離散コサイン変換を利用した音声ステガノグラフィ

脇山 正博・栗谷 康隆*・日高 康展・河口 英二**

A Prototype of Audio Steganography System Using Discrete Cosine Transform

Masahiro WAKIYAMA, Yasutaka KURIYA*, Yasunobu HITAKA and Eiji Kawaguchi**

Abstract

In recent years, the broadband is spreading quickly, and various troubles concerning leaks of personal information are increasing. For the purpose of resolving them, we study an information hiding system using steganography. It is a technique to prevent people from noticing the existence of secret data itself. Our research objective is to make experimental system that protects massive secret data by embedding it into dummy data. We propose a new embedding method which covers secret data into Wave data using a DCT method. We designed embedding/extracting algorithms and developed them. We experimented with 200 examinees by listening audio data for the purpose of researching the audio distortion tendency by DCT embedding parameter table, Consequently, we found a rough tendency of audio distortion by DCT embedding parameter table and a potentiality for some improvement.

Keywords : *data hiding, Audio Steganography, Embedding, DCT, WAVE, GUI, DLL*

1. はじめに

近年の情報技術の急速な進歩に伴い，様々な業務のオンライン化が進んでいる．しかし，個人情報などの秘密情報が外部に流出するといった問題が増加している．秘密情報の防衛手段としては，管理方法の厳格化や情報の暗号化が用いられている．

本稿では情報ハインディングの方法として，音声メディアへのステガノグラフィを用いた手法を提唱する[1]．この手法を用いて情報が第三者に流出しても情報の存在自体に気づかれないようにすることを目的とする．

本稿では，これまであまり研究されていない大容量の情報秘匿を目的とし，秘匿媒体として非圧縮デジタル音声メディア，秘匿手法として DCT または DWT 周波数領域への埋め込みを採用したステガノグラフィシステムを設計・試作し，実験による評価を行う．第 2 章ではステガノグラフィ技術について述べ，第 3 章および第 4 章ではシステムの設計について述べる．第 5 章では具体的なソフトウェアの開発手法について述べ，第 6 章では試作システムの評価実験と考察を行い，第 7 章で結論を述べる．

2. 秘密情報伝送システム

ディジタル情報の伝送手段として今日最もよく利用されているインターネットは，その性質として第三者による盗聴を防ぐことが難しい．そのため，秘密情報を安全に伝達するには秘密情報の防衛技術が必須となる．ステガノグラフィを利用した秘密情報伝送システムの概略を，図 1 に示す．秘密情報の送信者は日常ありふれた音声情報に対して，その内容を人間には識別できない程度に改変することにより，秘密情報を埋め込み伝送する．カバーデータに秘密情報の埋め込み処理を施したものを，ステゴデータと呼ぶ．受信者は予め取り決められた手順によってステゴデータより秘密情報を復元する．今日のインターネットではマルチメディア情報が無数にやりとりされている．そのため，暗号文とは違いステガノグラフィ処理の行われたファイルは見た目には通常マルチメディアファイルと同等のため，秘密情報が隠蔽されている情報を発見すること自体が困難となる．

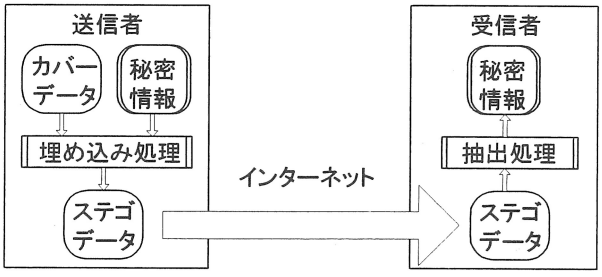

図 1　秘密情報伝送システム

3. 音声ステガノグラフィ

音声ステガノグラフィの概要を図2に示す．音声ステガノグラフィは，埋め込み機能と復号機能からなる．カバーデータにはWindowsプラットフォーム上で用いられるWAVE音声データを用いる．

埋め込み機能における入力はカバーデータとして WAVE 音声ファイル，秘密情報として任意形式の単一ファイル，埋め込みの挙動を決定する埋め込みパラメータファイルからなる．出力(ステゴデータ)として秘密情報入り WAVE ファイルが得られる．

復号機能ではステゴデータと埋め込みパラメータファイルを入力とし，埋め込みと逆の手順によって秘密情報を

* 九州工業大学大学院生命体工学研究科脳情報専攻 2 年
**九州工業大学名誉教授

復号する．埋め込みパラメータファイルとは，埋め込み処理の挙動を決定するパラメータをまとめたファイルであり，復号には埋め込み時と同じものを必要とする．そのため鍵としての役割も持つ．埋め込みパラメータとは，埋め込み処理の挙動を決定する情報であり，復号に必要な鍵の役割を持つ．

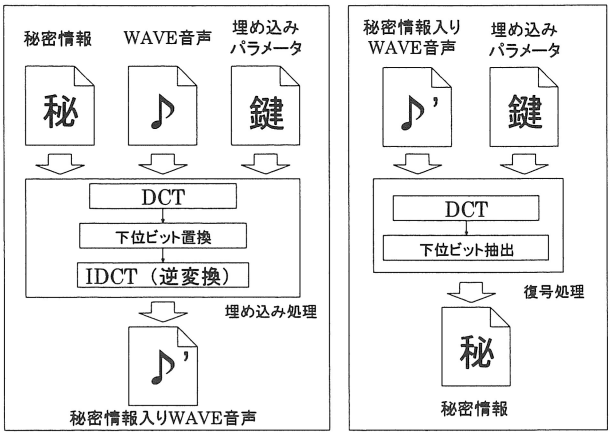

図2　音声ステガノグラフィ

4. 離散コサイン変換を用いた埋め込み・抽出

4.1　離散コサイン変換(DCT)

DCTとは信号処理で広く用いられる直交変換の一種である．本稿ではDCTとして式(1)を用いる．要素数Nの数列xを入力として式(1)を適用すると，出力としてxに唯一対応する周波数成分Xが得られる．周波数成分Xはxと同様に要素数Nの数列である．式(2)は式(1)に対する逆変換であり，IDCT(離散コサイン逆変換)と呼ぶ．

IDCTをXに対して適用することによりxが得られる．$X[0]$は直流分であり，$X[k]$はkが小さいほど低域，大きいほど高域となる．要素数Nは2以上の任意の自然数である．

$$X[k]=\begin{cases}\dfrac{\sqrt{2}}{N}\displaystyle\sum_{n=0}^{N-1}x[n] & (k=0)\\ \dfrac{2}{N}\displaystyle\sum_{n=0}^{N-1}x[n]\cos\left[\dfrac{(2n+1)k\pi}{2N}\right] & (k\neq 0)\end{cases} \quad \cdots(1)$$

$$x[k]=\frac{1}{\sqrt{2}}X[0]+\sum_{n=1}^{N-1}X[n]\cos\left[\frac{(2k+1)n\pi}{2N}\right] \quad \cdots(2)$$

4.2　DCT を用いた埋め込み

一般に，DCTによって得られた周波数成分を少量変化させてからIDCTを適用しても，元の情報はほとんど変化しないことが知られており，画像圧縮などに応用されている．

本稿で提唱する秘密情報埋め込みの手順を図3に，復号の手順を図4 示す．入力信号として音声情報xにDCTを適用し，得られた周波数成分Xに対して下位ビット置換を適用することによって秘密情報ビット列Sの埋め込みを行う．

秘密情報が埋め込まれた周波数成分をX'と定義する．X'にIDCTを適用し音声情報の形に変換し出力信号とする．ここで埋め込みパラメータとは，要素数Nおよび埋め込みビット数テーブルDである．Dについては4.3節で述べる．

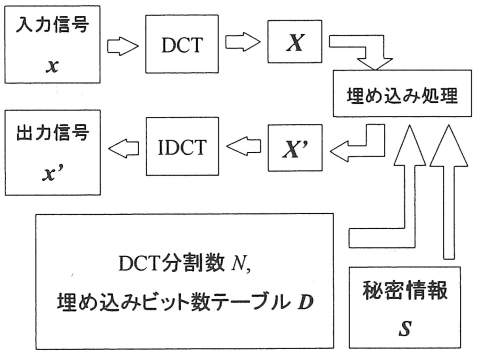

図3　埋め込み処理

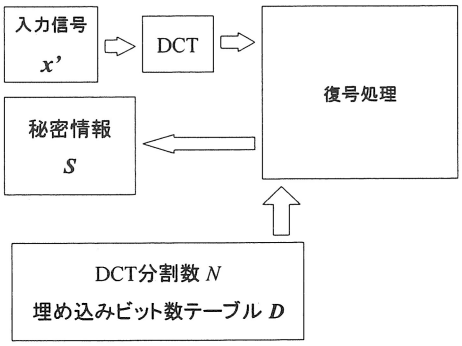

図4　複号処理

4.3　埋め込みビット数テーブル D

Xの各要素に埋め込む秘密情報ビット深度を数列Dと表す．Dは要素数Nの数列であり，$D[k]$ は$X[k]$ に埋め込む秘密情報ビット深度である．数列Dを用いた下位ビット置換を図5に示す．要素数Nおよび数列Dは埋め込みによる音声変化の周波数特性を決定する重要なパラメータとなるため，これらをテキストファイルにしたものを埋め込みパラメータファイルとして用いる．また，復号処理には埋め込み処理で用いたものと同じ埋め込みパラメータが必要となるため，埋め込みパラメータは鍵としての役割も持つ．Dは音声変化の挙動を決定する重要な埋め込みパラメータとなる．

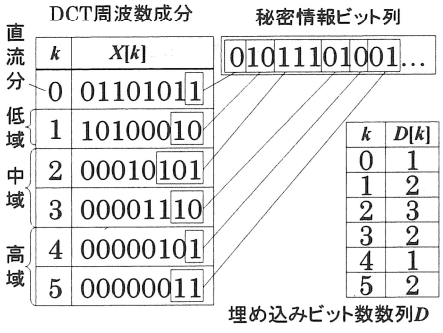

図5 DCT周波数領域への下位ビット置換(N =6の場合)

5．システム設計

5.1 処理方法

システム全体のフローチャートを図6および図7に，システム全体で用いられる変数の一覧を表1に示す．太線で囲まれた処理の詳細は次節で解説する．

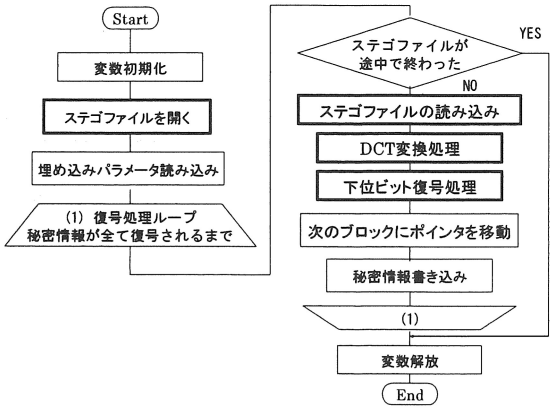

図6　埋め込み機能全体のフローチャート

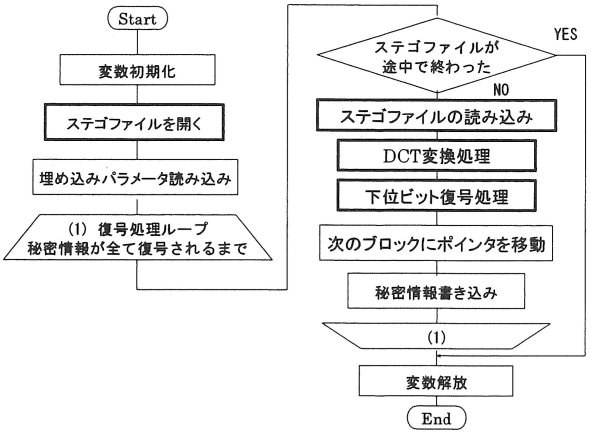

図7　復号機能全体のフローチャート

5.2 入出力データおよび変数定義

埋め込みパラメータファイルはデータ要素がカンマまたは改行で区切られた単純なテキストファイルであり，テキストエディタや表計算ソフトウェアで編集を行う．ファイルの先頭に要素数 N を書き，埋め込みビット深度 D の要素を順番に N 個書く．表1に使用した変数名を記す．

5.3 WAVE ファイルの分割処理

WAVE ファイルは非常に大きなサイズをとることがあるため，適当なサイズのバッファを用意して順次分割処理を行う．音声チャンネル数 nChannels が2以上(ステレオ音声など)の場合，音声データは同時刻に再生すべきサンプル値がチャンネル順に並んでいる．そのため，ファイルから読み込まれた音声データを blockBuffer に一時格納し，

表1　変数表

(a)　入出力関係

変数名	解説
N	埋め込みパラメータ： DCT 変換要素数 N
D	埋め込みパラメータ： 埋め込みビット深度数列 D
fpSecret	秘密情報ファイルへのポインタ
lenSecret	秘密情報ファイルのファイル長
pSecret	秘密情報ファイルの現読み込み位置
fpCover	カバーファイルへのポインタ
lenCover	音声データのファイル長
pCover	音声データの読み込み位置

(b)　WAVE ファイルヘッダー

変数名	解説
nBitsPerSample	1サンプルあたりのビット数
nBytesPerSample	1サンプルあたりのバイト数
nSamplesPerSec	1秒あたりのサンプル数
nChannels	音声チャンネル数

(c)　一時変数およびバッファ

変数名	解説
blockBuffer	入出力する音声データを格納するバッファ
BLOCKSIZE	blockBuffer の要素数
ptrD	blockBuffer の現在位置
ptrChannel	現在処理中のチャンネル番号を示す
secretBuffer	入出力する秘密情報を一時格納するバッファ
SIZES	secretBuffer の要素数
ptrS	secretBuffer の現在位置
ptrSb	ptrS のビット単位の現在位置(0~7)
sndBuffer	DCT/DWT 処理する音声データ
dctBuffer	DCT/DWT 変換された sndBuffer
k, j	汎用ループカウンター
C	バイナリ処理用一時変数
Ptr	ポインタ処理用一時変数
phase	復号処理用フラグ

同チャンネルのサンプル値だけを blockBuffer から sndBuffer に取り出して DCT 変換を行う．取り出すチャンネルのカウントは変数 ptrChannel を用いて行う．blockBuffer のデータサイズは，サンプル数 N に nChannels と1サンプルあたりのバイト数 nBytesPerSample を乗算した値を変数 BLOCKSIZE に格納して用いる．

5.4 DCT・IDCT 処理

DCT および IDCT 処理はサブルーチンとして定義され，変数は外部から独立したものを使用する．DCT 処理での入力変数は sndBuffer，出力変数は dctBuffer に格納される．IDCT 処理では入力が dctBuffer で出力が sndBuffer となる．

5.5 下位ビット置換・復号処理

DCT 処理により出力された変数配列 dctBuffer に対して下位ビット埋め込み・復号処理を行う．処理中に秘密情報

ポインタ ptrS が秘密情報バッファサイズ SIZES に達すると処理を中断し，バッファを更新する．ptrSb は秘密情報のビット単位のポインタで，初期値は 0 である．1 ビット埋め込むごとに 1 加算され，8 に達すると ptrS を 1 増加させて ptrSb は 0 に戻る．また，秘密情報を全て埋め込みまたは復号した時点で処理を終了する．

復号処理では埋め込まれた秘密情報データサイズが必要となるため，埋め込み処理ループ開始前に配列 secret の先頭 4 バイトに秘密情報サイズ lenS を挿入しておく．

復号処理では変数 phase を用いて処理段階の管理を行っている．初期段階である phase:0 では秘密情報サイズ lenS を取り出し，4 バイトを取り出した時点で phase:1（秘密情報の復号）に移行する．秘密情報はバッファ secretBuffer に格納され，バッファが埋まるとファイルに順次出力される．

6．システム評価実験

6.1　試作システム

本稿では試作システムを Windows プラットフォーム上の C 言語で開発し，Visual C++により実行形式および DLL(Dynamic Link Library)形式にコンパイルする．C 言語で開発しているため，Linux やマイコンなどの他のプラットフォームで広く利用可能である．また，1 つのソースを複数の研究生で引継ぎ，拡張されることを考慮し，できるだけ単純な文法で構成される．しかしながら，利便性の代償として動作速度は若干低下する．

6.2　ＧＵＩ試作システム画面

GUI とは操作を視覚的に行うインタフェースである．GUI によりアプリケーションの操作が容易となる．本稿では図 8 に示すような Windows プラットフォーム上で実行可能な GUI を開発した．ファイル指定操作が容易に行うことが可能で，ファイルや埋め込みに関する情報が表示される．実験などで何度も処理を行う場合，命令の入力操作時間を大幅に短縮可能である．

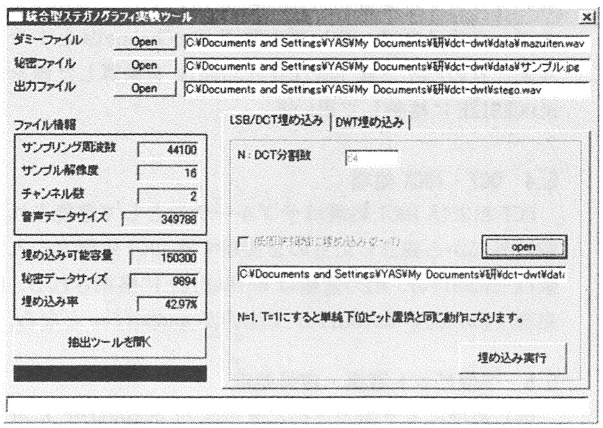

図 8　GUI を用いた埋め込み・復号アプリケーション

6.3　実験方法

要素数N=64として，埋め込みビット数数列Dを表2 に示す条件でそれぞれ埋め込みを行った音声を用意し，被験者に埋め込み前後の音声を順不同に聞き比べさせ，「どちらの音声に秘密情報が入っているか？」というクイズ形式で調査を行った．被験者は約200名である．この調査方式では，埋め込みによる音声変化が知覚しにくければ50%に近い値となり，知覚しやすければ100%に近い値となることが期待できる．

表2　実験条件

条件名	埋め込みビット数列D
[1] 低域のみ7ビット	$D[k]=\begin{cases}7 & (1\le K\le 21)\\ 0 & (elsewhere)\end{cases}$
[2] 中域のみ7ビット	$D[k]=\begin{cases}7 & (22\le K\le 42)\\ 0 & (elsewhere)\end{cases}$
[3] 高域のみ7ビット	$D[k]=\begin{cases}7 & (43\le K\le 63)\\ 0 & (elsewhere)\end{cases}$
[4] 低域のみ6ビット	$D[k]=\begin{cases}6 & (1\le K\le 21)\\ 0 & (elsewhere)\end{cases}$
[5] 中域のみ6ビット	$D[k]=\begin{cases}6 & (22\le K\le 42)\\ 0 & (elsewhere)\end{cases}$
[6] 高域のみ6ビット	$D[k]=\begin{cases}6 & (43\le K\le 63)\\ 0 & (elsewhere)\end{cases}$

6.4　実験結果・考察

実験結果を表 3 に示す．実験結果より，低域へ埋め込んだ場合が最も正答率が低く，中域に埋め込んだ場合が最も高かった．これにより，埋め込みビット数数列 D は低域を最も大きく，中域を最も小さく，高域を中間程度とすることで音声変化を知覚しにくい埋め込みが可能となると考えられる．

表3　実験結果

条件名	正答率
[1] 低域のみ7ビット	67%
[2] 中域のみ7ビット	100%
[3] 高域のみ7ビット	90%
[4] 低域のみ6ビット	50%
[5] 中域のみ6ビット	82%
[6] 高域のみ6ビット	80%

7．おわりに

DCTを用いたディジタル音声ステガノグラフィシステムの試作を行った．埋め込みパラメータによる音声変化特性の検証実験を行い，パラメータの傾向を得た．今後は，カバー音声の内容に応じたパラメータの決定方法を考案する予定である

参考文献

[1] Invitation to BPCS-Steganography:
http://www.datahide.com/BPCSj/

(2008年10月10日　受理)