、セキュリティの確保および料金の支払いを行うための仕組みである。

本稿では、第2節において、今回スクランブル化を行う楽音符号化方式として採用したTwinVQ方式についての簡単な説明および、採用した理由について述べる。また、第3節においては、スクランブル化の方法および、スクランブルの程度の制御方法について述べる。第4節においては、TwinVQの各係数にスクランブルをかけた結果、どのような効果がみられたかについて述べる。第5節においては、第4節のスクランブルに対し、主観評価を行った結果を示す。

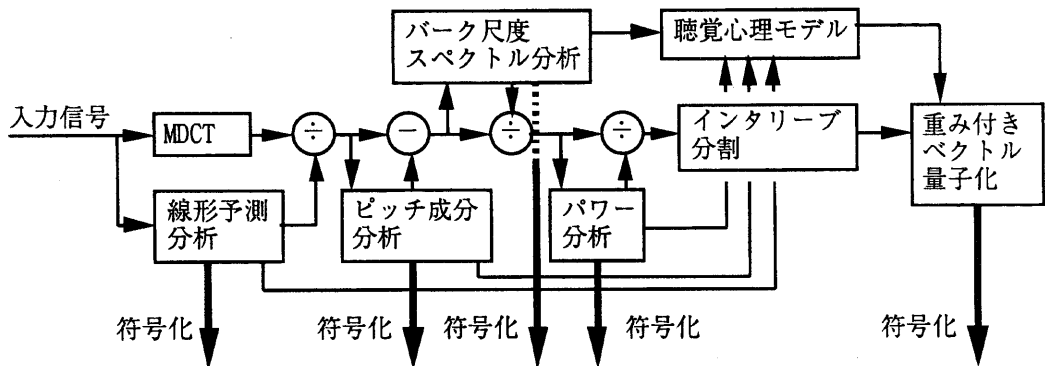


図 1

## 2. TwinVQ 方式

今回、スクランブルをかける楽音符号化方式としては、TwinVQ方式を採用した。TwinVQ方式は、例えば、CDのような16bit/44.1kHz/ステレオサンプリングの楽音を、96kbit/secの情報に圧縮できるなど、非常に圧縮効率のよい楽音符号

化方式で、MPEG4 オーディオの候補になっている。符号化の特徴としては下記が挙げられる。

- 変形離散コサイン変換 (MDCT) により周波数変換を行っている。

- 上記 MDCT 係数から、線形予測パラメータ、ピッチ成分、バーク尺度によるスペクトル包絡、パワー成分の各特徴値を抽出し、それぞれを適当

な方法により符号化する。

●上記各特徴値を取り去った後の残差の MDCT 係数を、聴覚心理モデルを考慮に入れた重み付けインターリーブベクトル量子化を用いて符号化する。

なお、TwinVQ 方式の符号化、復号化方式について、細かい説明をここで述べることはせず、図 1 を示すにとどめる。

この方式のように、楽音信号から複数の特徴値を抽出し、各特徴値と残差値を別々に符号化する方式に対しては、各特徴値および残差値に対し、異なった強さでスクランブルをかけることが可能で、例えば特徴値の一部を完全に残したスクランブルなどといったスクランブルのかけかたも考えられる。

TwinVQ が、半開示システムに適した符号化方式であること、また、将来標準的な楽音符号化方式になる可能性があるという観点から、今回、TwinVQ 形式を採用した。

### 3. TwinVQ 方式に対するスクランブル

本節においては、TwinVQ 方式に対するスクランブルの方法について解説する。

#### 3.1 フレーム単位でのスクランブル強度の制御

音の情報においては、時間軸方向へのつながりが重要な役割を持っている。従って、曲の一部に全くスクランブルのかけられていない部分があっても、その他の部分にスクランブルがかけられていれば、十分半開示としての役割を果たすことが可能である。逆にいえば、原音の音質判断のために、部分的にスクランブルの全くかけられていない部分を残すということも可能性としてありうる。

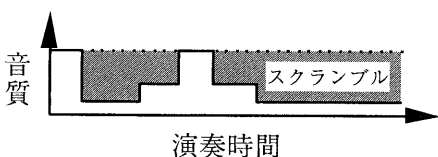


図 2

TwinVQ では、PCM 方式でデジタル化された楽音の、2048 サンプルを 1 フレームとして、フレーム単位で MDCT 変換を行っている。サンプリング周波数が 44.1kHz の場合、1 フレームは約 0.05 秒に相当する。スクランブルもこのフレーム単位で行うので、スクランブル強度は、時間にして約 0.05 秒単位で制御することが可能である。普通に人間が聴く楽音に対して、このスクランブル化の最小単位は、十分に小さい単位であると言える。

#### 3.2 フレーム内でのスクランブルの制御

##### 3.2.1 スクランブルをかける係数の選択

フレーム (2048サンプルに相当)

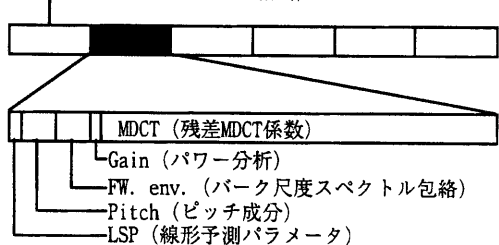


図 3

TwinVQ 形式のファイルでは、原音の 2048 サンプル分に相当するフレーム単位で逐次的に TwinVQ 係数が並んでいる。1 フレーム分の TwinVQ 係数は、さらに、LSP (線形予測パラメータを符号化したもの)、Pitch (ピッチ成分を符号化したもの) Forward envelope (バーク尺度によるスペクトル包絡を符号化したもの。Fw. env. と略す)、Gain (パワー分析部を符号化したもの) そして、MDCT (残差 MDCT 係数を符号化したもの) に分けることができる。スクランブルは、この各部分に対して独立に行われる。

LSP、Pitch、FW. Env.、Gain、MDCT の各係数はさらに、一般に図 4 によって表されるような部分係数グループおよび部分係数から構成されている。

スクランブルをかける際は、以下の様にしてどの係数に対してスクランブルをかけるかを指定

する。

●LSP、Pitch、FW. Env.、Gain、MDCT のうちどの係数にスクランブルをかけ、どの係数にスクランブルをかけないか。

●各係数において、どの部分係数グループにスクランブルをかけるか。

●各部分係数グループのなかで、どの部分係数にスクランブルをかけるか。

スクランブルをかける部分係数の数が増えるほど、スクランブルの度合いは大きくなる。

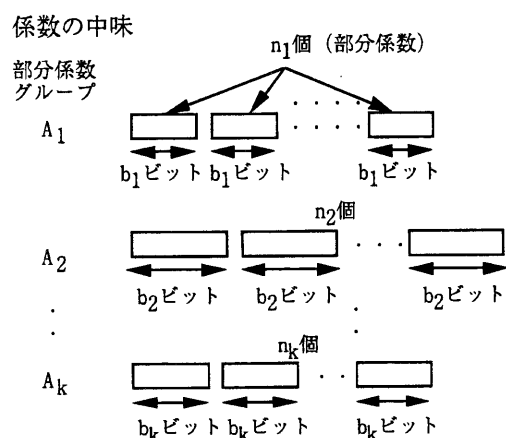


図4

### 3.2.2 係数に対するスクランブルのかけかた

実際にスクランブルをかける方法はいく通りも考えることができる。ここでは、その一例を示すことにする。

スクランブルをかける部分係数に対し、その係数の何ビット目にスクランブルをかけるかを決める。同じ部分係数であっても、どのビットにスクランブルをかけるかによって、スクランブルの質が変わってくる。

スクランブルをかけるビットが決まったら、乱数発生器により乱数のビット列を発生させ、そのビットに対し、ビット毎の排他的論理和をとる。

復号化を行う際は、スクランブル化されたビットに対して、暗号化の際と同じ乱数ビット列を発生

させ、排他的論理和をとれば、元の係数を復元することが可能である。なお、この暗号化の方式については、デジタル映像情報に対するスクランブル方式<sup>2)</sup>を参考にしている。

スクランブル化の方法としては、この他にも、スクランブルをかける部分係数に対して、決められた定数を足したり、引いたりするという方法が考えられる。この場合、例えば、乱数ビットを発生させ、そのビットによって足すか引くかを制御する等の方法により暗号化が可能である。

上記に挙げた方法は、いずれも TwinVQ 形式のビット列自身に直接作用させることが可能であり、かつ、高速での処理が可能である。このスクランブルの復号処理器を、TwinVQ の再生器の中に組み込むことで、オリジナルの楽音情報をファイルとして残すことのない、オンザフライ再生を実現でき、オリジナルの違法コピーを防ぐことが可能となる。

## 4. 各係数とスクランブルの質について

本節においては、TwinVQ の各係数が、楽音信号のどのような特徴値を抽出したものであるかについて説明し、その係数にスクランブルをかけた結果音質にどのような変化をもたらすかについて述べる。

TwinVQ の係数のうち、LSP、Fw. env.、MDCT は、全周波数領域に渡った波形の形状を記述するものである。LSP が最も大まかな波形の概要を表しており、従って、少ないビット数にスクランブルをかけることで、音質を大きく劣化させることができる。Fw. env. は、人間の聴覚の特性に従って低周波ほど間隔を狭くした、非線形の周波数目盛りであるバーク尺度によって区切られたスペクトラムで波形の概要を表したもので、低い音に対して大きな影響を与えることができる。MDCT は、実際は、元の信号を MDCT 変換したのから、各特徴値を抽出した後の残差の部分で、微妙なニュアンスを変化させることができる。

一方、Pitch は、周波数スペクトルの中で等間

隔のピークを形成するピッチ成分を抽出したものである。TwinVQ の符号化においては、Pitch 成分の少ない信号に対しては、この係数を抽出しないような仕組みになっている。また、Gain は、残差 MDCT のパワーを計算したもので、この係数のスクランブル化は、音量の変動をもたらす。

上記の特徴、および、実際にサンプルの信号音に対してスクランブルをかけた結果分かった、各係数に対するスクランブルの効果を以下に列記する。

●LSP … 大きなスケールで音がぼやける。元の音が、ポワッ、ポワッとした感じで歪む。

●Pitch … 元のピッチに対して、高調波の成分が加わるので、金属音を加えたような効果がある。

●Fw. env. … 音が部分的にぼやける。特に、重低音に特徴のある曲においては、重低音がほとんど消えてしまう効果がある。

●Gain … 音量がゆらぐ。音が割れたり、途切れたようになる。

●MDCT … 音の鋭さが失われ、音がにじんだような感じになる。

なお、上記において注記すべきなのは、各係数の符号化は、ベクトル量子化し、量子化テーブルを引いて符号化するという方式をとっているため、例えば、単純に係数のある部分だけをいじれば、高い音のみが歪むなどといった制御は困難であるということである。各係数の表す特徴の一部を劣化させようとする、その係数の表す特徴全体に影響を与えてしまうということである。

## 5. スクランブルに対する評価

本節においては、サンプルの楽音に対してスクランブルをかけ、主観評価を行った結果を示し、その考察を行う。

### 5.1 目標とするスクランブル

今回の評価試験では、実際に、以下の二つの条件を満たすようなスクランブルが、そもそも作成可能かどうかを調べることを目標とした。

●スクランブルの条件のひとつとして、元の楽

音と比較して音質が明らかに劣化していることが分かるものを作成することを目標とした。これは、著作権保護の観点から、元の音楽との違いがはっきり分かる必要があるからである。

●スクランブルの条件の二つ目として、適度な不快感を与えるものを作成することを目標とした。スクランブルをかけているにもかかわらず、全く不快感を与えることがなければ、わざわざ料金を支払ってスクランブルを解除したいと思う人が減少する。また、極度な不快感を与える様では、プロモーションファイルとしても、全く聞いてもらえなくなってしまう。

### 5.2 評価試験の方法およびその結果

実際の評価試験は、以下の様にして行った。

●スクランブルは、各係数に独立してかけ、それぞれ評価を行う。その上で、どの係数にスクランブルをかけた場合でも、5.1 節の二つの条件を満たすスクランブルのかけかたが可能かどうかを調べる。なお、複数の係数に同時にスクランブルをかけることは、今後の課題として残す。

●3.1 節で述べたように、スクランブルの強度を途中で変えることはせず、強度を一定に保ったまま評価試験を行う。これは、強度を途中で変えると、評価の仕方が複雑になるからである。

●各係数にかけるスクランブルの度合いであるが、係数の中でも、LSP、Pitch、Gain は、係数全体をスクランブルかすると明らかに不快な音となることが分かった。これらについては、さまざまな強度でスクランブルをかけたサンプルを作成し、その中で、筆者が最も目的に沿っていると思われる強度のスクランブルを用いた。なお、Fw. env.、MDCT については、係数全体をスクランブルした。

●評価試験としては、被験者に、まずスクランブルのかけられていない原曲を聞かせ、その後、各係数をスクランブルしたものを一回ずつ聞いてもらう。そうした上で、以下の二つの質問に対する回答を求めた。

- 1) 原曲との違いを、次の 3 段階の中から選択せよ。1=違いは明白である。2=違いは

分かりづらい。3=違いは分からない。

- 2) スクランプル化された曲を聴いたときの不快感を、次の4段階の中から選択せよ。  
 1=聞くに堪えない。2=やや不快である。  
 3=不快でないが奇妙である。4=それほど不快でない。5=快適である。

以下に、20人の被験者に対して行った評価試験の結果を示す。

評価	原曲との違い			不快感				
	1	2	3	1	2	3	4	5
LSP	18	2	0	11	8	1	0	0
Pitch	17	3	0	1	7	10	2	0
Fw. env.	13	6	1	0	4	10	6	0
Gain	16	4	0	0	12	5	3	0
MDCT	16	4	0	2	11	6	1	0

表1 スクランプルの主観評価

上記の結果から、Pitch、Fw. env.、Gain、MDCTの各係数については、おおよそ8割の人が、原曲と違うことを明白に分かり、かつ、適度な不快感を感じるスクランブルが可能であることが分かる。また、LSPに対するスクランブルは、歪みが大きく、聴くものに大きな不快感を与えないスクランブルを作成するのが難しいということが分かった。

## 6. まとめと今後の検討課題

以上の節をまとめ、以下を結論づける。

●楽音符号化方式の、TwinVQ形式に対する半開示システムを提案した。この半開示システムは、以下の特徴を持っている。

- 楽音の特徴の一部だけにスクランブルをかけることが可能。従って、オリジナルの内容確認が可能である。
- フレーム単位でのスクランブル化が可能なので、音質判断のために曲の一部をスクランブルをかけずに残すなどの、柔軟な処理が可能。
- 暗号化、復号化は、TwinVQ形式のビット

ストリームに直接作用させることができ、しかも処理が高速なので、TwinVQ再生器に組み込むことでオンザフライ再生が可能。

●半開示スクランブルを実際の曲に適用し、主観評価を行った。その結果、以下のことが分かった。

- 各係数に独立でスクランブルをかけたところ、LSP以外の係数については、8割以上の被験者が、元の曲との違いを明白に識別でき、かつ、適度に不快と感じるスクランブルを実現可能である。

今後の課題として、以下を検討したい。

●様々な曲調の曲に対して半開示スクランブルを適用し、どういった曲調の曲に対しては、どの係数をどの程度の強度でスクランブル化するのが適当かの関係を調べる。

●複数の係数に対するスクランブルを組み合わせた場合の効果について調べる。

●フレーム単位でのスクランブル化について、曲のどういった部分に、どの程度の強さのスクランブルをかけるのが適当かを調べる。

今回提案した半開示システムを実際に応用するにあたっては、以上についての検討が必要となってくるであろう。

## 謝辞

本研究遂行にあたり、音声情報研究部の方々には大変お世話になりました。ここに深く感謝の意を表したいと思います。また、データベース研究部の方々には様々なご指導を頂きました。深く感謝致します。

<sup>i</sup> 岩上・守谷・三樹：周波数領域重み付けインターリーブベクトル量子化 (TwinVQ) によるオーディオ符号化, 日本音響学会講演論文集, pp.339-340, 1994.

<sup>ii</sup> 藤井・山中：著作権保護のためのデジタル映像データ・スクランブル方式, SCIS96-9A, 1996.