

音信号への蝸牛遅延特性に基づく 情報ハイディング法に関する検討

小杉 敏三 (0910023)

北陸先端科学技術大学院大学 情報科学研究科

2011 年 2 月 8 日

キーワード: 蝸牛遅延特性, 並列型構成, 縦続型構成, 複合型構成, 音声ステガノグラフィ.

1 はじめに

近年, 情報通信技術の発達やインフラ整備が急速に進められたことにより, インターネットを利用したマルチメディアデジタルコンテンツ(テキスト, 音, 静止・動画像)の利用が盛んになっている. 静止・動画像の利用はもちろんのこと, 最近では音楽コンテンツや音声通信 (VoIP) の利用が急激に伸びている. そのため, インターネット上の音情報伝送の重要性が非常に高まっており, その利用の安全性を高めるために, コンテンツ保護(著作権保護など)や機密情報保護の中核技術が求められている. マルチメディア情報ハイディング (MIH) 技術は, これらの中核技術を担う, 暗号技術とは異なる情報保護技術として最近注目を集めている.

この MIH 技術を完全なものとして確立するためには, 次の四つの要求項目を満たす必要がある. (1) 埋め込み情報が利用者に知覚されず, 埋め込みによる原音の知覚可能な歪みが一切生じないこと(検知不可能性). (2) 情報が埋め込まれていることを利用者に気づかせないこと(秘匿性). (3) 通常の信号変換処理や悪意のある攻撃に対して影響を受けないこと(頑健性). (4) 埋め込み可能な情報を増やし, 汎用性を高めること(埋め込み情報量). 従来手法では, これら四つの要求項目を全て満たす手法はない. 一方, Unoki & Hamada が提案した蝸牛遅延特性に基づく情報ハイディング法は, 四つの要求項目の内検知不可能性と頑健性を満たす手法である. 本研究では, 検知不可能性と頑健性を満たす蝸牛遅延特性に基づく情報ハイディング法を用いて四つ目の要求項目である埋め込み情報量を満たすため, 蝸牛遅延フィルタの新しい構成法を提案する.

2 多段複合型フィルタ構成

蝸牛遅延を模擬した1次IIRフィルタを蝸牛遅延フィルタと呼ぶ．蝸牛遅延フィルタを2つ用いるフィルタ構成を基本型とし，蝸牛遅延フィルタを並列に構成するものを並列型，蝸牛遅延フィルタを縦続に構成するものを縦続型とする．多段複合型構成は，縦続型の各フィルタを並列型に拡張することで，1フレーム当たり複数ビットの割り当てを可能とする．多段複合型で構成した蝸牛遅延フィルタ $H_{\text{Cmp}}(z)$ とそのフィルタの群遅延 $\tau_{\text{Cmp}}(\omega)$ は，

$$H_{\text{Cmp}}(z) := \prod_{\ell=1}^L H_{\ell,m}(z) = \prod_{\ell=1}^L \frac{-b_{\ell,m} + z^{-1}}{1 - b_{\ell,m}z^{-1}} \quad (1)$$

$$\tau_{\ell,m}(\omega) = -\frac{d \arg(H_{\ell,m}(e^{j\omega}))}{d\omega} \quad (2)$$

$$\tau_{\text{Cmp}}(\omega) = \sum_{\ell=1}^L \tau_{\ell,m}(\omega) \quad (3)$$

となる．ただし， $\ell = 1, 2, \dots, L$ と $m = 0, 1, \dots, M-1$ である．ここで， $\ell = 1$ と固定し， $m = 0, 1, \dots, M-1$ とした場合が並列型構成であり， $\ell = 1, 2, \dots, L$ とし， $m = 0, 1$ とした場合が縦続型構成となる．これらを用いて情報の埋め込みを実現している．また，原信号と埋め込み信号の位相スペクトルの差分を取ることで情報の検出（ノンブラインド検出）を行っている．

3 客観評価結果

並列型，縦続型，複合型構成を提案し，音楽信号と音声信号における客観評価実験を行った．音楽信号における検知不可能性に関する評価結果より，現状提案している多段フィルタ構成の中では， $L = 1, N = 2$ と $L = 2, N = 1$ が一番最適な組み合わせだと言える．このフィルタ構成における最大埋め込み情報量は，1024 bps（512 fps の時1フレームで2 bits 表現）であった．音声信号における検知不可能性に関する評価結果より，現状提案している多段フィルタ構成の中では， $L = 1, N = 2$ が一番最適な組み合わせだと言える．このフィルタ構成における最大埋め込み情報量は，512 bps（256 fps の時1フレームで2 bits 表現）であった．次に，音楽信号と音声信号における頑健性に関する客観評価実験を行った．音楽信号における頑健性に関する評価結果より，現状提案している多段フィルタ構成の中では， $L = 2, N = 2$ が一番最適な組み合わせだと言える．このフィルタ構成における最大埋め込み情報量は，256 bps（64 fps の時1フレームで4 bits 表現）であった．音声信号における頑健性に関する評価結果より， $L = 1, N = 3$ が最良であることがわかる．このフィルタ構成における最大埋め込み情報量は，384 bps（128 fps の時1フレームで3 bits 表現）であった．さらに，音楽信号における各フィルタ構成の頑健性に関する客観評価として，StirMark ベンチマークテストを行った．StirMark ベンチマークテストの結果から，提案法は雑音を加えられる処理，振幅操作，ビット処理，フィルタリング処

理に関しては頑健性が保証できることがわかった．しかし，データ置換処理，位相操作，残響を加えられる処理に対しては頑健性が保証できないことがわかった．

提案法の応用として，音声ステガノグラフィーとして利用可能であるかを実際に 32×32 bitmap 画像を埋め込み，検出を行った．画像情報を埋め込んだステゴデータから画像を検出した結果，ビット検出率は 97.5% であり，26 bit の誤りが生じた．画像情報を埋め込んだ音声信号を PCM コーデックを用いて符号化・複合化してから画像を検出した結果，ビット検出率は 90.9% であり，93 bit の誤りが生じた．検知不可能性だけではなく頑健性も非常に重要であり，この二つを満たす提案法は音声ステガノグラフィーとして利用できることがわかった．

4 おわりに

本稿では，検知不可能性と頑健性を満たす蝸牛遅延特性に基づく情報ハイディング法を用いて四つ目の要求項目である埋め込み情報量を満たすため，蝸牛遅延フィルタの構成法を提案した．検知不可能性と頑健性に関する客観評価結果より，検知不可能性と頑健性を満たした上で埋め込み情報量を高めることができるとわかった．また，音声ステガノグラフィーの検討より，音声ステガノグラフィーとしても利用可能であることがわかり，マルチメディア情報ハイディング技術として，有効な手法であることがわかった．