

遺伝的アルゴリズムを用いた MP3 圧縮に耐性を有する 音声電子透かし埋め込み法

(知能情報システム学) 田仲 雅和

1. はじめに

デジタルコンテンツの著作権保護が社会問題となりつつある。現在、音声に対する電子透かしの研究は多数行われているが、その埋め込み位置について議論されたことは少ない。そこで本研究ではこの点に着目し、電子透かし埋め込み位置の最適化を検討した。

今回は、対数振幅を用いた音楽電子透かし法^[1]を用いて埋め込み位置を決定する多目的最適化問題を作成し、遺伝的アルゴリズム (GA) を用いて、近似最適解求めた。

2. 方法

2.1. 透かし埋め込み方法

[1]の手法では、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を方形領域 (ブロック) に分割する。ブロックは周波数方向に B_H 行、時間方向に B_w 列で合計 $B_w B_H$ 個のタイルから構成されている。タイルごとに、DFT係数の振幅を増減させることでそれぞれのブロックに透かしの埋め込み。振幅の変化量は心理聴覚モデルを用いて計算される。

2.2. 対象とする問題の定式化

MP3 圧縮後に復号した音声における電子透かしの検出率に制約を課し、電子透かし埋め込み後の音質と検出率を目的関数とした多目的最適化問題を作成した。ここで音質の評価には信号対量子化雑音比 (SNR) を改良して主観評価との対応関係を向上したSNRseg (Segmental SNR) ^[2]を用いた。そして、検出率とSNRsegを共に最大化させることを目的とした。

2.3. GA におけるコード化

固体の遺伝子を 0,1 で表現し、上位ビットでブロックの位置を表現し、遺伝子が 1 の場合埋め込みを行う。下位ビットで埋め込み強度をあらわす。

3. 実験

埋め込み対象の音声として、RWC研究用音楽データベース (音楽ジャンル)^[3] から、各ジャンルの先頭の音声 11 ファイルを選択し、モノラルに変換して途中の 10 秒を切り取って実験に用いた。また、10 秒の音声に対して 94 bitの透かしの埋め込みを行った。

周波数軸に 30 行、時間軸に 9 列の 1 ブロックに対して、270 タイルのうち 120 タイルを用いて 4 ビットを埋め込んだ。本実験では 270 タイルを時間軸で 3 つに分割し、90 タイルから 40 タイルを埋め込み位置として選択することとした。また、埋め込み強度は心理聴覚モデルを用いて計算した値からの減少量として 3 ビットで表現したので、遺伝子長は 93 ビットとなる。

GAの条件として、世代数 200、個体数 400、ルーレット戦略、2 点交差率 0.6、突然変異率

0.1 を用いた。突然変異は、埋め込み量 40 以下のランダムな数だけ埋め込み位置に対応する 90 ビットから 1 と 0 を交換し、埋め込み強度に対応する 3 ビットについては 1 点突然変異とした。適応度には $(\text{検出率} - 50)^3 / (96 - \text{SNRseg})$ を用い、初期個体における検出率の平均値を下回る個体は致死遺伝子とした。また、MP3 圧縮のビットレートは 96kbps とした。

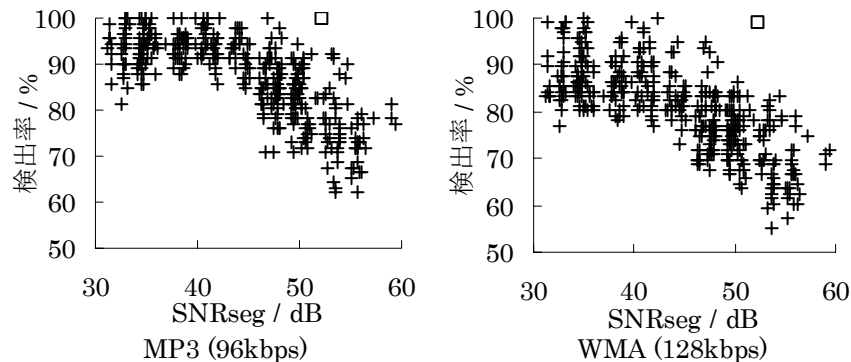


図 1 ランダムな埋め込みと本法の SNRseg と圧縮後の検出率

ジャズの音声 (RWC-MDB-G-2001 No.28) を用いた実験結果を図 1 に示す。ランダムに埋め込み位置を決めた場合の検出率と SNRseg と比較して、本法の結果 (グラフの□) は高い検出率 (100%) と高い SNRseg を実現することができた (図 1 の左)。また、他の音声ファイルにおいても本法の効果を確認した。

STEP2001^[4]から選んだ主な信号処理に対する耐性を表 1 に示す。MP3 (96kbps) 以外の処理においても、本法によって求められた近似最適解はランダムに埋め込み位置を決めた場合に比べ、高い耐性と SNRseg を得ることができた (図 1 の右)。

表 1. 信号処理に対する耐性

信号処理	検出率
Original watermark	100%
MP3 128 kbps	98%
MP3 96 kbps	100%
MP3 64 kbps	80%
AAC 128 kbps	89%
AAC 96 kbps	76%
WMA 128 kbps	99%
WMA 64 kbps	79%
Requantization 8bit	79%
Resampling 16kHz	74%

4. まとめ

音声の DFT 展開係数の配列において透かし情報埋め込み位置を決定する多目的最適化問題を作成し、GA を用いて近似最適解を求めた。本法により、音声ごとに最適に近い埋め込み位置を決定することができ、MP3 圧縮に対して高い耐性を維持したまま、より高音質で透かしの埋め込みを行うことができる。

5. 参考文献

- [1] R. Tachibana, S. Shimizu, S. Kobayashi, and T. Nakamura: "Au audio watermarking method robust against time- and frequency fluctuation", Proc. SPOE Int. Conf. On Security and Watermarking of Multimedia Contents III, vol.4314, pp.104-115, 2001.
- [2] 古井 貞熙: "音響・音声工学", 近代科学社, pp.159-160, 1992.
- [3] 後藤 真孝, 橋口 博樹, 西村 拓一, 岡 隆一: "RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース", 情報処理学会論文誌, vol.45, no.3, pp.728-738, 2004.
- [4] Jasrac: STEP2001 技術の評価仕様の概要, http://www.jasrac.or.jp/release/01/10_2-2.html, 2001.