

音楽データへのベクトル合成法を利用した電子透かしによる ブロックノイズの削減

齋藤 雅之 菊池 浩明 中西 祥八郎

東海大学

〒259-1292 神奈川県平塚市北金目 1117

email: masa2goh@ep.u-tokai.ac.jp

概要: ステガノグラフィや電子透かしの手法は時間領域に入れる手法と周波数領域に入れる手法の二通りに分けることができる。前者は計算処理は少ないが各種攻撃に弱く、後者は攻撃には耐性を持つが計算処理が膨大になる。特に音楽データに対して周波数領域への変換処理を行う場合、DCTやDFTを用い周波数領域に変換を行うと、ブロックノイズが顕著に表れ音質を劣化させてしまう。これに対して、岩切によって周波数領域の透かしデータを前もって時間領域上へ展開しておき音楽データに合成するベクトル合成法が提案されている [1]。岩切は、ブロックノイズを考慮してMDCTを用いてベクトル合成法を実現したが、我々は、ベクトル合成法ではブロックノイズの効果が小さいことに注目し、DCTを用いて実現することを提案する。提案方式は、MDCTを用いた従来の方式に比べて半分の計算量で埋め込みを行う。

キーワード：電子透かし，DCT，ブロックノイズ

Reduction in Block Noises Using the Vector Synthesis Method in Watermarking to Digital Audio Data

Masayuki Saito Hiroaki Kikuchi Shohachiro Nakanishi

Tokai university

1117 Kitakaname, Hiratuka City, Kanagawa 259-1292 JAPAN

email: masa2goh@ep.u-tokai.ac.jp

Abstract: In watermarking and steganography, there are mainly two approaches to embed a watermark data: frequency domain approach which is resistant to any attacks in the cost of processing time, a time domain approach which is fragile but requires less computing time. The DCT or DFT, which is used to change a digital audio into the frequency domain, generally, causes a block noise and deteriorates the quality. In order to deal with the issue, Iwakiri proposed a VSM (Vector Synthesis Method), in which watermarking data on a frequency domain is translated into a time domain using the MDCT before synthesizing. In this paper, we point out that the VSM causes a smaller block noise than the existing DCT method, and propose a lightweight watermarking method which uses the DCT instead of the MDCT.

Keywords : watermarking, DCT, block noise

1 はじめに

近年、コンピュータの普及により音楽データの利用形態も CD プレイヤーからコンピュータの上へと様変わりしてきている。しかし、音楽データはそのままでは容量が大きくなってしまいうため、通常は周波数領域変換を利用した圧縮を行う。この周波数領域変換はにおいて、DCT(離散コサイン変換)ではブロックノイズが顕著に聴覚されてしまうために、MDCT(変形離散コサイン変換)を利用するのが一般的である。この MDCT を使った技術は電子透かしにも利用されている。

電子透かしを二種類に分けると、時間領域への電子透かしと周波数領域への電子透かしに分けることができる。時間領域への電子透かしは、計算処理は早い各種攻撃への耐性が弱い。周波数領域への電子透かしは、攻撃には耐性を持つが計算処理の時間が大きくなる。それに対し、岩切は MDCT とベクトル合成法を用いた電子透かしを提案した [1]。この方式では、周波数領域の透かしデータを事前に展開しておき、時間領域への埋め込み処理だけで周波数領域への埋め込みを実現した。しかし、この手法では埋め込み時間を短縮することはできたが、電子透かし情報を検出する際にも MDCT の計算を必要とするため、検出にかかる計算量は依然として大きい。

我々は、このベクトル合成法では、音楽データ自体に周波数領域変換を行わない点に注目し、MDCT ではなく DCT を利用して埋め込み処理を行ってもブロックノイズが顕著に聴覚されないのではないかと考えた。本論文では、この仮説を検証するため DCT を音楽データにかけた際のブロックノイズを定量化し、そのベクトル合成法と DCT を利用した電子透かしのブロックノイズの差を実験的に明らかにすることを目的としている。

本実験では、まずブロックノイズを明らかにするために、単一周波数のコサイン波形に対して DCT 処理をかけ量子化をしブロックノイズの確認をする。また、実際の音楽にも同様の実験を行い、ブロックノイズの定量化を試みる。ブロックノイズの大きさに加えて、SNR(量子化雑音比)の観察を行い、実際に聴覚での音質の確認を試みる。

2 章では準備として DCT、ブロックノイズ、ベクトル合成法について述べ、3 章ではブロックノイズの定量化について述べる。4 章ではベクトル合成法でのブロックノイズの削減について述べ、5 章で考察を述

べていく。

2 準備

2.1 DCT

ここでは DCT について述べる。N をフレーム長 $x(n)$ を時刻 n におけるデータ、 $X(k)$ を基底 k における周波数成分とする。DCT と逆 DCT は各々次式で与えられる。

$$X_i(k) = \frac{2}{N} u_k \sum_{n=0}^{N-1} x(n) \cos \frac{(2n+1)k\pi}{2N} \quad (1)$$

$$x'_i(n) = \sum_{k=0}^{N-1} u_k X(k) \cos \frac{(2k+1)k\pi}{2N} \quad (2)$$

ここで

$$u_k = \begin{cases} 1/\sqrt{2} & \text{if } k = 0, \\ 1 & \text{if } k \neq 0. \end{cases} \quad (3)$$

とする。

2.2 ブロックノイズ

ここではブロックノイズについて述べる。

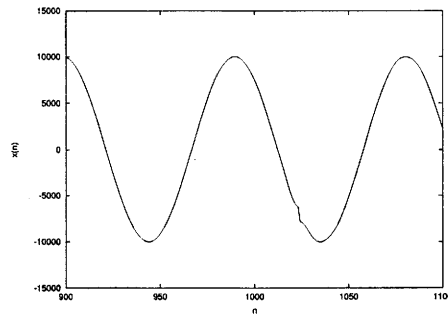


図 1: フレーム内平均差分の分布

図 1 は、単一周波数 $f = 550[\text{Hz}]$ のコサイン波形のデータに対して DCT で周波数領域に変換した後、量子化を行い、ブロックノイズを際立たせたものである。 $n = 1024$ でフレームの境界があり、データ $x(n)$ が大きく飛んでおり、これがノイズとして知覚される。

DCT ではフレームごとに計算を行うため、フレームの終点と次フレームの始点の間での連続性が失われるためにこの現象が発生する。MDCT では連続フレームを重ね合わせることでブロックノイズを低減させている。

2.3 ベクトル合成法を利用した差分拡散法

ここでは岩切によって提案されたベクトル合成法について述べる。以下にベクトル合成法と MDCT を利用した差分拡散法について記す。

2.3.1 埋め込み

STEP 1. オーディオデータ $x(n)$ を長さ N のフレームで分割する。

STEP 2. 雑乱数列 R を生成し、フレーム i の周波数成分 X_i を、埋め込み頻度 f の 2 倍の $2f$ 個の乱数 a_j について

$$X_i(a_j) \in [-p, p] \quad (4)$$

となるようにランダムに選ぶ。

STEP 3. X_1, X_2, \dots, X_M を各々逆 DCT 変換し、埋め込みベクトル w_1, w_2, \dots, w_M を求める。

STEP 4. 入力データ $x(n)$ に $w(n)$ を加え、埋め込みデータ $x'(n)$ を得る。

2.3.2 検出

STEP 1. $x'(n)$ から DCT を用いて、 $X'(k)$ を求める。

STEP 2. 埋め込みに用いたものと同じ雑乱数列 R から a_1, a_2, \dots, a_{2f} を抽出し、 f 個の差分

$$d_j = X'(a_j) - X'(a_{j+f}) \quad (5)$$

を求める。この分布の分散が平均的な分散よりも大きいかどうかで透かしの埋め込みを検出する。

図 2 に、この差分値の度数分布例を示す。埋め込みと同じ乱数列 R を用いた時の分布 (watermarked) と異なる R' を用いた時、すなわち、平均的な差分 (original) を表している。明らかに、正しい R を用いたときは、 d_i の分布が大きくなっていることが観察できる。

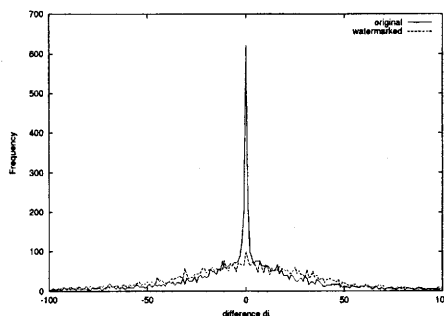


図 2: 周波数成分の差分値の分布 ($N = 256, f = 20, p = 50$)

3 ブロックノイズの定量化

3.1 ブロックノイズの定義

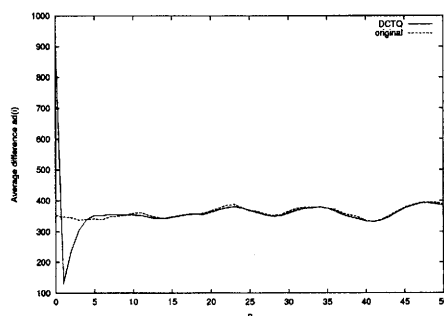


図 3: ブロックノイズ内の平均差分値の分布

ブロックノイズの大きさを次のように求める。 i 番目のフレームの最初のサンプルと $i-1$ 番目のフレームの最後のサンプルの差分、すなわち、要素 N の i 番目のフレームの差分 Δx_i は、

$$\Delta x_i = |x(N_i) - x(N_{i-1})| \quad (6)$$

で与えられる。この差分の全 M フレームの平均値

$$BN = \frac{1}{M} \sum_{i=0}^N \Delta x_i \quad (7)$$

をブロックノイズとする。フレームの境界で激しくデータが変化すれば、 BN が大きくなる。

ここで、サンプリングレート 16[bit], サンプリング周波数 44.1[kHz], モノラルでフォーマットされた POPS の音楽をフレームの大きさ $N = 1024$ で DCT と量子化誤差の強調処理 (Quantization) を行った結果を図 3 に示す。ここで、フレーム内の時間領域における i 番目と $i - 1$ 番目の差分の平均値の分布, すなわち、フレーム内のインデックス i における平均差分 (Average difference)

$$ad(i) = \frac{1}{M} \sum_{n=0}^N |x(nN + i + 1) - x(nN + i)|. \quad (8)$$

を表している。図 3 において、 $i = 0$ の $ad(0)$ は前フレームとの差分であり、この値が原曲と大きく異なっている。これをブロックノイズ $BN = ad(0)$ とする。また、フレーム境界以外での平均差分はブロックノイズに比べ極端に小さく、DCT におけるノイズの大半がブロックノイズに起因していると言える。表 1 はブロックノイズの平均、標準偏差を表している。ブロックノイズは他の平均差分と比べて、3 倍も大きいことがわかる。

表 1: ブロックノイズの平均, 標準偏差

	original	DCTQ
BN の平均	355.812	918.954
BN の標準偏差	358.045	821.918

4 ベクトル合成法でのブロックノイズ削減

本章では、2 章で示したベクトル合成法に DCT を用いて差分拡散法での電子透かしを埋め込み、そのブロックノイズを求める。実験には、3 章で用いた音楽データに対して DCT 変換をした後、量子化を行いブロックノイズを際立たせ、それと同程度の変化量の透かしデータを埋め込んだ。

4.1 評価法

透かしデータの強さ (大きさ) を示すために、周波数領域上での MSE (Mean Squared Error) を用いる。MSE は原曲と変更を加えた曲の周波数領域上での値の平方の差分の平均値である。MSE は入力 $W(k)$ と

表 2: DCT と VSM における埋め込みの音質とブロックノイズ値

	MSE	SNR[dB]	BN
DCT	19	46.1	431
DCTQ	20	45.5	756
VSM	19	44.2	431
original	N/A	N/A	352

埋め込み後 $W'(k)$ と全サンプル数 NM を用いて、

$$MSE = \frac{1}{NM} \sum_{k=0}^{NM} (W(k) - W'(k))^2 \quad (9)$$

と定める。

次に音質を評価する指標に SNR (信号対量子化雑音比) を用いる。SNR は次のように定義される。透かしや量子化による雑音を含むデータを $x'(i)$ とするとき、入力データを $x(i)$ とし、 $SNR_{[dB]}$ は

$$SNR = 10 \log_{10} \frac{\sum_i x(i)^2}{\sum_j (x(i) - x'(i))^2} \quad (10)$$

で与えられる。

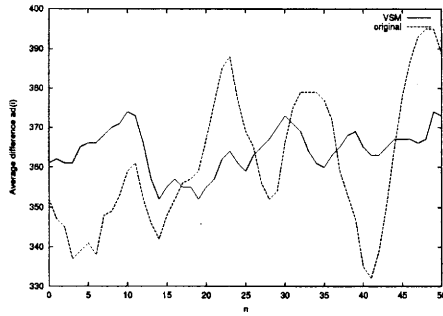


図 4: オリジナルデータと埋め込みデータのブロックノイズの比較

4.2 実験結果

音楽データに対して、フレーム長 $N = 256$, 埋め込み強度 $p = 100$, 埋め込み頻度 $f = 50$ [回/フレーム] の差分拡散法で電子透かしを埋め込み、DCT によるデータとベクトル合成法のデータを比較した。プ

ロックノイズの影響を強調するため、DCT のデータでは、DCT 係数を周波数で割り、

$$X_q(j + Ni) = \frac{X(j + Ni)}{j} \quad (11)$$

によって量子化誤差を引き起こしたデータ DCTQ(DCT with Quantization) も求めた。一方、ベクトル合成法 (VSM: Vector Synthesis Method) では、この誤差は生じ得ない。以上の3データ (1. DCT による透かし入り, 2. DCT と量子化誤差 (DCTQ), 3. ベクトル合成法 (VSM)) について、透かしの強さ (MSE), 音質 (SNR), ブロックノイズ値 ($ad(0)$) を表 2 に示す。

表 2 より、MSE がほぼ同じであることから、周波数領域上での変化量は DCT, DCTQ, VSM の間でほぼ差がないといえる。SNR も3つの間での差は少なくなっているが、ブロックノイズ値 $ad(0)$ の値は DCTQ が突出して大きくなっている。また、VSM と DCT の差がほとんどない。

原曲と DCT 及び、原曲と VSM によるフレーム内での平均差分 $ad(i)$ を図 3 と 4 に示す。図 3 において、量子化誤差を強調した DCT はフレーム境界にあたる $ad(0)$ において、フレーム内の平均的な振幅と比べ特異な振幅を生じている。これがブロックノイズの特徴であると考えられる。それに対し、VSM においては、 $ad(0)$ でも平均的な振幅の範囲内に収まっている。すなわち、ブロックノイズがほとんど発生していないと考えられる。

また、ベクトル合成法と組み合わせることによって DCT を用いてもブロックノイズを生じずに電子透かしの埋め込み出来る可能性が示された。DCT は MDCT に比べフレームに用いるサンプル数が半分で済むため、透かしの検出を半分以下の計算量に削減することができる。

5 終わりに

本研究では音楽データの周波数領域の変換に DCT を使った際に生じるブロックノイズを定量化した。また、DCT とベクトル合成法を用いてブロックノイズを聴覚できない程度に削減する電子透かしの手法を提案した。これにより、電子透かしの検出の際の計算量を従来の MDCT を使う手法の半分以下にする。今後は検出時間などの厳密な時間などを考慮した、より実用向けの電子透かしの開発を試みたい。

参考文献

- [1] 岩切宗利：ベクトル合成法によるデジタル楽への電子透かし，コンピュータセキュリティシンポジウム 2003，pp. 617-622 (2003).
- [2] 松井甲子雄，岩切宗利：デジタル音楽への変形離散コサイン変換と差分拡散法による電子透かし，情報処理学会論文誌，Vol. 44, No. 4, pp. 1101-1110 (2003).
- [3] 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，森北出版 (1998).