

埋込み情報の組合せ符号化法による情報ハイディング

山本 紘太郎

岩切 宗利

g44034@nda.ac.jp iwak@nda.ac.jp

防衛大学校 情報工学科

239-8686 横須賀市走水 1-10-20

あらまし 近年のデジタルコンテンツの普及に伴い、コンテンツに情報を密かに隠す情報ハイディング技術が注目されている。情報ハイディングは、コンテンツ保護のための電子透かしや、通信秘匿のためのステガノグラフィに応用される。現在の情報ハイディング技術では、埋込み情報量の増大に伴い、コンテンツに対する制御量が増加する。これは品質の低下や符号量の増大という問題の原因になり得る。本報告では、情報埋込みの効率を向上させるため、埋込み制御に選択肢を与え、その組合せを用いた組合せ符号化法について提案する。本方式を用いることによって、コンテンツへの制御量を局限しつつ、従来より多量の情報を埋め込むことが可能になる。

A Consideration of Combination Coding for Information Hiding

Kotaro YAMAMOTO

Munetoshi IWAKIRI

g44034@nda.ac.jp iwak@nda.ac.jp

National Defense Academy

1-10-20, Hashirimizu, Yokosuka-shi, Kanagawa-pref 239-8686, Japan

Abstract This report presents a combination coding method for information hiding. Information hiding is an important technology for digital contents. This is used to digital watermark and steganography. If embedding secret messages to one contents, the contents must be controlled (rewrites a local field, or inserts another information). In previous method, increase of embedding data involves increase of control to contents. To use combination coding method, the problem can solve. Combination coding uses signal-pattern of embedding, and improves efficiency of embedding.

1. はじめに

情報化社会の本格的な到来により、現在、さまざまなマルチメディアコンテンツがネットワークや大容量の記憶媒体を介して利用されている。それに伴い、情報ハイディングと呼ばれる技術が研究されている。これは、あるコンテンツに対し、密かに別の情報を隠す技術の総称である。情報ハイディング技術は、主に2つの応用分野で用いられている。1つはコンテンツ保護のた

めの電子透かし [1]、もう1つはコンテンツを媒体とした通信秘匿（ステガノグラフィ [2]）である。

情報ハイディングをこれらの分野に用いるときには、それぞれに要求される特性が異なる。特に、埋込み可能な情報量と、埋込みがコンテンツにおよぼす影響の関係は、重要な特性であり、従来からさまざまな方式で検討されてきた [1, 3, 4, 5, 6]。

これらの検討は、1つの埋込み領域を処理単

位としているため、埋込み量に比例してコンテンツが劣化する。

本報告では、埋込み処理の選択肢それぞれに情報を対応させ、高い埋込み効率を実現する組合せ符号化法を提案する。

第2章では、従来の埋込み情報の表現法について示し、その埋込み効率について考察する。第3章では、組合せ符号化法を提案し、第4章にその理論特性を示す。また、提案方式をXML情報ハイディング方式に応用した一例を第5章に示す。

2. 従来方式

2.1 情報表現法

一般的な情報ハイディングは、動画、音声、画像、電子文書などをカバーデータに用いる。これらの情報ハイディング技術は、各種メディアの特性に応じて、埋込み領域および情報挿入法を決定する。たとえば、画像データを対象とした情報ハイディングとしては、各画素値の下位ビットを埋込み情報に置き換える方法や、圧縮符号列を変換する方法などがある [1, 2]。ほかにも、電子文書を対象とした情報ハイディング技術として、XML[7]などの構造化文書の表記法に冗長性を見出す方法、自然言語の表現の置換により情報を表現する方法などがある [8]。

これらの方式は、通常、埋込み情報と埋込み処理が一对一対応である。

2.2 埋込み制御量とその効率

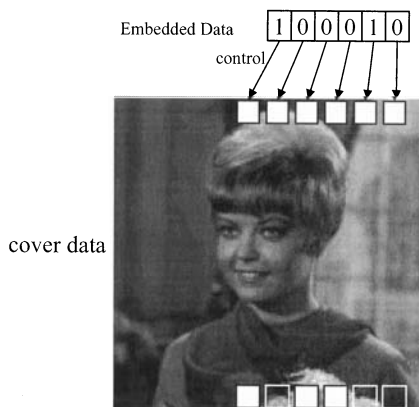
情報の埋め込みに必要なコンテンツの変化量を総制御量 C とし、そのときの埋込み情報量を P とすると、総制御量に対する情報埋込みの効率は、

$$E = \frac{P}{C} \quad (1)$$

になる。

1つの情報埋込みに制御量 a が必要ならば、従来手法による R 個の情報埋込みに、 $a \times R$ の総制御量が必要である。すなわち、埋込み情

(a) previous method



(b) proposal method

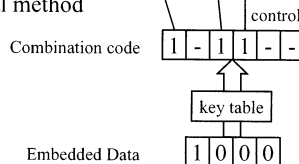


図1: 組合せ符号化法の概要

報量と総制御量が線形関係にある。したがって、制御量当りの情報挿入量（埋込み効率）は、

$$E = \frac{R}{a \times R} = \frac{1}{a} \quad (2)$$

である。

3. 提案方式

図1に示した提案方式では、複数領域の埋込み制御パターン組合せを1つの符号とみなし、それらに情報を対応付けることにより情報ハイディングを実現したものである。ここでは、本提案方式の基本的なアイデアとその実現法について示す。

3.1 組合せ符号による情報表現

情報埋込みに利用できる領域数（以下、領域数と呼ぶ）を F とし、実際に情報を埋め込む領域数（以下、制御数と呼ぶ）を r とする。このとき、その制御パターン数は、 ${}_r C_F$ である。提

表 1: key table の一例

Field size	Control	Combination code	⇔	Data
2	1	1 -	⇔	0
		- 1	⇔	1
4	2	-- 1 1	⇔	00
		- 1 - 1	⇔	01
		- 1 1 -	⇔	10
		1 - - 1	⇔	11
		1 - 1 -	⇔	unused
		1 1 - -	⇔	unused
5	1	---- 1	⇔	00
		--- 1 -	⇔	01
		-- 1 - -	⇔	10
		- 1 - - -	⇔	11
		1 - - - -	⇔	unused

案方式では、この選択肢を利用し、 $\lceil \log_2 F C_r \rceil$ ビットの情報を表現する。すなわち、本方式は、組合せそれぞれに情報を対応させ、その種類数に応じた量の情報埋込みを実現するものである。ただし、埋め込まれた情報の読み取りには、各パターンから情報への変換が必要である。

3.2 key table

本研究では、システムを単純化するために、各制御パターンと情報の対応をあらかじめ key table (変換表) として共有することを前提にした。表 1 の Combination code (組合せ符号) “- 1 1” は、4 つある領域のうち前 2 つを無視し、後半のみ制御することを示している。また、これに対応する Data (埋込み情報) は、“00” である。

3.3 組合せ符号の構成

提案方式による組合せ符号の生成手順は、次のとおりである。

- (1) カバーデータから埋込み可能領域数 F を求め、それと、実際に埋込み制御する領域数 r の関係から、埋込み量

$$M = \lceil \log_2 F C_r \rceil \quad [\text{bit}] \quad (3)$$

を求める。

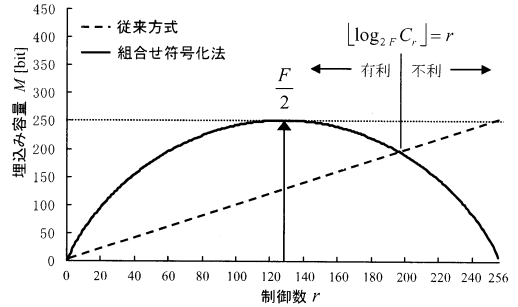


図 2: 制御数と埋込み量 ($F = 256$)

- (2) 埋込み情報列 (Data) から M ビット取出し、領域数 F (Field size) と制御数 r (Control) の対応から key table を調べ、組合せ符号 (Combination code) を決定する。

4. 提案方式の理論的特性

4.1 埋込み容量

図 2 は、 $F = 256$ の条件で従来方式と提案方式の埋込み量と制御数を比較したものである。これによると、提案方式は、 $r = F/2$ のときに埋込み容量が最大になる。また、同一制御数の埋込みでは、 $r \leq \lceil \log_2 F C_r \rceil$ のとき、提案方式の方が埋込み容量が大きい。

また、埋込み容量を最大にするとき r は、従来方式で $r = F$ であり、提案方式では $r = F/2$ である。図 3 は、このときの F と埋込み容量 M との関係を示したものである。この結果は、従来方式と比べ、提案方式の最大埋込み容量がわずかに劣ることを示している。

よって、埋込み容量を大きくしたい場合は、従来方式を用いる方がよい。ただし、高い埋込み効率を実現したい場合は、 $r = F/2$ で提案方式を利用することが望ましい。

4.2 埋込み効率

F , r が埋込み効率におよぼす影響について調べた。組合せ符号化法の埋込み効率は、式 (1)

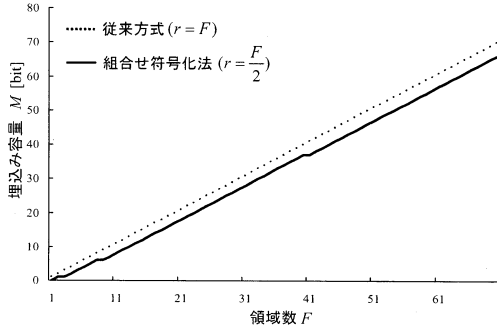


図 3: 領域数と埋込み量

より,

$$E = \frac{P}{C} = \frac{\lfloor \log_2 F C_r \rfloor}{a \times r} \quad (4)$$

と表せる.

1つの情報埋込みに必要な制御量を a とし, $r = 1$ とした場合の F と E の関係を図 4 に, $F = 256$ の場合の r と E の関係を図 5 にそれぞれ示す.

これらの図から, 次のことがわかる.

- (1) 図 4 は, F が大きくなるほど提案方式の埋込み効率が向上することを示している.
- (2) 図 5 は, $r = 1$ のときに提案方式の埋込み効率が最大であり, r が大きくなるにしたがって低下することを示している.

すなわち, 提案方式の埋込み効率は, F が大きく r が小さいほど良くなるといえる. また, 埋込み効率が最大のとき, その埋込み効率は,

$$E = \frac{\lfloor \log_2 F C_1 \rfloor}{a \times 1} = \frac{\lfloor \log_2 F \rfloor}{a} \quad (5)$$

である. このとき埋込み効率は, F の値が持つ情報量に比例する.

5. 情報ハイディングへの応用例

提案方式を実際に情報ハイディングへ応用し, その効果から提案方式を評価する. 本研究では XML 文書をカバーデータとして, XML ステガノグラフィ [8] に提案方式を適用した.

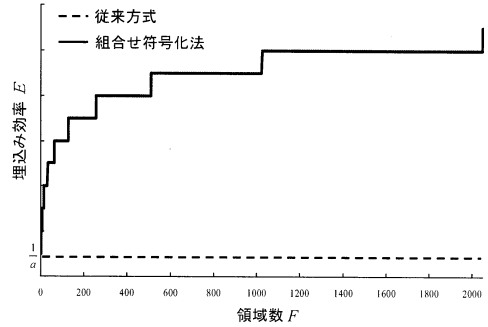


図 4: 領域数 F と埋込み効率 E ($r = 1$)

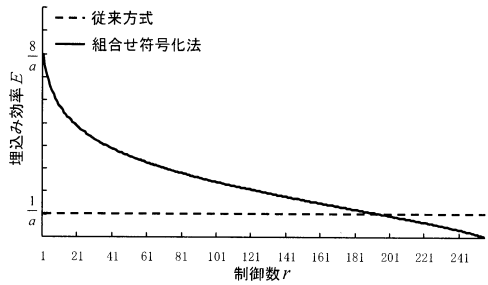


図 5: 領域数 r と埋込み効率 E ($F = 256$)

5.1 XML ステガノグラフィへの適用

文献 [8] では, 複数の XML ステガノグラフィを提案している. 本研究では, 組合せ符号の実装に XML タグ中の空白文字を利用する方式を利用した. この方式は, XML のタグを閉じるブラケット “>” の前に空白 (スペース) を配置できるという表記上の自由度を利用する方式である. すなわち, 閉ブラケット直前のスペースの有無により, 埋込み情報を表現する.

情報の埋込み処理は, 次の通りである.

- (1) XML のタグ数を算出し, あらかじめ定めた領域数 r と, 対応する key table から組合せ符号を構成する.
- (2) 組合せ符号により定まる閉ブラケットの直前にスペースを挿入する.

埋め込んだ情報は, ステゴデータから組合せ符号を抽出し, 共有 key table の対応する値を読み取ることにより復号できる.

表 2: 実験用 SVG データ

sample	orca	hawk	build
File size [byte]	15723	79948	484957
Tags	109	373	944



(a) orca



(b) hawk



(c) build

図 6: 実験用 SVG 画像

5.2 実験結果と考察

本実験では、表 2 に示した 3 種類の実験用カバードータ (SVG 形式の XML ファイル) に埋め込みを施し、組合せ符号化法を評価した。各 SVG の表示状態を図 6 に示す。図 7 は、SVG データ orca の一部を示したものである。

表 3 は、埋め込みビット数を調べた結果である。表 4 は、その埋め込み効率を示したものである。これらの結果から、次の考察を得た。

- (1) 表 3(a) は、同じ制御数でもファイル中のタグ数 (領域数) が多いほど埋め込み容量が増大することを示している。
- (2) 表 3(b) は、埋め込み容量が最大となる $r =$

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- Created with Inkscape (http://www.inkscape.org/)
--><svg height="297mm" id="svg1"
inkscape:version="0.41+cv5"
sodipodi:docbase="/home/matthew/media/image/my_pictures/svg"
sodipodi:docname="orca.svg" sodipodi:version="0.32"
width="210mm" xmlns="http://www.w3.org/2000/svg"
xmlns:cc="http://web.resource.org/cc/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:inkscape="http://www.inkscape.org/namespaces/inkscape"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:sodipodi="http://inkscape.sourceforge.net/DTD/sodipod
i-0.dtd" xmlns:svg="http://www.w3.org/2000/svg"
xmlns:xlink="http://www.w3.org/1999/xlink">
<metadata>
<rdf:RDF xmlns:cc="http://web.resource.org/cc/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" />
</work rdf:about="">
```

図 7: orca の XML 表示状態

表 3: 組合せ符号化法の埋込み結果 [bit]

(a) $r = 10, 20, 30$

r	proposal method			previous method
	10	20	30	
orca	45	71	88	109
hawk	63	109	146	373
build	76	136	188	944

(b) $r = F/2$

r	proposal method			previous method
	54	186	472	
orca	105	-	-	109
hawk	-	368	-	373
build	-	-	938	944

$F/2$ で埋め込みを施した結果である。この結果は、最大埋め込み容量において従来方式の方がわずかに高くなることを示している。

- (3) 表 4 は、タグ数 (領域数) が多く、制御数が少ないほど埋め込み効率が高いことを示している。また、全ての制御数において、埋め込み効率は向上し、総制御量を抑制できることがわかる。

以上の結果から、提案方式が埋め込み効率の向上に対して有効であり、4 章に示した理論的特性を満たすことを確認できた。

表 4: 組合せ符号化法の埋込み効率

r	proposal method			previous method
	10	20	30	
orca	0.56	0.44	0.37	0.13
hawk	0.79	0.68	0.61	0.13
build	0.95	0.85	0.78	0.13

6. おわりに

本報告では、情報ハイディングの埋込み効率を向上させる組合せ符号化法について提案した。また、XML ステガノグラフィに適用し、提案方式が有効であることを確認した。提案方式を利用すれば、情報ハイディングの埋込み効率を向上し、従来よりも少ない総制御量で多量の情報埋込みを実現できる。

今後、コンテンツの大容量化により領域数はさらに増加していくと考えられる。これは、領域数が大きいほど効率が良くなるという提案方式にとって有利である。また、XML ステガノグラフィへの応用は、大規模なXML データベースに対する未編集符号として活用できる可能性も考えられる。すなわち、本提案の組合せ符号は、複数の領域を用いて1つの情報を表現している。そのため埋込み情報の抽出時には、埋込みに利用した領域全てが存在しており、かつその内容が変化していないことが求められる。よって、組合せ符号の正しい抽出には、領域の完全性が必要である。この特性を利用すれば、コンテンツの原本性保証 [9] のための未編集符号や、フラジャイル電子透かし [10] 等への応用が可能である。

また、本方式は、制御数を制限できるためコンテンツに対する総制御量を少なくできる。この特徴は、ステゴ解析 [11] 対策を実現しやすいことを示している。

今後の課題として、埋込み領域に関する特性を利用した各種応用の検討、さまざまなコンテンツに対する適用が挙げられる。

参考文献

- [1] 松井甲子雄：電子透かしの基礎，森北出版 (1998).
- [2] 情報処理振興事業協会：「インフォメーションハイディングの技術調査」報告書 (1998).
- [3] 本吉史明, 内田理, 中西祥八郎：人間の視覚特性を考慮した Wavelet 変換に基づく電子透かし，2006 年コンピュータセキュリティシンポジウム，7A-3(2006).
- [4] M.Barni, F.Bartolini, and A.Piva：Improved wavelet-based watermarking through pixelwise masking, IEEE Trans.Image Process., Vol.10, No.5, pp.783-791(2001).
- [5] L.Boney, A.H.Tewfik, and K.H.Hamdy：Digital watermarks for audio signals, IEEE Int. Conf. on Multimedia Computing and Systems, pp.473-480(1996).
- [6] 越前功, 吉浦裕, 田口順一, 中野和典, 佐々木良一：動き検出に基づく動画用電子透かしの画質維持方式, 情報処理学会論文誌, Vol.43, No.8, pp.2529-2541(2002).
- [7] JIS X 4159:2005：拡張可能なマーク付け言語 (XML) 1.0(2005).
- [8] 井上信吾, 村瀬一郎, 滝澤修, 松本勉, 中川裕志：XML におけるステガノグラフィ手法の提案，2002 年暗号と情報セキュリティシンポジウム予稿集，pp.301-306(2002).
- [9] 社団法人日本土木工業協会：電子文書原本性に関する調査研究 (2003).
- [10] 石原希実, 阿部公輝：画像認証と改ざん検出可能なセミフラジャイル電子透かしの一手法，2004 年コンピュータセキュリティシンポジウム，2C-5(2004).
- [11] 井上大介, 鈴木雅貴, 岡庭隆文, 松本勉：演奏データファイル SMF に関するステゴ解析，2002 年暗号と情報セキュリティシンポジウム予稿集，Vol.I, pp.313-318(2002).