

## 周波数ホッピングと変形離散コサイン変換による デジタル音楽への電子透かし法

岩 切 宗 利<sup>†</sup> 松 井 甲 子 雄<sup>†</sup>

この論文では、スペクトル拡散の一方式である周波数ホッピングを用いて、高品質なデジタル音楽ソフトに電子透かしを埋め込む手法を提案する。その原理は、音楽データを変形離散コサイン変換 (MDCT) で周波数係数値へ変換し、その特定成分に透かしを埋め込むものである。すなわち、ランダムな系列として準備した鍵データによるホッピングパターンによって各音声フレームごとに異なる周波数帯域へ埋込みを施す。この鍵データを所有しない者には、その埋込みパターンを特定することが困難であるため、第三者から透かしの存在を秘匿することができる。本方式によれば、透かし成分の一部は帯域通過フィルタの影響を受けず、残存できる利点がある。さらに、MDCT 係数値への埋込みをその隣接成分で適応化することにより、波形スペクトルのレベル変調にも強い透かしを実現した。本手法を用いた実験では、40 dB 以上の高音質を維持しつつ、MPEG オーディオなどの高効率圧縮を施しても透かしが消失しないことを確かめている。

### Digital Watermarking Scheme to Audio Data by Frequency Hopping Method and Modified Discrete Cosine Transform

MUNETOSHI IWAKIRI<sup>†</sup> and KINEO MATSUI<sup>†</sup>

This paper presents a digital watermark scheme for high quality audio data. It is constituted with frequency hopping by random key sequence and embedding of watermark by the modified discrete cosine transform. The watermark signal is uniformly spread over the whole window frame when it is embedded to sound, and the watermarked noise is suppressed in very low level. Therefore, the sound quality is attained more than 40 dB in the segmental SN ratio by our scheme. The frequency hopping scheme brings us a robust watermark for copyright protection, thus the watermark is proof against such attacks as bandpass filtering, amplification and lossy compression.

#### 1. はじめに

音響データは、通常、サンプリング定理に基づいて音声を標準化し、量子化して線形パルス符号化 (PCM: Pulse Code Modulation)<sup>1)</sup>によりデジタル化されている。特に、音楽ソフトは、サンプリングレート 44.1 kHz, 16 bit 量子化の PCM によりデジタル化される。この方法によると、人間の可聴周波数帯域をすべてカバーできるため、聴感的に高い音質を保ったままデジタル化できる。

一方、このデジタルデータは完全な形式で容易に複製できる特徴がある。この忠実な再現性はデジタル化の大きな利点である反面、デジタル著作物の著作権保護が必須の要件になる。この対策として、著作権情報

を人間が知覚できない形式で音楽ソフトに埋め込む電子透かし<sup>2)</sup>が注目されている。

音楽コンテンツへの電子透かしやそれに類する試みとして、Boney らによる聴感的マスキングを利用する手法<sup>3)</sup>、松井らの量子化雑音に見せかける手法<sup>4)</sup>、岩切らの圧縮音声符号に直接埋め込む手法<sup>5)~7)</sup>やスペクトル拡散法<sup>8)</sup>、富岡らの音源定位制御法<sup>9)</sup>および松本らのシンセサイザ符号 (SMF) への埋込み法<sup>10)</sup>などが検討されている。これらの手法によれば、聴感的な音質をほとんど劣化することなく透かし信号を埋め込むことができる。また、スペクトル拡散<sup>11)</sup>を応用した電子透かし技術は、拡散符号列を鍵とすることで高い秘匿性を実現できる手法と考えられる。しかし、従来、電子透かしとして検討されたスペクトル拡散は、直接拡散 (DS) 方式<sup>11)</sup>を原理<sup>8)</sup>としたものである。本研究では、スペクトル拡散のもう 1 つの手法である周波数ホッピング (FH) 方式<sup>11)</sup>に基づいた音楽ソフトへ

<sup>†</sup> 防衛大学校情報工学科  
Department of Computer Science, National Defense  
Academy

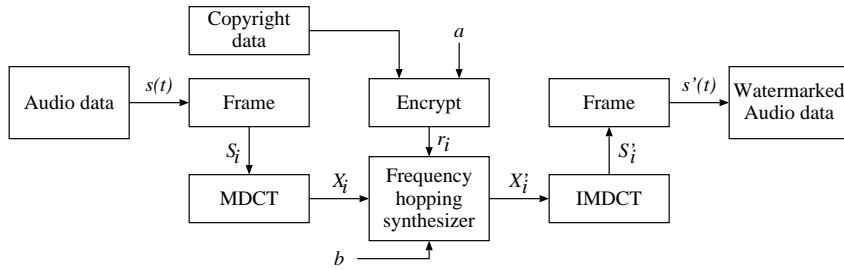


図1 透かしの埋込み処理の流れ

Fig. 1 Watermark embedding diagram.

の電子透かし法について検討した。

その方法は、各コンテンツごとに異なる疑似乱数系列による周波数ホッピングパターン<sup>11)</sup>(FHパターン)を鍵データとして準備し、電子透かしの存在を秘匿するために埋込みを施す周波数成分を時間軸上でランダムに変化させるものである。この単位時間あたりの埋込み成分の変化の頻度を高くすれば、狭帯域の透かし信号が時間軸上でランダムに広帯域へ拡散された状態になり、そのパターンを知らない第三者による埋込み信号の不正な抽出を困難にできる。また、音楽データの占有帯域に対して透かしの埋込み帯域幅が狭く、その埋込み位置の分布がランダムで一様なほど、デジタル信号処理による影響を受けにくいと考えられる。

2章ではFH方式のスペクトル拡散および変形離散コサイン変換を用いた透かしの埋込み方法を示す。3章ではFH方式を原理とする埋込みの特徴について示す。4章では、この原理に基づいたシステムを構成し、2, 3の実験を行った結果を示す。

## 2. 透かしの埋込みと復号

### 2.1 埋込み法

本手法を用いた電子透かしの処理ブロックを図1に示す。まず、音楽ソフトから抽出された音声信号を変形離散コサイン変換(MDCT)<sup>12)</sup>する。MDCT<sup>12)</sup>は、図2に示すように \$M\$ 個のスペクトル係数を求めるために \$2M\$ 個の時系列サンプルを用いる。これは周波数分離度を高くし、かつ、フレームを互いに重複させてフレーム間歪みを抑制する処置である。ここで、フレームとは、処理単位ごとに抽出された時系列サンプルのデータ列を意味する。

サンプリング時刻 \$t\$ における音声信号を \$s(t)\$ としたとき、\$i\$ 番目の音声フレーム \$S\_i\$ のMDCT係数 \$X\_i(k)\$ は

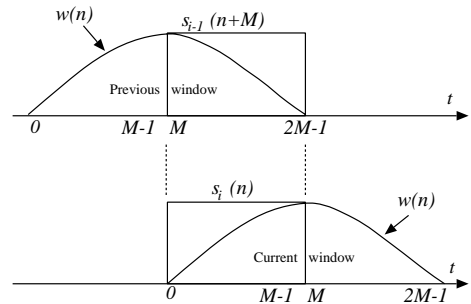


図2 MDCTの窓関数

Fig. 2 MDCT window to signal interference.

$$X_i(k) = \frac{2}{M} \sum_{n=0}^{2M-1} w(n) \cdot c(k, n) \cdot s(n + iM) \quad (1)$$

$$0 \leq k \leq M-1, 0 \leq n \leq 2M-1$$

により求まる。窓関数 \$w(n)\$ および MDCT 基底 \$c(k, n)\$ は、それぞれ

$$w(n) = \sin\left(\frac{\pi(2n+1)}{4M}\right) \quad (2)$$

$$0 \leq n \leq 2M-1$$

$$c(k, n) = \cos\left(\frac{\pi(2k+1)(2n+M+1)}{4M}\right) \quad (3)$$

$$0 \leq k \leq M-1, 0 \leq n \leq 2M-1$$

である。ここで、電子透かしの秘匿性を確実にするためにスペクトル拡散法の一方式である周波数ホッピング(FH: Frequency Hopping)<sup>11)</sup>を用いる。まず、鍵 \$a\$ を用いて暗号化した著作権データからビット列 \$r\_i\$ を各 MDCT フレームごとに抽出し、

$$p_i = r_i \pmod{M-2} + 1 \quad (4)$$

のように FH パターン<sup>11)</sup> \$P = \{p\_i | i = 1, 2, \dots\}\$ を定める。これは、埋込み処理を隣接する成分の強さの平均値

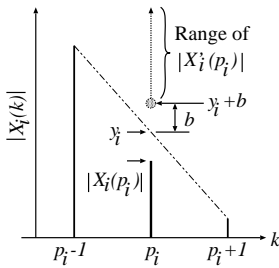


図3 埋込み規則

Fig. 3 Watermark embedding rule.

$$y_i = \frac{|X_i(p_i - 1)| + |X_i(p_i + 1)|}{2} \quad (5)$$

$$1 \leq p_i \leq M - 2$$

で適応化するための準備である．ここで，図3のように透かしの強度を制御する値  $b$  を導入し

$$|X'_i(p_i)| = \begin{cases} |X_i(p_i)|, & |X_i(p_i)| > y_i + b \\ y_i + b, & |X_i(p_i)| \leq y_i + b \end{cases} \quad (6)$$

とする．この  $|X'_i(p_i)|$  を埋込み処理済みの信号成分とする．さらに，オリジナルの  $X_i(p_i)$  と同じ正負符号の  $X'_i(p_i)$  を置き換えることで MDCT 係数上での埋込み処理を完了する．この MDCT 係数  $X'_i(k)$  を逆変換すれば，透かしが埋め込まれた状態の音声フレーム  $S'_i$  を生成できる．同様の埋込み処理を繰り返し，各フレームごとに異なる周波数成分へ埋込みを施したうえで図2のように隣接するフレームを干渉させながら署名済みの音楽データを再生する．この方法によれば，音声全体の波形スペクトルに不正なレベル変調を受けても，各成分の相対的な大小関係が保たれるため不正処理に対して高い耐性を実現できる．また，本方式では各変換フレームごとに異なる周波数成分を制御するためフレーム間歪みを生ずるが，MDCT 変換法により隣接フレームを互いに干渉させることで，この歪みを低減できる長所もある．さらに，本手法によれば周波数ホッピングパターンに用いたランダム系列  $r_i$  を鍵データとして透かしの存在を秘匿できる．

## 2.2 透かしの復号

透かしの検出は，次の手順で行う．まず，音楽データから抽出した  $2M$  サンプルの標本値  $S'_i$  を MDCT 係数  $X'_i(k)$  へ変換する．次に，埋込みに用いた FH パターンによって定まる周波数成分の強さ  $|X'_i(p_i)|$  と隣接成分  $X'_i(p_i - 1), X'_i(p_i + 1)$  から式(5)に従いこれらの平均値  $y_i$  を求め，その大小関係を調べる．その結果， $y_i$  が対象とする  $|X'_i(p_i)|$  に対して大きいかな否かにより，次式で透かし出力  $d_i$  を決定する．

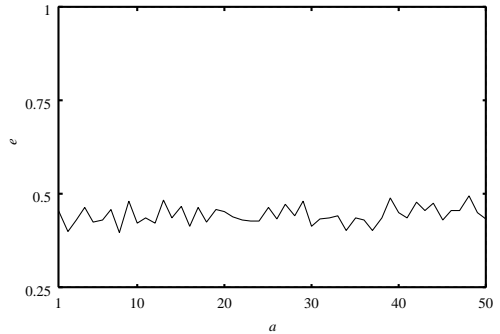


図4 ランダムな FH パターンによる検出例

Fig. 4 Detection rate by random pattern  $p_i$ .

$$d_i = \begin{cases} 1, & |X'_i(p_i)| \geq y_i \\ 0, & |X'_i(p_i)| < y_i \end{cases} \quad (7)$$

この操作をすべての音声フレーム  $S_i$  について実行し， $d_i = 1$  の分布を調べる．

まず，全音楽データから得られる音声フレーム  $S_i$  のフレーム数を  $I$  とする．このとき，透かし出力  $d_i = 1$  の検出率  $e$  は，

$$e = \frac{\sum_{i=1}^I d_i}{I} \quad (8)$$

として算出できる．ここで埋込みを施した周波数成分  $X'_i(p_i)$  をすべて特定できたならば，つねに  $|X'_i(p_i)| \geq y_i$  の関係が成立する．すなわち，正しい FH パターン  $P$  を知る者のみが，透かしの出現率  $e = 1.0$  を得られるのである．一方，FH パターンを知らない者がランダムに選んだ周波数成分から透かし出力を得ようとしても  $|X'_i(p_i)|$  と  $y_i$  の大小関係が不定になるため，図4のように約 0.5 以下の検出率になる．また，この図から隣接する周波数成分の大小関係は，一般にランダムであることも分かる．したがって，埋込みに用いた FH パターン  $P$  を知る者でなければ，正しい透かし出力を得ることは難しい．本方式では，透かしを検出するための FH パターン  $P$  の生成に暗号化のアルゴリズムと鍵および著作権データが必要であることに注意する．

## 3. FH 方式の特徴

### 3.1 透かしの秘匿性

スペクトル拡散を原理とする電子透かしとして，すでに岩切らによる直接拡散 (DS) 方式を用いた手法<sup>8)</sup>が提案されている．その方法によれば，各音声フレームごとに狭帯域な透かしの周波数スペクトルを帯域全体へ平坦な状態に拡散するため，埋込みに用いた乱数

系列がなければ透かしのスペクトルを元の状態に集中できない。すなわち、DS方式では、この乱数系列が透かし情報を復号するための鍵となっている。たとえば、その系列長が  $M$  ビットならば、 $2^M$  個の鍵が存在することになる。

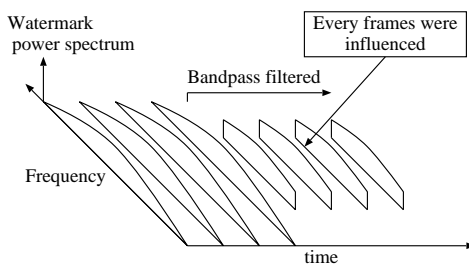
一方、ここで提案する周波数ホッピング (FH) 方式は、埋込み周波数成分を高い頻度でランダムに変化させ、時間軸上で周波数スペクトルの分布を平坦にするものである。よって、埋込みを施した周波数成分を正しく特定できなければ、高い検出率  $e$  を得られない。すなわち、FH方式では時間軸上の埋込み周波数成分の変化 (FHパターン) が透かしを復号する鍵となる。たとえば、ある音声フレームから  $M$  個の周波数係数が生成された場合、埋込み成分の選択肢は  $M - 2$  通り存在することになる。このとき、 $I$  個の音声フレームから構成される音楽データへ透かしを埋め込み、復号するための鍵の種類数は  $(M - 2)^I$  である。たとえば、 $M = 1024$ 、 $I = 50$  とするとき鍵の総数は約  $10^{150}$  となるので、埋込みに用いた FH パターンを知らない第三者にとって、透かしが埋め込まれた周波数成分を特定することは困難であると思われる。

したがって、これらのスペクトル拡散を原理とする埋込み手法は、高い秘匿性が要求される電子透かしに適していると考えられる。

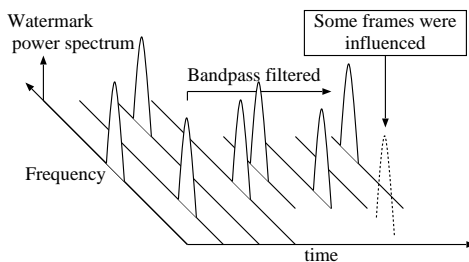
### 3.2 帯域通過フィルタ耐性

DS方式による電子透かしは、図5(a)のように帯域全体に透かし信号のスペクトルが存在するため、帯域通過フィルタによって攻撃を受けるといずれの音声フレームも影響を受ける。一方、FH方式では、この攻撃によって同図(b)のように影響を受ける音声区間が限定される。一般に、通過帯域幅は、制限帯域に比して広く設定される。よって、FH方式を原理とした埋込み方法によれば、大部分の透かし成分が帯域フィルタによる影響を受けずにすむ利点があると考えられる。その結果、FH方式では、周波数スペクトルの分布を観察することで制限帯域を判別し、不適切な周波数帯を無視することにより、透かし検出に及ぼす攻撃の影響を抑制できることなどが明らかとなる。

さらに、通過帯域内の透かし信号は、フィルタ処理の影響をほとんど受けないため、誤り訂正符号の導入による効果も大きく、影響を受けにくい周波数帯に埋込みを集中する工夫や、聴感的音響モデルによる埋込み強度の適応化も容易である。これらの特性から、FH方式による電子透かしは、帯域通過フィルタに耐性を実現できる手法であると思われる。



(a) DS scheme



(b) FH scheme

図5 帯域通過フィルタの影響  
Fig. 5 Influence on watermark by bandpass filtering.

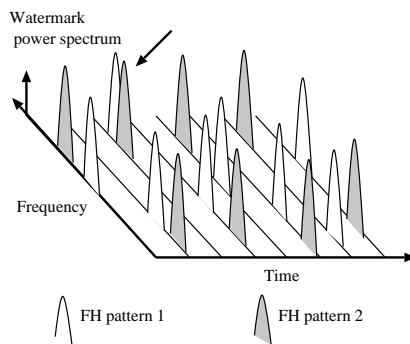


図6 透かしの多重化と独立性  
Fig. 6 Watermarks multiplying.

### 3.3 多重埋込み

FH方式による電子透かしは、埋込みに用いるFHパターンが異なる複数の透かしを多重化できる。たとえば、図6のように、2つの異なるFHパターンの埋込みを多重に施しても互いに影響を与えることはない。このように互いに異なる周波数成分へ透かしを埋め込むFHパターンを用いれば、同一のコンテンツに複数の透かしを多重に埋め込むことができる。また、図6の矢印成分のように埋込みの一部が偶然に重複することも考えられる。たとえば、ある音声フレームに埋込み成分の選択肢が  $M - 2$  個存在した場合、多重化した透かしの埋込み位置が重複する確率は  $1/(M - 2)$  である。さらに、2カ所の音声フレームで埋込み成分

表 1 実験音声

Table 1 Sound for experiment.

Name	Samples	Sec
Classic	440,832	10
Jazz	440,832	10
Dance	440,832	10

が重複する確率は  $1/(M-2)^2$  である。同様に、複数の音声フレームで埋込み成分の重複が生起する確率は一般に低いと考えられる。したがって、本方式によれば、複数の透かしを同一のコンテンツへ多重に埋め込んでも問題ないと考えられる。ただし、FH パターンの生成アルゴリズムを設計する際に、透かしの埋込み位置が重複しない系列を発生させる手法を適用する着意は必要である。

#### 4. 実験結果と考察

高音質な音楽ソフトに電子透かしを埋め込む際、音質の劣化が少なく音楽データを高効率圧縮しても透かしが消失しにくいことが望ましい。本実験では、高音質な音楽データに埋込みを施したとき、音質に与える影響および MPEG Audio Layer III (MP3)<sup>3)</sup> による高効率圧縮などのデジタル信号処理が透かしに与える影響について検討した。

##### 4.1 実験音声

表 1 の実験データは、音楽ソフトの再生音を 44.1 kHz, 16bit でデジタル化したものである。通常、これらの音楽ソフトはステレオ音であるが、議論を簡単にするため、その片側成分のみを抽出して実験データとした。また、実験に用いる FH パターンの生成には、あらかじめ 255 種類の疑似乱数データ (No.1~255) を準備した。

##### 4.2 音質の評価法

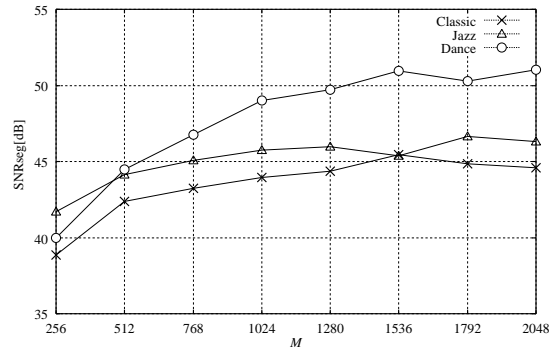
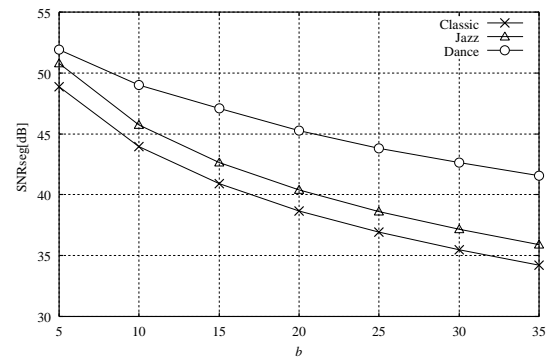
音質の客観的な評価尺度として最も基本的なものに信号対量子化雑音比 (SNR: Signal to quantization Noise Ratio) がある。SNR [dB] の評価式は、入力音声  $So(m)$  とその量子化誤差  $Er(m)$  を用いて次のように定義される<sup>1)</sup>。

$$\text{SNR} = 10 \log_{10} \frac{\sum_m So^2(m)}{\sum_m Er^2(m)} \quad (9)$$

ここでは、SNR を改良して主観評価との対応関係を向上した  $\text{SNR}_{\text{seg}}$  (Segmental SNR) を用いた<sup>1)</sup>。

$$\text{SNR}_{\text{seg}} = \frac{1}{N_f} \sum_{f=1}^{N_f} \text{SNR}_f \quad [\text{dB}] \quad (10)$$

$N_f$  は測定区間のフレーム数を表し、 $\text{SNR}_f$  は、 $f$  フレームにおける SNR である。本実験では、1 フレーム

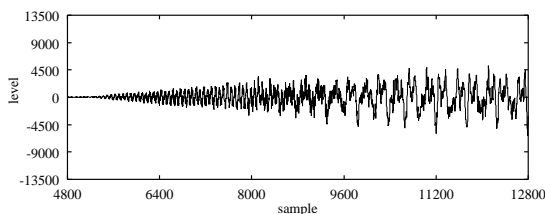
図 7  $M$  に対する  $\text{SNR}_{\text{seg}}$  ( $a = 1, b = 10$ )Fig. 7  $\text{SNR}_{\text{seg}}$  to  $M$  ( $a = 1, b = 10$ ).図 8  $b$  に対する  $\text{SNR}_{\text{seg}}$  ( $M = 1024, a = 1$ )Fig. 8  $\text{SNR}_{\text{seg}}$  to  $b$  ( $M = 1024, a = 1$ ).

の長さを 32 ms とした。また、誤差のない音声フレーム、すなわち  $\text{SNR}_f = \infty$  の音声フレームを除外して測定した。

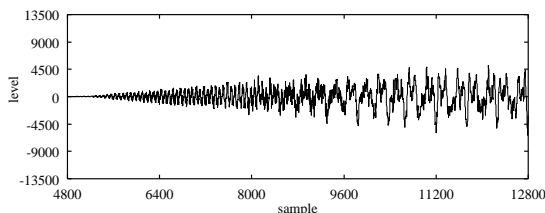
##### 4.3 音質への影響

表 1 に示した実験音声へ透かしを埋め込んだときの音質を調べた。ここでは、拡散帯域幅を決定する MDCT 変換のフレーム長  $M$  と埋込み強度のパラメータ  $b$  を変化させた結果をそれぞれ図 7 および図 8 に示す。これらから、ある程度の大きさの  $M$  までは、埋込み周波数の広帯域化によって埋込み頻度が減少し音質が向上した。また、透かし強度  $b$  が弱いほど音質に与える影響を少なくできることも分かる。これは埋込みの際に条件を満たさない周波数成分への制御量が少なくなることに起因する。

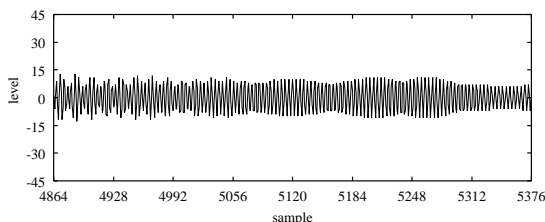
さらに、本方式では音声フレームごとに異なる周波数成分を操作するため、フレーム間に予測できない歪を生じる可能性がある。図 9 にオリジナルの波形と透かしを埋め込んだ波形を比較し、透かし信号の埋込みによる音声波形への影響を観察した。図 9(a) に Classic のバイオリン演奏音の一部 (8,000 サンプル) を抽



(a) Decoded waveform of piano sound without watermark



(b) Decoded waveform of piano sound with watermark



(c) Difference waveform

図9 音声波形の比較

Fig. 9 Comparison of sound waveform.

出した波形, 図 9(b) に  $M = 1024$ ,  $a = 1$ ,  $b = 10$  による埋込みを施した波形, 図 9(c) にそれらの差分波形を示した. ここで, 図 9(c) は, フレームの隣接部分を抽出し, 縦軸を 300 倍に拡大してある. これらの波形の比較から, 音声フレーム間に不自然な波形歪が生じないことを確認できた. これは, 埋込みのための周波数変換に MDCT を用いることで, 隣接するフレームどうしを干渉させたためであると考えられる. また, 著者らの聴取によれば, 埋込みを施した再生音声とオリジナルの音質に差異を感じられなかった.

#### 4.4 透かしの多重化と検出

本手法によれば, 透かし信号の埋込みに用いた周波数ホッピングのパターンを鍵とすることで高い秘匿性を実現できる. ここでは, 複数の検出鍵 (FH パターン) をランダムに準備し, それらの擬似鍵で不正に透かしが検出されないかどうかを確かめた. まず, Classic に 2 種類の埋込み鍵  $a = 64, 128$  で  $b = 10$ ,  $M = 1024$  による透かしの埋め込み, それら以外のランダムな 253 種類の擬似鍵  $a$  による検出結果を図 10 に示す. この図は, 横軸に検出鍵の種類を表し, 縦軸にその鍵による透かしの検出値を示している. 実験の

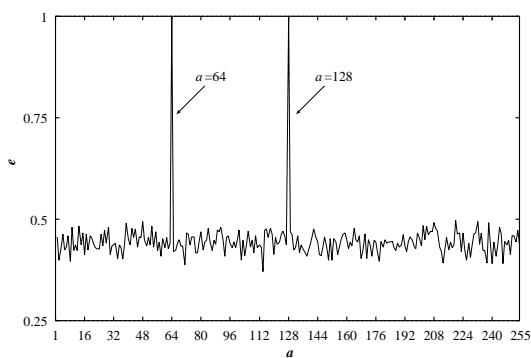


図 10 多重透かしの検出率 ( $b = 10$ ,  $M = 1024$ )

Fig. 10 Detection rate of multiplex watermarks.

表 2 レベル変調による影響

Table 2 Influence on watermark by amplification.

$\alpha$	0.50	0.75	1.25	1.50
$e$	1.00	1.00	1.00	1.00

結果 (図 10 参照) から, 擬似鍵による検出率は約 0.4 程度でほぼランダムになることが分かった. これは, 検出に用いた周波数成分  $X'_i(p_i)$  とその隣接成分の平均値  $y_i$  の大小関係がほぼランダムであることを示している. 一方, 正しい埋込み鍵で透かしを検出すると埋込み周波数成分  $X'_i(p_i)$  がその隣接成分の平均値  $y_i$  よりつねに大きく (検出値 “1”) なる. よって, 本方式によれば検査対象のコンテンツと透かしの埋込みに用いた FH パターン (検出鍵) が一致しなければ高い検出率を得られないことが分かる. したがって, 鍵情報を保有しない第三者が透かしの存在を不正な手段で特定し, 音質をほとんど劣化することなく透かし情報のみを除去することは難しい.

また, この実験結果 (図 10 参照) から, 異なる FH パターンを用いた複数の透かしの同一のコンテンツへ多重に埋め込み, 検出できることも分かった.

#### 4.5 レベル変調耐性

透かしの埋込みを隣接する周波数成分で適応化する手法は, 音声のダイナミックレンジを変化させる増幅や減衰などの処理に耐性を持つと考えられる. そこで,  $a = 1$ ,  $b = 10$ ,  $M = 1024$  の条件で埋込みを施した Classic の波形スペクトルを定数  $\alpha = 0.5 \sim 1.5$  で増幅 (減衰) し透かしが消失しないかどうかを調べた. その結果得られた表 2 の検出率から, 本方式によれば音声波形の波形振幅レベルの変調による影響を受けにくいことを確認できた.

#### 4.6 帯域通過フィルタ耐性

周波数ホッピング方式を用いて透かしの埋め込む利

表 3 帯域通過フィルタによる影響

Table 3 Influence on watermark by bandpass filtering.

Band[kHz]	Classic	Jazz	Dance
1.0 ~ 21.0	0.87	0.87	0.87
2.0 ~ 20.0	0.84	0.82	0.83
3.0 ~ 19.0	0.77	0.77	0.75
4.0 ~ 18.0	0.73	0.75	0.74
5.0 ~ 17.0	0.70	0.69	0.70

点として、帯域通過フィルタへの耐性がある。DS方式では、広い周波数帯域へ透かしがフラットな弱いスペクトルとして拡散するため、すべての音声区間で帯域通過フィルタの影響を受けることになる。その対策として、透かしの強度を高くしたうえで、高音質を維持するために拡散帯域幅を広くする処置が必要であった。一方、FH方式によれば制限帯域外に埋め込まれた透かし成分は原理的にフィルタリングの影響を受けにくいと考えられる。そこで、一般に利用されている離散フーリエ変換による帯域通過フィルタ<sup>1)</sup>を用いて再生音楽の帯域幅を制限し、透かし信号が消失しないかどうかを調べた。表3に  $a = 1, b = 10, M = 1024$  として埋込みを施し、帯域制限を施した音楽データの周波数帯域と、それらからの透かし検出率を示した。この結果から、FH方式を原理とする電子透かしは帯域通過フィルタに対して高い耐性を有することが分かった。さらに、帯域通過フィルタ処理を受けた周波数成分を判別できれば、透かしの検出率を向上できると考える。

#### 4.7 高能率圧縮符号化の影響

周波数変換やサブバンド符号化などを原理とした高能率圧縮符号化方式である MPEG Audio Layer III (MP3) のアルゴリズム<sup>13)</sup>を用いて高度なデジタル信号処理への耐性を調べた。MP3による符号圧縮は、インターネットを利用した高音質な音楽の配信ツールとして普及しており、その不正目的の利用について著作権保護上の問題が指摘されている。本実験では、インターネットで標準的に利用されている圧縮率が約 1/10 のビットレートの MP3 を用いた。図 11 は、 $a = 1, b = 10$  として  $M$  を変化させながら埋込みを施した実験データを MP3 で圧縮伸張し、その再生音楽から透かしを検出した結果である。この図では、疑似鍵を用いて得られる埋込みのない状態の検出率を黒色の点(●)で示し、埋込み鍵を用いて得られる埋込みのある状態の検出率を白色の点(○)で示した。すなわち、黒点の集合と白点の分離性が高いほど、透かしの検出は容易となる。この結果から、MP3 圧縮符号化による冗長成分の除去によって透かしの検出率

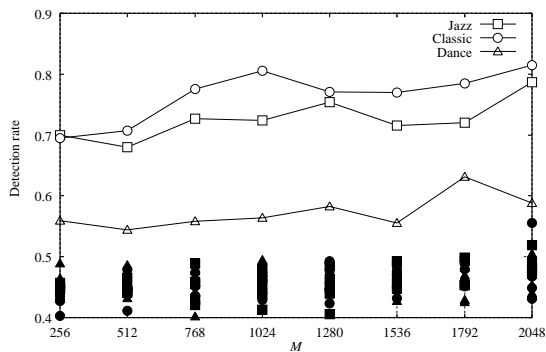


図 11 MPEG 圧縮による影響

Fig. 11 Influence on watermark by MP3 compression.

がやや低下するものの、埋込みのある状態の検出値が一般に高いことを確認できた。また、Danceの周波数スペクトルは、他の実験音声と比べ広い帯域に分布している。このような音声の周波数スペクトルは、聴感的マスキング効果を利用する MPEG 符号化によって広い帯域で変化する(図 12 参照)。そのため、Danceの検出率は、他の実験音声と比べ、低い値になったと考えられる。

#### 4.8 Jitter Attack

音楽データに埋め込まれた電子透かしを破壊する攻撃法として知られる Jitter Attack<sup>14)</sup>によって、本手法による透かしが消失しないかどうかを調べた。Jitter Attack は、音楽の周波数帯域を不特定に 0.1%程度増減する攻撃法である。これは、人間の聴覚が微小なピッチの変動を知覚できない特徴を利用したものである。まず、実験データに  $a = 1, b = 10$  の鍵で  $M$  を変化させながら透かしを埋め込み、Jitter Attack を施した再生音楽データを作成した。これらから得られた検出率を、前節と同様に図 13 で比較する。この図から、 $M$  を大きな値にするほど、透かしの検出率が低下することが分かる。これは、Jitter Attack がランダムな再サンプリングによりピッチを微小変動する手法であるため、検出処理のフレーム区間が長いほど、1 フレームあたりに受ける攻撃の密度が増大することに起因する。したがって、Jitter Attack への耐性が必要ならば埋込み処理の MDCT 変換フレーム長  $M$  を小さな値とするのが望ましい。また、微小なピッチ変動の影響を受けにくい低周波帯域に埋込み範囲を限定する工夫や、誤り訂正符号の導入による検出率の向上も必要であると考えられる。

#### 4.9 従来方式との比較

Boney らの手法<sup>3)</sup>は、原音声を必要とする方式であり、原音声を必要としない提案手法とは本質的に異なる。

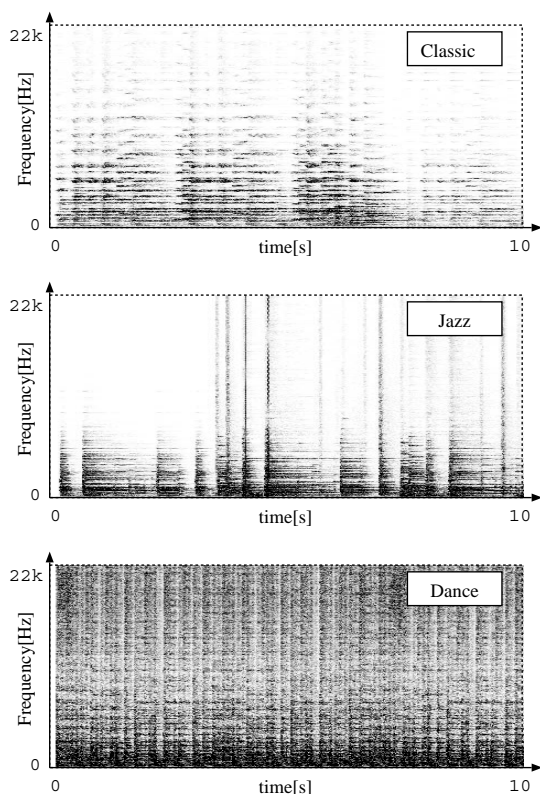


図 12 MPEG 圧縮による周波数分布の変化

Fig. 12 Influence on spectrograph by MP3 compression.

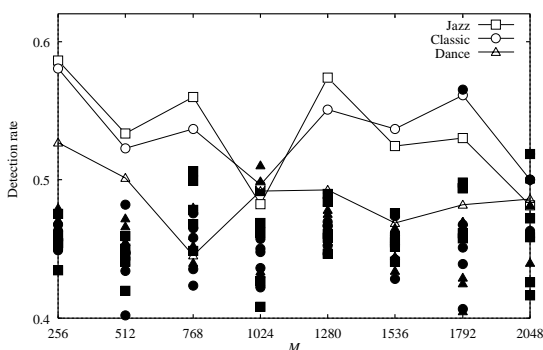


図 13 Jitter attack による影響

Fig. 13 Influence on watermark by jitter attack.

る．また，符号圧縮の過程で埋込みを施す方式<sup>5)~7)</sup>や，MIDI 符号を埋込み対象とする方式<sup>10)</sup>などは，提案方式とは異なる分類に属する技術と見なせる．ほかにも，いくつかの手法が提案されているが，ここでは，同じスペクトル拡散を原理とする DS 方式<sup>8)</sup>と提案方式を比較し，その違いを明確にする．

文献 8) に示されている DS 方式は，原理的な提案であり，攻撃耐性を高めるための工夫がなされてい

い．たとえば，埋込み法として 1 つの周波数成分を量子化する方式を採用している点あげられる．固定の刻み幅を用いた量子化による埋込み信号は，レベル変調のような簡易な攻撃でも消失しやすい問題があると考えられる．

まず，DS 方式<sup>8)</sup>の処理フレーム長を 1,024 サンプルとして，提案方式と同等の音質 (図 7 参照) になる埋込みを Classic に施した．ここで生成した実験データの音質 SNR<sub>seg</sub> は，約 44.19[dB] である．この実験データを用いて，表 2 と同様にレベル変調耐性を調べた結果，0.38~0.62 の検出率が得られた．この結果から，提案手法のように，信号レベルの相対的な大小関係を用いた埋込み手法の方が，電子透かしの埋込み技術として適しているといえる (表 2 参照)．また，この実験データを帯域通過フィルタ (通過帯域 1,000~21,000[Hz]) で処理したときの DS 方式による検出率は 0.54 であった．提案方式の検出率が 0.87 (表 3) であることから，FH 方式の方が帯域通過フィルタの影響を受けにくいことを，実験的にも確認できた．

DS 方式による電子透かしは，符号圧縮などの高度な信号処理による影響を受けやすい．特に，本研究の耐性評価に用いた Jitter Attack のように，符号同期を崩す攻撃には脆弱であると考えられる．ここでは，DS 方式の処理フレーム長を 512 サンプルとして，Classic に埋込みを施した実験データを準備した．この実験データの音質は，提案方式と同等 (図 7 参照) の 43.03[dB] とした．まず，準備した実験データを MP3 で圧縮伸長し，透かしを検出すると 0.49 の検出率が得られた．一方，Jitter Attack で処理し，透かしを検出したときの検出率は 0.51 であった．これらの実験結果と DS 方式の埋込みのない状態<sup>8)</sup>の期待値が 0.5 であることから，DS 方式の場合，いずれの攻撃においても透かしが完全に消失することが分かった．

## 5. おわりに

この論文では，音楽ソフトに用いられる高品質音声データに透かし情報を密かに埋込む方法を提案した．本手法によれば，音楽ソフトの音質をほとんど損なうことなく，ある程度の符号圧縮やデジタル信号処理にも耐える電子透かしの埋め込むことができる．よって，不正コピーされた音楽ソフトを調べて透かし情報を検出することで不正行為の事実を特定できる．

一方，利用者は，透かしの存在を知覚できないうえに不正な手段でその存在を知ることも難しい．よって，不正行為を心理的に抑止する効果も期待できる．



今後は、聴感的音響モデルや誤り訂正符号の導入による音質や攻撃耐性の向上について検討したい。

### 参 考 文 献

- 1) 小澤一範：デジタル移動通信のための高能率音声符号化技術，トリケップス (1992)。
- 2) 松井甲子雄：電子透かしの基礎—マルチメディアのニュープロテクト技術，第7章，森北出版 (1998)。
- 3) Boney, L., Tewfik, A.H. and Hamdy, K.N.: Digital watermarks for audio signals, *Proc. International Conference on Multimedia Computing and Systems*, pp.473–480 (1996).
- 4) 松井甲子雄，中村康弘，ナタウットサムパイブーン：音声通信への文字情報の埋め込み，第18回情報理論とその応用シンポジウム，pp.389–392 (1995)。
- 5) 岩切宗利，松井甲子雄：適応差分PCM符号化における音声符号へのテキスト情報の埋め込み，情報処理学会論文誌，Vol.38, No.10, pp.2053–2061 (1997)。
- 6) 松井甲子雄，岩切宗利：低遅延符号励振線形予測符号化による音声符号への電子透かし，画像電子学会誌，Vol.27, No.5, pp.475–482 (1998)。
- 7) 岩切宗利，松井甲子雄：共役構造代数符号励振線形予測による音声符号へのテキスト情報の埋め込み，情報処理学会論文誌，Vol.39, No.9, pp.2623–2630 (1998)。
- 8) 岩切宗利，松井甲子雄：スペクトル拡散と変形離散コサイン変換による高品質デジタル音声のための電子透かし法，情報処理学会論文誌，Vol.39, No.9, pp.2631–2637 (1998)。
- 9) 富岡淳樹，中村高雄，小川 宏，高嶋洋一：マルチチャンネルデジタルオーディオに対する電子透かし，1998年電子情報通信学会情報・システムソサイエティ大会，D-14-4, p.323 (1998)。
- 10) 松本 勉，井上大介，北林創太：演奏データファイルSMFへの情報ハイディング方式，2000年暗号と情報セキュリティシンポジウム，SCIS2000-C03 (2000)。
- 11) 山内雪路：スペクトラム拡散通信，東京電機大学出版 (1994)。
- 12) 筒井京弥：楽音・音声圧縮方式—ATRAC2，インタフェース，Vol.23, No.7, pp.134–142, CQ出版 (1997)。
- 13) Rao, K.R. and Hwang, J.J.: *Techniques and Standards for Image, Video, and Audio Coding*, Prentice Hall (1996). 安田 浩，藤原 洋 (監訳)：デジタル放送・インターネットのための情報圧縮技術，共立出版 (1998)。
- 14) Petitcolas, F.A.P., Anderson, R.J. and Kuhn, M.G.: Attacks on Copyright Marking Systems, *2nd Workshop on Information Hiding*, pp.218–238 (1998).

(平成12年5月1日受付)

(平成13年12月18日採録)



岩切 宗利 (正会員)

昭和45年生。平成5年防衛大学校情報工学科卒業。平成10年防衛大学校理工学研究科情報数理専攻修了。平成11年防衛大学校情報工学科助手。マルチメディアと情報セキュリティに関する研究に従事。電子情報通信学会，日本音響学会，画像電子学会，映像情報メディア学会会員。



松井甲子雄 (正会員)

昭和14年生。昭和36年防衛大学校電気工学科卒業。昭和40年九州大学大学院工学研究科電子専攻修了。昭和56年防衛大学校電気工学科教授。平成元年同大情報工学科教授。この間暗号学，情報セキュリティ，電子透かし，音声・画像データの符号化に関する研究に従事。工学博士。電子情報通信学会，画像電子学会，映像情報メディア学会会員。著書「電子透かしの基礎」(森北出版)で第15回電気通信普及財団賞受賞。