

ADPCM 符号化方式に適用可能なデータ埋め込み方式の検討

近藤 和弘^{†a)} 金澤 英樹^{†*} 吉井 亜美^{†**} 加藤 浩[†]
清野 智之^{†***} 中川 清司[†]

On Data Embedding Methods Applicable to ADPCM Family of Codecs

Kazuhiro KONDO^{†a)}, Hideki KANAZAWA^{†*}, Ami YOSHII^{†**}, Hiroshi KATO[†],
Tomoyuki SEINO^{†***}, and Kiyoshi NAKAGAWA[†]

あらまし 抽出に原音が不要な軽量、実時間処理可能なデジタルデータ埋め込み方式について検討した。主に通信用途を想定しており、ADPCM 符号化を用いる通信用音声符号化標準 G.726 や 727, および 7kHz 帯域オーディオ符号化標準 G.722 に適用できる。既に提案されている予測値極性符号のランレングスに応じて ADPCM 符号最下位ビットにテキストを埋め込む方式に対し、逆量子化極性符号のランレングスに応じて ADPCM 符号最下位ビットに埋め込む方式を提案した。また、ビットエラー等により容易に破壊される前者の埋め込み方式と、頑強な後者の方式を同時に埋め込む 2 重データ埋め込み方式も合わせて提案した。提案方式によりデータを埋め込んでも音質への影響はほとんどないことを確認し、埋め込み可能量、ビットエラーに対する頑強性について検討した。

キーワード ADPCM 符号化, 7kHz オーディオ符号化, G.726, G.722, 電子透かし

1. はじめに

近年、データ・ネットワークの高速化、普及により音声通信も同ネットワークを通して行われることが当たり前となりつつある。デジタル化した音声信号は符号化してもなお相当量の冗長性を含んでおり、この冗長性を利用してデータを埋め込むことが検討されている [1]。さらに、防衛大の岩切らは通信用音声符号化方式の国際標準である適応差分符号化方式 (ADPCM) 方式 [2] [3] へのテキスト情報の埋め込みについて検討している [4]。この方式は極めて低処理量で簡単にデータを埋め込むことができ、また受信側でのデータ抽出に原音を用いる必要がない。よって原音の利用が困難な実時間通信における、なりすまし防止の電子透かしの他に、キャプション、あるいは音声に付随するテキストデータとして利用できる可能性がある。

岩切らの方法では ADPCM 符号化における予測符号の極性の連続数 (ランレングス) がランダム分布することに着目し、この連続数が一定範囲内にある時、ADPCM 符号の最下位ビットにいわゆる“ビットスチール”を行い、テキスト情報を 1 ビットずつ埋め込む。極性符号のランダム性により埋め込み位置が広く分散され音質への影響も少なく、かつ連続数の許容範囲が既知でないと正確にテキストを抽出できない。このように優れた特徴を持った埋め込み方法である。しかし、送信側、受信側の予測値は基本的に後ろ向き適応動作を行っているため、もしビットエラーや改ざん、第 3 者による攻撃等何らかの原因で基準としている ADPCM 符号が送信側と受信側で食い違くと、適応動作も大きく食い違うこととなり、その差は何らかの対策を採らない限り収束しない。よって予測値極性を基準としている埋め込みテキストの抽出への影響も大きい。

そこで、本論文では埋め込み基準を食い違いの影響が持続しない逆量子化差分値とする第 2 のテキスト埋め込み方法を提案する [5]。さらに提案方式と岩切らの方式による 2 種類のテキストを同時に埋め込む方式を提案し、頑強な提案方式と改ざんに弱い岩切らの方

[†] 山形大学工学部, 山形県

Faculty of Engineering, Yamagata University, 4-3-16 Jonan,
Yonezawa, Yamagata 992-8510, Japan

* 現在、(株) テクニカルヒューマンリソース, 東京都

** 現在、(株) 富士通ハイパーソフトテクノロジー, 神奈川県

*** 現在、(株) エフコム, 福島県

a) E-mail: kkondo@yz.yamagata-u.ac.jp

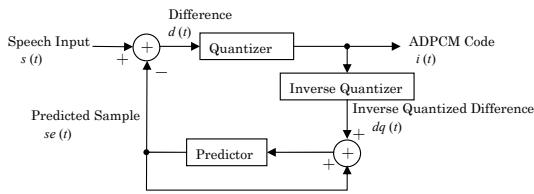


図 1 ADPCM 符号化方式 (送信部)

Fig. 1 Configuration of the ADPCM Codec (Transmitter)

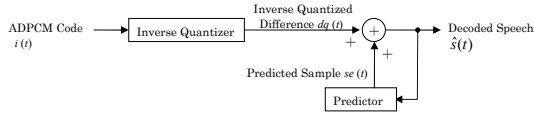


図 2 ADPCM 符号化方式 (受信部)

Fig. 2 Configuration of the ADPCM Codec (Receiver)

方式の特性を同時に利用する方式を提案する。また、本方式は ADPCM 符号化の様な後ろ向き適応動作を用いる方式に簡単に適応できる。よって、電話品質音声用の ADPCM 符号化のみではなく、やはり ADPCM 符号化を内包する 7kHz 帯域オーディオ符号化を目的としたサブバンド ADPCM 符号化方式 [6] にも適用できることを示す [7]。

以下、第 2 章では電話品質の ADPCM 方式への提案テキスト埋め込み方式について説明し、その評価結果を示す。さらに第 3 章では 7kHz 帯域オーディオ符号化 ADPCM 方式へのテキスト埋め込み方式およびその評価結果について示す。最後にまとめと今後の課題を述べる。

2. 電話品質 ADPCM 符号化へのテキスト埋め込みの検討

2.1 電話品質 ADPCM 符号化方式の概要

図 1, 2 に国際標準 G.726 [2] に準拠した ADPCM 符号化方式の送信部, 受信部を示す。過去のサンプルより線型予測したサンプル $se(t)$ と入力 $s(t)$ の差分信号 $d(t) = s(t) - se(t)$ を適応量子化し ($i(t)$)、伝送する。受信側では、量子化差分信号 $i(t)$ を逆量子化し ($dq(t)$)、これに予測信号 $se(t)$ を加えて音声信号 $\hat{s}(t) = dq(t) + se(t)$ を復号する。G.726 では後ろ向きの適応予測と適応量子化を採用している。過去の逆量子化差分値 $dq(t-n)$ より予測係数を更新し、また過去の量子化差分信号 (ADPCM 符号) $i(t-m)$ により適応量子化ステップを更新する。送信部も受信部も

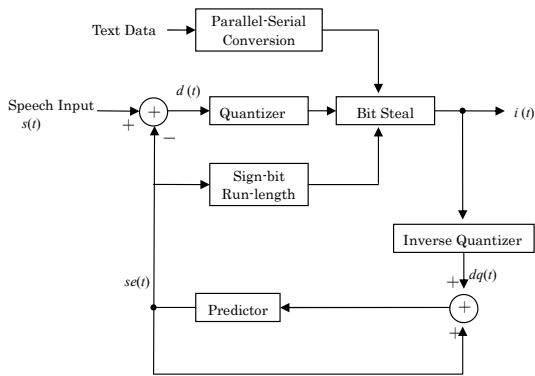


図 3 岩切らによるテキスト埋め込み ADPCM 符号化方式 (送信部)

Fig. 3 Configuration of the Text-Embedded ADPCM Codec by Iwakiri et. al.(Transmitter)

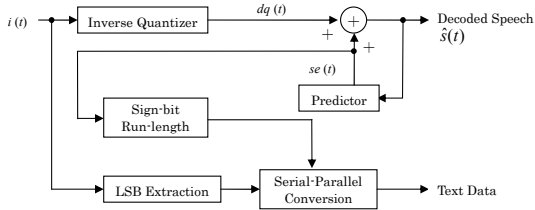


図 4 岩切らによるテキスト埋め込み ADPCM 符号化方式 (受信部)

Fig. 4 Configuration of the Text-Embedded ADPCM Codec by Iwakiri et. al.(Receiver)

適応動作には ADPCM 符号 $i(t)$ あるいはこれより一意的に決定できる信号のみを基準にしているの、エラーがない限り適応動作は一致する。よって予測サンプル $se(t)$ も送信側と受信側で完全に一致する。

2.2 電話品質 ADPCM 符号化方式へのテキスト埋め込み

岩切らは上記 G.726 ADPCM 方式に対し、予測信号 $se(t)$ の極性符号の連続値 (ランレングス) がランダムに存在することに着目し、極性変化時のランレングスが設定範囲内に入る場合のみ ADPCM 符号の最下位にテキストを 1 ビット埋め込んでいる [4]。この方式の送信部および受信部を図 3, 4 に示す。

この方式でも音声符号化の動作は G.726 と同じである。送信部では適応予測サンプル $se(t)$ の極性変化を監視し、同じ極性符号が連続するサンプル数、すなわちそのランレングスを算出する。このランレングスが設定範囲内にある場合、並列-直列変換したテキストデータの 1 ビットずつ極性変化したサンプルの ADPCM 符号 $i(t)$ 最下位ビットをスチールし、埋め

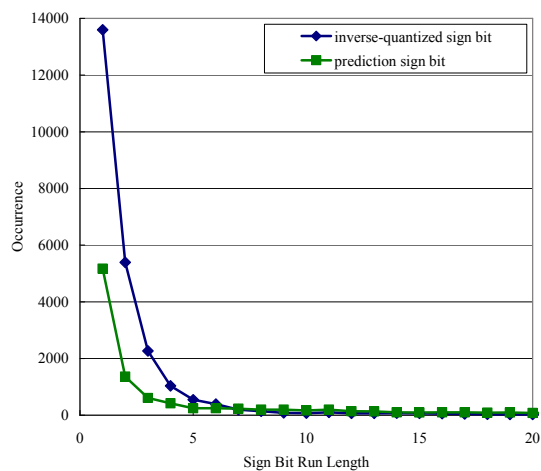


図 5 予測値、および逆量子化差分値極性符号ランレングスの頻度分布

Fig. 5 Prediction and Inverse-quantized Sample Sign Run-length Occurrence

込む。すなわち最下位のビットを並列 - 直列変換したテキストデータ 1 ビットに書き換える。この時、適応動作を乱さないために、このテキストを埋め込んだ ADPCM 符号を元に適応動作を行っている。

一方、受信部でも $se(t)$ の極性符号ランレングスを算出し、同じ設定範囲内にある時、極性変化点の ADPCM 符号最下位のビットを抽出し、これを順に直並列変換して埋め込みテキストを再現する。前述のように ADPCM 符号 $i(t)$ にビットエラーが生じない限り送信部と受信部の $se(t)$ は一致するので、埋め込みテキスト抽出のタイミングも一致し、テキストを再現することが可能となる。この方法はテキストの埋め込みをある程度ランダムな間隔で行なうことになるので、音質に影響をほとんど与えずに実現可能であることが示された。しかし伝送中に ADPCM 符号に誤りが生じたり、攻撃があった場合は、送信側と受信側で適応動作に差が生じ、これが長期間に渡り収束しない。よってその極性符号にも持続的に影響があり、抽出されたテキストにも大きな影響がある。そこで、同じく極性符号の連続値がランダムに分布する逆量子化差分値 $dq(t)$ のランレングスに注目する。図 5 に予測値 $se(t)$ および $dq(t)$ の極性符号ランレングスの頻度分布を示す。

このランレングスは同一の日本語読み上げ文について算出した値である。双方とも指数分布に近い分布を示しており、双方とも極性変化点は [4] で述べられて

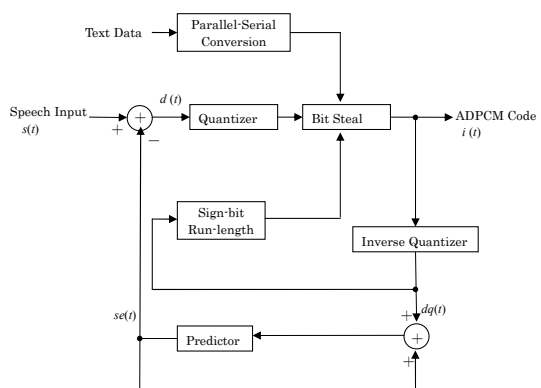


図 6 提案テキスト埋め込み ADPCM 符号化方式 (送信部)

Fig. 6 Configuration of the Proposed Text-Embedded ADPCM Codec (Transmitter)

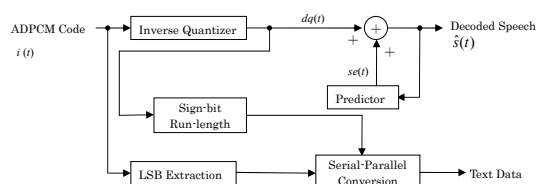


図 7 提案テキスト埋め込み ADPCM 符号化方式 (受信部)

Fig. 7 Configuration of the Proposed Text-Embedded ADPCM Codec (Receiver)

いるように、音声信号に広く分散することが期待できる。逆量子化差分値極性ランレングスには短いランレングスが多く見られるのに対し、予測値極性は比較的広範囲に分布している。全ラン中、ランレングスが 1、すなわち孤立した極性変化の割合は逆量子化値極性が 55% (全逆量子化値極性ラン数 24516 中 13599)、予測値極性が 48% (全予測値極性ラン数 10828 中 5168)、であった。

以上のように、逆量子化差分値極性も予測値極性と類似の特性を示すことが分かった。そこで、逆量子化差分値 $dq(t)$ の極性符号ラン長が一定範囲内にある時に ADPCM 符号 $i(t)$ の最下位にテキストを埋め込むことを検討する [5]。これにより、仮に ADPCM 符号 $i(t)$ に誤りが生じてもその影響は誤りがあったサンプルのみで、その後の逆量子化差分値の極性には影響しない。

図 6, 7 に提案方式の送信部、および受信部を示す。送信部では逆量子化差分値 $dq(t)$ の極性変化を監視し、極性符号のランレングスを算出する。このランレ

グスが設定範囲内にある場合、並列 - 直列変換したテキストデータの 1 ビットずつ極性変化したサンプルの ADPCM 符号 $i(t)$ 最下位ビットをスチールし、埋め込む。この時、岩切らの方式と同様に、適応動作を乱さないために、このテキストを埋め込んだ ADPCM 符号を元に適応動作を行っている。

一方、受信部でも $dq(t)$ の極性符号ランレングスを算出し、同じ設定範囲内にある時、極性変化点の ADPCM 符号最下位のビットを抽出し、これを順に直並列変換して埋め込みテキストを再現する。やはり ADPCM 符号 $i(t)$ にビットエラーが生じない限り送信部と受信部の $dq(t)$ は一致するので、埋め込みテキスト抽出のタイミングも一致し、テキストを再現することが可能である。

以下、提案方式により逆量子化差分符号極性ランレングスを参照して埋め込まれたテキストを Type I、岩切らによって提案された予測値極性符号ランレングスを参照して埋め込まれたテキストを Type II と呼ぶことにする。

2.3 シミュレーションによる基本性能評価

シミュレーションにより提案方式の基本性能を評価した。評価には音響学会の研究用音声データベース [8] のうち ATR 音素バランス文を読み上げた男女の音声を用いた。原音は 16kHz 標本化、16bit なので、これを 8kHz にダウンサンプルして用いた。

2.3.1 テキスト埋め込み率

本節では各ランレングス設定値に対するテキスト埋め込み率 (ADPCM 符号量に対する埋め込みテキスト量の比) を評価する。以降、 se_{kmin} 、 se_{kmax} を予測値極性ランレングスの許容最小値および最大値、また dq_{kmin} 、 dq_{kmax} を逆量子化値極性ランレングスの許容最小値および最大値とする。

表 1 に埋め込み Type 別ランレングス最小値と埋め込みビット数、および埋め込み率を示す。Type I ではランレングス最小値は dq_{kmin} 、Type II では se_{kmin} となる。最大値 dq_{kmax} 、 se_{kmax} はいずれも 250 としたが、これは実測された $se(t)$ 、 $dq(t)$ の最大ランレングスより十分大きく、実質最大値を設定しないのに等しい。よって最小値を 1 とすれば、どのようなランレングスでも極性変化点に埋め込みを許可することになり、最大埋め込み量が得られることになる。

表 1 に示すように、この条件では Type I は Type II の約 2.5 倍 (率では約 2 倍) の埋め込みが可能であ

表 1 ランレングス最小値対テキスト埋め込み量
Table 1 Minimum Run Length vs. Embedded Text Bits

Speech Sample	Text Type	Minimum Run Length	Embedded Bits (Text Ratio [%])
Male	I	1	24,516 (9.56)
	II		10,828 (4.22)
	I	2	10,850 (4.23)
	II		5,580 (2.18)
	I	3	5,574 (2.17)
	II		4,208 (1.64)
	I	10	873 (0.34)
II	2,178 (0.85)		
Female	I	1	26,375 (10.39)
	II		10,650 (4.16)
	I	2	11,777 (4.60)
	II		6,122 (2.39)
	I	3	6,529 (2.55)
	II		4,781 (1.87)
	I	10	867 (0.34)
II	2,728 (1.06)		

る。これは図 5 に示したように、逆量子化差分値の極性ランレングスが短い値に集中しており、極性変化自体が多いためによる。したがって、表 1 に示す様に最小値 dq_{kmin} を増加させると、Type I の埋め込み量は急激に減少する。

一方、予測値極性ランレングスは広く分布しているので、最小値 se_{kmin} を増加させてもそれほど Type II の埋め込み量は減少せず、最小値 3 程度で Type II の埋め込み量の方が Type I より多くなる。また、同一 Type では埋め込み量の話者依存性は少なく、安定した埋め込み量が見込める。

2.3.2 テキスト埋め込みの音質への影響

テキスト埋め込みの増加とともに音質が悪化すると考えられるが、今回評価した範囲内では主観的にはほとんど影響を受けることはなかった。音質に関しては、3 章で提案する広帯域音声符号化方式において詳細に評価する。

2.3.3 伝送ノイズ・改ざんに対する耐性

図 8 に ADPCM 符号にランダムにエラーを発生したときに Type I および II のテキストに生じたエラー率を示す。埋め込みに用いたパラメータは $dq_{kmin} = se_{kmin} = 1$ 、 $dq_{kmax} = se_{kmax} = 250$ である。エラー耐性はエラー発生位置の影響を受けると考えられるため、3 つの異なる乱数シードを用いた同一エラー率の 3 種の異なるエラーパターン (Bit Error Pattern 1, 2, 3) を用いて各々評価した。ビットエラーにより極性変化位置が影響を受けるため生じる置換誤

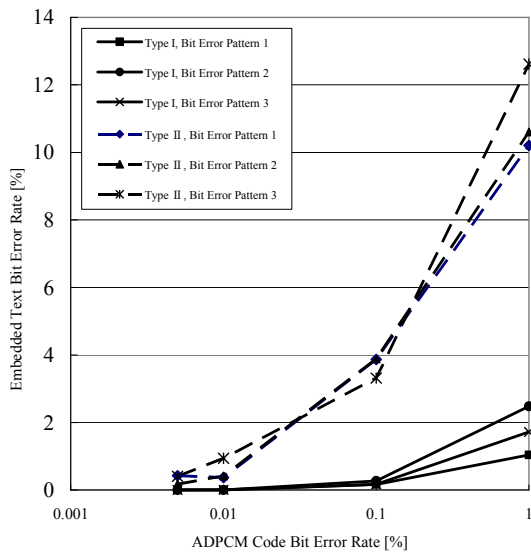


図 8 ADPCM 符号化方式に対する埋め込みテキストのビットエラー耐性
Fig. 8 Bit Error Resilience of Embedded Text for the ADPCM Codec

り (Substituted Bits) に加え、極性変化が新たに生じたことによる挿入誤り (Inserted Bits), また極性変化が消滅することによる脱落誤り (Deleted Bits) も生じるので、全体のビットエラー率を以下で定義した.

$$BER [\%] = 100 \frac{\text{Substituted Bits} + \text{Deleted Bits} + \text{Inserted Bits}}{\text{Total Evaluated Bits}}$$

ここで、Total Evaluated Bits とは埋め込んだテキストのビット数を示す.

予想通り Type II のテキストはビットエラーの大きな影響を受ける. 一方、Type I はほとんど影響が見られない. このビットエラーの原因となっている逆量子化極性および予測値極性のビットエラー率を図 9 に示す. 予測値極性のビットエラーは ADPCM 符号のエラー率の 4 倍程度まで増幅されるが、これは送信側と受信側の予測器適応動作に食い違いが生じたためである. 一方、逆量子化極性は、逆量子化値が適応ループには基本的に無関係なので、ADPCM 符号のエラー率より極めて少ない. しかしいずれの Type のテキストも極性符号の誤りがそれ以降のランレングスの算出に影響するので、極性符号のエラー率に比べればテキストのエラー率は増幅される傾向にある.

2.4 2重テキスト埋め込み

以上のように極めて類似した基準で極めて特性の異なるテキストの埋め込みを行う事が出来る. Type II

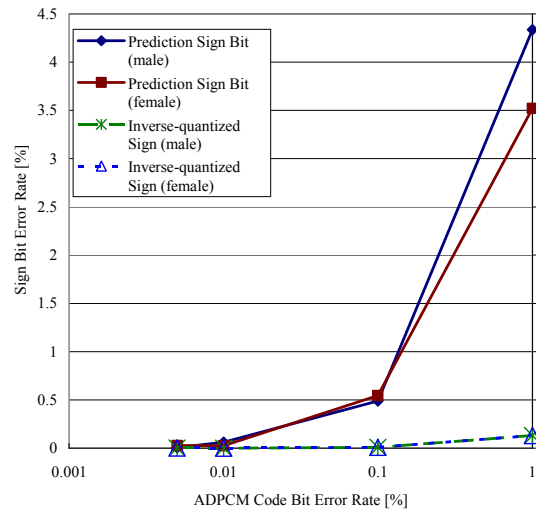


図 9 ADPCM 符号化方式における極性符号のビットエラー耐性
Fig. 9 Sign Bit Error Resilience of the ADPCM Codec

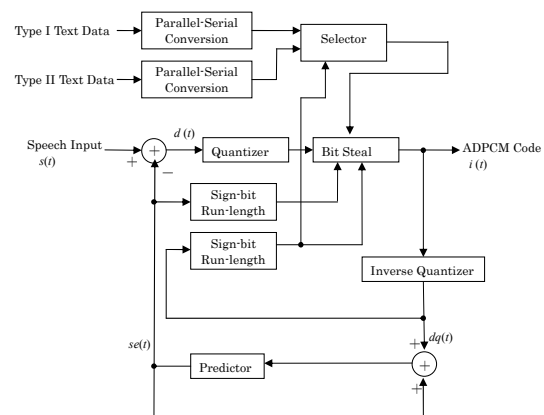


図 10 2重テキスト埋め込み ADPCM 符号化方式 (送信部)
Fig. 10 Configuration of the Proposed Dual Text-Embedded ADPCM Codec (Transmitter)

は少しでも ADPCM 符号に誤りが生じると破壊され、Type I はあまり影響がないか、あるいはあってもその影響が短時間で消滅する. そこでこの二つの特性を生かし、それぞれの基準で同時に 2 種類のテキストを埋め込む方式を提案する. 図 10, 11 に提案方式の送信部、および受信部を示す. 上記 2 種類の埋め込み条件が同時に成立する場合も考えられるが、この時は Type I を優先させることとする. もちろん逆に Type II を優先させることも可能である.

送信部では、図 3 に示す岩切らの方式と同様に Type

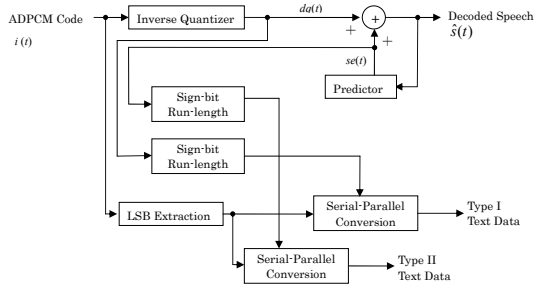


図 11 2重テキスト埋め込み ADPCM 符号化方式 (受信部)

Fig. 11 Configuration of the Proposed Dual Text-Embedded ADPCM Codec(Receiver)

II のテキスト埋め込みタイミング判定用に予測値 $se(t)$ の極性ランレングスを算出し、同時に Type I テキスト埋め込みタイミングの判定用に逆量子化差分値 $dq(t)$ の極性ランレングスも算出する。 $se(t)$ あるいは $dq(t)$ に極性変化が発生し、かつランレングスが設定範囲内である場合は、量子化器出力の最下位に各々の Type のテキストを埋め込む。同時に両方の極性が変化し、設定範囲内にある場合は Type I を優先し、Type I のテキストを埋め込む。Type II のテキストはここでは埋め込まず、次に条件を満たすサンプルで埋め込む。

埋め込みテキストとしては、Type I および Type II のテキストを別々に並列-直列変換し、これを Selector により一方を選択する。各々の Type のデータ 1 ビットを埋め込む度に各テキストをビットシフトし、次のビットを用意する。

受信部は送信と同じ判定動作を行ない、ADPCM 符号 $i(t)$ の最下位ビットをシフト入力して直列-並列変換し、埋め込まれたテキストを得る。 $dq(t)$ の極性ランレングスが範囲内にある場合は、Type I の直列-並列変換器に ADPCM 符号最下位ビットをシフト入力し、 $se(t)$ の極性ランレングスが範囲内にある場合は Type II の直並列変換器に ADPCM 符号最下位ビットをシフト入力する。 $se(t)$, $dq(t)$ が同時に変化し、ランレングスが設定範囲内にある場合は Type I を優先し、最下位を Type I の直並列変換器へのみビットシフトする。ここでもビットエラーがない限り送信側、受信側で各々 $se(t)$, $dq(t)$ は一致するので、その極性変化は一致し、テキストの埋め込み、および抽出のタイミングは一致するはずである。また同時に埋め込み条件を満たした場合でも、優先する種類を送信部、受信部において統一しておけば、テキストの抽出は可能である。

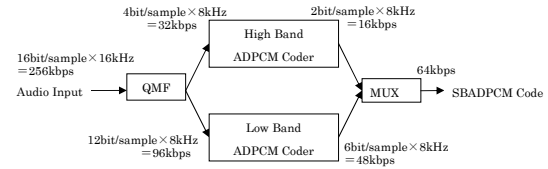


図 12 広帯域 ADPCM 符号化方式 (送信部)

Fig. 12 Configuration of the Sub Band ADPCM Codec (Transmitter)

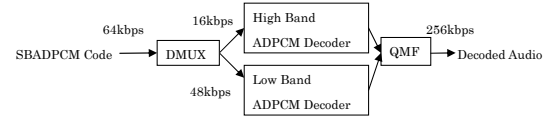


図 13 広帯域 ADPCM 符号化方式 (受信部)

Fig. 13 Configuration of the Sub Band ADPCM Codec (Receiver)

この 2重テキスト埋め込み方式のビットエラー耐性の評価も行なっているが、Type I, II を各々単独に埋め込んだ場合とほとんど性能が変わらないことを確認している。

3. 広帯域会議音声用符号化方式へのテキスト埋め込みの検討

3.1 広帯域 ADPCM 符号化方式の概要

図 12, 13 に音声会議等に用いられることを想定した 7kHz 帯域オーディオ符号化標準 G.722 の構成を示す [6]. G.722 は入力音声を高低の 2 帯域に分割し、それぞれに ADPCM 符号化を適用するサブバンド ADPCM 方式を採用している。聴覚特性を考慮し、高域は 2bit/sample、低域は 6bit/sample で符号化する。ADPCM 符号化の動作は前章で説明したものと同じであり、後ろ向きの適応予測と適応量子化を採用している。よって前章と同じ様に逆量子化差分 $dq(t)$ あるいは予測値 $se(t)$ の極性ランレングスに応じてテキストを埋め込むことができる。

3.2 広帯域 ADPCM 符号化方式へのテキスト埋め込み

前章と同じように予測値極性符号ランレングスと逆量子化差分極性ランレングスを基準として ADPCM 符号最下位を“ビットスチール”してテキストを埋め込む方式を検討する。高域は 2bit しかないので、テキストの埋め込みにより音質への影響が大きいと考えられる。一方、低域は 6bit の符号を出力するのでテキストを埋め込んでも影響は少ないと考えられるため、低域のみに前章と同じ方式でテキストを埋め込む。

3.3 シミュレーションによる基本性能評価

シミュレーションにより提案方式の基本性能を評価した。サブバンド ADPCM 方式は帯域が 7kHz であり、広範なオーディオ信号の符号化を想定している。よって評価サンプルは音声だけではなく、音楽も含めることとした。用いたサンプルは MPEG オーディオ符号化方式の音質評価にも用いられたヨーロッパ放送連盟 (EBU) により製作された SQAM (Sound Quality Assessment Material) [9] より英文読み上げ男女各 1 発話、ピアノ・ソロ 1 曲を用いた。さらに Oregon Graduate Institute より提供されている、音声強調方式評価用に用意された SpEAR (Speech Enhancement Assessment) [10] より女性歌唱 (アカペラ) 1 例を使用した。原音はすべて 16kHz 標本化、16bit である。

3.3.1 テキスト埋め込み率

前章と同様、予測値極性ランレングス最大値、最小値 se_kmax , se_kmin , および逆量子化差分値極性ランレングス最大値、最小値 dq_kmax , dq_kmin の各パラメータを変化させ、テキスト埋め込み率への影響を評価した。

電話帯域 ADPCM 符号化方式と同様に最小値を増加させると埋め込みテキスト量は単調減少し、また最大値を減少させるとテキスト量は単調減少する。また Type I の埋め込み量は Type II の埋め込み量より多い傾向にある。

3.3.2 テキスト埋め込みの音質への影響

図 14 に Type I のみ埋め込んだ場合の埋め込みテキスト率に対する segmental SNR を示す。逆量子化差分値極性ランレングス最大値 dq_kmax は 100 で固定し、最小値 dq_kmin を 1 から増加させて、Type I 埋め込みテキスト率を変化させた。ここで符号化前の音声を原信号とし、ノイズは符号化 (テキスト埋め込みを含む)、復号処理した音声信号と原信号との差とした。従って、テキストを埋め込まない音声の segmental SNR は単純に SB-ADPCM の符号化雑音対原音の比となる。図よりテキスト埋め込みを増加させても音質の劣化は極めて緩やかであることが分かる。また図 15 に埋め込みテキスト率に対する主観音質 (Mean Opinion Score) を示す。この実験では日本音響学会の研究用音声データベース [8] より ATR 音素バランス文読み上げ音声を男女各 1 発話用いた。被験者数は健聴男子 10 名とした。この図からも埋め込みテキスト率の増加に対する音質劣化の傾向は特に見られない。

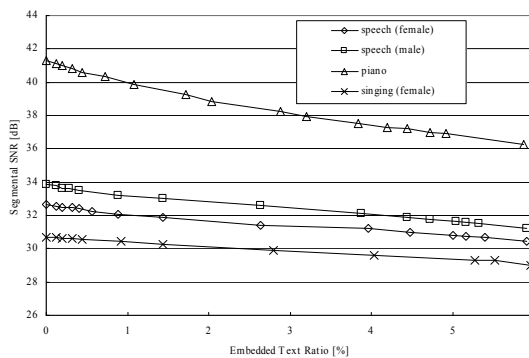


図 14 広帯域 ADPCM 符号化方式におけるテキスト (Type I) 埋め込み率対 Segmental SNR
Fig. 14 Embedded Text (Type I) Ratio vs. Segmental SNR for the SB-ADPCM Codec

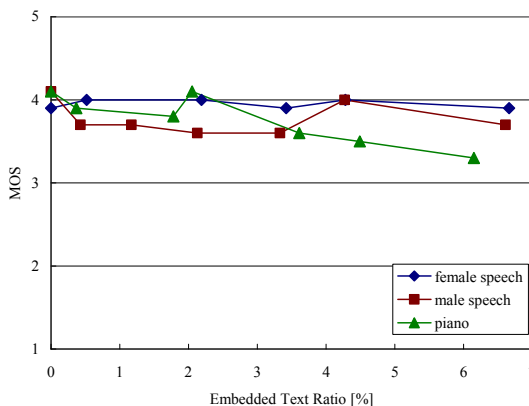


図 15 広帯域 ADPCM 符号化方式におけるテキスト (Type I) 埋め込み率対主観音質 (MOS)
Fig. 15 Embedded Text (Type I) Ratio vs. Mean Opinion Score for the SB-ADPCM Codec

3.3.3 伝送ノイズ・改ざんに対する耐性

図 16 にサブバンド ADPCM 符号にランダムにエラーを発生したときに Type I および II のテキストに生じたエラー率を示す。埋め込みに用いたパラメータは $dq_kmin = se_kmin = 1$, $dq_kmax = 100$, $se_kmax = 250$ である。ここでも Type II のテキストはビットエラーの大きな影響を受ける。一方、Type I にはほとんど影響が見られない。大体においてこの傾向の音源依存性はないが、piano の Type II テキストのみ 10% のエラーをサブバンド ADPCM 符号に加えるとエラー率が極めて低くなった。しかし今のところこの組み合わせ以外では例外は認められていない。

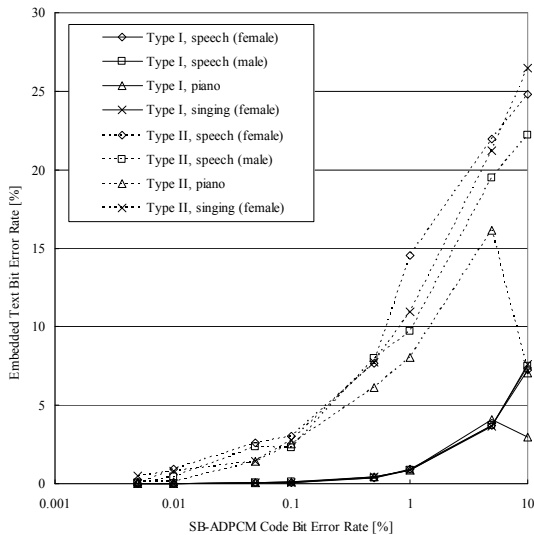


図 16 広帯域 ADPCM 符号化方式に対する埋め込みテキストのビットエラー耐性

Fig. 16 Bit Error Resilience of Embedded Text for the SB-ADPCM Codec

4. むすび

テキストの抽出に原音が不要な軽量、実時間処理可能なテキスト埋め込み方式を提案した。主に通信用途を想定しており、ADPCM 符号化を用いる通信用音声符号化標準 G.726 や 727, および 7kHz 帯域オーディオ符号化標準 G.722 に適用できる。予測値極性符号のランレングスに応じて圧縮符号最下位ビットにテキストを埋め込む方式に加え、逆量子化値極性符号のランレングスに応じて圧縮符号最下位ビットに埋め込む方式を提案し、さらにこの 2 種類の埋め込みを同時に行なう 2 重のテキスト埋め込み方式を提案した。前者は音声符号のビットエラーや改ざんにより高いエラー率を示す一方、後者は極めて低いエラー率を示し、対照的な特性を示すことを確認した。さらに極性符号のランレングス許容範囲を調整することで、2 種類の埋め込みテキスト合計量を制御できることもわかった。また、本方式によりデータを埋め込んでも音質への影響はほとんどないことを確認した。今後はより圧縮率の高い、後ろ向き適応予測を用いる音声符号化方式への展開を検討していく。

謝辞 本編は山形大学工学部、電子情報工学科における長年の卒業研究の成果をまとめたものである。度々の議論、ご支援頂いた電子情報工学科の中川研究

室の学生、職員一同に感謝する。

文 献

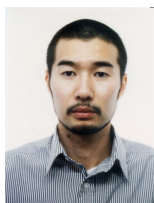
- [1] L. Boney, A. H. Tewfik, K. N. Hamdy, "Digital Watermarks for Audio Signals," Proc. IEEE International Conf. on Multimedia Computing and Systems, pp.473-480, Hiroshima, Japan, Jun. 1996.
- [2] "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," ITU-T Recommendation G.726, 1990.
- [3] "5, 4, 3 and 2 Bit/sample Embedded Adaptive Differential Pulse Code Modulation (ADPCM)," ITU-T Recommendation G.727, 1990.
- [4] 岩切宗利, 松井甲子雄, "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め込み," 情処学論, vol. 38, no. 10, pp.2053-2061, Oct. 1997.
- [5] 金澤英樹, 近藤和弘, 吉井亜美, 中川清司, "予測符号化音声信号用電子透かし方式の一検討," 電気関係学会東北支部連合大会, 2A4, p.19, Aug. 2000.
- [6] "7kHz Audio-Coding within 64 kbit/s," ITU-T Recommendation, 1988.
- [7] 清野智之, 近藤和弘, 中川清司, "広帯域音声符号化に適用可能な電子透かしの検討," 第 4 回情処学東北支部研究会, 02-4-A3, Mar. 2003.
- [8] 日本音響学会, "音響学会研究用連続音声データベース," 1991.
- [9] European Broadcasting Union, "SQAM(Sound Quality Assessment Material) - Recordings for Subjective Tests," April 1988.
- [10] CSLU, Oregon Graduate Institute of Science and Technology, "Speech Enhancement Assessment Resource (SpEAR) Database," http://ee.ogi.edu/NSEL_beta_release_v1.0.

(平成 xx 年 xx 月 xx 日受付)



近藤 和弘 (正員)

昭 57 早大・理工・電子通信卒。昭 59 同大学院理工学研究科修士課程了。同年、(株)日立製作所入社。同中央研究所で通信用音声信号処理システム、およびデジタルビデオ用画像符号化システムの研究開発に従事。平 4 (株) テキサス・インスツルメンツ筑波研究開発センター入社。平 8 米国テキサス・インスツルメンツ入社, Media Technologies Laboratory, Member of Technical Staff. 平 11 山形大・工・助教授。主として通信用音声・音響信号処理システムの研究教育に従事。工博。日本音響学会, IEEE 各会員。



金澤 英樹

平 12 山形大・工・電子情報卒. 同年 Vantan J College 入学. 平 13 テクニカル・ヒューマンリソース入社. 映像事業部. 在学中 ADPCM 符号化方式への電子透かし埋め込み方式の研究に従事.



吉井 亜美

平 13 山形大・工・電子情報卒. 同年 (株) 富士通ハイパーソフトテクノロジー入社. 富士通 (株) ソフトウェア事業本部へ出向. 在学中 ADPCM 符号化方式への電子透かし埋め込み方式の研究に従事.



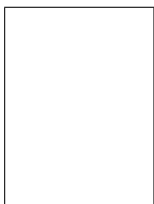
加藤 浩

山形大・工・電子情報在学中. ADPCM 符号化方式への電子透かし埋め込み方式の研究に従事.



清野 智之

平 15 山形大・工・電子情報卒. 同年 (株) エフコム入社. 在学中広帯域 ADPCM 符号化方式への電子透かし埋め込み方式の研究に従事.



中川 清司 (正員:フェロー)

昭 43 阪大・通信卒. 昭 45 同大大学院修士課程了. 同年電信電話公社 (現 NTT) 電気通信研究所入所. ミリ波導波管伝送方式および光ファイバディジタル中継伝送方式の研究実用化に従事. 平 8 より山形大・工・教授. 情報通信システムの研究・教育に従事. 工博 (阪大). 昭 52 年度本学会学術奨励賞, 昭 53, 平 2 英国 IEE 論文賞, 昭 59 および平 3 本学会業績賞. 著書「光増幅器とその応用」(オーム社, 共著) など. IEEE 会員. 本学会フェロー.

Abstract We investigated on a low-complexity digital data embedding method applicable to ITU-T standards G.726, G.727 ADPCM and G.722 Sub Band ADPCM. We proposed a method which embeds text data in the least significant bit of the ADPCM output code when the inverse-quantized difference signal sign bit run-length is within its predetermined range. The proposed embedded text was proven to be robust to bit errors, making it suitable to embed data which need to remain intact, *e.g.* copyright information, while its effect on the objective as well as subjective speech quality was found to be negligible.

Key words ADPCM, 7kHz Audio Coding, G.726, G.722, Digital Watermarks