

Audio data hiding based on amplitude modulation and its application

西村, 明
Faculty of Informatics, Tokyo University of Information Sciences

<https://doi.org/10.15017/18879>

出版情報 : 九州大学, 2010, 博士 (芸術工学), 論文博士
バージョン :
権利関係 :

第7章 結論

本論文は、埋め込み前の信号を検出時に必要としない、ブラインド検出を可能とする帯域分割と振幅変調に基づいた新しい音響情報秘匿技術を提案し、その性能を検証した上で、技術の新しい利用法を示し評価することを目的とした。

第1章では、音響情報秘匿技術について概観し、技術への要求点、すなわちブラインド検出、低音質劣化、秘匿性、耐性(頑強性)、適用性、大容量、空間伝搬耐性を挙げ、従来の研究においては、技術への幅広い要求点を満たしているかどうかの検証が少ないことを説明した。

第2章では、音響情報秘匿技術に関して、用いられる用語や概念などをはじめに説明し、技術への要求とそれに対応した評価方法を示した。次に、過去の研究において提案されてきた代表的な情報秘匿手法を解説し、その特徴を簡単にまとめた。さらに、性能向上のため補助的に用いられる技術も説明した。最後に、電子透かしとしての技術利用の現状について示した。

その結果、技術の評価方法に関しては、埋め込み情報量や、検出率(エラー率)によって評価する耐性など、定量的に評価できる指標もあるが、これらの結果は対象とする音響信号に依存するため、少数の音楽信号を用いただけでは、公正な評価が行われているとは言いがたいことを主張した。秘匿性については定量的な評価方法がまだ確立しておらず、音質劣化に関しても主観評価実験に問題のある従来研究が多いことを示した。さらに、聴覚モデルを用いた客観音質劣化評価法(PEAQ)は、知覚符号化音響信号を対象としており、情報秘匿に伴う音質劣化度合を評価できるかどうかについては明らかではないことも主張した。そして、これまで提案されてきた多くの音響情報秘匿技術には一長一短があり、利用場面における技術への要求に合うかどうかを検証する必要があることを述べた。一方、実用化段階の技術は、その情報秘匿および検出手法が明らかにされていないものが多く、技術に対する秘匿性、耐性、音質などの評価は、技術の提供企業の宣伝内容を信じるしかない状態であると述べた。

第3章では、隣接する帯域信号ペアに与える逆相の振幅変調を秘匿情報のキャリアとす

る新しい音響情報秘匿手法を開発した。この手法の特徴は、すでにホスト信号に存在する振幅変動成分によって生じる変調マスキングによって、キャリアとなる振幅変動成分が聴感上目立ちにくい、という点である。また、ホスト信号に含まれる振幅変動成分の強さを元にして変調強度を設定することで、様々なジャンルの音楽信号に適用できることを目指した。そして、音響電子透かし用途を前提として、RWC 音楽ジャンルデータベース [13] の 100 曲を用いて、情報埋め込みおよび検出シミュレーション実験を行った。

その結果、4.8 bps の情報を 0 dB の強度で埋め込んだ場合、ステゴ信号に対する 21~32 kbps/ch の知覚符号化と復号化、残響時間 1.5 秒、SNR 20 dB のノイズ付加、 $\pm 2\%$ の時間長変換などの変形に対して、90%以上の楽曲で 90%以上の情報検出が可能であった。-5 dB の強度で情報を埋め込んだ場合、32~48 kbps/ch の知覚符号化と復号化、残響時間 0.5 秒、SNR 40 dB のノイズ付加、 $\pm 1\%$ の時間長変換に対して、90%以上の楽曲で 90%以上の情報検出が可能であった。一方、この情報秘匿手法はピッチ変換に対して脆弱であった。

また、RWC 音楽ジャンルデータベースの中から、情報秘匿に伴う音質変化を検知しやすい楽曲を著者が選別し、情報秘匿に伴う音質変化の検知限と、検知限以上の変調強度を与えた場合の、音質劣化度合を調べた。音質変化検知訓練を積んだ被験者 4 名に対して、音質変化の検知限となる埋め込み強度は、もっとも検知しやすい楽曲において -10 dB 程度であった。MP3 128 kbps で符号化し復号化した音楽信号の原音に対する音質変化を検知できる被験者 5 名を対象として、ITU-R BS.1116-1 に準拠して行った音質劣化度合を評価する実験によって得られた主観差分等級 (SDG) より、-5 dB の埋め込み、MP3 128 kbps、0 dB の埋め込み、MP3 96 kbps の順に音質劣化度合が増すことが分かった。また、-5 dB の情報埋め込みに伴う音質劣化度合は“違いが分かるが気にならない”程度であることが分かった。

さらに、情報秘匿に伴う音質劣化の主観評価と対応する客観評価を行うため、知覚符号化信号の音質劣化を客観評価する PEAQ を取り上げ、PEAQ の出力する ODG 値と、主観評価の結果である SDG 値との対応を調べた。その結果、SDG と ODG には有意な相関が認められ、SDG 値に個人差はあるものの、平均 SDG 値は ODG 値より大きく外れることは無かった。このため、今後は PEAQ を振幅変調に基づく情報秘匿に起因する音質劣化の客観評価法として用いることとした。

第 4 章では、前章で示した技術を、残響と背景雑音が重畳する空間伝搬条件において利用することを検討した。まず、スピーカ拡声されるアナウンス音声にデータを埋め込み、

利用者の手元の機器で復号化と表示を行う利用を前提とした。そのため、男女合計 22 名の広帯域日本語音声信号に対して 48 あるいは 64 bps にてデータを埋め込み、客観音声品質劣化度合と、残響および背景雑音下での検出率をシミュレーション実験により調べた。

その結果、48 bps のデータを振幅変調度 0.4 で埋め込むと、SNR 10 dB の場合でも 90% の条件で 84% 以上の検出率が得られることが分かった。また、情報秘匿に伴う音声品質劣化は、広帯域 PESQ による客観評価の結果、振幅変調度 0.4 で埋め込みを行うと平均的に“Poor (劣っている)”程度に劣化することが分かったが、VCV 音節識別実験からは、SNR 10 dB の環境でも 95 % 以上の明瞭度が得られることが分かり、発話内容を伝達するには問題ないことも明らかになった。

次に、埋め込み情報の検出を、マイクロホンにより受信した端末で行うのではなく、携帯電話の音声通話により接続する携帯電話音声ネットワークの先にあるサーバコンピュータにおいて実行するため、携帯電話の音声符号化を経ても情報伝達が可能かどうかを調べた。対象は、音声データベースおよび音楽ジャンルデータベースの楽曲 100 曲であり、8 bps にて情報埋め込みを行った場合のシミュレーション実験および実室内環境における実験を行った。また、この利用形態におけるデータ埋め込みに伴う客観音質劣化度合を、サンプリング周波数 8kHz の電話帯域音声信号に対しては PESQ を用いて、広帯域音楽信号に対しては PEAQ を用いて調べた。

その結果、音声信号に対しては、SNR 20 dB 以上の背景雑音のみが重畳される場合は、6.7 kbps 以上の AMR ビットレートにおいて、90% 以上の音声信号に対して 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 90% 以上の音声信号に対して 85% 以上のビット検出率が得られた。また、音楽信号の 4kHz 以下に埋め込んだ場合には、音声信号より約 5 ポイント程度検出率が高いことが分かった。客観音質劣化度合は、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、残響や背景雑音が存在しても、AMR コーデックのビットレートが高ければ、品質を大きく劣化させずに振幅変調に基づく情報秘匿による携帯電話音声ネットワークを通じた情報伝送が可能であることが分かった。

第 5 章では、振幅変調に基づく音響情報秘匿手法と、従来から提案されているエコー拡散法による情報秘匿手法 [15] について、シミュレーション実験によってその性能を比較した。電子透かし用途としての性能を比較するために、双方の情報秘匿手法の埋め込み情報量を 4.8 bps で揃え、PEAQ によってほぼ同等の客観音質劣化度合が得られる埋め込み

強度にて、それぞれの手法により情報埋め込みを行った。そして、第 3.6 節において行った、RWC 音楽ジャンルデータベースに含まれる様々なジャンルの楽曲 100 曲を対象とした、データ埋め込みおよび検出シミュレーション実験を再び行った。

その結果、知覚符号化と復号化、残響、ノイズ付加などの変形に対して、双方の情報秘匿手法は平均的には十分な耐性を持つことが分かった。しかし、エコー拡散法の方が、多くの条件において、全 100 曲に対する 10 パーセント検出率および最低検出率が低く、秘匿情報検出が困難となる楽曲が存在することが分かった。言い換えると、振幅変調法の方が、あらゆる楽曲に対して適用できる性能を備えていることが分かった。

また、音声信号に対して 48 bps の情報を秘匿し、スピーカ再生を前提として、残響と環境背景雑音が付加される環境をシミュレーションし、双方の手法を比較した。その結果、振幅変調法は最も検出性能が低い妨害条件においても、半分以上の音声信号で 90% 以上の検出率を示す一方、エコー拡散法は検出率が最も高い妨害条件でも検出率 90% に達することはなかった。よって、音声信号に多くのデータを埋め込んでスピーカ再生し、マイクロホン受信した音からデータを検出して利用するような応用場面においても、振幅変調法の方が適していることが明らかになった。

第 6 章では、振幅変調に基づく音響情報秘匿技術を用いて、スピーカから再生される音に同期して情報を呈示するシステムを提案した。ここでは、埋め込む情報にブロック符号化によるエラー訂正を施し、残響および付加雑音環境下において、エラー訂正の範囲に収まる正検出率をシミュレーション実験により求めた。また、カラオケの伴奏音楽に歌詞を同期呈示するための情報を埋め込んでおき、スピーカ再生される伴奏音楽に同期してリアルタイムに歌詞を表示させるシステムを構築した。そして、情報の同期呈示に必要な埋め込み情報検出における時間精度と、埋め込みデータ検出への残響や歌唱音の影響も調べた。

その結果、伴奏音楽に対して背景雑音および振幅クリッピング歪あるいは残響の重畳を経ても、音楽信号に埋め込まれたデータは十分検出可能であることが分かった。また、SNR 0 dB で歌唱音が付加されても、90% 以上の楽曲の、90% 以上の歌唱音を含むフレームでデータ検出が可能なが分かった。また、歌詞表示の時間制御の基となるデータフレーム境界時刻の検出精度も十分であることが分かった。PEAQ を用いた客観音質劣化評価の結果、伴奏音楽の音質劣化は MP3 48 kbps/ch と 64 kbps/ch で符号化された音楽信号の中間程度であることが分かった。

第 7 章では、本研究の成果と今後の課題をまとめた。

本研究で提案した振幅変調に基づく音響情報秘匿技術に対する今後の課題は、各章の考察において示した。すなわち、電子透かし用途としては、知覚検知閾に基づいた変調強度設定、ピッチ変換および差分攻撃に対する耐性の向上である。また、空間伝搬用途においては、情報埋め込み実時間ソフトウェアの実装、検出処理負荷の低減と検出ソフトウェアのPDA等携帯機器への実装である。この振幅変調に基づく音響情報秘匿技術は、電子透かしから空間伝搬のステガノグラフィまで、様々な場面において利用可能であることを本論文は示してきた。これら今後の課題を解決することによって、より一層利用可能な範囲が広がり、実用に繋がることが期待される。

さらに、音響情報秘匿技術全般に対する今後の課題は、技術の新たな利用場面の提案とその実用化であると言えよう。情報秘匿技術の利用は、著作権保護/管理や秘匿通信といった従来重要視されてきた分野だけに限られる訳ではない。本論文では、空間伝搬を前提とした応用、すなわちアナウンス音声やBGMを用いた聴覚障害者や外国人を含む利用者への情報提供、カラオケ歌詞表示などの利用場면을提案してきた。そのような本論文で試みてきた以外の音響情報秘匿技術の利用としては、既にいくつかの技術と新しい利用場面が提案されている。それらは、音響データ通信路におけるパケット損失の隠蔽や高品質化 [7, 4, 73, 6]、メタデータの記録と検索 [8]、改ざん検出 [74] などを含む情報法科学的利用などである。本研究やさらに新たな研究における音響情報秘匿技術は、利用場面の要求に合った性能を満たすだけでなく、その実用化を通じて社会の利益に貢献する方策も積極的に探っていくべきであろう。