

適応差分 PCM 符号化における音声符号への テキスト情報の埋込み

岩 切 宗 利[†] 松 井 甲 子 雄[†]

この論文では、デジタル音声の符号化方式である適応差分 PCM 方式において音声符号の一部を利用してテキスト情報を密かに伝送する一方法を提案する。その原理は、音声の符号化に際し、近接相関に基づいて得られる適応予測値の極性符号の変動現象に着目して、テキストから生成したランダムビット系列で音声符号を制御しながら密かに文字情報を埋込み伝送するものである。この方法によると文字情報が埋め込まれた伝送符号をそのまま音声として再生しても聴感的にまったく違和感を与えないで済ますことができる。この論文では、数十バイト/秒以上の埋め込まれたテキスト情報の存在を第三者に知られることなく秘匿伝送できることを実験的に示す。この方法は秘匿通信のほかに、音声ソフト等の著作権保護のための電子透かしの埋込みにも適用できるものである。

Embedding a Text into Audio Codes under ADPCM Quantizer

MUNETOSHI IWAKIRI[†] and KINEO MATSUI[†]

In this paper, we propose how to embed a text secretly into digital voice data under Adaptive Differential PCM quantizer. When input voice is sampled and quantized, a difference signal is obtained by subtracting an estimate of the input signal itself. The polarity sign of the estimates often fluctuates between plus and minus according to the input signal. By using this fluctuation, we embed a sequence of random bit of text into the output PCM codes, secretly. The multiplexed signals are audible as good sound for us and thus anyone is not aware that the text is being simultaneously transferred through the voice channel. The amount of the text is estimated about twenty bytes per second in 32 kbit/s ADPCM system. This method would be also available for digital watermarking.

1. はじめに

音声は、意図を伝達する手段として重要な役割を果たしている。従来、音声はアナログ波形を忠実に再現する方法により伝送されてきた。しかし、高性能な DSP の出現で信号の高速処理が可能となり、各種情報端末のデジタル化が進み、音声を伝送または蓄積するためのデジタル音声符号化技術の重要性は高まっている。また、音声ではその情報伝達量に比して冗長度が高く、この削減処理による通信効率の向上のためにはデジタル信号の方が実行容易であり、その立場から数多くの研究成果が報告されている^{1)~5)}。

音声符号化方式は波形符号化方式、ハイブリッド符号化方式および分析合成符号化方式に大別される。波形符号化方式の基本的なものに波形の振幅をサンプリング定理に基づいて数値化し、それを量子化する線形パルス

符号化 (Pulse Code Modulation: PCM) 方式がある。これは、通称 A/D 変換と呼ばれている。これに近接相関に基づいた適応予測と適応量子化を行う手法を組み入れた適応差分 PCM (Adaptive Differential PCM: ADPCM) 符号化方式がある¹⁾。この ADPCM 符号化方式は、ITU-T (International Telecommunication Union Telecommunication Sector) 勧告 G.726¹⁰⁾として標準化されており、現用の通信においても広く採用されている。

ここで、音声情報とテキスト情報を併送することを考えてみよう。一般には、送信に必要となる帯域幅が増大するため通信効率が悪くなる。そこで、音声の冗長性を利用して文字情報を音声符号中に埋め込み伝送する方法が考えられる。この方法では帯域幅を増大させずに、密かに文字情報を併送することができる。このような観点からアナログ通信において、すでに二、三の研究がなされている。Steel⁶⁾は、音声スクランブラの特徴を巧みに利用して文字情報を埋め込み、Wong⁷⁾は、音声の位相を切り替えることによりピッ

[†] 防衛大学校情報工学科

Department of Computer Science, National Defense Academy

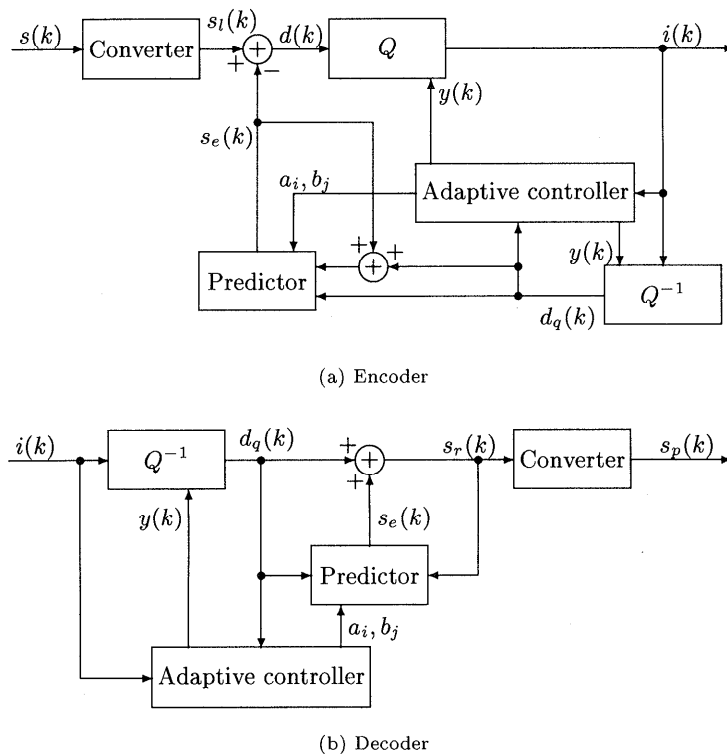


図1 G.726 32 kb/s ADPCM
Fig.1 G.726 32 kb/s ADPCM.

ト情報を埋め込む方法を提案している。松井ら⁸⁾はデジタル音声通信の量子化歪に見せかけて文字情報を埋め込み伝送することを試みている。また、Boney⁹⁾らは音声信号に著作権保護のための特殊信号すなわち、電子透かしを埋め込むことを提案している。

これらの方法では、第三者にその埋込み位置が判明しやすく、埋め込まれたビット系列を解析されテキスト情報を解読される危険性がある。そこで、この論文では国際標準規格であるG.726に対象を限定し、簡単な仕掛けで埋込みビット位置を秘匿する一方法を提案する。

その基本的なアイデアは、ADPCM方式における近接相関に基づいた予測値の同じ極性符号の連続数がランダムに分布することに注目し、それを尺度にしてテキストビットの埋込み位置を選択する方法である。この方法によりテキスト情報を出力符号系列の不特定箇所に分散配置することができるので、その埋込みの存在を隠すことが可能となる。

まず、2章ではADPCM方式を要約する。3章にADPCM符号化過程においてテキスト情報を音声符号に埋め込む方法を述べ、4章にその合成手順と復号手順の詳細を示す。5章では、ITU-T勧告G.726の

32 kb/s ADPCM符号化手順に従って音声符号化システムを構成し、その実験結果を示す。

2. ADPCM符号化方式の概要

ADPCM方式は、適応予測と適応量子化の構造から前向き型 (forward type) と後向き型 (backward type) の2つに分類される¹⁾。前向き型は、音声を固定長のフレームごとに線形予測分析 (LPC) を行い、それによる予測残差信号と予測フィルタの係数を符号化して伝送する。また、後向き型では、予測係数を予測残差信号や予測信号値を用いて各音声サンプルごとに修正する。この方式では、前向き型のように予測係数を伝送する必要がないため、伝送ビットレートの増大や線形予測分析による処理遅延が生じない利点がある。しかし、伝送誤りが生じた場合に送信側と受信側で予測係数が異なってしまう再生音質が劣化する欠点がある。この後向き型適応予測方式の1つが、ITU-T勧告G.726の32 kb/s ADPCM方式である。この符号化手順の処理ブロックの概要を図1に示す。

この符号化手順の特徴として、2次の全極型予測器と6次の全零型予測器を用いて伝送路誤りに強く設計されていることがあげられる¹⁾。それらの予測器を用

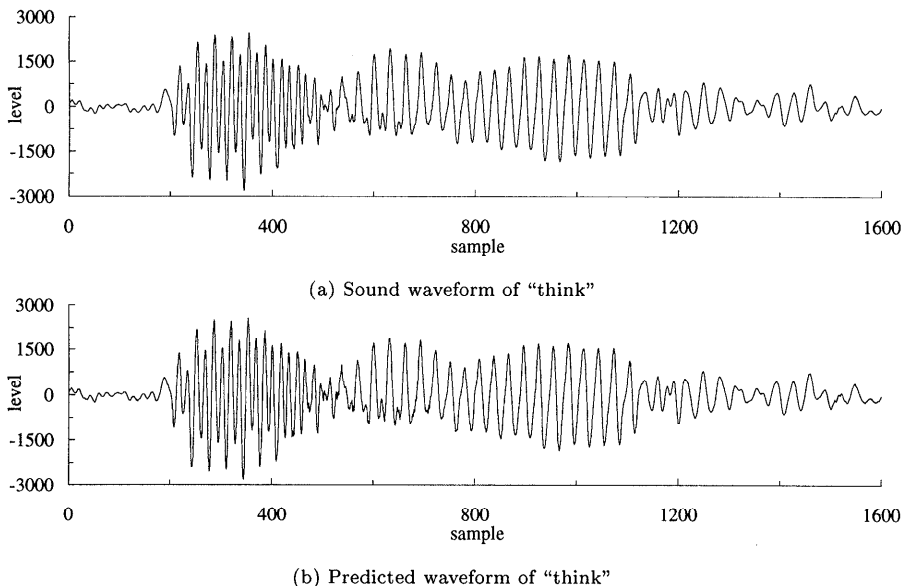


図2 “think”の音声波形と予測波形
Fig.2 Sound and predicted waveforms of “think”.

いた k 番日の予測値 $s_e(k)$ は、全極型予測係数 $a_i(k)$ を用いて次の方法で求められる (図1参照)。

$$s_e(k) = \sum_{i=1}^2 a_i(k-1)s_r(k-i) + s_{ez}(k) \quad (1)$$

ここで、 $s_{ez}(k)$ は全零型予測値を表し、また $s_r(k-i)$ は予測値 $s_e(k-i)$ と逆量子化残差値 $d_q(k-i)$ の和で表され、それぞれ次式で求められる。

$$s_{ez}(k) = \sum_{j=1}^6 b_j(k-1)d_q(k-j) \quad (2)$$

$$s_r(k-i) = s_e(k-i) + d_q(k-i) \quad (3)$$

ただし、全極型予測係数 $a_1(k)$ 、 $a_2(k)$ は、次式により更新される。

$$a_1(k) = (1 - 2^{-8})a_1(k-1) + (3 \cdot 2^{-8})\text{sgn}[p(k)]\text{sgn}[p(k-1)] \quad (4)$$

$$a_2(k) = (1 - 2^{-7})a_2(k-1) + 2^{-7}\{\text{sgn}[p(k)]\text{sgn}[p(k-2)] - f[a_1(k-1)]\text{sgn}[p(k)]\text{sgn}[p(k-1)]\} \quad (5)$$

ここに関数 sgn は、変数値が正のとき $+1$ で、負のとき -1 となる。 $p(k)$ は予測値と量子化残差を加えたもので次式により求まる。

$$p(k) = d_q(k) + s_{ez}(k) \quad (6)$$

また、関数 f は経験的な関数で次式により求めら

$$f(a_1) = \begin{cases} |a_1| \leq 2^{-1} & \Rightarrow 4a_1 \\ |a_1| > 2^{-1} & \Rightarrow 2\text{sgn}(a_1) \end{cases} \quad (7)$$

ただし、各係数の更新範囲は次のように限定される。

$$|a_2(k)| \leq 0.75 \quad (8)$$

$$|a_1(k)| \leq 1 - 2^{-4} - a_2(k) \quad (9)$$

また、全零型予測係数 $b_j(k)$ は、次式により更新される。

$$b_j(k) = (1 - 2^{-8})b_j(k-1) + 2^{-7}\text{sgn}[d_q(k)]\text{sgn}[d_q(k-j)] \quad (10)$$

$j = 1, 2, \dots, 6$

3. 合成原理

一般に伝送符号の特定ビットをテキストのビット情報で制御するとノイズが発生し再生音質の劣化を招くことになる。よって、テキスト情報を埋め込む際には、その埋込み位置、頻度ならびに秘匿性を十分考慮しなければならない。図1(a)の予測器による英語音声の発音“think”の入力波形およびその予測波形を図2に示す。図2の予測波形(b)は入力波形(a)によく追従しており、入力音声に依存して不規則に振動していることが分かる。このように波形予測値の極性符号の変化点は、多くの場合入力音声に依存して広く分散する。そこで、この予測値の変化点間の極性符号の連続数を予測符号ランレングスとよぶことにし、 l_{se} で表す。

図3に男声、女声による英語および日本語の音声を用いて各 l_{se} の平均出現率を調べた結果を示す。ただし、 \times 印は実測値、実線は推定指数分布曲線である。この結果から l_{se} は、ほぼ指数分布に従い、音声中にランダムに出現すると見なすことができる。これらの特性はテキストデータの埋込み位置を不特定個所に分散するのに好ましい性質と考えられる。

3.1 埋込み法

ADPCM方式における予測波形を利用して埋め込む方法を述べる。まず、音声をADPCMで符号化する際に、予測符号ランレングス l_{se} を計測する。 l_{se} が埋込み条件を満たすならばテキストビットを極性変化点に埋め込む。条件を満たさないならば、埋込み処理をせず、そのまま伝送する。もし、 l_{se} がランダムに発生しているならば不特定箇所にテキストビットを分

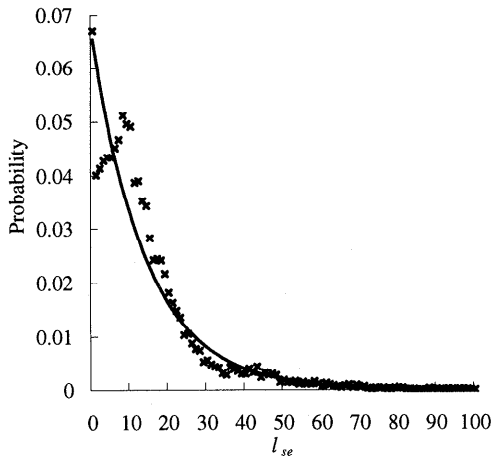


図3 l_{se} の出現率
Fig. 3 Appearance probability of l_{se} .

散配置することが可能となる。この埋込み条件の指定にはさまざまな方式が考えられるが、ここでは、上限値 k_{max} と下限値 k_{min} を用いて次の条件式を使う。

$$k_{min} \leq l_{se} \leq k_{max} \tag{11}$$

ただし、伝送符号にテキストのビット系列を埋め込むには、その埋込みの処理要領を慎重に考慮しなければ必要以上に音質を劣化させる結果になる。そこで、予測係数の適応化処理に及ぼす影響を最小限にするために図4の埋込み処理ブロックを用いる。この処理ブロックでは、量子化された音声符号に対しテキストビットを埋込み伝送後、その制御された伝送符号をフィードバックして各処理ブロックを適応化する。このようにシステムを構成すれば埋込みの影響が蓄積しないため、適応化処理ブロックに与える影響を至短時間に抑えることができ、送信側と受信側それぞれの適応化係数の同期を容易に確保できる。埋込み処理に関しては、先の条件式により選定された伝送符号 $i(k)$ の最下位ビット $i_{lsb}(k)$ をテキストビット $t(x)$ により1ビットずつ表1に示す方法で制御する。この論文では、再生音声に与える影響が最も少ない最下位ビット $i_{lsb}(k)$ に埋込み位置を設定していることに注意する。

3.2 テキスト抽出法

テキストのビット系列は、3.1節に示した方法に従い伝送符号のビット系列中に埋め込まれている。よって、受信側では埋込み位置を確実に特定しなければテキストビット系列を抽出することはできない。その処理ブロックを図5に示す。この処理ブロックでは、受

表1 埋込み処理要領
Table 1 Embedding scheme.

$Emb[i(k), t(x)]$	$t(x) = 0$	$t(x) = 1$
$i_{lsb}(k) = 0$	No change	$i_{lsb}(k) \leftarrow 1$
$i_{lsb}(k) = 1$	$i_{lsb}(k) \leftarrow 0$	No change

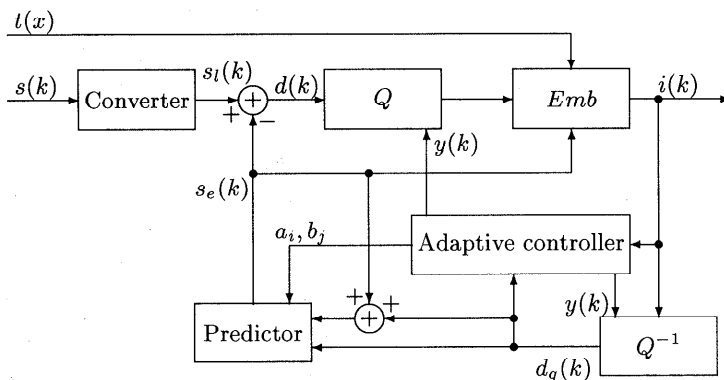


図4 埋込み処理ブロック図
Fig. 4 Embedding system block diagram.

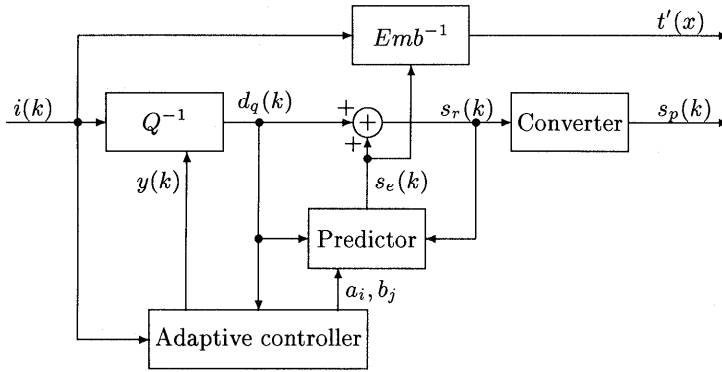


図5 復号処理ブロック図
Fig. 5 Decoding system block diagram.

信符号により各処理ブロックの適応化がなされ、出力音声の予測値 $s_e(k)$ が得られる。 $s_e(k)$ から予測符号ランレングス l_{se} を計測して、送信側と同じ埋込み条件式を満たす受信符号を特定する。その特定した受信符号からテキストのビット系列で制御されているビット $i_{sb}(k)$ を抽出してテキストを復号する。

また、図5 から明らかなように、音声の再生には、特別な符号処理を必要としない。よって、音声は、特別な処理構造を持たない普通の受信機でも再生できる。これは、不正に情報を得ようとする第三者から、埋込みテキストの存在を隠すうえで重要な特徴である。

4. アルゴリズム

前述の合成原理を符号化手順および復号化手順として具体的に記述する。まず、使用する各変数を次のように定義する。

k : 入力音声符号の順序ラベル, x : テキストの順序ラベル, $s(k)$: 入力音声, $s_l(k)$: 線形 PCM 信号, $t(x)$: テキストのビット系列, $d(k)$: 差分信号, $y(k)$: 適応量子化パラメータ, $i(k)$: 量子化差分信号, $d_q(k)$: 逆量子化差分信号, $s_r(k)$: 再構成信号, $s_p(k)$: 出力音声, $s_e(k)$: 適応予測値, $a_i(k)$, $i = 1, 2$: 2次全極型適応予測係数, $b_j(k)$, $j = 1, 2, 3, 4, 5, 6$: 6次全零型予測係数, $Q[d(k), y(k)]$: 量子化関数, $Q^{-1}[i(k), y(k)]$: 逆量子化関数, l_{se} : 予測符号ランレングス数, $Emb[i(k), t(x)]$: 埋込み処理関数 ($i(k)$ に $t(x)$ を埋め込む), $Emb^{-1}[i(k)]$: ビット抽出関数 ($i(k)$ の特定ビットを抽出する), $t'(x)$: 抽出ビット系列。

4.1 符号化手順

- (Step 1) $s(k) \rightarrow s_l(k)$
- (Step 2) $s_e(k)$ を求める。

- (Step 3) $s_l(k) - s_e(k) \rightarrow d(k)$
- (Step 4) $Q[d(k), y(k)] \rightarrow i(k)$
- (Step 5) $sgn[s_e(k)] = sgn[s_e(k-1)]$ のとき $l_{se} + 1 \rightarrow l_{se}$
- (Step 6) $sgn[s_e(k)] \neq sgn[s_e(k-1)]$ かつ $k_{min} \leq l_{se} \leq k_{max}$ のとき $Emb[i(k), t(x)]$, $x + 1 \rightarrow x$
- (Step 7) $sgn[s_e(k)] \neq sgn[s_e(k-1)]$ のとき $1 \rightarrow l_{se}$
- (Step 8) $i(k)$ を送信する。
- (Step 9) $Q^{-1}[i(k), y(k)] \rightarrow d_q(k)$
- (Step 10) $a_i(k+1), b_j(k+1), y(k+1)$ を求める。
- (Step 11) $k + 1 \rightarrow k$
- (Step 12) 通信終了でなければ (Step 1) に戻る。
- (Step 13) 終了。

4.2 復号化手順

- (Step 1) $i(k)$ を受信する。
- (Step 2) $s_e(k)$ を求める。
- (Step 3) $sgn[s_e(k)] = sgn[s_e(k-1)]$ のとき $l_{se} + 1 \rightarrow l_{se}$
- (Step 4) $sgn[s_e(k)] \neq sgn[s_e(k-1)]$ かつ $k_{min} \leq l_{se} \leq k_{max}$ のとき $Emb^{-1}[i(k)] \rightarrow t'(x)$, $x + 1 \rightarrow x$
- (Step 5) $sgn[s_e(k)] \neq sgn[s_e(k-1)]$ のとき $1 \rightarrow l_{se}$
- (Step 6) $Q^{-1}[i(k), y(k)] \rightarrow d_q(k)$
- (Step 7) $d_q(k) + s_e(k) \rightarrow s_r(k)$
- (Step 8) $s_r \rightarrow s_p(k)$
- (Step 9) $a_i(k+1), b_j(k+1), y(k+1)$ を求める。
- (Step 10) $k + 1 \rightarrow k$

- (Step11) 通信終了でなければ (Step 1) に戻る.
- (Step12) 抽出系列からテキストを再生する.
- (Step13) 終了.

5. 実験結果

5.1 実験システムの概要

ITU-T 勧告 G.726 の 32 kb/s ADPCM 符号化方式に基づき、実験システムを構成してシミュレーションを実施した。実験音声は、表 2 に示す男声、女声それぞれの日本語および英語を用いた。音声源は、FM ラジオならびに会話テープから抽出したものをを用いた。

5.2 音質の評価法

客観的な音質の評価尺度として最も基本的なものに信号対量子化雑音比 (SNR: Signal-to-quantization Noise Ratio) がある。SNR の評価式は、入力音声 $s(n)$ とその量子化誤差 $e(n)$ を用いて、式 (12) で定義される¹⁾。

$$SNR = 10 \log_{10} \left\{ \frac{\sum_n s^2(n)}{\sum_n e^2(n)} \right\} \text{ [dB]} \quad (12)$$

本実験では、客観的評価法として上式で示した SNR およびそれを改良して主観評価との対応関係を向上させたセグメンタル SNR (SNR_{seg}) を用いた。これは、式 (13) で定義される¹⁾。 N_f は測定区間のフレーム数を表し、本実験では 1 フレームの長さを 32 ms とした。

$$SNR_{seg} = 1/N_f \sum_{i=1}^{N_f} SNR_i \text{ [dB]} \quad (13)$$

また、主観的評価法として、評価者の絶対判断によるオピニオン評価を用いた¹⁾。これは、複数の評価者に音質を 5~9 段階に絶対評価させ、得られた評価値から平均オピニオン値 (MOS: Mean Opinion Score) を求めるものである。本実験では下に示す 5 段階評価とした。また、被験者の先入観による影響を避けるた

め、埋込みのない音声と埋込みのある音声を無作為に提示して、それぞれの平均オピニオン値 (MOS) を比較した。

「オピニオン評価の基準」

- 非常に良い : 5
- 良い : 4
- 普通 : 3
- 悪い : 2
- 非常に悪い : 1

5.3 実験結果と考察

表 2 の 8 種類の音声を用いて実験した結果を示す。まず、図 3 から l_{se} は、短い値ほど多く出現することが分かる。また、実験音声 Em を用いて $k_{min} = 0$ とし k_{max} を逐次増大させた場合における埋込みバイト量の変化を図 6 に示した。これは埋込み条件の値域を広くするほど大量に埋め込むことができることを示している。これらを考慮して埋込みの条件 k_{min} , k_{max} を定めればテキスト情報を埋め込む時間的密度を操作することが可能となる。ただし、高い密度で埋込みを施すと第三者による解読の可能性も高くなるため、埋込み条件の決定は慎重に行う必要がある。

次に、Em を用いて、埋め込むテキストのバイト量に対する必要ビット操作量を調べた結果を図 7 に示す。これからテキスト情報伝送に必要なビット量の約半分程度のビット操作で埋込みが可能なが分かる。これは埋込み操作が先に述べた表 1 に従う点からも容易に推測できる。

ここで、埋込み処理ブロック (図 4) について考察する。無作為に抽出した伝送符号の最下位ビットを適応化フィードバック処理後にランダム系列で制御した場合と適応化フィードバック処理前に制御した場合の

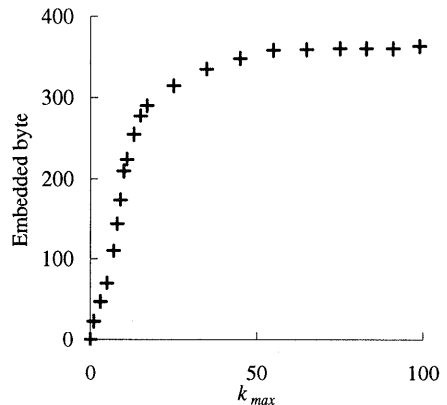


図 6 埋込み条件 k_{max} によるバイト量の変化
Fig. 6 Change of bytes embedded for k_{max} .

表 2 実験用音声データ

Table 2 Sound data for experiment.

Data	Language	type	samples	sec
Jm	Japanese	male	40,000	5
Jw	Japanese	female	40,000	5
Em	English	male	40,000	5
Ew	English	female	40,000	5
JmL	Japanese	male	480,000	60
JwL	Japanese	female	480,000	60
EmL	English	male	480,000	60
EwL	English	female	480,000	60

音質劣化状況を図 8 に示す。ただし、音声データは Em である。この図から、適応化フィードバック処理前に埋め込めば、音質に与える影響を非常に小さくできることが分かる。よって、図 4 の処理ブロックは、埋込みに適した構造の 1 つであると考えられる。

実際に音声波形の一部を取り出して観察してみる。図 9 に埋込みのない再生音声波形と埋込み条件を $k_{min} = 0, k_{max} = 500$ として大量の埋込みを施した再生音声波形およびそれらの差分波形を示した。この波形は、Em における発音 “think” に対応する部分である。この結果から、埋込みによる波形の大きな歪みは観察されない。また、同じ条件で埋込みを施した場合の埋込み密度と再生音質を表 3 に示す。この表より、数十バイト/秒の大量な埋込みを施しても、大きな歪みを生じないことが確認できる。

次に、各再生音質を主観的評価により検討する。実験は被験者 12 名の協力を得て行った。実験音声として、埋込みなしで符号化した再生音声と、埋込み条件

を $k_{min} = 0, k_{max} = 500$ として大量の埋込みを施した再生音声を用意した。実験の結果を表 4 に示す。この表には各音声の平均オピニオン値 (MOS) および SNR_{seg} を示した。埋込みのない再生音声の MOS は約 3.15 で SNR_{seg} は約 39.31 [dB] となり、埋込みのある再生音声の MOS は約 3.04 で SNR_{seg} は約 37.89 [dB] となる。これらを比較すると、埋込みにより音質はやや劣化するが、埋込みのある場合とない場合の再生音質は、ほぼ同程度と考えられる。それぞれの MOS が 3 程度の値を示したのは、G.726 の大幅な符号圧縮および 8 kHz の波形標本化により、再生音声がやや不明瞭になったためと考えられる。

これらの結果から、原音声信号や埋込みのない再生音声を持たない第三者が、通話音声を盗聴するとき、埋込みのある音声を特定し、さらに、その伝送符号から埋め込まれたテキストを解読することは難しいと推察できる。

5.4 今後の課題

バックワード型の ADPCM 通信方式は伝送誤りが生じると音質が大きく劣化する。そのため、本手法における埋込み情報も伝送路誤りにより受ける影響は大きい。そこで、伝送符号に誤りが生じても正しい符号を再現できる工夫を施す必要がある。その対策として誤り訂正符号の導入が考えられる。たとえばハミング符号や巡回符号および BCH 符号などが適用できる。これらの方法で、ある程度の伝送路誤りには対処可能であろう。さらに、送信側と受信側で定期的に内部状

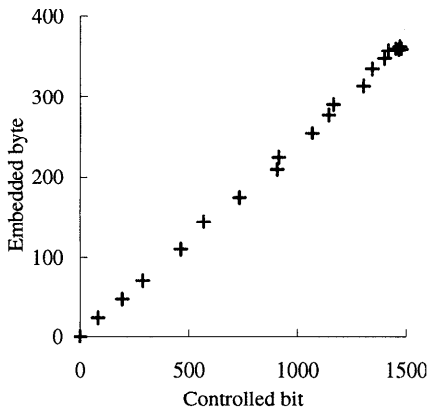


図 7 操作ビット数に応じた埋込みバイト量

Fig. 7 Amount of embedded bytes depending on controlled bit.

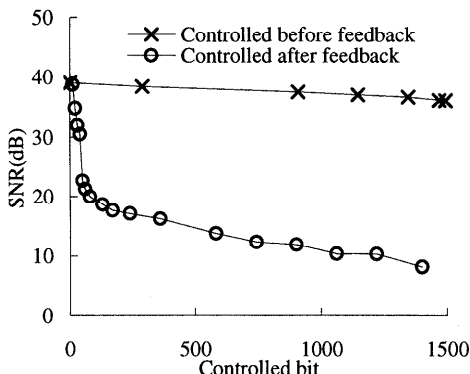


図 8 ビット操作が音質に与える影響

Fig. 8 Influence on SNR by bit control.

表 3 再生音質
Table 3 SNR of reproduced sound.

Data	Embedded rate (byte/sec)	SNR [dB]	SNR_{seg} [dB]
Jm	0	37.17	39.18
	56.2	35.57	37.77
Jw	0	37.06	38.03
	66.2	34.53	36.29
Em	0	39.04	39.77
	72.8	36.10	37.91
Ew	0	38.30	37.85
	70.4	36.75	36.33

表 4 平均オピニオン値および SNR_{seg}
Table 4 Mean opinion score and SNR_{seg} .

Data	Sound without text		Sound with text	
	MOS	SNR_{seg} [dB]	MOS	SNR_{seg} [dB]
JmL	3.17	39.43	3.08	38.01
JwL	3.00	39.20	2.67	37.92
EmL	3.75	39.07	3.33	37.83
EwL	2.67	39.53	3.08	37.79
Average	3.15	39.31	3.04	37.89

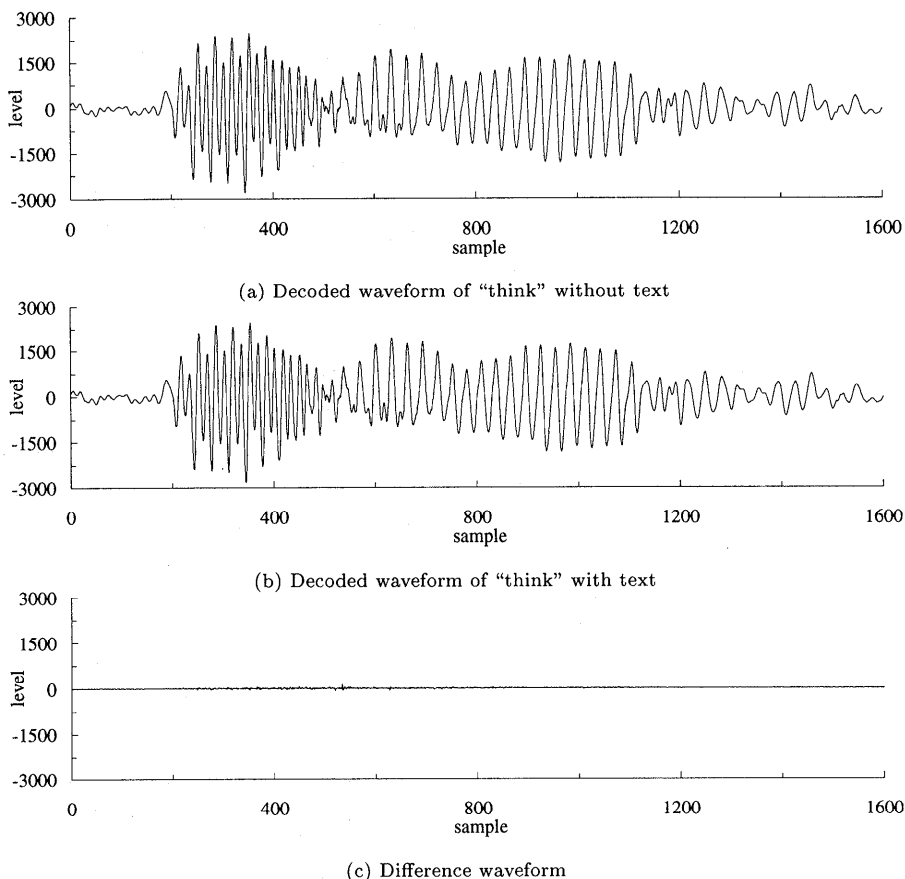


図9 音声波形の比較
Fig. 9 Comparison of sound waveform.

態を初期化する対策なども考慮しておくことと万全である。また、この埋込み方法を第三者に察知されて、テキストのビット系列が復号されることを防止するために、あらかじめテキストを暗号化したランダム系列として準備するならば秘匿性はきわめて大きくなる。実用化に際しては、このような配慮が重要である。

6. む す び

この論文では、ADPCM方式の伝送符号にテキスト情報を密かに埋め込む手法を提案した。本手法を用いれば、受信符号にテキスト情報が埋め込まれた状態でも受信者に聴感的な違和感を与えない良好な音声を再生できる。よって、帯域幅を増大せず音声とテキスト情報を併送することが可能となり回線資源を有効利用できる。また、受信者は通話中であっても受信符号から即座に重要な埋込み情報を取り出して表示・保存することができる。これは一般の音声通話で効率良く意図を伝達するための補助的な情報表示を可能にする。一方、送信者は特定相手のみに知らせたい非公開

情報を暗号化して音声符号に密かに埋め込み、第三者にその存在を知られることなく伝送できるため、将来の秘匿通信の一手段としての応用も考えられる。さらに、マルチメディアの著作権保護対策の1つとして注目される音声の電子透かしにも容易に適用可能である。

参 考 文 献

- 1) 小澤一範：デジタル移動通信のための高効率音声符号化技術，トリケップス (1992)。
- 2) Rabiner, L.R. and Schafer, R.W.: *Digital Processing of Speech Signals*, Prentice Hall (1978). 鈴木久喜 (訳)：音声デジタル信号処理 (上)，コロナ社 (1983)。
- 3) 大賀寿朗，山崎芳男，金田 豊：音響システムとデジタル処理，電子情報通信学会 (1995)。
- 4) 中田和男：音声の高効率符号化，森北出版 (1986)。
- 5) 安田 浩：マルチメディア符号化の国際標準，丸善 (1991)。
- 6) Steel, R. and Vitello, D.: *Simultaneous Transmission of Speech and Data Using Code-*

breaking Techniques, *The Bell System Technical Journal*, Vol.60, No.9, pp.2081-2105 (1981).

- 7) Wong, W.C., Steel, R. and Xydeas, C.S.: Transmitting Data on the Phase of Speech Signals, *The Bell System Technical Journal*, Vol.61, No.10, pp.2947-2970 (1982).
- 8) 松井甲子雄, 中村康弘, ナタウトサムパイプーン: 音声通信への文字情報の埋め込み, 第18回情報理論とその応用シンポジウム, pp.389-392 (1995).
- 9) Boney, L., Tewfik, A.H. and Hamdy, K.N.: Digital Watermarks for Audio Signals, *Proc. International Conference on Multimedia Computing and Systems*, pp.473-480 (1996).
- 10) Recommendation G.726, ITU (1990).

(平成9年1月6日受付)

(平成9年7月1日採録)



岩切 宗利

昭和45年生, 平成5年防衛大学校情報工学科卒業. 現在防衛大学校理工学研究科情報数理専攻在学中. 音声符号化と情報セキュリティに関する研究に従事.



松井甲子雄 (正会員)

昭和14年生, 昭和36年防衛大学校電気工学科卒業. 昭和40年九州大学大学院工学研究科電子専攻修了. 昭和56年防衛大学校電気工学科教授. 平成元年同大学校情報工学科教授. この間暗号学, 情報セキュリティ, 電子透かし, 音声・画像データの符号化に関する研究に従事. 著書「画像深層暗号」(森北出版). 工学博士. 電子情報通信学会, 画像電子学会, 映像情報メディア学会会員.