

# 振幅変調に基づくデータハイディング手法の 携帯電話音声コーデックへの耐性

西村 明<sup>†</sup>

<sup>†</sup> 東京情報大学 総合情報学部 情報文化学科  
〒 265-8501 千葉県若葉区谷当町 1200-2  
TEL 043-236-4658, akira@rsch.tuis.ac.jp

あらまし 帯域分割と振幅変調に基づくデータハイディング手法が、残響と背景雑音、音声コーデックに対してどの程度耐性を持つかについて、コンピュータ・シミュレーションによって調べた。22 話者の音声信号と、100 曲の様々なジャンルの音楽信号をホスト信号として用いた。結果は、残響と背景雑音が重畳され、12.2 kbps の AMR コーデックを経た場合において、8 bps でデータ埋め込みを行なった音声および音楽信号からは、少なくとも 80% 以上のビット検出率が得られることが分かった。また、データ埋め込みに伴う音質劣化について、PESQ と PEAQ を用いて客観的劣化度合の測定を行なった結果、それらの劣化度合は、主観的評価の「まあ良い」や「劣化がやや気になる」よりやや悪い程度の評価に相当した。

キーワード 応用音響、音声符号化、背景雑音、電話通信、残響

## Robustness against speech codecs on an audio data hiding based on amplitude modulation

Akira NISHIMURA<sup>†</sup>

<sup>†</sup> Department of Media and Cultural Studies, Faculty of Informatics,  
Tokyo University of Information Sciences  
1200-1, Yatoh-cho, Wakaba-ku, Chiba-city, Chiba 265-8501, Japan  
TEL +81-43-236-4658, akira@rsch.tuis.ac.jp

**Abstract** A data hiding technique into audio signals using subband amplitude modulation was evaluated by a computer simulation in terms of robustness against reverberations, background noises and the speech codec. Speech signals of 22 speakers and 100 tunes of music signals in various genres were served to the host audio data for embedding. Computer simulation showed that the speech and music signals with a background noise and reverberations were able to transmit at least 80% of embedded data at 8 bps when using AMR bitrate of 12.2 kbps. Objective measurement of sound quality degradation induced by data hiding was performed by PESQ and PEAQ algorithms. The resultant objective scores corresponded to the subjective evaluation of slightly degraded from ‘fair’ or ‘slightly annoying’.

**Key words** Acoustic applications, Speech codecs, Acoustic noise, Telephony, Reverberation

### 1. はじめに

音響信号にデータを埋め込み検出する利用法は、音響信号の内容に関連するデータを埋め込む電子透かしと、音響信号の内容に関連の薄いデータを埋め込むステガノグラフィに大きく分類できる。

近年、データを埋め込んだ音響信号をスピーカから再生し、マイクロホンで受音してデータを検出する応用がいくつか考

られている。この場合の受信器としては、携帯電話や PDA などが挙げられる。埋め込むデータと利用方法をいくつか挙げてみると、まず、商品やサービスに関する情報を、音響信号 (例えば CM 音楽) の聴取者に与えて広告宣伝効果を高めるものがある [1]。また、聴覚によって音響信号 (例えば公共空間でのアナウンス音声) の内容を把握できない聴覚障害者に対して、音響信号の内容と関連の深い情報を与える福祉用途 [2] がある。さらに、カラオケの伴奏音楽に歌詞表示情報を埋め込み、音楽と

同期して歌詞を表示する娯楽用途も考えられている [3]。

上記のような用途は、従来の電子透かしやステガノグラフィを包含するものであり、まさに音響信号へのデータハイディングと呼ぶべき利用法であろう。このような用途には、従来の電子透かしやステガノグラフィで用いられていた技術に無かった要求も現れてくる。つまり、スピーカ再生しマイクロホン収音することを経ても、データ埋め込み音響信号は以下に挙げる変形に耐性を持つ必要がある。

- (1) スピーカおよびマイクロホンにおける伝送周波数特性の歪
- (2) スピーカからマイクロホンまでの音響経路で生じる反射音や残響
- (3) マイクロホン受音時に付加される背景雑音
- (4) マイクロホン受音時の入力超過による振幅クリッピング歪

その一方で、従来の音響透かし用途において重要視されていた、知覚符号化 (MP3 や AAC など) や、悪意ある攻撃 (ピッチ変換、時間長変換、データ切り取りなど) への耐性については、結果として耐性が高くなることも十分あるとは言え、あまり考慮する必要はない。また、埋め込み済み音響信号の音質劣化を用途によって許容される程度まで認めつつ、必要な耐性を確保し埋め込みデータ量を高める必要がある。

スピーカ再生とマイクロホン受音に適したデータ埋め込み手法もいくつか提案されているが、マイクロホンがスピーカに近接することを要求したり [4]、埋め込み情報量が少なかったり [2], [5]、埋め込み情報量は多くても上記のような変形に対する耐性が定量的に調べられていなかったりする [6]。また、多くの音響電子透かし手法の耐性や埋め込みデータ量は、音響信号の内容に依存することが知られているが、音声から音楽までの様々な音響信号に対して有効なデータ埋め込みと検出ができるかについての検証例は少ない。

著者の考案した振幅変調に基づく音響信号へのデータハイディング手法は、上記 1, 2, 3 [7] および 4 [3] を満たすことが、多くの音楽信号と音声信号に対して確認されており、スピーカ再生とマイクロホン受音を前提とした音響信号へのデータハイディングに適しているといえる。この手法を用いると、よく響く室内である 1.3 秒程度の残響が付加されかつ雑音を信号音に対して  $-15\text{dB}$  で付加した比較的困難な条件下において、エラー訂正符号を併用すると 10 bps 程度 [3] である。これは、例えば音声信号を文字化した時の情報伝送量 (60 ~ 100 bps 程度) には、まだ及ばない。

よって、伝送したいデータをそのまま埋め込まずに、データに対するインデックス情報を埋め込み、伝送情報量を削減することが当面必要である。そのためには、伝送/表示したい情報を表示端末に事前に蓄積しておき、検出したインデックスに対応する情報を表示する方法 [3] と、埋め込まれた情報を表示端末では検出せずに、マイクで受信した音響信号をサーバコンピュータに伝送して、サーバコンピュータ側で埋め込みインデックス情報の検出と表示情報の端末への伝送を行なう方法が考えられる。

携帯電話を利用者の端末とする場合は、後者のほうが、新たなデータ検出ソフトウェアを端末にインストールする必要がない点で適している。この場合、マイクロホンで収音された埋め込み済み音響信号は、音声コーデックによって符号化の後、公衆電話ネットワークを経由して復号化され、サーバコンピュータの検出プログラムに入力される。このため、データハイディング済みの音響信号の、携帯電話音声コーデックによる符号化と復号化に対する耐性が必要となる。

本研究では、これまでの研究で効果が検証されてきた振幅変調に基づくデータハイディング手法が、音声コーデックに対して、どの程度の耐性を持つかを調べ、データ埋め込み済み信号の携帯電話による伝送の可能性について検討を行なう。また、データ埋め込みに伴う音声品質の劣化度合を評価するために、客観的音声品質評価手法のひとつである PESQ を用いる。また、音楽信号の劣化度合の評価には、客観的音楽品質評価手法として PEAQ [8] を用いて評価を行なう。

## 2. 振幅変調に基づく音響信号へのデータハイディング

本方式では、2 つの隣接する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 (図 1 参照)。透かし埋め込み周波数帯域を全て帯域分割したものについて、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、そのグループ間の振幅変調位相差に透かし情報を埋め込む。埋め込みデータは一定時間フレーム毎に更新され、この埋め込みフレーム周期毎に基本となる帯域グループの変調初期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、埋め込み時の暗号鍵によって決定される疑似乱数系列によって初期変調位相差があらかじめ与えられ、かつどの隣接帯域同士が帯域ペアとなるかについても暗号鍵によって決定されるため、どの帯域ペア間の変調位相差に透かしが埋め込まれているかを容易に知られることはない。

透かしデータの位相差への変換は、2 ビットの透かしデータ値について隣接する値間で 1 ビットのみ異なるようなグレイ符号化を施した後、パイロットグループとの変調位相差  $\pi/2$  毎に値を割り当てる 4 値の PSK 方式をとる。帯域群の数を  $n$ 、埋め込みフレーム時間長を  $T_s$  とし、 $M$  個の異なる変調周波数を同時に用いて埋め込みを行えば、埋め込み情報量のビットレートは、 $2M(n-1)/T$  bps となる。

埋め込みの強度は振幅変調度に対応するが、ホスト信号に対する検出演算後に検出される変調周波数と同じ変動のパワーを基準として強度を設定する方法を用いると、ホスト信号に合わせた埋め込み強度を設定できることが分かっている [9]。しかし、この方法では、一定時間波形を分析して強度を決定する必要があるため、リアルタイムなデータ埋め込みに適していない。そこで、今回はリアルタイムな埋め込みの可能な、あらかじめ振幅変調度を定めておき、そのまま一定として可変としない方法をとることにした。

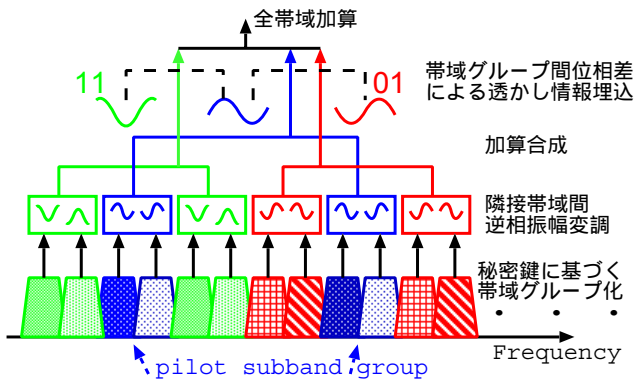


図 1 帯域ペアへの逆相振幅変調と、秘密鍵による帯域グループ化、帯域グループとパイロットグループとの変調位相差への情報埋め込み。

## 2.1 検出方法

データ検出時には、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。また、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。同期加算後の振幅変動波形をFFTして帯域グループ間の振幅変動位相差を算出し、復号化してデータを検出する。

個々の帯域毎に振幅変調を与え、それらを個別に検出することは、音響信号には一般に振幅変動が多く含まれているため、それに打ち勝つ大きい振幅変調を与える必要があり、聴感上音質変化が問題になりやすい。しかし、一般に隣接する帯域間の振幅変動には高い相関があり、検出時には隣接する帯域間で同相の変動成分がキャンセルされるため、透かし埋め込み時に与える振幅変調強度が弱くて済むという利点がある。

さらに、データ埋め込み済み信号に対するフィルタリングやイコライジングといった周波数特性上の変形によって、帯域間の振幅比が変化しても、隣接帯域に生じた強度差は振幅変調検出処理の結果定数項となり、検出に影響を及ぼさない。

## 2.2 埋め込み区間同期検出方法

透かし入り音響信号の任意の時間区間からデータ検出を行うためには、検出時に埋め込みフレームの同期をとる必要がある。変調位相が埋め込み周期毎に反転している基本帯域群から抽出された振幅変動波形に対して、埋め込みフレーム周期長の振幅変動波形 ( $W_a$ ) を 1 周期おきに同期加算し、さらに次の 1 周期おきの振幅変動波形 ( $W_b$ ) を同期加算したものを減算して、1 埋め込み周期分の累積変動波形を得る。この累積変動波形は、その同期加算開始時刻がちょうど埋め込み時のフレーム周期開始時刻に一致したときに、埋め込み変調周波数において観測される変動パワーが最大となる。よって、同期加算開始時刻を 1 埋め込みフレーム周期時間だけずらしながら得られた累積変動波形のうち、埋め込み変調周波数において最大の変動パワーが観測される同期加算開始時刻を、埋め込み区間の先頭として検出時のフレーム同期を実現する (図 2 参照)。

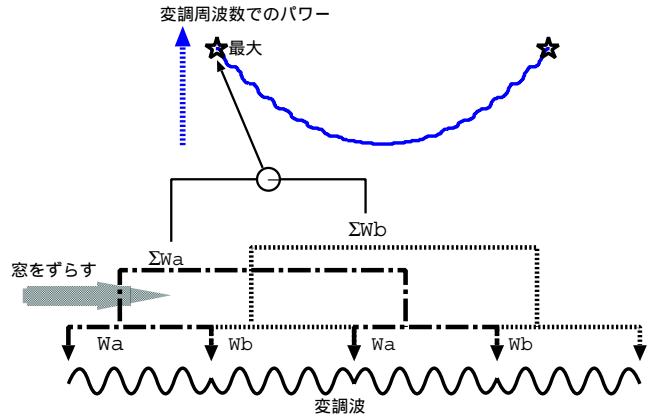


図 2 埋め込み区間同期のための累積変動波形から最大パワーを検出した時刻をフレーム開始時刻とする。

## 3. 耐性シミュレーション

音声及び音楽信号の 4kHz 以下の帯域にランダムなデータを埋め込み、残響や背景雑音を付加した後、AMR (Adaptive Multi Rate) コーデックによる符号化および復号化を経た後の波形に対して、検出処理を行なった。データ埋め込みビットレートは 8 bps とし、パラメータ値は、表 1 に示した。埋め込み強度である、振幅変調度は 0.4 としたが、これは 4. 節での客観的音質劣化度合の評価を元に決定した。

音声信号としては日本音響学会研究用連続音声データベース Vol.1 に収録されている音声を、話者ごとに連結して、1 話者あたり 36 秒分の音声信号を 22 名分 (男性 10 名、女性 12 名) 作成して用いた。これらの信号はサンプリング周波数 16 kHz であったが、8 kHz に変換して用いた。

音楽信号としては、RWC 研究用音楽データベース RWC-MDB-G-2001 [10] に収録された様々な音楽ジャンルの 100 曲の左チャンネル冒頭 60 秒を用いた。これらの信号はデータ埋め込み時にはサンプリング周波数 44.1 kHz であり、埋め込み後に 8 kHz に変換して用いた。

### 3.1 スピーカ放射に伴う残響と背景雑音

半分のシミュレーション条件では、残響のある室内において、データ埋め込み済み信号がスピーカ再生され、マイクロホン受音されることを模擬するために、RWCP 実環境音声・音響データベース [11] より選んだ、残響時間 1.3 秒の可変残響室で収録されたインパルス応答 (ファイル名: ir130.dat) を畳み込んだ。その後、室内の背景雑音に似たスペクトルを持つ Hoth ノイズ [12] を、SNR 10, 20, 30 dB のいずれかで付加した。なお、残響と背景雑音の無い条件も加えた。

### 3.2 携帯電話の音声コーデック

第三世代 3GPP (3rd Generation Partnership Project) 携帯電話におけるデジタル音声信号符号化方式の代表的なものとして、AMR 方式が用いられている [13]。音声符号化は、音声生成の要である声帯振動 (励振源) と声道共鳴 (フィルタ) を、それぞれ表現するパラメータを符号化することで情報圧縮を実現する。AMR では、あらかじめ単位振幅のパルスの取りうる

Parameters	Values
AM depth	0.4
bit rate	8 bps
subband pairs	63
subband groups	7
frame period	6 s
mod. freq. [Hz]	1.17, 1.67, 2.17, 2.67

位置と極性をお互いに少数に限定して決めておき、それら数本のベクトルの和で励振源を表現する。そして各パルスの位置の最適な組み合わせを歪の評価で選択する。伝送されるパラメータ値は、LSP(Line Spectral Pair)、ピッチ、コードベクトルとゲインである。

復号時には、コードベクトルをそれぞれのゲインで調整した後加算されて生成した励振源を、線形予測フィルタに通じて音声信号を合成する。よって、ベクトル符号帳の情報量は少なく、効率のよい情報圧縮が可能である一方、入力波形に存在する微細な時間波形情報はこのような分析合成によって失われるため、エコー法やスペクトル拡散法によって情報を埋め込んだ音響信号の耐性は低い。

また、AMR 方式の特徴としては、8000Hz サンプリングかつ 13bit 直線量子化 (8bit A-law あるいは  $\mu$ -law 圧縮) された音声波形に対して、短い時間フレーム (160 サンプル、0.02 秒に相当) 毎に 4.75 ~ 12.2 kbps の広い範囲でビットレートを可変して伝送ができる点である。さらに、有音無音検出機能、背景雑音生成機能、フレーム誤り隠ぺい機能、Wide-band (16000Hz サンプリング) 対応などがあるが、これらの機能はここでは扱わない。

AMR 符号化および復号化ソフトウェアは、3GPP TS 26.073 [14] に付属している ANSI-C コードをコンパイルして用いた。シミュレーション時の AMR コーデックのビットレートは、4.75 ~ 12.2 kbps のビットレートが 2 フレーム毎に 28 フレーム (0.56 秒) の周期で連続的に変化する条件、6.7 kbps、12.2 kbps の 3 種類とした。

### 3.3 結果

結果は埋め込んだビット値に対して得られた正しいビット値の割合を検出率とした。図 3 には、22 名の音声信号での平均検出率を示した。エラーバーは、22 条件中の 10 から 90 パーセントの範囲を示している。図 4 には、100 曲の音楽信号での平均検出率を示した。エラーバーは、100 条件中の 10 から 90 パーセントの範囲を示している。

これらの結果より、残響が無い場合は、SNR 20dB 以下では、6.7 kbps の AMR コーデックを経ても 80 % 以上の検出率が得られた。残響が付加されると、検出率は大きく落ち、6.7 kbps の AMR コーデックでは検出率が 80% を下回ることが多く、事実上利用できないに等しい。しかし、12.2 kbps の AMR コーデックであれば、85%以上の検出率が得られることが分かった。また、音楽信号の方が、残響のある条件での検出率が高いことも確認できた。これは、音声信号の成分が時間周波数

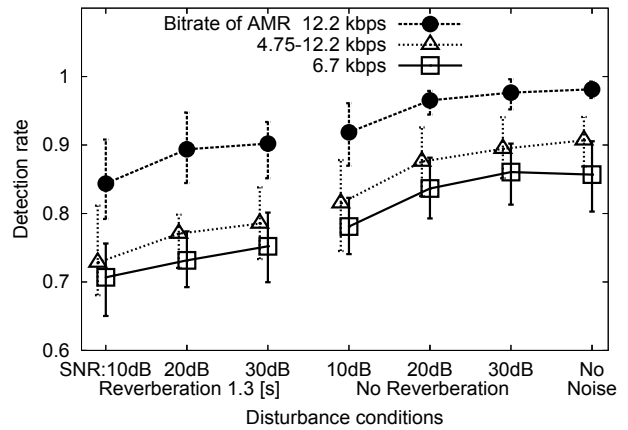


図 3 22 名の音声信号に埋め込まれたデータの平均ビット検出率。エラーバーは 10 から 90 パーセントの範囲を示す。

的に粗い分布をしているのに対し、音楽信号は密に分布しているため、埋め込みの効率が高く変形に対して頑強になるのが理由である。

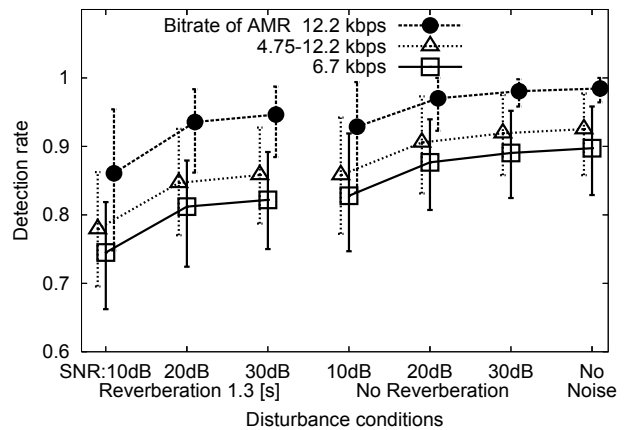


図 4 RWC-MDB-G-2001 100 曲に埋め込まれたデータの平均ビット検出率。エラーバーは 100 曲それぞれの検出率の 10 から 90 パーセントの範囲を示す。

## 4. 音質劣化の客観的評価

### 4.1 PESQ による音声品質劣化度合測定

ITU-T 勧告 P.862 PESQ (Perceptual Evaluation of Speech Quality) は、電話帯域音声や音声コーデックの品質劣化度合を測定するためのアルゴリズムである [15]。PEAQ は原信号と音声コーデックを経た後の信号を比較し、心理音響特性に基づいた信号の内的表現の差分を、音質劣化度合として報告する。PESQ の結果は、人間を被験者として測定した主観的な劣化度合評価値である MOS (Mean Opinion Score) に対応することが分かっている。

データ埋め込み後の音声信号の音質は、音声コーデックを経た音質と似ているので、その劣化度合を PESQ により測定した。耐性シミュレーションでも用いた、日本音響学会研究用連続音声データベース Vol.1 より、22 名の話者による 1200 の音素バランス文を 2 つづつ繋げて 600 文の 8 秒前後の音声信号とし、ITU-T より提供されるソースコードをコンパイルした

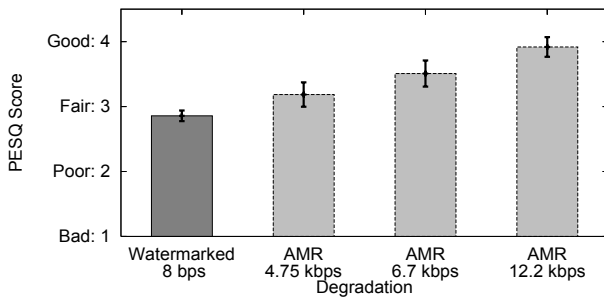


図 5 PESQ スコア

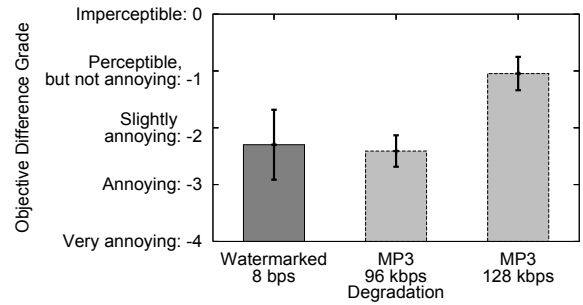


図 6 PEAQ による客観的音質劣化度合

PESQ ソフトウェアを用いて評価を行なった。

結果は、図 5 に示した。音質劣化の比較参考のため、AMR コーデックによる 4.75、6.7 および 12.2 kbps にて符号化および復号化した後の音声信号についても、同様に PESQ スコアを算出して示した。この結果から、データを埋め込んだ音声信号の品質劣化は、携帯電話よりやや音質が悪い程度であろうと予測される。また、同じ埋め込み強度 (振幅変調度 0.4) にてデータ埋め込みを行なった前報 [7] では、VCV 音節の明瞭度試験を行っており、この場合 SNR 10 dB の条件でも平均で 93% の正答率となった。このことから、データ埋め込み済み音声信号の品質劣化は明らかなもの、音声情報を伝えるには十分であろうと考えられる。

#### 4.2 PEAQ による音質劣化度合測定

広帯域音響信号の音質劣化を客観的に測定する手法のひとつとして、ITU-R 勧告 BS.1387-1 PEAQ(Perceptual Evaluation of Audio Quality) がある。これは原音 (劣化なし音) 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を計算し、主観評価実験 (ITU-R BS.1116-1) で得られる劣化度合を予測する手法である。

ここでは、Kabel [8] による PEAQ の基本バージョンの実装を用いて、データ埋め込み済み音楽の音質劣化度合を測定した。図 6 には、今回の条件でデータを埋め込んだ場合の劣化度合と、比較対象として、MP3 の 96kbps(48kbps/ch)、128kbps(64kbps/ch) で符号化し復号化した音楽信号についての、音質劣化度合の平均値と ±1 標準偏差の値をプロットした。この結果から、データ埋め込みに伴う音質劣化は、平均的には「劣化がやや気になる」よりやや悪い程度であることが分かった。また、MP3 によって符号化した音楽信号の劣化度合と比較すれば、96kbps 程度であることが分かった。データ埋め込み済み音楽はスピーカ再生され、そこに背景雑音加わることが必至である使用条件を前提とすると、データ埋め込みに伴う音質劣化は問題の無い程度であろうと考えられる。

### 5. 考 察

#### 5.1 携帯電話ネットワークにおけるパケットロスの影響

AMR コーデックは、伝送経路におけるパケットロスを隠蔽するような機構を必須要件として含んでいる。この主な仕組み

は、伝送エラーが起きたフレームの前後のフレームのパラメータ値から補間を行なって、エラーフレームのパラメータを推定し、復号化するものである。

今回はこのパケットロス隠蔽のアルゴリズムは用いなかったが、AMR 符号化後のデータに対して、1 フレーム (20 ms) 単位での抜き取り、ゼロデータフレームとの置換、ゼロデータフレームの挿入の 3 種類のパケットロスを等確率でランダムに起こすような変形を加えた後、復号化を行うシミュレーション実験も実施した。

その結果、3%程度のパケットロスでは、いずれの条件でも 1~2%程度の検出率の低下しか見られず、単純なパケットロスに対しては、ある程度の耐性を持つことが分かった。AMR コーデックによるパケットロスの隠蔽が行なわれた場合は、さらに検出率の低下は起こりにくくなるものと考えられる。

#### 5.2 実効データ伝送量と使用環境への頑強性

今回のシミュレーションでは、残響、背景雑音、音声コーデックの 3 つの妨害要因に対する耐性を持たせるため、データ埋め込み量を 8 bps と少なくした。エラー訂正符号を埋め込み時に用いることは無かったが、技術を実用化するには、なんらかのエラー訂正符号を用いる必要がある。

ここで、8 秒分のデータ (64bit) に対して、BCH(63,36,5) 符号化と軟判定復号法 [16] を併用したとすると、エラー訂正限界は 9bit 程度となる。この場合、86%のビット検出率が得られれば、36bit 分の情報伝送が可能となり、実効データ伝送量は、4.5 bps となる。今回の残響を重畳したシミュレーションにおいて、約 90%以上の信号条件で、86%のビット検出率を満たすのは、SNR 20 dB 以上での 12.2 kbps のみであった。残響が無い場合は、6.7 kbps の AMR コーデックによって、約 90%以上の信号条件で 80%のビット検出率を満たすことができたが、この場合はさらにエラー訂正限界を向上させる必要があり、実効データ伝送量は、上記の半分の 2~3 bps が妥当な線であろう。

データ埋め込み強度である振幅変調度を、今回用いた 0.4 から 0.6 に上げることによって、検出率は軒並 5~6% 程度向上するが、PESQ スコアも 2.85 から 2.40 まで低下する。一方、振幅変調度 0.6 で埋め込みを行っても、音節明瞭度としては、SNR 10 dB の条件において平均で 86%程度を得ており [7]、文章了解度にはほとんど問題ないと思われる。よって、より困難な使用環境において、さらなる音質の劣化が許容できる場合には、振幅変調度を上げて頑強度を高めることも考えられる。

### 5.3 客観的音質劣化度合

PEAQ による測定値は、MP3 などの知覚符合化圧縮に伴う音質劣化を比較的良好に予測できることは明らかになっているが、振幅変調に伴う音質劣化も同等に予測できるかどうかについては、いまだ明らかになっていない。前報 [9] において、主観的劣化度合と PEAQ による客観的劣化度合との相関を調べた結果、MP3 符号化を経た音楽の場合は相関が有意であったが、振幅変調により劣化した音楽信号の場合は、有意な相関は得られなかった。その後、被験者を増やした結果、一応有意な相関が得られている ( $R = 0.435, p = 0.030$ )。今後主観的劣化と客観的劣化指標との対応については、さらに検証を進めるべきである。

### 6. ま と め

データ埋め込み済み音響信号をスピーカから再生し、携帯電話のマイクで受信し AMR 音声コーデックを経て後、データ検出を行なう利用法を前提とした。そのため、振幅変調に基づくデータ埋め込みを行なった音響信号が、残響と背景雑音そして AMR 音声コーデックに対して耐性をどの程度もつかをコンピュータ・シミュレーションにより調べた。22 名が発話した音声信号と 100 曲の様々なジャンルの音楽信号をホスト信号として、8 bps でデータ埋め込みを行なった。その結果、背景雑音のみが重畳される場合は、6.7kbps 以上の AMR ビットレートにおいて、80% 以上のビット検出率が得られた。さらに残響が重畳する場合は、12.2 kbps の AMR ビットレートにおいて 80% 以上のビット検出率が得られた。

データ埋め込みに伴う客観的音質劣化度合を、音声信号に対しては PESQ を用いて、音楽信号に対しては PEAQ を用いて調べた。その結果、音声信号は、「まあよい (fair)」よりやや悪く、音楽信号は「劣化がわずかに気になる (slightly annoying)」よりやや悪い程度であった。

これらの結果から、残響や背景雑音が多い条件でも、AMR コーデックのビットレートが高ければ、品質を保ったまま十分にデータハイディングによる情報伝送が可能であることが分かった。

謝辞

本研究の一部には平成 19 年度東京情報大学共同研究 2、および東北大学電気通信研究所における共同プロジェクト研究による補助を受けた。

### 文 献

- [1] 松岡保静, 中島悠輔, 吉村健, “音響 OFDM - 音のバーコード -,” 第 1 回 マルチメディア情報ハイディング研究会資料, 61—64, (2007).
- [2] 東北大学電気通信研究所, 国立特殊教育総合研究所, “空気伝搬音響透かしの福祉工学への応用,” 第 1 回 マルチメディア情報ハイディング研究会資料, 103—106, (2007).
- [3] 西村明, 坂本真一, “スピーカ再生音に同期した音響電子透かしを用いる情報提示 - カラオケ歌詞表示システム -,” 情報処理学会 音楽情報科学研究報告, MUS-71-2, 7—12 (2007).
- [4] 茂出木敏雄, “音響空間のコピキタス化に向けた電子透かし埋込み容量の拡大技術,” 電気学会論文誌 C, 127, 1013—1021 (2007).
- [5] 松岡保静, “サブバンド位相シフトを用いた音響電子透かし埋込み法,” 電子情報通信学会技術研究報告, 529—533 (2006).

- [6] 松岡保静, 中島悠輔, 吉村健, “可聴帯域における音波情報伝送技術: 音響 OFDM,” 電子情報通信学会技術研究報告, 25—29 (2006).
- [7] Akira Nishimura, “Data hiding for speech sounds using sub-band amplitude modulation robust against reverberations and background noise,” in *Proceedings of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 7—10 IEEE, (2006).
- [8] 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).
- [9] Akira Nishimura, “Subjective and objective quality evaluation for audio watermarking based on sinusoidal amplitude modulation,” in *Proceedings of the 19th International Congress on Acoustics*, (2007).
- [10] 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).
- [11] 中村, 比屋根, 浅野, 遠藤, “実環境における音響シーンデータベースの構築,” 日本音響学会講演論文集, 秋, 621—622 9 2007.
- [12] IEEE Standard, “Draft Standard Methods for Measuring Transmission Performance of Analog and Digital Telephone Sets, Handsets and Headsets,” **269-2002**, (2002).
- [13] 3rd Generation Partnership Project, “Mandatory Speech Codec speech processing functions AMR Speech Codec; General Description,” **26.071**, (2001).
- [14] 3rd Generation Partnership Project, “ANSI-C code for the Adaptive Multi Rate speech codec,” **26.073**, (2001).
- [15] ITU-T Recommendation, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” **P.862**, (2001).
- [16] 藤井康広, 越前功, 山田隆亮, 手塚悟, 吉浦裕, “電子透かし検出に適した誤り訂正符号の拡張方式,” 情報処理学会論文誌, 45, No. 8, 1980—1997 8 2004.