

リアルタイム埋め込みのための音楽電子透かし
AUDIO WATERMARKING FOR REALTIME EMBEDDING立花 隆輝[†]
Ryuki Tachibana

1. はじめに

コピー制御や放送モニタリングのために音楽電子透かしの研究は盛んに行われている。特に近年は周波数領域で原音¹の振幅を操作すると位置ずれや伸縮に対する耐性を持たせ易いことが判明し、様々な手法が提案されている [1, 2, 3, 4]。

電子透かしの用途の一つにキューシートの自動作成がある。テレビやラジオで放送する音楽や宣伝に電子透かしを先に埋め込んでおいて、放送時に受信した信号から電子透かしを検出し自動的に放送のログ(キューシート)を作成するという用途である。生放送についてもキューシートの自動作成を行うには、リアルタイムでの電子透かし埋め込みが必要となる。

本論文ではまず、振幅操作による音楽電子透かし手法をリアルタイム埋め込みに応用する際に生じる新しい問題を示し、次にその問題を回避する方法と実験結果について報告する。

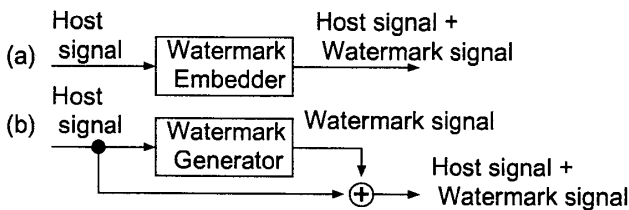


Fig. 1. (a) Conventional watermark embedder, and (b) Watermark Generator proposed in the paper

2. リアルタイム埋め込みに伴う問題

一般に音楽電子透かしの埋め込み装置はソフトウェアにせよハードウェアにせよ、入力として原音を受取り、原音と透かし信号の合計を出力する(図 1a)。これを応用して PC でリアルタイム埋め込み装置を実装するなら、サウンドカードのラインイン端子に入力された音楽信号に対して埋め込みを行った結果を、サウンドカードのラインアウト端子から出力することになる。この方法には二つの問題点がある。

P1 サウンドカードの入出力と透かし信号の計算にかかる時間だけ、原音に遅延が生じる。

P2 リアルタイム埋め込みを行って即座に放送を行う状況において埋め込み装置に故障が生じた場合に、原音の出力にも問題が生じてしまう。

そこで考えられるのが装置からは透かし信号のみ出力し、外部で原音と透かし信号の重畳を行う構成である(図 1b)。この構成では原音に遅延は生じず、万一故障が発生しても原音の放送

は止まらない。しかし、振幅操作を行う手法 [2] でこれを実現するには問題がある。

P3 透かし信号が原音より遅延し、完全な同期が不可能になる。

[2] では周波数領域において、原音の位相を保ち振幅を増減することによって電子透かしの埋め込みを行っている(図 2a)。しかし埋め込み信号が最低でも 1 フレーム遅れる(512 サンプル=約 12msec) ことによって、埋め込み信号と原音の位相はほぼ独立になってしまう。特に疑似乱数の符号 s が負で振幅を減らすべきだった周波数(図 2a2)においては、ノイズが重畳されるのと同じことになり振幅が逆に増えることになってしまうので、検出に逆効果になる。

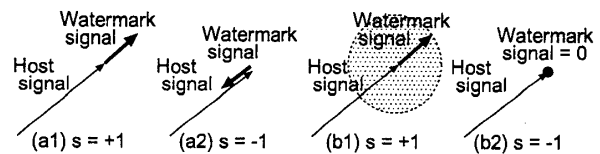


Fig. 2. The host signal and the watermark signal for the conventional method (a) and for the proposed method (b)

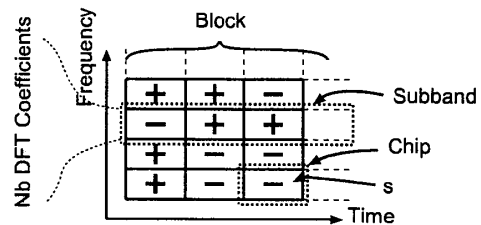


Fig. 3. Pattern block in the time-frequency plane

3. 方法

[2] の埋め込みアルゴリズムでは原音を離散フーリエ変換(DFT)を用いて短時間分析し、得られる時間周波数平面で矩形領域に分割し、個々の矩形領域に疑似乱数を割り当てる(図 3)。そして矩形領域に含まれる周波数成分の振幅を疑似乱数 s に応じて増減する。

本論文では透かし信号と原音の間に遅延がある状況でも振幅の増減を可能にするために、この埋め込みアルゴリズムを次のように修正した。疑似乱数の符号が正だった場合には透かし信号の位相は観測した原音の位相と同じにする。実際には遅延のために透かし信号の位相をランダムに決めるのと検出結果はさほど変わらない(図 2b1)。疑似乱数の符号が負だった場合には、その周波数帯には埋め込みは行わない(図 2b2)。この方法によって疑似乱数が正の部分と負の部分で振幅に偏りが生じ、電子透かしの検出が可能になる。ただし効率的な振幅の操作は

[†]日本アイ・ピー・エム株式会社 東京基礎研究所
神奈川県大和市下鶴間 1623-14

¹Host Signal。電子透かし埋め込み前の音楽。

不可能になるので検出成績の著しい低下は避けられない。そこで最小可聴値も計算し [2] よりも大きな透かし信号を用いることとした。

4. 実験

実験では約 400Hz から約 15kHz を等間隔に 24 の周波数帯に分割し、ブロックには $8 \times 4 = 32$ フレーム = 16384 サンプル = 0.37 秒の時間方向長さを持たせた。1 ブロックには同期信号と 7 ビットを埋め込んだ。30 秒間には 81 ユニットの存在するので 567 符号ビットを検出することができるが、エラー訂正符号を用いて 64 情報ビットを埋め込んだ。実験にはオーケストラ 3 曲、単楽器 3 曲(ピアノ、バイオリン、チェンバロ)、ポップス 3 曲の計 9 曲の 44.1kHz 16 ビットサンプリングで各 100 秒のモノラル音楽を用いた。透かし信号の大きさは原音に対して平均して 34.77dB 小さい音量となった。ある時間位置での透かし信号と原音の周波数スペクトルの例は図 4 の通りである。実験では作成した透かし信号を、原音に対して N_s サンプル遅れさせて重畳し、そこから検出したビット検出値を観測した。図 5 の横軸は遅れ N_s 、縦軸はビット検出値の平均値である。コンテンツに埋め込みがない時にはビット検出値の分布は標準正規分布に近似できる。[2] の方法(図 5a)では 8 サンプル程度のわずかなずれから検出がまったく不可能になったのに対して、本論文の方法(図 5b)では遅れに対して検出値は変化しなかった。ただし、ポップスについてはたとえ遅れがなくても検出値が大きくなる傾向があった。これはポップスでは原音の音量が大きく図 4c のように最小可聴値を用いた透かし信号を大きく上回っているからである。音質に問題を及ぼさないのもっと大きな透かし信号を加えるように心理聴覚モデルが出力を大きくすべきであるのだが、その調整は今後の課題である。また、実験では 64 ビットのメッセージを正しく検出できた回数も調べた。オーケストラと単楽器については標準の埋め込み強さで全て検出が可能だった。ポップスについては標準の倍の大きさの透かし信号を埋めた時に半数以上の検出が可能になった。図 6 は透かし信号の大きさとビット検出値の平均の関係である。

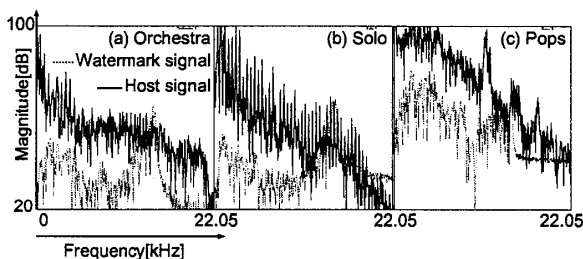


Fig. 4. Sample of the watermark and host signal for (a) Orchestra, (b) Solo, (c) Pops

5. おわりに

本論文では音楽電子透かしをリアルタイム埋め込みで用いた時に生まれる新たな要求を明らかにし、この要求を満たす方法を紹介した。この方法は [2] の埋め込みアルゴリズムのみを修正するものであり、検出アルゴリズムは [2] と同じもので検出が可能である。そのため、たとえば放送モニタリングを行っている時に、本論文の手法を用いてリアルタイム埋め込みされ

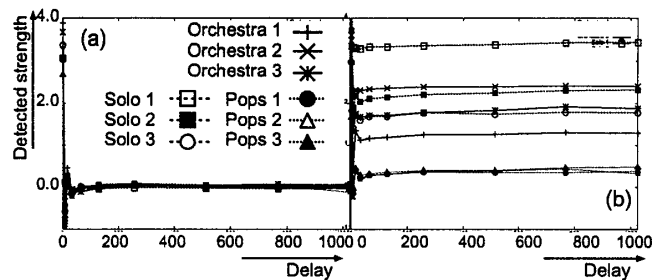


Fig. 5. Detected watermark strength

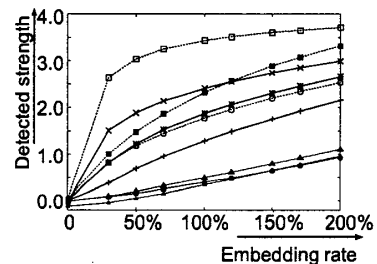


Fig. 6. Detected watermark strength

たコンテンツと、従来の手法を用いて埋め込まれたコンテンツが混在していても、一つの検出アルゴリズムを用いて高速に透かしの検出を行うことができる。

しかし本論文の方法はまだ耐性、音質ともに実用には不十分なレベルであり、リアルタイム埋め込みに適した心理聴覚モデルの作成が特に不可欠である。

6. 参考文献

- [1] C-P. Wu, P-C. Su, and C-C. J. Kuo, "Robust and efficient digital audio watermarking using audio content analysis," in *SPIE Int. Conf. on Security and Watermarking of Multimedia Contents II*, San Jose, USA, January 2000, vol. 3971, pp. 382-392.
- [2] R. Tachibana, S. Shimizu, T. Nakamura, and S. Kobayashi, "An audio watermarking method robust against time- and frequency-fluctuation," in *SPIE Conf. on Security and Watermarking of Multimedia Contents III*, San Jose, USA, January 2001, vol. 4314, pp. 104-115.
- [3] J. Haitsma, M. van der Veen, F. Bruekers, and T. Kalker, "Audio Watermarking for Monitoring and Copy Protection," in *Proc. of ACM Multimedia*, Marina del Rey, USA, November 2000, pp. 119-122.
- [4] D. Kirovski and H. Malvar, "Spread-spectrum audio watermarking: requirements, applications, and limitations," in *IEEE Forth Workshop on Multimedia Signal Processing*, Cannes, France, October 2001, pp. 219-224.