振幅変調を用いた音声信号へのデータハイディング*

○西村 明 (東京情報大)

1 はじめに

音響信号に、電子透かし技術を用いてデータを埋め込み利用する場面として、スピーカから再生された透かし入り信号をマイクロホンで収音し、埋め込まれた透かしデータを検出する利用が考えられる。例えば、公共機関でのアナウンス音声にそのテキスト情報を埋め込み、手もとの携帯端末で検出した透かし情報を閲覧することにより、音声信号を用いて健聴者と同じ情報を難聴者に提供することができる。

しかし、音声信号は音楽信号に比べてエネルギーが時間/周波数的に疎であるため、透かしを埋め込む余地は少ない。また、透かし入り音声の空間伝搬に伴い、反射音や残響、背景雑音が重畳され、これらは透かしの検出性能を劣化させる。

著者が開発した振幅変調に基づく音響電子透かし手法 [1] は、残響に耐性があり、透かし検出時に埋め込み後に付加された雑音を打ち消す特徴を持つ。本報告では、この手法を用いて音声信号へのデータハイディングを行い、残響と背景雑音が付加される環境でも有効にデータ検出ができることをシミュレーション実験により示す。また、透かし埋め込みに伴う音声品質の劣化が少ないことを、VCV 音節識別実験によって調べる。

2 振幅変調に基づく音声信号へのデータハイディング

データ埋め込みおよび検出手法は、従来著者が音楽信号を対象として検討した音響透かし手法 [1, 2] と同じである。

2.1 埋め込み方法

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。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す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。

透かし強度となる振幅変調度は、音声信号を分析し適応的に最適な変調度を与えることも可能だが、リアルタイムに埋め込むためには分析長を長く取れないため、変調度は全帯域で同一の値とし、固定した。

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。

2.3 埋め込み区間同期検出方法

透かし埋め込み前の原音無しで透かし検出を行うためには、検出時に埋め込みデータ区間の同期をとる必要がある。抽出された振幅変動波形に対して、埋め込み周期おきに加算と減算と繰り返して得られた 1 周期分の累積変動波形を、フレーム開始時刻を埋め込み周期分だけずらしながら求め、埋め込み変調周波数において最大変動パワーが得られた開始時刻を埋め込み区間の先頭として区間同期を実現する。

3 透かし埋め込み・検出シミュレーション

3.1 実験条件

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。

音声信号への透かし埋め込み条件は、Table 1 に示した。振幅変調度は、0.3〜0.7 の範囲とした。

Table 1 Watermarking conditions.

Parameters	Values	
bit rate	64 bps	48 bps
sampl. freq.	22050 Hz	←
freq. region	≤ 6034 Hz	←
subband pairs	140	←
subband groups	25	←
frame period	3 s	4 s
mod. freq. [Hz]	1, 1.67, 2.33, 3	1, 1.5, 2, 2.5

残響のある環境を想定して、透かし入り音声信号に、RWC 実環境音声・音響データベースより選んだ、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、背景雑音として 4 種類の環境騒音 (収録場所: 駅のホーム、地下連絡通路、空港ロビー、混雑した交差点)、あるいはローパスノイズ (カットオフ 500 Hz、–9 dB/oct. : 他の環境騒音の平均的スペクトルに近い) のいずれかを付加した後、透かし情報

*Data hiding in speech signal based on amplitude modulation
Akira Nishimura (Tokyo University of Information Sciences)

を検出する処理を行った。5 種類の背景雑音は、オーバーオールの音声信号パワーに対して、信号対雑音比 (SNR) は 10dB と 20dB の 2 通りとした。

全ての音響信号は、サンプリング周波数 22.05kHz に変換後、処理を行った。よって、22 種の透かし入り音声と 5 種の背景雑音を組み合わせて 110 条件が、5 段階の透かし強度、2 段階のノイズ強度、2 つの埋め込みビットレート条件の組み合わせに対してシミュレーションされた。

3.2 結果

Fig.1 は、SNR およびビットレート毎に、埋め込んだビット値と同じビット値が検出された率を、110 埋め込み条件の中央値と、誤差棒により 10〜90 パーセンタイル値で示した。この結果から、振幅変調度 0.6 かつ 64bps で透かし埋め込みを行えば、90%の模擬条件において、84%以上のビット検出率が得られることが分かった。

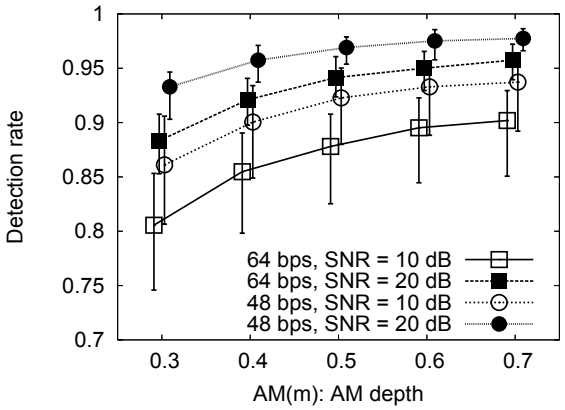


Fig. 1 Medians of bit detection rate. Error bars denote 10th to 90th percentile of the detection rate obtained in 110 conditions.

4 透かし入り VCV 音節明瞭度試験

4.1 実験条件

音響透かし技術にとって、音質劣化の少ないことは重要であるが、1 節で述べたような応用場面では、音声了解度が十分であれば、多少の音質劣化は許容されよう。そこでまず、透かしが音声了解度にどの程度影響を与えるのかを調べる前段階として、透かしを入れた VCV 音節の明瞭度試験を行った。

VCV 音節は、a,i,u,e,o の先行 5 母音と 25 の日本語子音、母音 a によって構成される 125 種とした。埋め込み条件は Table 1 と同じで、透かし強度は振幅変調度 0.4、0.6 および透かし無し条件とした。雑音付加時の明瞭度も調べるために、3 節で用いたローパスノイズを SNR 10 dB で付加する条件も加えた。被験者は正常な聴力を持つ 5 名であり、防音室内でヘッドホン両耳聴 (片耳あたり音圧 72dB) にて聴取した音声をパソコンに入力し回答した。音節、SNR 条件、埋め込み強度条件の組み合わせをランダムに 1 巡するセットを 2 回異なる日に繰り返した。

4.2 結果

Fig.2 に、被験者間の平均正答率を示した。誤差棒は被験者間での最小と最大正答率を示している。この結果より、透かしを埋め込むと、雑音環境下で音節明瞭度が比較的低下することが分かる。しかし、振幅変調度 0.6 で埋め込みを行った場合でも、雑音環境下で 83%以上の正答率が得られており、文章了解度は音節明瞭度より一般的に高いゆえ、透かし埋め込み時も充分な文章了解度が得られるだろうと考えられる。

異聴表を用いた分析では、どの条件でも先行五母音と第二母音には誤答がほとんど見られず、先行母音に依存した子音の異聴が目立った。これは透かしによる振幅変調が、先行母音と子音との調音結合部に影響を与えるためと考えられる。異聴傾向の詳細な分析は、帯域毎の適正な透かし強度決定などの、音節明瞭度を保つ埋め込みアルゴリズムの改善に繋がるだろう。

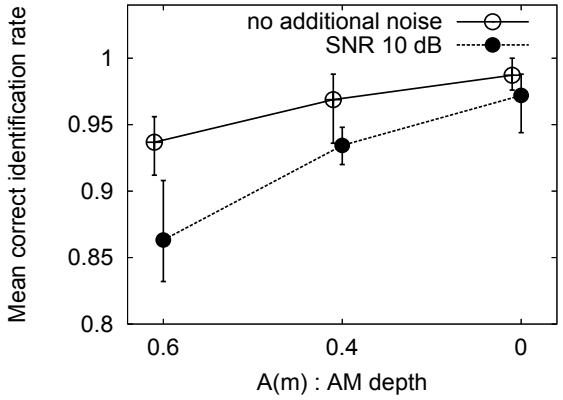


Fig. 2 Mean articulation scores of 125 VCV syllables. Error bars denote minimum and maximum among the five subjects.

5 考察

発話のテキスト情報をそのまま埋め込むには、100bps 前後の実効データ伝送速度が望まれる。今回のシミュレーションではエラー訂正は考慮に入れていないが、埋め込み時にブロック符号化を用いて検出時にエラー訂正を行う場合、実効データ伝送速度は 20bps 程度しか達成できない。今後は埋め込み強度の帯域間適正化や検出手法の改善により、データ伝送効率をさらに高める必要があろう。

参考文献

[1] Akira Nishimura, “Audio watermarking based on sinusoidal amplitude modulation,” in *Proceedings of ICASSP 2006, IV*, 797—800 IEEE, (2006).

[2] 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).