

無限既約コードを用いたデルタ変調法について

中村八束* 西田秀次**

(昭和55年10月31日受理)

On the Delta Modulation Method by the Infinite Irreducible Codes

Yatsuka NAKAMURA and Hideji NISHIDA

Here is proposed a new coding method of waveforms utilizing the infinite irreducible codes in noiseless coding theory. This method, called here the II-code-method, is a generalization of the well-known delta modulation method.

As to the facility of this method, the good results are shown by applying the method to the actual voice waveforms.

First, it is proved that the method is optimal when the waveform is given by an additive random process, i. e., the average word length is equal to the entropy of the waveform.

Secondly, the notion of the infinite irreducible code is defined, and the infinite probability vectors, for which this code is optimal, are shown as examples.

Lastly, the method is applied to an actual voice waveform and it becomes clear that the method gives an enough approximation. As a voice waveform has a repeat structure with a pitch time, it is possible to compress the data steeply by using such a repeat structure and the II-code-method.

1 ま え が き

近年マイクロコンピュータの発達にもなって音声波形をはじめとする種々のアナログ波形をデジタルデータで蓄積し再生する技術が重要度を増している。その際必要とされる機能は、1. 波形が忠実に再生されること、2. 手法が簡単であること、3. データ圧縮率が高いこと、4. より多くの種類の波形に適用可能であること、等があげられる。従来、線形予測法 (LPC)¹⁾、デルタ変調方式 (DM)²⁾、差分パルス符号変調方式 (DPCM)³⁾ 等が知られている、しかし LPC 法は3については優れているが2、4の点で欠点があり、DM 法は2、3、4については優れているが1について欠点があり、DPCM 法は2、4については優れているが1、3については同時には満足できない。

* 情報工学教室 教授

** 大学院修士課程

我々はDM法のデータ圧縮率をそれほどひき下げないで改良する方法を提案する。これはよく知られている既約コード⁴⁾を無限個の通報の場合に拡張したコードを利用するもので比較的簡単な方法である。我々の方法は、波形が時間的に一様な加法過程で与えられるときはデータ圧縮率が最大であることが証明できる。だが実際の波形は理想的な加法過程ではないからかならずしも最適ではないが、くり返し構造等をうまく利用することで、ある場合にはLPC法よりもデータ圧縮率を高めることすら可能である。また音声以外の波形に対しても適用可能であるという一般性を有している。

2 確率過程としてのアナログ波形

今、波形の周波数帯域が制限されている等の理由により Shannon の標本化定理等が適用できて時間は離散的なものとして扱えると仮定する。今 n 番目の標本点での値を x_n とすると、

$$\Delta_n = x_{n+1} - x_n \quad (2.1)$$

なる差分は互いに独立で同じ確率密度分布 p_Δ を持つ確率過程になるとする。即ち x_n は時間に対して一様な加法過程をなすとする。

一般に精度に対する要求が ε の誤差で与えられている場合の波形の情報量は次のような計算で与えられる。即ち各 x_n の値域を長さ ε の区間 Δx に分ける。するとこの区間上の確率によって標本点 1 より標本点 n までのエントロピーを考えると、

$$H_\varepsilon^n \doteq - \sum_{\Delta x_1 \dots \Delta x_n} p(x_1, x_2, \dots, x_n) \Delta x_1 \Delta x_2 \dots \Delta x_n \log p(x_1, x_2, \dots, x_n) \Delta x_1 \Delta x_2 \dots \Delta x_n. \quad (2.2)$$

ここで $x_i \in \Delta x_i$ ($i=1, 2, \dots, n$) とし、 p は確率過程 x_n の同時確率密度とする。これは ε が十分小さければ、

$$(2.2) \doteq - \int \dots \int p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) dx_1 \dots dx_n - n \log \varepsilon \quad (2.3)$$

となる。第一項は Shannon の微分エントロピー⁵⁾である。ところで加法過程 x_n は Markov 過程であるから、

$$p(x_l, x_{l+1}, \dots, x_m | x_1, x_2, \dots, x_{l-1}) = p(x_l, x_{l+1}, \dots, x_m | x_{l-1}) \quad (2.4)$$

がいえる。従って式 (2.3) の第 1 項を $H(x_1, x_2, \dots, x_n)$ と書き、

$$\begin{aligned} H(x_l, x_{l+1}, \dots, x_m | x_1, x_2, \dots, x_{l-1}) \\ = - \int \dots \int p(x_1, \dots, x_m) \log p(x_l, \dots, x_m | x_1, \dots, x_{l-1}) dx_1 dx_2 \dots dx_m \end{aligned} \quad (2.5)$$

とすると、

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1}{n} H(x_1, \dots, x_n) \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \{H(x_1) + H(x_2, \dots, x_n | x_1)\} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \{H(x_1) + H(x_2 | x_1) + H(x_3, \dots, x_n | x_1, x_2)\} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \{H(x_1) + \sum_{i=2}^n H(x_i | x_1, x_2, \dots, x_{i-1})\} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \{H(x_1) + \sum_{i=2}^n H(x_i | x_{i-1})\}. \tag{2. 6}
\end{aligned}$$

ここで,

$$p(x_i | x_{i-1}) = p_d(x_i - x_{i-1}) \tag{2. 7}$$

がいえるから式(2. 6)は,

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \sum_{i=2}^n - \iint p(x_{i-1}, x_i) \log p_d(x_i - x_{i-1}) dx_{i-1} dx_i \right\} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \sum_{i=2}^n - \int p(x_{i-1}) \int p_d(x_i - x_{i-1}) \log p_d(x_i - x_{i-1}) dx_i dx_{i-1} \right\} \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \sum_{i=2}^n - \int p(x_{i-1}) \int p_d(t) \log p_d(t) dt dx_{i-1} \right\} \\
&= - \int p_d(t) \log p_d(t) dt. \tag{2. 8}
\end{aligned}$$

よって精度 ε の平均エントロピー \bar{H}_ε は,

$$\bar{H}_\varepsilon \doteq - \int p_d(t) \log p_d(t) dt - \log \varepsilon \tag{2. 9}$$

$$\doteq - \sum_{dt} p_d(t) dt \log p_d(t) dt \tag{2. 10}$$

となる。ここで dt は長さ ε の区間である。

今、過程 x_n に対してその近似過程 z_n を次のようにして作る。まず実数軸 R^1 の任意の可測分割 $\mathcal{A} = \{A_1, A_2, \dots\}$ を考え、この分割の要素の数は高々可付番であるとする。各 A_i の中に代表元 a_i を考え、この選択関数を r とする。即ち、

$$r(A_i) = a_i \in R^1. \tag{2. 11}$$

今、分割 \mathcal{A} と選択関数 r に依存して実数軸上の関数 $f(\cdot|\mathcal{A}, r)$ を次のように定義する。 f の値域は離散的である。

$$f(t|\mathcal{A}, r) = r(A) \quad (\text{ただし } t \in A \text{ とする}). \quad (2.12)$$

次に確率過程 d_n と z_n を、

$$\begin{cases} d_n = f(x_{n+1} - z_n | \mathcal{A}, r) \\ z_n = z_{n-1} + d_{n-1} \\ z_1 = f(x_1 | \mathcal{A}, r) \end{cases} \quad (2.13)$$

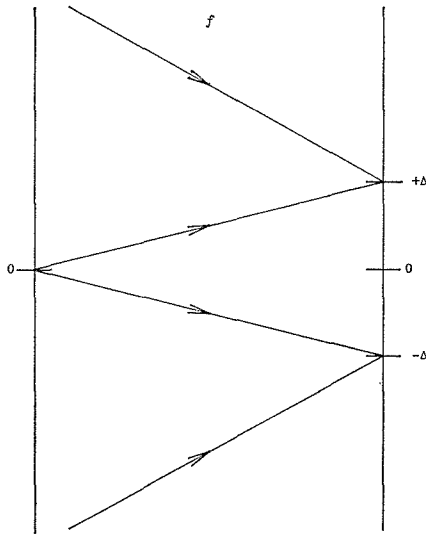


Fig. 1 : An example of $f(\cdot|\mathcal{A}, r)$

によって帰納的に定義する。この z_n なる過程を x_n の近似過程として用いる。 $\mathcal{A} = \{R_1^+, R_1^-\}$, f を Fig. 1 のように写像とすると, z_n は通常の DM 法による波形となる (R_1^+ は R_1 の非負部分, R_1^- は負部分をあらわす)。従って DM 法は式(2.13) で与えられる過程の特殊な場合にすぎないことになる。 f の形から見ても解るように近似としては非常に誤差の大きなものである。そこで分割の個数をもっと増やしてみる。

$$\mathcal{A} = \left\{ \left(n + \frac{1}{2} \right) D, \left(n + \frac{3}{2} \right) D : n = 0, \pm 1, \pm 2, \dots \right\} \quad (2.14)$$

とし, f は各区間の中点を与えるものとしよう (Fig. 2)。この方法は従来の DM 法に比べてビット数は増えるが, lD ($l=0, \pm 1, \pm 2, \dots$) という値を使うことができるので slope overload⁶⁾ をなくすることができる, また 0 という値を持つために granular noise⁷⁾ をなくすることもできる。この方法がたしかに誤差 ε の近似過程になっていることは ($D=\varepsilon$ とお

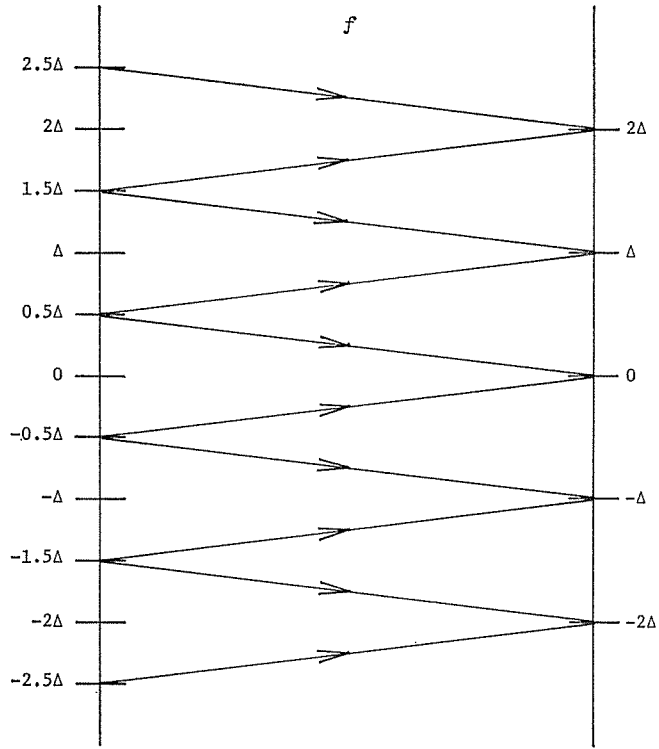


Fig. 2 : An example of $f(\cdot|\mathcal{A}, \gamma)$

いて),

$$\begin{aligned}
 & |x_{n+1} - z_{n+1}| \\
 &= |x_{n+1} - z_n - d_n| \\
 &= |(x_{n+1} - z_n) - f(x_{n+1} - z_n|\mathcal{A}, \gamma)| \leq \varepsilon
 \end{aligned} \tag{2. 16}$$

が式(2. 13) よりいえるから, z_n は x_n の誤差 ε の過程になっている. 式(2. 14)で \mathcal{A} を与えたときのこの近似方法は, 実はよく知られている DPCM 法でも用いている. 従って式(2. 13)のような枠組みの中で DPCM 法と DM 法は同種のものであるといえる.

式(2. 14)で \mathcal{A} と f を与えると, f は良い性質を持っている. それを次に示そう.

定理1 分割 $\mathcal{A} = \{A_i; i=0, \pm 1, \pm 2 \dots\}$ に対し,

(C) $t \in A_i \iff$ ある $t_0 \in A_0$ に対し

$$t = t_0 + \gamma(A_i) = t_0 + a_i$$

(D) $a_i = i\Delta$ (適当な Δ に対して)

がいえるならば,

$$f(t+a_i|\mathscr{A}, \gamma) = f(t|\mathscr{A}, \gamma) + a_i \quad (2.17)$$

が全ての実数 t について成立する.

証明 $t \in A_0$ ならば式(2.17)の右辺 $= f(t|\mathscr{A}, \gamma) + a_i = a_i$. また式(2.17)の左辺 $= f(t+a_i|\mathscr{A}, \gamma) = a_i$ が(C)よりいえる. よって式(2.17)が成立する.

$t \in A_j$ のときは(C)より, ある $t_0 \in A_0$ に対して,

$$t = t_0 + a_j.$$

よって

$$\begin{aligned} & f(t+a_i|\mathscr{A}, \gamma) \\ &= f(t_0+a_j+a_i|\mathscr{A}, \gamma) \\ &= f(t_0+(i+j)A|\mathscr{A}, \gamma) \\ &= f(t_0+a_{i+j}|\mathscr{A}, \gamma) \\ &= a_{i+j} \\ &= f(t_0+a_j|\mathscr{A}, \gamma) + a_i \\ &= f(t|\mathscr{A}, \gamma) + a_i. \end{aligned} \quad (2.18)$$

Q. E. D.

明らかに式(2.14)で与えられる \mathscr{A} と f は条件(C)と(D)を満たしている. 従って式(2.17)が成立する. 式(2.17)が成立するとき, 次の定理が成立する.

定理2 式(2.17)が \mathscr{A} や f について成り立つとき式(2.13)で与えられる近似過程 z_n について,

$$z_n = f(x_n|\mathscr{A}, \gamma) \quad n = 1, 2, \dots \quad (2.19)$$

が成立する.

証明 $n=1$ のときは明らか. $n>1$ のときは,

$$\begin{aligned} z_n &= z_{n-1} + d_{n-1} \\ &= z_{n-1} + f(x_n - z_{n-1}|\mathscr{A}, \gamma) \\ &= z_{n-1} + f(x_n + (-1 \cdot z_{n-1})|\mathscr{A}, \gamma) \\ &= z_{n-1} + f(x_n|\mathscr{A}, \gamma) - z_{n-1} \\ &= f(x_n|\mathscr{A}, \gamma) \end{aligned}$$

Q. E. D.

定理2より, 式(2.17)が成立するような近似過程は, 最初の波形を差分をとらずに直接精度 Δ で近似したものに等しいことになる. 従って $\Delta = \varepsilon$ であれば, 式(2.13)のように差分過程 d_n を使って与えられる z_n は元の波形と誤差 ε 以内にあることになる. 式(2.2)で H_ε^n を計算した時に行なったような x_n の値域を長さ ε の区間 Δx に分ける近

似方法を見れば、実は H_s^p の計算は近似過程 z_n のエントロピーを計算していることになる。よって z_n の平均エントロピーもまた式(2. 10)で与えられ、 z_n のエントロピーは差分過程 d_n の1次エントロピーに等しいことになる。(差分過程 d_n は一般には独立過程ではないので、1次エントロピーで与えられることは一見妙であるが、実は今までのエントロピーの計算に式(2. 2), (2. 3), (2. 13)の各段階で誤差が入っている。)

以上から平均長が差分過程 d_n の1次エントロピーに等しいような符号化の方法があれば、それは元の波形 x_n の ε 近似として最適か最適に近いものであることになる。

3 無限既約コード

雑音の無い通信路における符号化として Huffman code はよく知られており、語頭(prefix)が他の語になっていないという既約性の過程を満たすコードの中では Huffman code は最適であることが知られている⁸⁾。このような既約符号をアルファベットが無限にある場合に拡張できないだろうか、そのために Kraft の不等式を拡張する。

Kraft の不等式 i 番目の通報 m_i ($i=1, 2, \dots$) に D 進コード列を割り振るとする。 D 進コード列の長さを l_i としてこのときこの長さを持つ既約符号系が存在する必要十分条件は、

$$\sum_{i=1}^{\infty} D^{-l_i} \leq 1 \quad (D \geq 2) \quad (3. 1)$$

が成立することである。

証明 必要性：今、長さが l のコードを持つ通報の数 M_l を考える。 $M_l \leq D^l$ であることは明らかであるが、更に D^l ケの符号のうち語頭が l より短い符号と一致するものは M_l ケの中には現れない。長さ l_i の符号を語頭に持つものの数は、

$$D^{l-l_i} \quad (3. 2)$$

よってこれを全部加えたものの数を D^l より引き去ると、

$$0 \leq M_l \leq D^l - \sum_{l_i < l} D^{l-l_i} \quad (3. 3)$$

が得られる。よって任意の l について、

$$\sum_{l_i < l} D^{-l_i} \leq 1. \quad (3. 4)$$

ここで l は任意だから式(3. 1)が得られる。

十分性： $\mathfrak{M}_k = \{m_1, \dots, m_k\}$ として、 \mathfrak{M}_k に対する既約符号系 \mathcal{S}_k が存在すると仮定する。もし \mathcal{S}_{k+1} における通報 m_i ($i \leq k$) の符号と \mathcal{S}_k における m_i の符号が等しい(仮定 E)とき、

$$\mathcal{S} \equiv \bigcup_{k=1}^{\infty} \mathcal{S}_k \quad (3. 5)$$

とおけば無限個の通報 m_1, m_2, \dots に対する既約符号系 \mathcal{S} が得られる。

そこで任意の k について仮定 E を満たすような符号系 \mathcal{S}_k が存在することを帰納法によって示す。あらかじめ、

$$0 < l_1 \leq l_2 \leq \dots \quad (3. 6)$$

としておいて一般性を失わない。 $k=2$ の時は長さ l_1 と l_2 は正であるので、

$$\mathcal{S}_2 \begin{cases} m_1 \text{ の符号 : } \underbrace{d_1 d_1 \cdots d_1}_{l_1} \\ m_2 \text{ の符号 : } \underbrace{d_1 d_1 \cdots d_2}_{l_1} \underbrace{d_1 d_1 \cdots d_1}_{l_2 - l_1} \end{cases}$$

とおけば \mathcal{S}_2 は既約符号系である。今、 k まで仮定 E を満たしながら符号系 \mathcal{S}_k が存在したとする。 $k+1$ のとき、

$$\sum_{i=1}^{k+1} D^{-l_i} \leq 1 \quad (3. 7)$$

であるから、 $l_{k+1} > 0$ より

$$\sum_{i=1}^k D^{-l_i} < 1. \quad (3. 8)$$

今長さ l_{k+1} の符号の中で語頭が \mathcal{S}_k の中で現れないものは、必要性のところの議論より、

$$D^{l_{k+1}} - \sum_{i=1}^k D^{l_{k+1} - l_i} \quad (3. 9)$$

だけの個数存在する。これは、

$$= D^{l_{k+1}} \left\{ 1 - \sum_{i=1}^k D^{-l_i} \right\} > 0 \quad (3. 10)$$

が式 (3. 8) を使っていえるから、従って少なくとも 1 つ以上長さ l_{k+1} のものが存在する。これを m_{k+1} の符号として \mathcal{S}_k に付け加えたものを \mathcal{S}_{k+1} とすれば、明らかに \mathcal{S}_{k+1} は仮定 E を満たす既約符号系である。

Q. E. D.

無限既約符号系の種々の例を Table 1 に示す。インデックスを並び換えると Ex. 2, 3 のように負の整数を許すこともある。Ex. 1, 2, 3 は符号に規則性があるため、符号化と復号化の作業は表を使わなくとも簡単なアルゴリズムによって行なえる。例えば Ex. 3

Table 1: Example of infinite irreducible code systems

Ex. 1	Ex. 2	Ex. 3	Ex. 4
$m_1 : 0$	\vdots	\vdots	$m_1 : 10$
$m_2 : 10$	$m_2 : 1100$	$m_3 : 1100$	$m_2 : 11$
$m_3 : 110$	$m_1 : 100$	$m_2 : 100$	$m_3 : 000$
$m_4 : 1110$	$m_0 : 0$	$m_1 : 00$	$m_4 : 010$
\vdots	$m_{-1} : 101$	$m_{-1} : 01$	$m_5 : 0110$
	$m_{-2} : 1101$	$m_{-2} : 101$	\vdots
	\vdots	$m_{-3} : 1101$	\vdots
		\vdots	

の場合のアルゴリズムは1の数を数えれば良いので簡単である。それに対して Ex. 4 のようなものは特に規則性が無い。今、無限既約符号を II-code と呼び、規則性のある(アルゴリズムで与えられる) II-code を algorithmic II-code と呼ぶことにする。我々が実際に取り扱うことが出来るのは algorithmic II-code のみである。

さて各通報に確率 p_i ($i=1, 2, \dots$) が与えられているものとし、 $p_1 \geq p_2 \geq \dots$ を仮定する(もしある番号から先が全て0ならば有限の場合に帰着する)。符号の長さの列が $\{l_i\}$ の II-code 系があるとき、その code の平均の長さ \bar{l} は、

$$\bar{l} = \sum_{i=1}^{\infty} l_i p_i \quad (3. 11)$$

で与えられる。符号系を変えた時 \bar{l} がいつ最小になるかを考える。

定理3 通報の確率が $\{p_i\}$ の時、もし $-\log_D p_i$ が全て整数ならば、

$$l_i = -\log_D p_i \quad (i = 1, 2, \dots) \quad (3. 12)$$

なる長さを持つ II-code が存在し、それは \bar{l} を最小にし、かつその最小値は $\{p_i\}$ のエントロピー

$$-\sum_{i=1}^{\infty} p_i \log p_i \quad (3. 13)$$

に一致する。

証明

$$\sum_{i=1}^{\infty} D^{-l_i} = \sum_{i=1}^{\infty} p_i = 1 \quad (3. 14)$$

であるから無限の場合の Kraft の不等式より、長さ l_i の II-code が存在することは明らか。

次に $\sum_{i=1}^{\infty} D^{-s_i} \leq 1$ なる長さの列 $\{s_i\}$ を持つ他の任意の II-code を考える(全ての i について $s_i > 0$)。このとき $q_i = D^{-s_i}$ とおくと $\sum_{i=1}^{\infty} q_i \leq 1$ であり、

$$\sum_{i=1}^{\infty} s_i p_i = - \sum_{i=1}^{\infty} p_i \log q_i. \quad (3. 15)$$

ところで良く知られたエントロピーの不等式⁹⁾より,

$$\sum_{i=1}^{\infty} p_i \log \frac{p_i}{q_i} \geq \left(\sum_{i=1}^{\infty} p_i \right) \log \frac{\left(\sum_{i=1}^{\infty} p_i \right)}{\left(\sum_{i=1}^{\infty} q_i \right)} = - \log \left(\sum_{i=1}^{\infty} q_i \right) \geq 0 \quad (3. 16)$$

がいえる。よって,

$$\sum_{i=1}^{\infty} p_i \log p_i \geq \sum_{i=1}^{\infty} p_i \log q_i \quad (3. 17)$$

であり, 式 (3. 15) と合わせて,

$$\sum_{i=1}^{\infty} s_i p_i \geq - \sum_{i=1}^{\infty} p_i \log p_i = \sum_{i=1}^{\infty} l_i p_i. \quad (3. 18)$$

即ち $\{p_i\}$ のエントロピーが $\{l_i\}$ の平均長に等しく, それが最小値を与えることがわかる。

Q. E. D.

よって Ex. 1 のような符号系は確率が $p_i = 2^{-i}$ ($i=1, 2, \dots$) のような確率の通報に対して最適な符号となり, Ex. 2 では,

$$p_i = \begin{cases} \frac{1}{2} & i=0 \text{ の時} \\ \frac{1}{2^{|i|}} \cdot \frac{1}{4} & \text{その他} \end{cases} \quad (3. 19)$$

のような確率の通報に対して最適な符号となる。 $-\log_D p_i$ が必ずしも整数とならない時には, どのような II-code が最適となるかは興味のある問題である。

前節での議論とこの節の定理とを合わせると, $-\log_D p_i$ が整数となるような誤差 ε の近似過程が加法過程であれば, そのエントロピーにちょうど等しい平均長を持つ II-code が存在することになる。即ちこの場合は最もデータ圧縮率の大きい方法が II-code によって与えられることになる (許容誤差 ε の時)。 II-code を使って差分過程を符号化する方法を II-code 法と呼び次節で具体的な波形に適用し, その得失を評価する。

4 音声データへの適用と他の方法との比較

II-code 法を $|a|, |i|, |u|, |e|, |o|$ の 5 母音データに適用しシュミレーションを行なった。なおここで用いた各音声データは 8 KHz でサンプリングしたものである。 II-code は Ex. 2 に示した algorithmic II-code である。 II-code 法の近似精度能力とデータ圧縮率を調べるために, 比較として DM 法と LPC 法の 1 種である PARCOR 法¹⁰⁾を取り

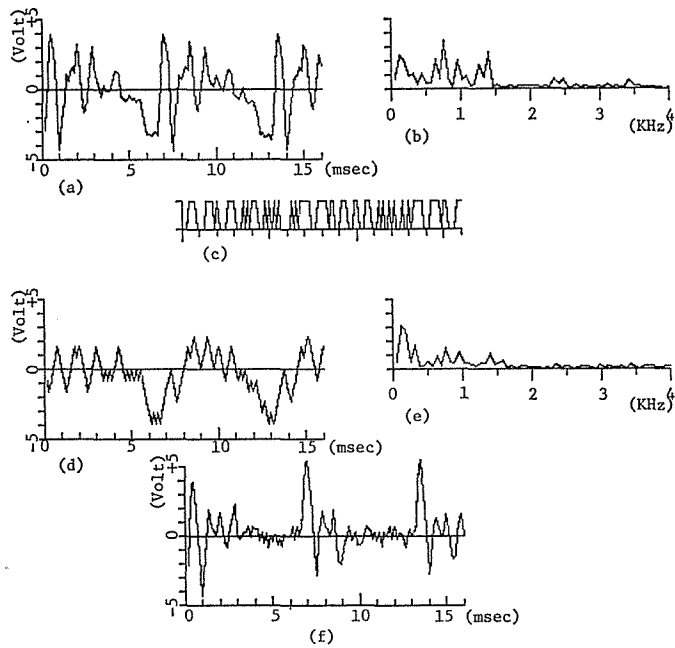


Fig. 3 : (a) An actual waveform of /a/ (b) A speech spectrum of (a)
 (c) A code sequence by DM (d) A predicted waveform by DM ($\Delta=20$)
 (e) A speech spectrum of (d) (f) An error sequence between (a) and (d)

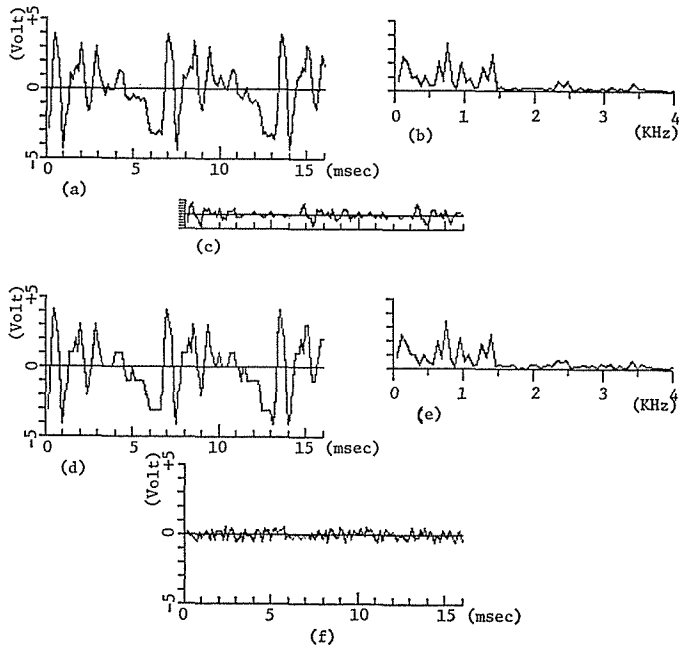


Fig. 4 : (a) An actual waveform of /a/ (b) A speech spectrum of (a)
 (c) A code sequence by II-code (d) A predicted waveform by II-code ($\Delta=26$)
 (e) A speech spectrum of (d) (f) An error sequence between (a) and (d)

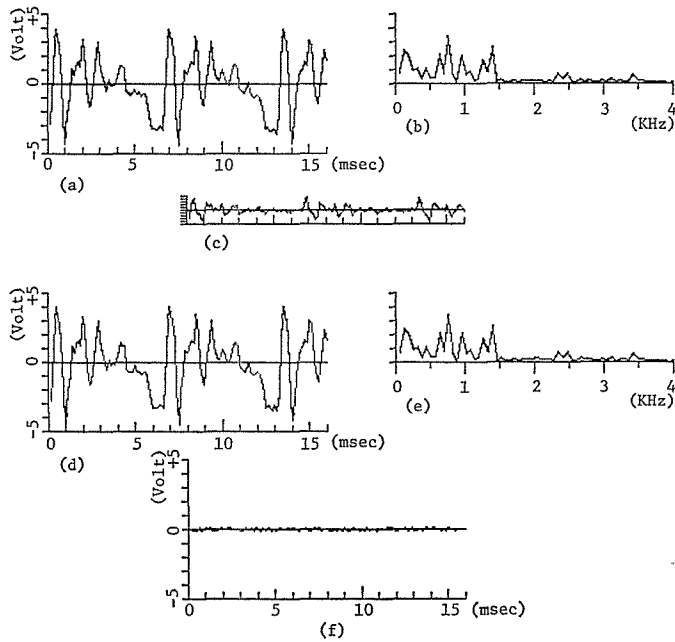


Fig. 5 : (a) An actual waveform of /a/ (b) A speech spectrum of (a)
 (c) A code sequence by II-code (d) A predicted waveform by II-code ($\Delta=6$)
 (e) A speech spectrum of (d) (f) An error sequence between (a) and (d)

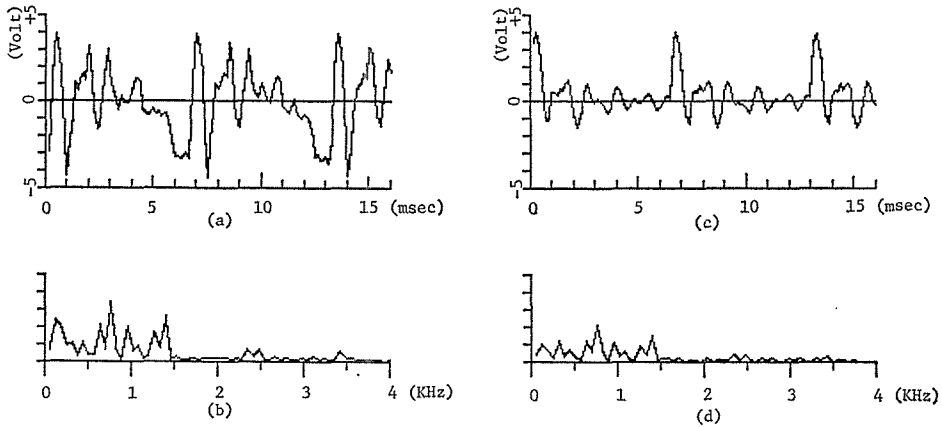


Fig. 6 : (a) An actual waveform of /a/ (b) A speech spectrum of (a)
 (c) A predicted waveform by PARCOR (d) A speech spectrum of (c)

上げてみる。

まず近似精度についてであるが DM 法の場合は Δ の値が一定であるために Δ を大きくすると granular noise, 小さくすると slope overload が相反して生じ Δ の値の決定が難しい (Fig. 3). だが II-code 法においては, Δ の値は $l\Delta$ ($l=0, \pm 0.5, \pm 1.5, \dots$) と可変であるので slope overload は完全に取り除くことができ, 差分 0 が許されているので

granular noise も取り除かれることより、より忠実に原音データが再現される (Fig. 4, 5). PARCOR 法においては、フレームごとにエンベロープを決定するパラメータ (k_1, \dots, k_{12}) と音源信号に関するパラメータ (A : 振幅, P : ピッチ周期, V : 有声度) を抽出し再現するものであるが、音源信号として有声音部ではピッチ周期に同期したインパルス列、無声音部では白色雑音 (擬似ランダム信号) を用いているため、どうしても再生データの劣化は避けられないが、周波数成分は忠実に再生されている (Fig. 6).

ここで近似精度の評価基準として最小 2 乗誤差 $\left(\sqrt{\frac{1}{N} \sum_{i=1}^N |f_i - \hat{f}_i|^2}\right)$, f_i : 原音データ値, \hat{f}_i : 予測値, $N=128$) を導入して |a|, |i|, |u|, |e|, |o| の各 5 母音に対して比較したものが Table 2 である.

Table 2: A comparison of fidelities of 3 coding methods

	/a/	/i/	/u/	/e/	/o/
DM ($d=20$)	38.72	22.01	12.29	37.25	46.97
II-code ($d=26$)	7.49	7.71	7.25	8.96	7.63
PARCOR	58.67	71.73	102.19	60.42	78.62

次にデータ圧縮率についてであるが、DM 法の場合は 1 つの音声信号に対して最低限の 1 ビットを与えるもので、1 秒間の原音データ記憶に 1 ビット $\times 8 \times 10^3 = 8000$ ビット必要となる。II-code 法の場合、例として |a| では $d=26$ で $\bar{l}=2.62$ ビットであるので、同様 1 秒間の原音データ記憶に 2.62 ビット $\times 8 \times 10^3 = 20960$ ビット必要となる。PARCOR 法の場合、PARCOR 係数の効果的な符号割り当てとして不均一ビット配分が有効な手段であることはよく知られている¹¹⁾。その一例を Table 3 に示す。今 8 KHz でサンプリングした原音データを 1 フレーム = 16 msec で分析するものとして、1 秒間で 62.5 フレーム必要となり、必要ビット数は 114 ビット $\times 62.5 = 7125$ ビットとなる。

Table 3: An example of PARCOR coefficients by variable coding

	PARCOR coefficients					source			total
	k_1	k_2	$k_3 \sim k_6$	$k_7 \sim k_{10}$	$k_{11} \sim k_{12}$	A	P	V	
bits	12	9	8	7	6	7	7	7	114

原音データは 1 サンプルに対して 8 ビット与えているので、1 秒間では 8 ビット $\times 8 \times 10^3 = 64000$ ビット必要でこれから考えると DM 法, PARCOR 法は大幅に情報圧縮されている。だが近似精度から考えると DM 法, PARCOR 法は情報圧縮のためどうしても再生データの劣化は避けられない。一方 II-code 法は近似精度がよいのでそこに注目し、データ圧縮率を上げるためくり返し構造を利用した方法を次節で述べる。

5 くり返し構造の圧縮

一般に楽器音、音声等の波形にはピッチ周期を持つかなりはっきりしたくり返し構造が存在する。例えば Fig. 7 は母音|a|の0.5秒間の波形であるが、発声の始めと終わりではその波形に18.89の平均2乗誤差しか生じていない。この誤差はPARCOR法を使った近似

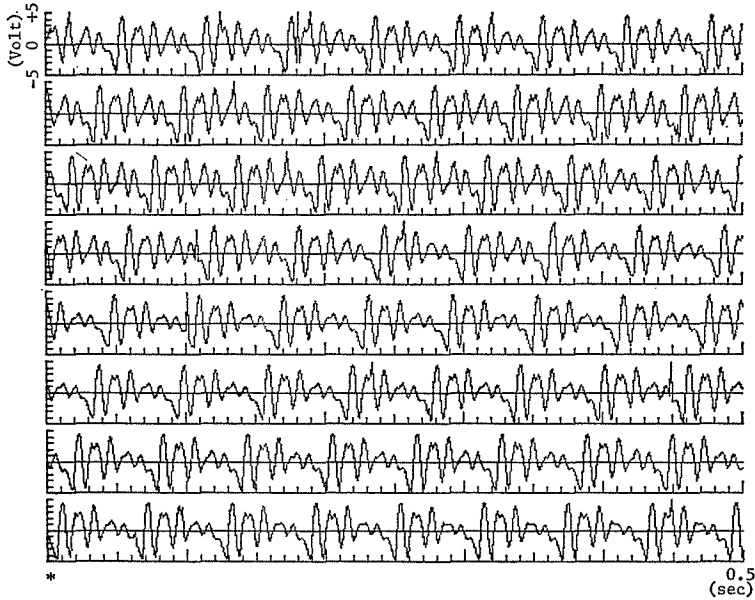


Fig. 7: A waveform of /a/ in 0.5 sec

誤差58.67よりはるかに小さい。PARCOR法はこのくり返しをピッチ周期情報としているが、くり返し構造の直接的利用では無いため、データ圧縮率もそれほど向上しない。例えば|a|ではくり返し構造を直接に利用した場合はII-code法を併用すると1秒間の原音データが159ビットにまで圧縮される。それはFig.7の波形においてピッチ周期は7.62msecであるので1ピッチ周期分の符号化により $\bar{I}=2.607$ ビットが得られるので全体では 2.607 ビット $\times 7.62 \times 10^{-3} \times 8 \times 10^3 \approx 159$ ビットとなる。

このような圧縮方法は波形を実際に観察してはじめて出来るものであり、real timeの圧縮処理で可能かどうかは不明であるが、少なくとも音声記憶には積極的に利用できるものである。例えば子音|sa| (Fig.8)については、A部分(子音部)、B部分(過渡部)、C部分(母音定常部)、D部分(母音減衰部)と分けた時、C部分にはピッチ周期を持ったくり返し構造が存在するので1ピッチ周期分の波形をC'とすると $C=23C'$ とあらわすことができる。従ってまた子音データ|sa|全体は(A, B, 23C', D)という形で表現することができる。このときII-code法による近似波形はFig.9のようになる。この時必要ビット数はわずかに3050ビットにすぎない。このようにII-code法におけるくり返し構造の利用は音声波形の圧縮には強力であると思われる。

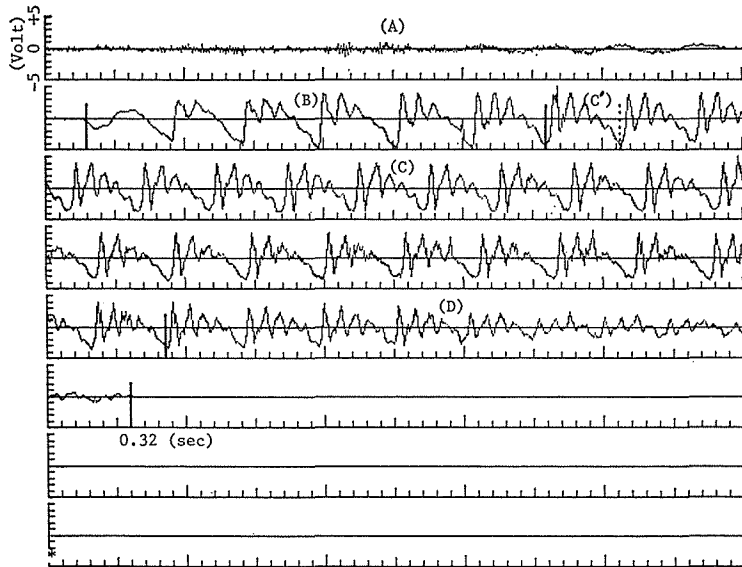


Fig. 8 : A waveform of /sa/ in 0.32 sec

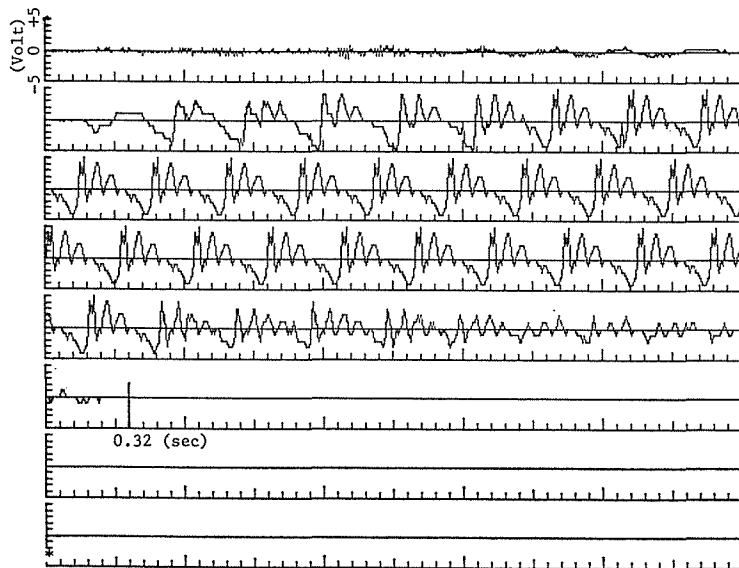


Fig. 9 : A predicted waveform by II-code utilizing the repeat structure of Fig. 8

6 おわりに

本論文では II-code を実際の音声波形に適用し, granular noise, slope overload 等のひずみの軽減をはかり, 十分な近似精度が得られる上に, くり返し構造との併用によってデータ圧縮率がより上げられることを示した.

最後に、東京エレクトロンの奥田博之氏、および信州大学工学部 花谷圭人助教授には本研究へ御協力と御指導をいただいたことを感謝する。

参 考 文 献

- 1) 中田和男：音声，p. 87，コロナ社，1977
- 2) R. W. Schafer・L. R. Rabiner：Digital Representations of Speech Signals, Proc. IEEE, 63-1 (1975), p. 665
- 3) 文献2), p. 665
- 4) F. M. Reza：確率・情報・コード(鶴見茂監訳)，p. 208，共立出版，1973
- 5) C. E. Shannon・W. Weaver：The Mathematical Theory of Communication, p. 87, Univ. of Illinois Press, 1949
- 6) 文献1), p. 65
- 7) 文献1), p. 65
- 8) R. M. Fano：情報理論(宇田川銈久訳)，p. 79，紀伊国屋書店，1965
- 9) M. S. Pinsker：Information and Information Stability of Random Variables and Processes (Translated and edited by A. Feinstein), p. 21, Holden-Day Inc., 1964
- 10) 文献1), p. 93
- 11) 安居院猛・中嶋正之：コンピュータ音声処理，p. 70，産報出版，1980