

振幅変調に基づく音響電子透かしによる音質劣化の主観的評価と客観的評価*

西村 明 (東京情報大)

1 はじめに

これまで、多くの音響電子透かし技術が公表されてきた。しかし、技術の開発にのみ留まり、広範囲な音響信号に適用可能かどうかの検証や、透かし埋め込みに伴う音質劣化の主観評価や客観評価が、定量的に行われていない技術も多い。

著者が開発した振幅変調に基づく音響電子透かし手法は、透かし埋め込み前の信号を検出時に必要とせず、聴覚の変調マスキング特性によって透かしのキャリアとなる振幅変調が知覚されにくい特徴を持つ。また、透かし検出に必要な振幅変調強度を自動的に設定する方法を用いて、幅広いジャンルの音楽信号への透かし埋め込みと検出が可能である [1]。

前報 [2] では、透かし埋め込みに伴う音質劣化の度合いが少ないことを主観評価実験によって示した。本報告では、聴覚興奮パターンのモデル化による比較を用いた客観的音質劣化評価を行い、主観評価結果との対応が見られるかを報告する。

2 振幅変調に基づく音響透かし

2.1 埋め込み方法

本方式では、2つの隣接する周波数帯域に分割された信号同士にそれぞれ逆位相の正弦振幅変調を与える。透かし埋め込み帯域を全て帯域分割し、このペアとなる隣接帯域を複数含む2つ以上のグループに各帯域を分類し、そのグループ間の変調位相差に透かし情報を埋め込む。埋め込みデータフレーム毎に、基本となるパイロット帯域グループの変調位相を反転させることによって、検出時に埋め込み区間の同期検出を可能とする。さらに、すべての帯域ペア間には、埋め込み時の鍵によってランダムに決定された初期変調位相差があらかじめ与えられる。透かしデータの符号化は、位相差 $\pi/2$ 毎に値を割り当てる4値のPSK方式をとる。

2.2 検出方法

透かし検出時には、ペアとなる隣接する帯域信号の振幅包絡をそれぞれ求め、それらの比の対数を振幅変動波形として抽出する。その後、埋め込み時の鍵に基づいて帯域間の変調位相差を補正して、埋め込み時に鍵で定めた帯域グループ毎に変動波形の同期加算を行い、変動成分を強調する。透かし情報は、同期加算後の振幅変動波形をFFTして帯域グループ間の変動位相差を算出して求める。

3 透かし埋め込みに伴う音質劣化度合いの測定

多くの音楽ジャンルの楽曲を含む RWC 研究用音楽ジャンルデータベース (RWC-MDB-G-2001)[3]100曲から、事前に聴取実験を行い、透かし埋め込みによる音質劣化を知覚しやすい曲として、No.45(レゲエ)、69(ギターと男性ボーカル)、89(ピアノ)を選んだ。また、透かしの耐性を調べたシミュレーションにおいて、最も耐性の低かった(同じ強度で埋め込んだ透かし情報の検出率が悪かった) No.99(雅楽)も加えた。

3.1 主観的音質評価

これまで透かし強度と変形への耐性の関係を調べた結果、0dBの強度で透かしの埋め込みをすれば、ピッチ変換および時間長変換を除き、知覚符合化 (MP3/32kbps, RealAudio8/21kbps) や、ノイズ付加 (SNR 20dB)、残響付加 (1sec) を経ても95%以上の楽曲で85%以上のビット検出率を達成できた。透かし強度が-5dBの時に同程度の検出率は、知覚符合化の場合44kbps/ch、SNRは30dB、残響時間は0.5sec程度で達成できる [1]。

ここでは、その音質劣化の度合いを調べるために、極めてわずかな音質劣化を検知する評価手法を定めたITU-R BS.1116-1に準拠する、隠れ基準付き二重盲検法を用いて評価実験を行った。この手法では、コンピュータ画面上で、音質劣化した信号と劣化の無い基準信号のペアが、実験者と被験者に分からないようにラベル付けされている。被験者はそれらと、明示されている基準信号の3種を、再生中の任意の時刻で自由に切替えて繰り返し聴取し、ペアのうち音質劣化が感じられる方に、5(違いが知覚できない)・4(違いが分かるが気にならない)・3(わずかに気になる)・2(気になる)・1(とても気になる)の間で0.1刻みの評価を与える。

評価対象のNo.99については、予備実験にて透かしの検知が困難であることが分かったため、透かし強度は0dB、+10dBの2条件とした。その他の3曲については、透かし強度は0dB、-5dBの2条件とした。さらに各4曲について、音質劣化度合いのアンカーとして、MP3でエンコード(128, 96kbps)した後デコードした2条件も評価対象とした。対象区間は、冒頭1分間のステレオ信号とし、ヘッドホン受聴(STAX Lambda Nova Classic)した。各4曲毎の4条件での音質評価を1セッションとし、事前に4時間の弁別訓

*Subjective and objective evaluation on quality degradation induced by audio watermarking based on amplitude modulation

Akira Nishimura (Tokyo University of Information Sciences)

練を積んだ4名の被験者が、曲順をランダムに5・6セッションの評価を繰り返して行った。

3.2 実験結果と考察

音質劣化音に与えられた評価から基準音に与えられた評価を差し引いた差分値 (Subjective Difference Grade; SDG 値) を算出した。t 検定の結果、4名の被験者は各全16条件のうち13条件以上で、危険率6%未満で音質劣化した音楽を正しく選択して評価していることが分かった。図1に、各曲および各条件毎のSDG値の被験者平均値と標準偏差をエラーバーで示した。

図1から、透かし検知が比較的容易な曲でも、0dBの強度の透かしを埋め込むことによる音質劣化は、MP3/96kbpsと128kbpsによる音質劣化の中間程度、-5dBの透かしによる音質劣化はMP3/128kbpsと同程度であることが分かった。また、-5dBの透かしを埋め込んだ場合、原音との違いは分かるが、気にならない程度であることも分かった。さらに、透かし情報検出率が低かったNo.99では、0dBの透かし強度を用いても音質劣化は少なく、検出率を高めるために透かし強度を高めかつ、透かしの知覚検出閾を反映したより適切な透かし強度設定が可能であると示唆された。

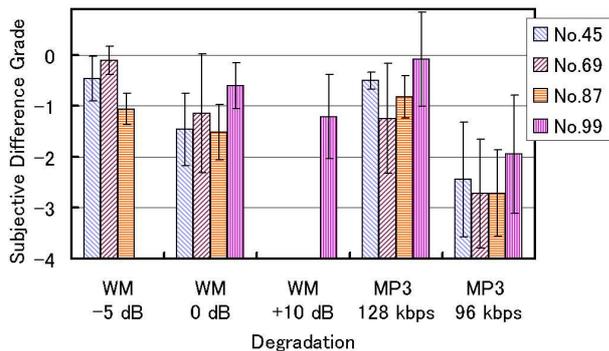


Fig. 1 Subjective difference grades.

3.3 客観的音質評価

音質劣化を客観的に測定する手法として、ITU-R BS.1387、いわゆるPEAQの基本バージョンの実装[4]を用いて、データ埋め込み済み音楽の音質劣化度合を測定した。これは原音と加工音(劣化あり音)をそれぞれ、聴覚フィルタを模したフィルタ群で帯域分割した上で、絶対閾値、周波数マスキングや時間マスキングを考慮した興奮パターン上での相違の度合から複数の指標(Model Output Value; MOV)を計算し、MOVに主観劣化評価結果とよく合うような重み付けを行って劣化度合を予測する手法である。得られる客観的劣化度合(ODG)は、ITU-R BS.1116の測定で得られるSDGに対応する値である。

図2には、横軸をSDG値、縦軸をODG値として、MP3への評価と、透かし埋め込みへ評価を示した。これらのSDGとODGとの相関を求めたところ、MP3

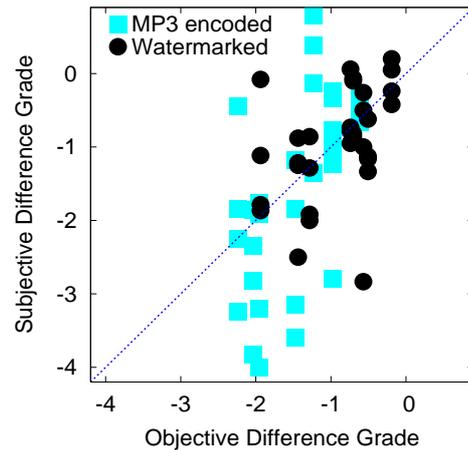


Fig. 2 Objective and subjective difference grade of watermarked and MP3 encoded music.

の場合は、 $R = 0.60 (p < 0.001)$ で高度に有意、透かしの場合は、 $R = 0.44 (p = 0.0056)$ で有意であった。よって、主観評価と客観評価はある程度対応することが分かった。また、透かしによる劣化の方が、主観評価と客観評価の相関は小さいが、図1からも分かるように同じODGに対する被験者間のSDGのパラつきは小さい。

4 考察とまとめ

今回の被験者は、十分な劣化検知訓練は積んでいたが、専門的に音質評価を行った経験は無かった。これが、主観劣化評価のパラつきを大きくする原因になった可能性がある。しかし、対象人数や劣化評価経験が少ないとはいえ、透かしによる劣化の主観評価値は、客観的劣化評価値から大きく外れることは少なかったため、PEAQアルゴリズムは、振幅変調に基づく透かしの劣化をある程度予測することが可能であろう。

主観評価実験参加者や対象曲を増やせば、PEAQの出力するODG値だけではなく、PEAQが用いる聴覚モデルの各段階での出力(MOV値)を用いて、より正確な客観劣化評価からの主観劣化評価の予測が可能になるとと思われる。

参考文献

- [1] 西村明, “帯域分割と振幅変調に基づく音響電子透かし,” 暗号と情報セキュリティシンポジウム 2006, No. 3F4-2 電子情報通信学会, (2006).
- [2] 西村明, “振幅変調に基づく音響電子透かしによる音質劣化の評価,” 日本音響学会講演論文集, 秋 2-1-10, 456-457 (2006).
- [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: 音楽ジャンルデータベースと楽器音データベース,” 音講論, 春 3-7-6, 843-844 (2003).
- [4] P. Kabal, “An Examination and Interpretation of ITU-R BS.1387: Perceptual Evaluation of Audio Quality,” TSP Lab Technical Report, Dept. Electrical & Computer Engineering (2002).