

高 能 率 符 号 化 の 動 向 *

山 崎 芳 男 **

(早稲田大学理工学研究所)

1. ま え が き

広帯域音響信号を対象とした高能率符号化の研究は早くから行われていたが実用に供されているのは、比較的圧縮率の低い非一様量子化やADPCMが一部に導入されている程度であった。最近人間の聴覚の特徴を巧みに利用して1チャンネル当たり 64 ~ 128kbit/s程度で符号化する様々な高能率符号化のアルゴリズムが提案されている。

高能率符号化は文字どおり何等かの手段により信号をできるだけ能率良く伝送するための符号化技術である。高能率符号化には信号の情報量に着目してその冗長度を除去することにより無歪で伝送容量の節約を図る方法と人間の感覚を利用して歪をできるだけ感知されにくくするよう工夫した圧縮とに大別される。一般にいったんデジタル化した信号をどこまで圧縮できるかといった側面のみから据えられがちであるが、高能率符号化はむしろ伝送路は節約したうえ、質の向上を図ろうという前向きな技術として捉えるべきものである。

本稿では音響信号を中心に高能率符号化技術の基本的事項とその最近の動向について解説する。

2. 信号の情報量に着目した高能率符号化

2.1 情報量と伝送容量

シャノンは信号の平均情報量 H と伝送路の伝送容量 C の関係が $H < C$ ならば無歪伝送が可能であり、 $H > C$ ならば歪を限りなく $C - H$ に近づけて伝送することが可能な符号化法が必ず存在するという美しい定理を導いている。

音響信号の平均情報量は信号や収音方法にも依存するので正確に把握するのは難しいがその情報量は1チャンネル当たり 120 ~ 400kbit/s以下であることが分かっている。これは少なくとも 120 ~ 400kbit/s の伝送容量で無歪伝送が期待できることを示している¹⁾。

ところで信号の平均情報量と伝送容量が等しい状態、

すなわち無歪伝送の限界をレート歪限界 (Rate Distortion Boundary) と呼ぶ。この限界を超えてある程度歪を許容すると大幅な伝送路の節約が可能である。

2.2 音響信号の情報量

人間が通常聴取する音響信号の周波数帯域は 15 ~ 20 kHz, ダイナミックレンジは 80 ~ 120dB に及ぶ。この信号を基本的な手法でそのままデジタル化するには 30 ~ 50kHz の標準化周波数と 14 ~ 20bit の量子化ステップが必要となる。すなわち、1チャンネル当たり 450 kbit/s ~ 1Mbit/s の伝送容量を必要とすることになる。

表 - 1 に標準化周波数 44.1kHz, 16bit 一様量子化でデジタル化したオーケストラ演奏の一部を示す。各標本の量子化値が 16bit 一様量子化で表現できる 2^{16} すなわち 65,536 の異なる値に等確率で無秩序に発生していれば、この信号を伝送するには必然的に $44.1k \times 16 = 705.6kbit/s$ の伝送容量を必要とする。しかるに明らかに "0" "1" の生起にはある種の規則性が見られる。図 - 1 にデジタル化された楽音信号の振幅分布と上位 12bit の使用された頻度、生起率を百分率で示す。また、図 - 2 にそのパワースペクトルを示す。このようにデジタル化された広帯域音響信号の各ビットの使用状況やレベル分布にはかなり偏りがある。更に楽音のスペクトル構造は高域で大幅に低下している。この偏りは前述のように信号の持つ冗長度に起因するものであり、これを利用して伝送路の節約、すなわちこの冗長度を何等かの方法で軽減することにより高能率符号化が可能である。

表 - 1 の各ビットの 0, 1 が独立にそれぞれ $p, q = 1 - p$ の確率で発生すると仮定すると、その平均情報量は各ビット当たり $-(p \log_2 p + q \log_2 q)$ [bit] となる。この方法で各ビットの平均情報量の総和を求めると 8 ~ 9bit/標本となる。これは各ビットの 0, 1 が独立に発生すると仮定した結果であるが、表 - 1 から観察されるように、各ビット特に上位のビットは独立に生起すなわち記憶の無い確率過程に属しているとは言い難い。そのうえ各ビット間の相関も観察される。そこで 16bit で一様量子化された信号を多重マルコフ過程として捉えその情報量を計算すると 1 標本当たり 4 ~ 6bit となる。これは後述のように冗長度を取り除いたうえエントロピー符

* Review of high efficiency coding.

** Yoshio Yamasaki (Science and Engineering Laboratory, Waseda University, Tokyo, 169)

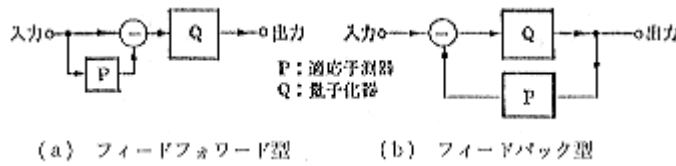


図 - 3 予測符号化の構成

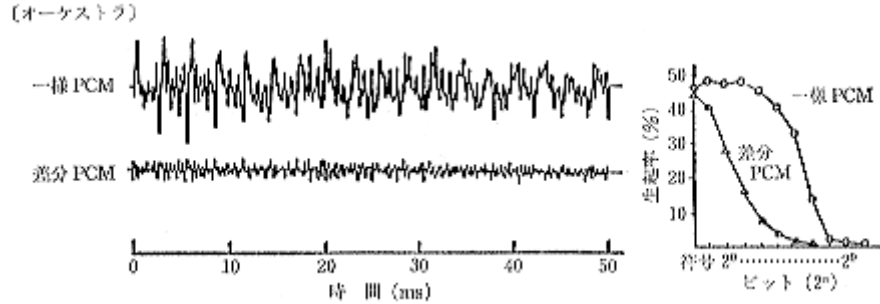


図 - 4 差分 PCM による波形と各ビットの生起率

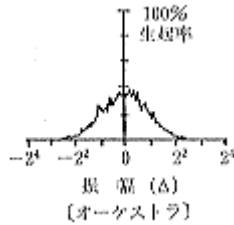


図 - 5 予測符号化における残差成分の振幅分布

報源シンボルには短い符号を、出現頻度の低いシンボルには長い符号を与える不等長符号化方法である。電信におけるモールス符号は典型的なエントロピー符号化である。エントロピー符号化としてハフマン符号、シャノンの符号化、ランレングスリミテッド符号化などが知られている。

ハフマン符号の符号化方法は各シンボルを出現確率の高い順に並べ、次に出現頻度の低い方から二つの組み合わせさせて符号 0, 1 を与え、その出現確率を加算したうえで、出現確率の高い順に並べ最初と同じ操作を繰り返して各シンボルの符号を決定する。その結果最も出現確率の高いシンボルには最短、出現確率の最も低いシンボルに最長の符号が割り当てられる。

前述のオーケストラ演奏の予測符号化の結果 5bit に納まった残差信号(図 - 5)に変調規則が簡単に復調の容易なハフマン符号を適用した例を表 - 2 に示す。1 標本当たりハフマン符号の平均情報量は 2.979bit で信号の情報量 2.94bit にごく近い値にまで達している。すなわちレートディストーション限界にほぼ一致している。ただしエントロピー符号は不等長符号であるので、伝送に使用する場合は大容量の一時的記憶素子と辞書(符号割当表)が必要であり、変復調には長い処理時間を要す

表 - 2 各振幅出現確率とハフマン符号

振幅	出現確率 $P(n)$	情報量 $P \log_2 P$	ハフマン符号	ビット数 M_n	$M_n \times P(n)$
1	0.4290	0.5238	0	1	0.4290
2	0.1716	0.4364	110	3	0.5148
3	0.0977	0.3278	100	3	0.2930
4	0.0644	0.2547	1011	4	0.2574
5	0.0460	0.2042	11110	5	0.2298
6	0.0345	0.1677	10101	5	0.1727
7	0.0259	0.1402	11111	6	0.1613
8	0.0215	0.1189	111011	6	0.1287
9	0.0175	0.1019	111001	6	0.1047
10	0.0144	0.0881	101000	6	0.0864
11	0.0120	0.0767	1111100	7	0.0841
12	0.0101	0.0671	1110100	7	0.0708
13	0.0086	0.0589	1110001	7	0.0609
14	0.0073	0.0519	1010011	7	0.0512
15	0.0063	0.0458	11111011	8	0.0500
16	0.0054	0.0405	11101011	8	0.0429
17	0.0046	0.0358	11100001	8	0.0368
18	0.0040	0.0316	11100001	8	0.0317
19	0.0034	0.0279	10100100	8	0.0272
20	0.0029	0.0245	111110100	9	0.0262
21	0.0025	0.0215	111010100	9	0.0223
22	0.0021	0.0187	101001011	9	0.0190
23	0.0018	0.0162	1111101011	10	0.0178
24	0.0015	0.0139	1111101010	10	0.0148
25	0.0012	0.0118	1110101010	10	0.0122
26	0.0010	0.0099	1010010101	10	0.0099
27	0.0008	0.0080	11101010111	11	0.0086
28	0.0006	0.0063	11101010110	11	0.0065
29	0.0004	0.0047	10100101001	11	0.0046
30	0.0003	0.0032	101001010001	12	0.0032
31	0.0001	0.0016	101001010000	12	0.0015
32	0.0	0.0	—	0	0.0
$\sum P(n)$	$= 1.000$	$\sum P \log_2 P$			$\sum (M_n \times P(n))$
		$= 2.9400$	(bit)		$= 2.9790$
					(bit)

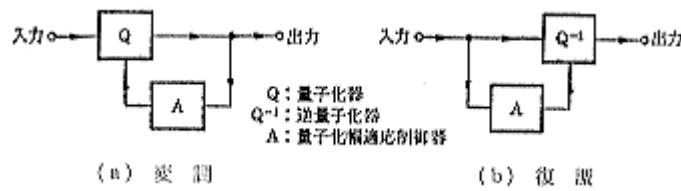


図 - 6 APCM の構成

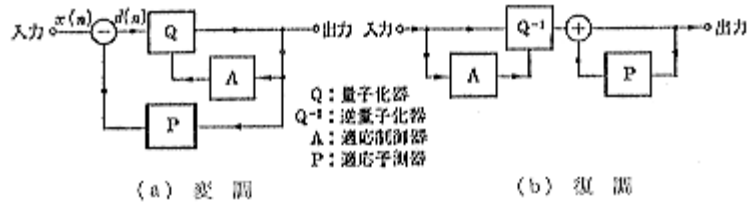


図 - 7 ADPCM の構成

3.4 適応型 PCM (APCM)

音響信号の振幅や周波数分布は時間と共に比較的穏やかにではあるが大幅に変化する。そこで近傍の信号の性質に応じて量子化ステップ幅を変化させる符号化が適応型 PCM (APCM) である。

図 - 6 に適応型 PCM (APCM) の構成例を示す。APCM では直前の符号により次の量子化ステップ幅を決定している。表 - 3 に 3bit APCM の量子化ステップの係数の例 (Jayant らによる) を示す²⁾。この例では、直前の標本の量子化値の振幅が 00 と 01 の場合には量子化ステップ幅に 0.9 を乗ずることにより量子化ステップ幅を小さくし、10 のときには 1.25、11 のときには 1.75 を乗ずることにより量子化ステップ幅を大きくしている。このように適応型符号化ではたとえ量子化ビット数が少なくともある程度時間をかければ様々な量子化ステップ幅が得られるので、低域信号のダイナミックレンジは拡大する。

表 - 3 3bit の APCM の量子化ステップの与え方の例

振 幅	係 数
11	1.75
10	1.25
01	0.9
00	0.9

表 - 4 適応量子化法におけるステップ幅係数 (Jayant による)

ビット数	APCM	ADPCM
2	0.6, 2.2	0.8, 1.6
3	0.85, 1, 1, 1.5	0.9, 0.9, 1.25, 1.75
4	0.8, 0.8, 0.8, 0.8, 1.2, 1.6, 2.0, 2.4	0.9, 0.9, 0.9, 0.9, 1.2, 1.6, 2.0, 2.4
5	0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6	0.90, 0.9, 0.9, 0.9, 0.95, 0.95, 0.95, 0.95, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3.0, 3.3

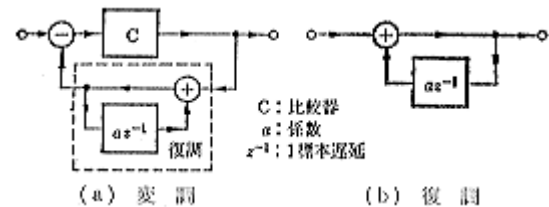


図 - 8 ΔM の構成

3.5 適応型差分 PCM (ADPCM)

適応型 DPCM は差分 PCM (DPCM) に適応形ステップ幅を導入したものである。すなわち、信号 $x(n)$ を直接量子化するのではなく、予測値 $\hat{x}(n)$ との差 $d(n)$ を適応量子化するもので、前述の APCM よりも能率がよく中程度の質の高効率符号化として大変有効な手法である。多機能電話機のメッセージ録音などの短時間のテープ等を使わない固体録音機や案内放送、自動販売機等の各種の音声サービスに盛んに使われている。

図 - 7 に ADPCM の構成、表 - 4 に係数の例を示す。APCM と同様に直前の符号の大小により量子化ステップ幅に乗ずる係数が異なり、次の量子化ステップ幅を決定する。

3.6 ΔM と適応型 ΔM

ΔM は信号を 1bit で量子化する符号化方法である。図 - 8 に最も簡単な ΔM の構成例を示す。この ΔM は前述の DPCM を固定量子化ステップ幅の 1bit にした

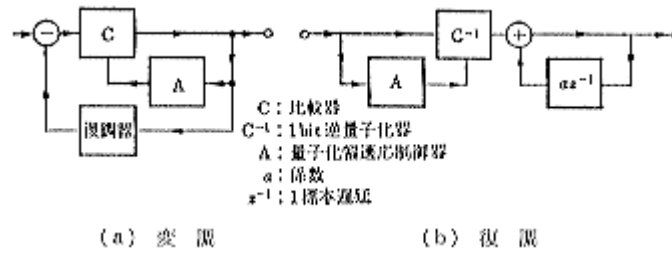


図 - 9 適応型 ΔM の構成

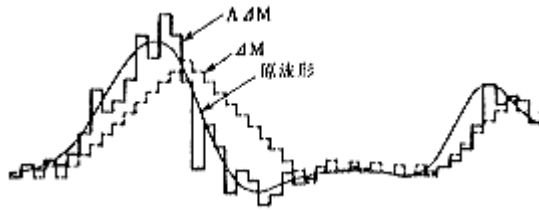


図 - 10 ΔM と適応型 ΔM の波形

もので、標本値が一つ前の標本値より大きい場合は1、小さい場合は0とする簡単な符号化方法である。

原理から明らかなように信号が急速に変化する場合には大きな歪が生ずる。この歪は標本化周波数を高く設定すれば減少するが、あまり高く設定したのでは節約の意味がなくなってしまう。例えば標本化周波数を10倍に設定しても元の標本化間隔に10量子化ステップ幅の変化にしか追従できない。PCMでは同じ10bit/標本で1,024量子化ステップ幅の変化にまで追従可能である。従って ΔM は簡単なシステムや特に高速を要求される場合にのみ有効な手法と言える。

一方、適応型 ΔM は図 - 9に示すように同じ符号が続く場合には係数 a を1.5、反転する場合には0.8というように量子化ステップ幅を信号に適応して変化させる、いわば1bitのADPCMである。図 - 10は ΔM と適応型 ΔM の信号への追従の様子を比較したものである。適応型 ΔM は簡単なハードウェアで比較的良好に信号に追従するので、留守番電話の録音・再生等には現在も使われている。

4. 高能率符号化の最近の動向

表 - 5に実際に使われている、あるいは規格化、提案されている高・中品質音響信号を対象とした各種の符号化方法を示す。

なお、CDの応用分野であるCD - に音楽の長時間化の手法として4 ~ 8bitのADPCMが採用されているが、これはむしろ294標本ごとに予測フィルタと量子化ステップ情報を伝送するブロック符号と呼んだ方が妥当だろう。

4.1 聴覚特性を利用した高能率符号化

移動体を対象としたデジタル放送やCDやDAT

等のパッケージメディアあるいはISDNを念頭においた画像や音響信号の高能率符号化の動きが盛んである。その代表的なものがミットラン大統領の提案によるユーレカ（欧州先端技術開発）計画の一プロジェクト、1986年デジタル音声地上放送システム開発プロジェクト（ユーレカ147 DAB プロジェクト）の開発したMUSICAM (Masking-pattern Universal Sub-band Integrated Coding and Multiplexing), ISO - IEC/SC 2/WG11 略称 MPEG (Moving Picture Experts Group) に提案された MUSICAM, ASPEC (Audio Spectral Perceptual Entropy Coding), ATAC (Adaptive Transform Audio Coding), SB - ADPCM (Sub-Band Adaptive Difference PCM), CCIRに提案されているNHKの低域予測型サブバンド符号化, Silence, 更に最近のフィリップス社のDCC (Digital Compact Cassette) やソニーのMD (Mini Disc) 等である。

これらに共通する最近注目されている技術が、マスキングや臨床帯域等の聴覚特性を利用してサブバンド符号化や変換符号化を行う方式である。

よく知られているように二つ以上の音が同時に存在するとき、一方が他方の存在により聞こえなくなる現象がマスキングと呼ばれる。二つの周波数の異なる純音によるマスキングはその周波数が近いほど顕著である。

一方、帯域雑音による純音のマスキング量は雑音の帯域幅を広げていくと最初は増加するが、ある幅以上になると増加しなくなる。この幅を臨界帯域幅と呼んでいる⁵⁾。

図 - 11にMPEG - 1に提案されたMASCAM (Masking-Pattern Adapted Sub-band Coding and Multiplexing) に使われた臨界帯域を500Hzと1kHz幅に近似した24帯域とマスキング特性から求めた必要なビット数を示す。

4.1.1 サブバンド符号化

聴覚特性を利用して信号を複数のサブバンドに分割したうえ、それぞれのバンドを異なる量子化特性で量子化する符号化方式である。バンド数は4 ~ 64, 分割方法としてはQMF (Quadrature Mirror Filter) 等がよく使われる。

表 - 5 各種の符号化法

方式	標準化 周波数 (kHz)	量子化 ビット数	信号帯域	伝送レート (kbit/s) 1CH 当たり	特徴
32k ADPCM (CCITT G 721)	8	4	300 Hz~3.4 kHz	32	適応予測 2次 IIR+6次 FIR 適応量子化 DLQ (Dynamic Loading Quantizer)
64k Sub-Band ADPCM (CCITT G 722)	16	4	50 Hz~7 kHz	64	Sub-Band 符号化 低域 (50 Hz~4 kHz) 6.5/4 bit 量子化 高域 (4~7 kHz) 2 bit 量子化
384k オーディオ符号化 (CMTT)	32	11/10	50 Hz~	384	14/11 bit A 測瞬時圧伸 14/10 bit 準瞬時圧伸 (32サンプルごと)
衛星TV規格	32	14/10	50 Hz~15 kHz	512	14/10 準瞬時圧伸 14 bit 直線
NHK衛星放送Aモード	32	14/10	50 Hz~15 kHz	768	14/10 準瞬時圧伸
FM東京FM多重放送	8	4	300 Hz~3.4 kHz	32	CCITT 32k ADPCM 準拠
CD-I	37.8	8	17 kHz	309	ADPCM
	37.8	4		459	
	18.9	4	8.5 kHz	89	
8mm VTR 音声	31.5	8	50 Hz~15 kHz	354	アナログ対数圧伸 (2:1) とデジタル 10→8 bit 9 折線圧伸の併用
放送中継 (NTT)	32	11	50 Hz~15 kHz	768	13→11 bit 7 折線準瞬時圧伸
DCC (デジタル コンパクトカセット)	32	44.1 48	5 Hz~20 kHz (44.1 kHz)	384	PASC (Precision Adaptive Sub-band Coding)
	44.1				
MD (ミニディスク)	44.1		5 Hz~20 kHz (44.1 kHz)	300	ATRAC (Adaptive Transform Acoustic Coding)

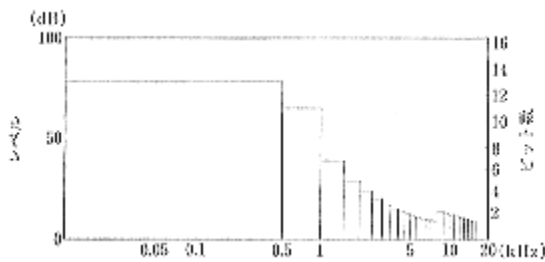


図 - 11 帯域ごとに必要なビット数

4.1.2 変換符号

時間信号を5~50msの窓で切り出したうえ、DCTやDFTにより周波数領域に変換したうえ、臨界帯域幅に対応するグループに分けて、マスキングを考慮して必要最小限の量子化ビット数で符号化する方式である。

図 - 12にASPECの構成を示す。16bitの入力信号を折り返し雑音を抑圧するTDAC (Time Domain Aliasing Cancellation) 手法を用いてMDCT (Modified Discrete Cosine Transfom) により周波数領域に変換す

る。一方FFTされた信号から臨界帯域ごとのマスキング閾値を決め、量子化、エントロピー符号化(ハフマン符号化)する。

4.1.5 量子化雑音のスペクトルに着目した符号化

サブバンド符号化も変換符号化もマスキングを利用した量子化雑音のスペクトル分布に着目した符号化と見ることが出来るが、必ずしも高能率符号化を目的としたものではないが、最近量子化雑音のスペクトル分布を聴覚特性を考慮して制御した符号化が提案されている。

変調のAD/DA交換や著者らが提案し1bit,高速標準化も量子化雑音のスペクトル分布に着目した符号化と見ることが出来る。

4.2 動画の高能率符号化

テレビ会議や電話用の動画の高能率符号化としては1990年にCCITTで成立した64kbit/sの整数(1~30)倍の動き補償付きのハイブリッドDCTによる勧告H.261がある。

CDやDATを対象とした1.5Mbit/sのMPEG-1では更に早送り、逆転再生を考慮してH.261を基本に

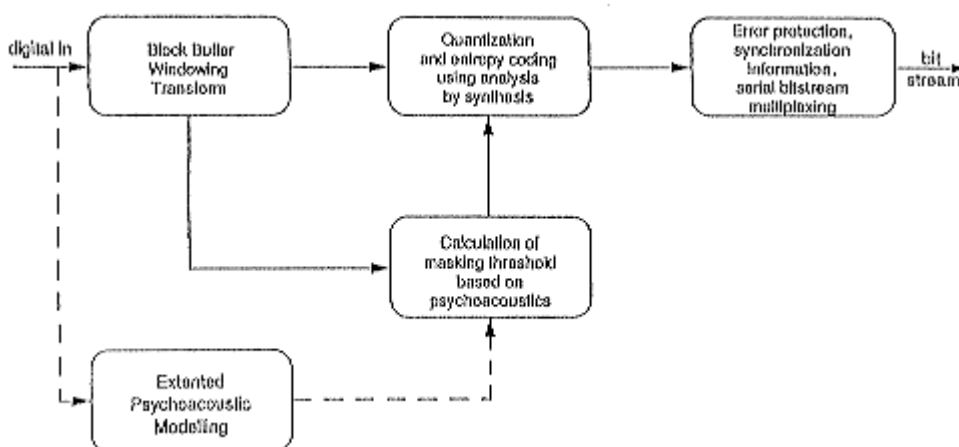


図 - 12 変換符号化 ASPEC エンコーダ

12 ~ 15 フレームを単位にフレーム内，順方向予測更に
双方向予測を加えた符号化方式を採用している。

5. む す び

以上音響信号を中心に高能率符号化を概観した。人間の聴覚・視覚特性を利用した高能率符号化が実用期を迎えようとしている。

音響信号の高能率符号化で多少気になるのはすべて現行の CD，DAT あるいは日本の衛星放送 B モードが出発点になっている点である。高能率符号化の出発点はたとえ目標の伝送容量が小さくとも、デジタル化される前の音響信号そのものに置くべきではないか。

ところで長島茂雄が「いわゆる…」と語った。と聞けば彼いや彼の声を知る多くの人はいの独特な声と言ひ返しを音として思いうかべることができる。また、S 社の CD35DH51 の 3 曲目と伝えれば、信号を実際に送らなくともその番号の CD を（持っていなければ買ったうえで）再生することにより、たった数文字分せいぜい 100bit 程度で膨大な信号を符号化、伝送したことになる。このように受信側に人間の記憶や知識、放送局の資

料室なみの膨大な記録の蓄積が備わっておれば超高能率符号化が期待できる。今日の記憶媒体の驚異的高密度化と価格や消費電力の低減傾向を見ると、案外近い将来、受け側の大容量記憶を前提としたこのような高能率符号化が実現するかも知れない。

末筆ながら本稿をまとめるにあたり、資料の提供及び数々のご助言を賜った NHK 放送技術研究所音響聴覚研究部の若栗尚氏に深謝する。

文 献

- 1) 山崎芳男, 伊藤 毅, "広帯域音響信号の PCM 伝送," テレビジョン学会録音録画研資 (1974. 10).
- 2) N. S. Jayant, "Adaptive delta modulation with one-bit memory," Bell Syst. Tech. J. 49, 321 - 342 (1970).
- 3) "MPEG/audio subjective assessments test report," ISO/IEC JTC1/SC2/WG11, MPEG90/196 (1990. 7).
- 4) J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," IEEE J. SAC 6 (2), 314 - 323 (1988).
- 5) G. Theile, G. Stoll and M. Link, "Low bit-rate coding of high-quality audio signals-An introduction to the MASCAM system," EBU Rev. No. 230, 158 - 181 (1988).