



Title	DVI-ADPCMに対するロスレスステガノグラフィの一手法
Author(s)	青木, 直史; Aoki, Naofumi
Citation	電子情報通信学会論文誌 A, J93-A(2), 104-106
Issue Date	2010-02-01
Doc URL	<a href="https://hdl.handle.net/2115/47386">https://hdl.handle.net/2115/47386</a>
Rights	copyright © (社) 電子情報通信学会 2010
Type	journal article
File Information	DTR93A-2_104-106.pdf



## DVI-ADPCM に対するロスレスステガノグラフィの一手法

青木 直史<sup>†</sup> (正員)

A Lossless Steganography Technique for DVI-ADPCM

Naofumi AOKI<sup>†</sup>, Member

<sup>†</sup> 北海道大学大学院情報科学研究科, 札幌市

Graduate School of Information Science and Technology,  
Hokkaido University

N14 W9, Kita-ku, Sapporo-shi, 060-0814 Japan

あらまし 本研究は, DVI-ADPCM に対するロスレス方式のステガノグラフィの可能性について検討している. 提案法は DVI-ADPCM が折返し 2 進符号によって音声データを符号化していることに着目したものであり, 折返し 2 進符号の冗長性を利用することで秘匿情報の埋込による音声データの劣化を引き起こさない手法となっている.

キーワード ステガノグラフィ, DVI-ADPCM, 折返し 2 進符号, VoIP

### 1. まえがき

音声データに秘匿情報を埋め込む手法として, G.711 に対するロスレス方式のステガノグラフィがこれまでに提案されている [1]. これは, G.711 が折返し 2 進符号によって音声データを符号化していることに着目したものであり, 折返し 2 進符号の冗長性を利用することで秘匿情報の埋込による音声データの劣化を引き起こさない手法となっている.

同様の手法を適用することで, 本研究は, DVI(Digital Video Interactive)-ADPCM に対するロスレス方式のステガノグラフィの可能性について検討している. DVI-ADPCM は, Windows や Macintosh など現在の標準的なコンピュータ環境がサポートする音声データの基本的なコーデックの一つであり, IP 電話のコーデックとしても利用されている [2].

提案法は DVI-ADPCM が折返し 2 進符号によって音声データを符号化していることに着目したものであり, 折返し 2 進符号の冗長性を利用することで秘匿情報の埋込による音声データの劣化を引き起こさない手法となっている. 本論文では, 提案法の原理について説明した後, 提案法のキャパシティについて述べる.

### 2. DVI-ADPCM における音声データの復号

DVI-ADPCM は, ADPCM(Adaptive Differential Pulse Code Modulation) 方式による音声データのコーデックの一つである. DVI-ADPCM は, 量子化精度

```
d = s >> 2;
if (c & 0x1) d += s >> 1;
if (c & 0x2) d += s;
if (c & 0x4) d = -d;
x += d;
```

図 1 DVI-ADPCM の復号手順 (3bit 符号化)

Fig.1 Decoding Procedure of DVI-ADPCM (3 bit coding).

```
d = s >> 3;
if (c & 0x1) d += s >> 2;
if (c & 0x2) d += s >> 1;
if (c & 0x4) d += s;
if (c & 0x8) d = -d;
x += d;
```

図 2 DVI-ADPCM の復号手順 (4bit 符号化)

Fig.2 Decoding Procedure of DVI-ADPCM (4 bit coding).

```
if (s == 7 && (c & 0x7) == 0)
{
    if (b == 0) c &= 0x7;
    if (b == 1) c |= 0x8;
}
```

図 3 提案法の埋込手順

Fig.3 Embedding procedure of the proposed technique.

が 16bit の音声データを 3bit または 4bit の圧縮データに符号化する 2 種類の方式を規定したものになっている [3].

3bit 符号化の場合, 圧縮データは, 1bit の正負の sign, 2bit の magnitude から構成された 3bit の折返し 2 進符号によって表される. 4bit 符号化の場合, 圧縮データは, 1bit の正負の sign, 3bit の magnitude から構成された 4bit の折返し 2 進符号によって表される.

3bit 符号化と 4bit 符号化の復号手順は, C 言語ではそれぞれ図 1 と図 2 のように表すことができる. ここで, c は圧縮データ, x は 1 時刻前の音声データ, d は差分, s はステップ幅を表している. これらの図に示すように, 1 時刻前の音声データを更新することで得られる現時刻の音声データ x は, 単純なビット演算を利用して計算することができる [3].

### 3. 提案法

DVI-ADPCM のステップ幅は, 3bit 符号化と 4bit

符号化のどちらの場合も、7から32767までの89段階の値をとり得ることが規定されている[3].

4bit符号化の場合、ステップ幅が7となり、かつ圧縮データのmagnitudeが0となる時、差分dは0となるため、圧縮データのsignの値にかかわらず復号結果は等しくなる。すなわち、4bit符号化については、この条件が成立する場合に限り、音声データに劣化を引き起こさずに秘匿情報を埋め込むことができる。

本研究の提案法は、4bit符号化を対象として、図3のように秘匿情報を埋め込むことで、ロスレス方式のステガノグラフィを実現している。ここで、bは0または1の秘匿情報を表している。

なお、3bit符号化についてはこの条件が成立しないため、本研究の提案法と同様の手法によってロスレス方式のステガノグラフィを実現することはできない。

IP電話のコーデックとして利用されるADPCM方式のコーデックには、DVI-ADPCMのほかにITU(International Telecommunication Union)勧告のG.726[4]がある。DVI-ADPCMと同様、G.726は折返し2進符号によって音声データを符号化している。

G.726は、音声データによっては、圧縮データのmagnitudeが0となる時、圧縮データのsignの値にかかわらず復号結果が等しくなる場合がある。そのため、本研究の提案法と同様、折返し2進符号の冗長性を利用してロスレス方式のステガノグラフィを実現できる可能性がある。

#### 4. 提案法のキャパシティ

本研究では、IP電話への適用を念頭において提案法のキャパシティを調べた。評価実験では、本研究のために開発したソフトウェアIP電話を利用し、ヘッドセットを接続したノートパソコンを端末として、標準化周波数8kHz、量子化精度16bitの条件で、実環境における電話音声の収録を行った。

なお、通話時における背景雑音による影響を考慮し、電話音声の収録は、個室と事務室の2箇所で行った。個室は話者1名だけの静かな環境、事務室は周囲がパソコンなどを使ってデスクワークを行っている中で話者が発話している環境となっている。

評価実験では、男性1名と女性1名の話者に参加してもらい、それぞれの環境で2回ずつ、合計8個の音声データの収録を行った。なお、できる限り実際に近い条件で提案法のキャパシティを調べるため、会話の内容は自由とし、それぞれ2分間程度の音声データを収録した。

表1 実環境の音声データ

Table 1 Speech data obtained from actual environment.

	speech	L (s)	R (%)	RMS (dB)
private room	m1	134	51	-56.55
	m2	126	52	-59.66
	f1	124	57	-56.42
	f2	126	59	-55.72
office room	m3	127	48	-49.73
	m4	133	44	-48.59
	f3	123	49	-49.41
	f4	134	54	-48.04

表1に、それぞれの音声データの長さL、発話区間の割合R、非発話区間における背景雑音のRMS(Root Mean Square)を示す。なお、背景雑音のRMSは、量子化精度が16bitの音声データがとり得る最大の振幅である $2^{15}$ を基準としてdBで表示している。

この表に示すように、個室の背景雑音は-55.72~-59.66dBの範囲、事務室の背景雑音は-48.04~-49.73dBの範囲であった。なお、発話区間と非発話区間の分類は、実際に音声データを試聴することで判断した。

表2に、提案法のキャパシティを示す。なお、この表には、比較のために、G.711に対するロスレス方式のステガノグラフィ[1]のキャパシティを併せて示している。結果として、G.711に対するロスレス方式のステガノグラフィよりも、提案法のキャパシティは小さくなる傾向を示すことが分かった。

また、提案法のキャパシティは背景雑音の強度に左右され、背景雑音が小さく抑えられた環境ではある程度のキャパシティが期待できるが、背景雑音が目立つ環境では提案法のキャパシティはほとんど0となってしまうことが分かった。

背景雑音が-55.72~-59.66dBの個室の場合、キャパシティは0.029~0.16%の範囲となり、2.32~12.8bit/sの秘匿情報を埋め込めるが、背景雑音が-48.04~-49.73dBの事務室の場合、キャパシティは0.0012~0.0032%の範囲となり、0.096~0.256bit/sの秘匿情報しか埋め込めないことが分かった。

#### 5. むすび

評価実験の結果から、G.711に対するロスレス方式のステガノグラフィと同様、折返し2進符号の冗長性を利用すると、DVI-ADPCMに対するロスレス方式のステガノグラフィを実現できる可能性があることが分かった。

ただし、評価実験の結果から、背景雑音が小さく

表 2 提案法のキャパシティ  
Table 2 Capacity of the proposed technique

		speech	G.711 (%)	DVI-ADPCM (%)
private room	m1		4.7	0.058
	m2		6.4	0.16
	f1		4.3	0.038
	f2		3.3	0.029
office room	m3		2.5	0.0032
	m4		2.4	0.0013
	f3		1.7	0.0017
	f4		2.4	0.0012

抑えられた環境ではある程度のキャパシティが期待できるが、背景雑音が目立つ環境では提案法のキャパシティはほとんど 0 となってしまうことが分かった。

提案法のキャパシティは背景雑音の強度に左右され、背景雑音が小さく抑えられた無音区間を多く含む音声データでは提案法のキャパシティは大きくなる可能性がある。ただし、こうした音声データの場合、IP 電話のオプション機能である無音圧縮 [5] が適用されると、提案法のキャパシティは小さくなってしまふ可能性がある。

無音圧縮を適用しない場合、提案法のキャパシティを改善するには、背景雑音を軽減する信号処理手法の併用が一つの方策となる可能性がある。今後の課題として検討していきたい。

## 文 献

- [1] N. Aoki, "A technique of lossless steganography for G.711," IEICE Trans. Commun., vol.E90-B, no.11, pp.3271-3273, Nov. 2007.
- [2] RFC1890, RTP Profile for Audio and Video Conferences with Minimal Control, 1996.
- [3] Microsoft, Multimedia Data Standards Update, 1994.
- [4] ITU-T G.726, Recommendation G.726, 1990.
- [5] D.J. Wright, Voice over Packet Networks, John Wiley & Sons, 2001.

(平成 21 年 5 月 6 日受付, 9 月 6 日再受付)