

# 混合 AR モデルとそのロスレス符号化に関する一検討

A Study on Lossless Coding Efficiency of Mixed AR Model

岩橋政宏<sup>1</sup>

貴家仁志<sup>2</sup>

<sup>1</sup>長岡技術科学大学 工学部 電気系

<sup>2</sup>首都大学東京 システムデザイン学部

Masahiro IWAHASHI<sup>1</sup>

Hitoshi KIYA<sup>2</sup>

<sup>1</sup>Nagaoka University of Technology

<sup>2</sup>Tokyo Metropolitan University

**アブストラクト** 映像の高品質化に伴い、より高いビット深度を持つ信号に適した、新しい符号化が求められている。高ビット深度の映像は、より多くの階調を表現できる。この信号に対して、従来の汎用画像に仮定された特性を、直接適用できるとは限らない。本報告では、混合 AR モデルを新たに提案し、その符号化効率を検討する。このモデルでは、画素値を上位ビットと下位ビットに分け、上位を従来の AR(1)モデル、下位を白色雑音として表現する。例として、DPCM によるロスレス符号化を取り扱う。結果、画素値を上位と下位に分け、上位にのみ DPCM を適用した後、それぞれをエントロピ符号化することで、符号化効率が改善されることを示す。

## 1 はじめに

映像信号のハイビジョン化には、画素数（空間解像度）およびフレームレート（時間解像度）の増加が、これまで用いられてきた。近年、更なる高画質化の手段として、階調数（ビット深度）の増加が注目されている[1,2]。高いビット深度を持つ画像は、ハイダイナミックレンジ (High Dynamic Range: HDR) 画像と呼ばれ、適切なトーンマッピングを施すことで[3-4]、ビット深度が 8 [bit]である従来の表示装置でも、これまでにない映像表現が可能となる。

高ビット深度な画像に対する符号化は、[5]において検討された。従来の符号化との互換性は、ビット深度スケラブル符号化として、[6,7]に提案された。国際標準である H.264 との互換性は、[8,9]にて検討された。近年では、量子化誤差の最小化や[10]、ロスレス符号化との互換性などが[11]、報告されている。しかし、高ビット深度の画像に対し、従来の汎用画像に仮定された特性を、直接適用できるとは限らない。

そこで本報告では、混合 AR モデルを新たに提案し、その符号化効率を検討する。このモデルでは、画素値を上位ビットと下位ビットに分け、上位を従来の AR(1)モデル、下位を白色雑音として表現する。例として、DPCM によるロスレス符号化を取り扱う。結果、画素値を上位と下位に分け、上位にのみ DPCM を適用した後、それぞれをエントロピ符号化することで、符号化効率が改善されることを示す。

## 2 混合 AR モデルの提案

従来の AR(1)モデルと白色雑音について述べた後、提案する混合 AR モデルについて説明する。

### 2.1 指定した分散を持つ AR(1) モデル

サンプル数が  $N$ 、値が  $x_A(n)$ ,  $n = 0, 1, \dots, N-1$  である AR(1)モデルを、 $z$ 変換により、

$$X_A(z) = \sum_{n=0}^{N-1} x_A(n)z^{-n} \quad (1)$$

と表すと、 $X_A(z)$ は、その自己相関係数  $\rho$  をパラメータとして、白色雑音  $W(z)$ により、

$$X_A(z) = \frac{W(z)}{1 - \rho \cdot z^{-1}} \quad (2)$$

として生成される。ここで、

$$|W(e^{j\omega})| = \frac{1}{\sqrt{\rho}} \quad , \quad \lambda = \rho + \rho^{-1} \quad (3)$$

とおくと、 $\rho \neq 1$  に対しては  $\lambda > 2$  なので、周波数振幅特性は、

$$|X_A(e^{j\omega})| = \frac{1}{\sqrt{\lambda - 2 \cos \omega}} \quad (4)$$

となる。従って、分散が  $\sigma_S^2$  である AR(1)モデル  $X_S(z)$  の周波数振幅特性は、

$$|X_S(e^{j\omega})| = \frac{1}{\sqrt{\lambda - 2 \cos \omega}} \cdot \frac{\sigma_S}{\sigma_A} \quad (5)$$

但し、

$$\sigma_A^2 = \frac{1}{N} \cdot \frac{1}{\sqrt{(\lambda+2)(\lambda-2)}} \left( 1 - \frac{1}{N} \cdot \sqrt{\frac{\lambda+2}{\lambda-2}} \right) \quad (6)$$

として記述される（付録 A を参照）。

### 2.2 指定した分散を持つ白色雑音

周波数特性が  $W(z)=1$  と記述される白色雑音の場合、サンプル数を  $N$  とすると、付録 A の式(A.2), (A.3)より、その分散は、

$$\sigma^2 = \frac{N-1}{N^2} \quad (7)$$

となる。従って、分散が  $\sigma_W^2$  である白色雑音  $X_W(z)$  の周波数振幅特性は、

$$|X_W(e^{j\omega})| = \frac{N}{\sqrt{N-1}} \cdot \sigma_W \quad (8)$$

として記述される。

### 2.3 下位ビットが白色雑音である混合 AR モデル

本報告では、上位  $U$  [bit] が AR(1) モデルであり、下位  $L$  [bit] が白色雑音であるモデル：

$$X_M(z) = 2^L \cdot X_U(z) + X_L(z) \quad (9)$$

を混合 AR モデルと呼ぶ。図 1 には、その概念を示す。

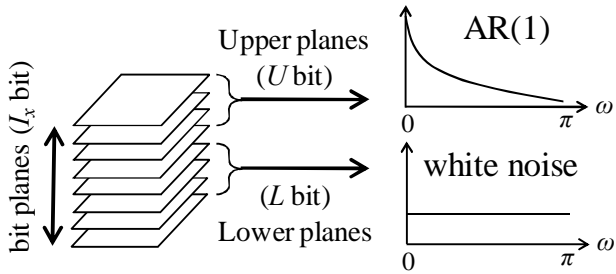


図 1 提案する混合 AR モデル

ここで、 $X_U(z)$  はビット深度が  $U$  [bit] の AR(1) モデルであり、その値を、

$$x_U(n) \in [0, 2^U) \in \mathbf{Z} \quad (10)$$

を満たす整数とする。また、 $x_U(n)$  の確率密度関数を、パラメータ  $r$  により制限されたガウス関数：

$$p(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m_x)^2}{2\sigma^2}\right) & , |x-m_x| < r\sigma \\ 0 & , |x-m_x| > r\sigma \end{cases} \quad (11)$$

により規定する。ここで、

$$r\sigma = 2^{U-1}, \quad m_x = 2^{U-1} \quad (12)$$

とおくことで、式(10)が満たされる。一方、 $X_L(z)$  はビット深度が  $L$  [bit] の白色雑音であり、その値を、

$$x_L(n) \in [0, 2^L) \in \mathbf{Z} \quad (13)$$

を満たす整数とする。このとき、それぞれの分散は、

$$\sigma_U^2 = \frac{2^{2U-1}}{4r^2}, \quad \sigma_L^2 = \frac{2^{2L}}{12} \quad (14)$$

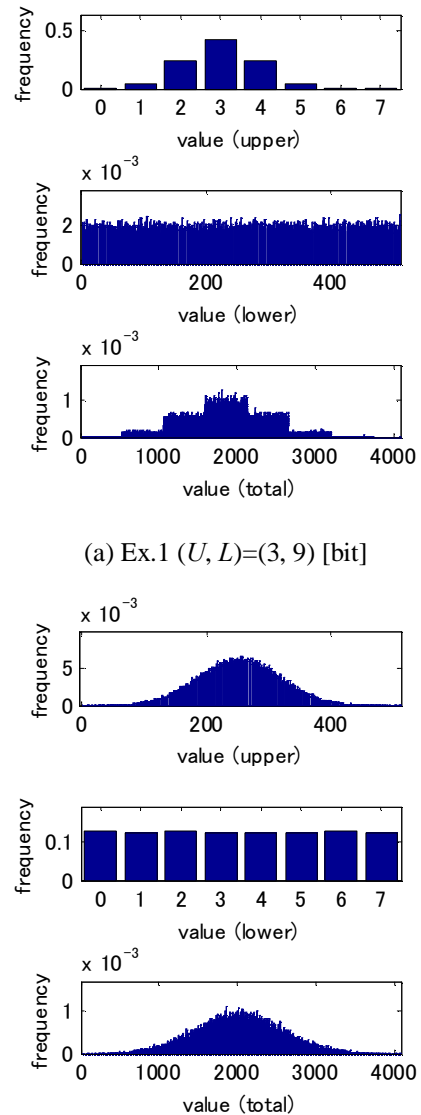
となる。以上より、式(14)の  $\sigma_U$  と  $\sigma_L$  をそれぞれ、式(5)の  $\sigma_S$  と式(8)の  $\sigma_W$  に代入することで、式(9)に定義した信号モデルの周波数振幅特性を、

$$|X_M(e^{j\omega})| \cong 2^L \cdot |X_U(e^{j\omega})| + |X_L(e^{j\omega})| \\ = \frac{2^L}{\sqrt{\lambda - 2\cos\omega}} \cdot \frac{\sigma_U}{\sigma_A} + \frac{N}{\sqrt{N-1}} \cdot \sigma_L \quad (15)$$

として理論的に記述できる。

### 2.4 混合 AR モデルの特性

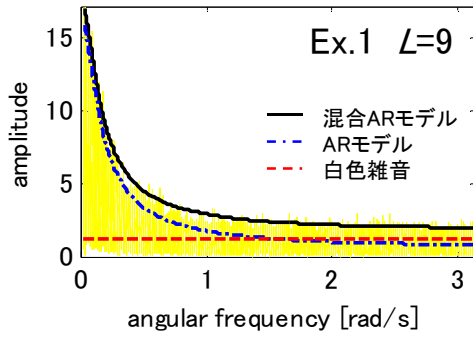
図 2 には、 $\rho=0.9$ 、 $r=4$  の例について、上位、下位、全体のビットプレーン、それぞれの値の確率密度を示す。上位がガウス型分布、下位が一様分布となっている。図 3 には、それぞれの周波数振幅特性を示す。式(5), (8), (15)の理論値が、実験値を良く近似していることが確認できる。



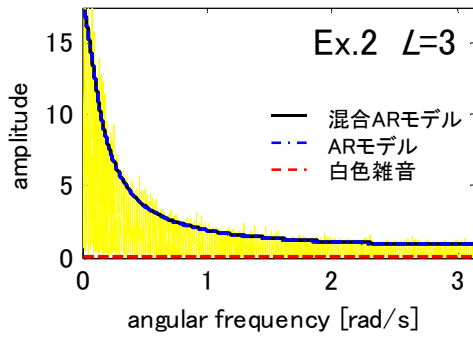
(a) Ex.1 ( $U, L$ )=(3, 9) [bit]

(b) Ex.2 ( $U, L$ )=(9, 3) [bit]

図 2 信号値のヒストグラム ( $U+L=12$ ,  $Pr=0.5$ )



(a) Ex.1 (U, L)=(3, 9) [bit]



(b) Ex.2 (U, L)=(9, 3) [bit]

図3 周波数振幅特性  
(黒、青、赤が理論値、黄が実験値)

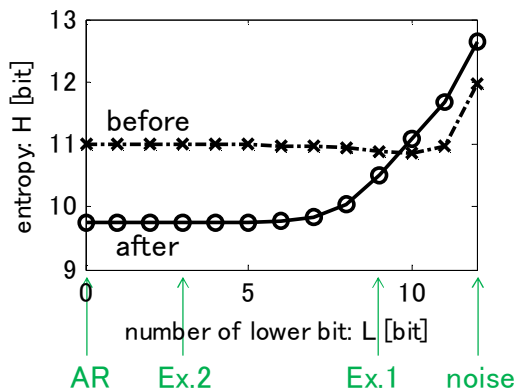


図4 従来の符号化の効率評価

### 3 従来の符号化とその効率評価

前節で定義した混合 AR モデルに対し、従来の符号化を適用し、その効率を調べる。本報告では、例として、DPCM によるロスレス符号化を扱う。

#### 3.1 従来の符号化の適用結果

図4には、混合 AR モデルに対し、DPCM を適用する前と後、それぞれにおけるエントロピを示す。ここでは、 $N=2^{16}$ 、 $\rho=0.9$ 、 $r=4$ 、 $L+U=12$  [bit] の 1次元信号を用い、DPCM の伝達関数は  $H(z)=1-z^{-1}$  とした。

同図より、全 12 [bit] 中の、下位 0~9 [bit] が白色信号である場合は、DPCM の適用により、エントロピを低減できることがわかる。しかし、下位 10 [bit] 以上が白色信号である場合は、DPCM の効果が無い。このときは、DPCM を適用せず、直接、エントロピ符号化の方が、符号化効率が良いことがわかる。

#### 3.2 符号化結果の理論解析

伝達関数が  $H(z)$  として表現される DPCM は、入力信号  $X(z)$  に対して、

$$Y(z) = X(z)H(z) \quad (16)$$

と作用する。これにより、信号の分散は、

$$\sigma_y^2 = E_y - m_y^2 \quad (17)$$

但し、

$$\begin{cases} E_y = \frac{1}{N} \cdot \frac{1}{2\pi} \int_0^{2\pi} |X(e^{j\omega})H(e^{j\omega})|^2 d\omega \\ m_y = \frac{1}{N} \cdot X(1)H(1) \end{cases} \quad (18)$$

に変化する。その後、エントロピ符号化を適用することで、ビットストリームが生成される。

本報告では、エントロピ符号化による圧縮の結果として得られるビットストリームの符号量を、1 次のエントロピである

$$H_y = - \int_{-\infty}^{+\infty} p(y) \log_2 p(y) dy \quad (19)$$

により評価する。ここで、 $p(y)$  は値  $y$  の生起確率（確率密度関数）である。分散は一般に、

$$\sigma_y^2 = \int_{-\infty}^{+\infty} p(y) \cdot y^2 dy \quad (20)$$

と計算できることから、エントロピと分散との間には、

$$H_y = \frac{1}{2} \log_2 \gamma_y \cdot \sigma_y^2 \quad (21)$$

なる関係がある。 $\gamma_y$  は確率密度関数により決まる定数である（付録 B を参照）

ここで、式(21)のエントロピが、

$$H_y = H_y^{(\gamma)} + H_y^{(\sigma)} \quad (22)$$

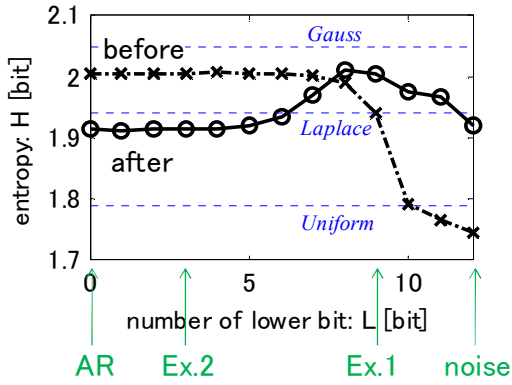
但し、

$$H_y^{(\gamma)} = \frac{1}{2} \log_2 \gamma_y, \quad H_y^{(\sigma)} = \frac{1}{2} \log_2 \sigma_y^2$$

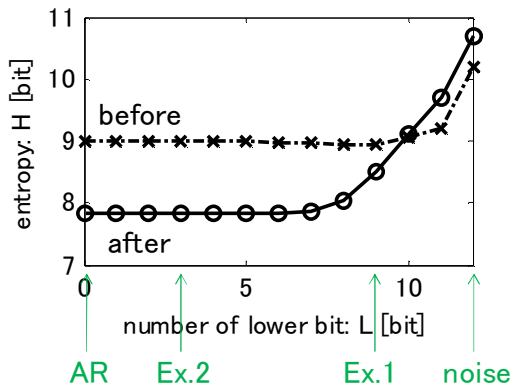
なる 2 つの項に分離できることに注目する。図4のエントロピについて、これらの内訳を調べた結果を図5に示す。同図(a)は、分布形状に依存する項  $H_y^{(\gamma)}$  を示す。白色雑音である下位ビットが  $L=6$  [bit] 以下であれば、DPCM の結果は Laplace 分布に近いが、 $L=8, 9$  [bit] では Gauss 分布に近づいている。一方、同図(b)に

示される分散に依存する項  $H_y^{(\sigma)}$  は、同図(a)の  $H_y^{(\gamma)}$  に比べて大きく支配的である。また、その形状が図 4 と類似しており、 $L \geq 8$  [bit] においてエントロピが増加している。

以上より、DPCM は白色信号に対する分散利得が大きいため、信号が白色信号に近い程、すなわち、 $L$  が大きい程、エントロピ低減効果が弱くなると言える。



(a) 分布形状に依存する項  $H_y^{(\gamma)}$



(b) 分散に依存する項  $H_y^{(\sigma)}$

図5 図4に示したエントロピの内訳

#### 4 混合 AR モデルのための新しい符号化

前節では、従来の符号化が混合 AR モデルには適さないことが示された。以下、この新しいモデルに適する符号化を提案し、その効率を調べる。

##### 4.1 提案する符号化の概要

図 6 には、混合 AR モデルに適した新しい符号化の概要を示す。入力信号を上位  $U$  [bit] と下位  $L$  [bit] に類別し、それぞれを異なる方法により符号化する。上位ビットの符号化 (図中の Encoder 1) には、AR(1)モデルに適した処理、例えば、DPCM などの予測符号化や、ウェーブレット変換などを適用して、信号を無相関化する。その後、エントロピ符号化する。

この無相関化が、白色雑音に対して 1 より大きな分散利得を持つ場合、例えば、ロスレス符号化のための DPCM や、双直交な整数ウェーブレット変換、整数離散コサイン変換などは、下位ビットの符号化 (図中の

Encoder 2) には用いずに、直接、エントロピ符号化を適用する。

従来の AR(1)モデルに対しては、このような類別符号化は、一般に、符号化効率を低下させることが知られている。しかし、本報告で定義された混合 AR モデルに対しては、その限りではないことを次節に示す。

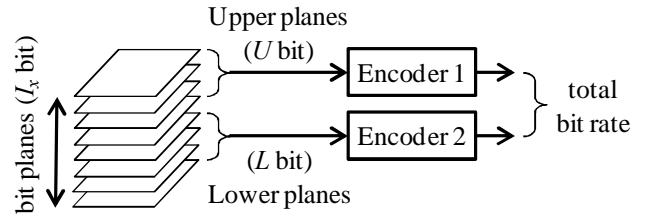


図6 ビットプレーンの類別符号化

##### 4.2 提案する符号化の適用結果

図 7 には、「proposed」として、図 6 の類別符号化によるエントロピを示す。上位ビットのみに DPCM を適用した。「existing」として、類別せずに DPCM を適用した場合を示す。同図より、下位ビットが 8 [bit] 以上の場合、類別符号化により 0.12 [bpp] 以上、エントロピを低減できることがわかる。

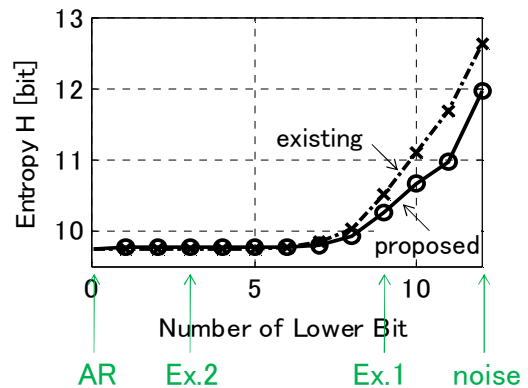


図7 提案する符号化の効率評価の結果

##### 4.2 下位ビットがスパースな場合

これまでの実験では、各ビットプレーンにおける値 (0 あるいは 1) が、1:1 の割合で生起している。ここで、値 1 の生起確率を  $Pr$  とすると、図 2 に示した Ex.1 および Ex.2 では、 $Pr=0.5$  であった。以下、下位ビットにおいて、 $Pr < 0.5$  なるスパースな場合について考察する。

例えば、Ex.3 として  $Pr=0.2$ 、 $U=9$ 、 $L=3$  とする。このときのヒストグラムを、図 8 に示す。また、 $Pr=0.5$  に対する結果である図 7 は、 $Pr=0.2$  に対しては図 9 のように変化する。同図より、下位ビットが 8 [bit] 未満であっても、類別した方が、符号化効率が高くなるのがわかる。具体的には、下位ビットが 1 [bit] 以上において、0.18 [bpp] 以上のエントロピ低減効果が得られる。

参考までに、Ex.1~3 における信号のパラメータを、表 1 にまとめる。これら以外の場合については、「existing」に対する「proposed」のエントロピ低減効果を、図 10 にまとめる。同図(a)より、下位ビットがスパースではない  $Pr=0.5$  の場合、全 12 [bit]中の 9 [bit]以上が白色雑音であれば、 $\rho=0$  であっても、類別符号化の効果があることがわかる。同図(b)より、下位ビットがスパースであれば、下位ビットの数に拘わらず、類別符号化の効果があることがわかる。

表 1 本報告で用いた入力信号

	白色雑音	混合ARモデル			従来のARモデル
		Ex.1	Ex.2	Ex.3	
上位ビット	U=0	U=3	U=9		U=12
下位ビット	L=12	L=9	L=3		L=0
		Pr=0.5		Pr=0.2	

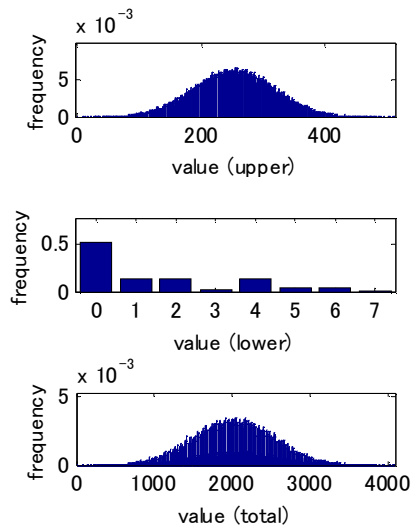


図 8 Ex.3 の信号値のヒストグラム ( $U, L$ )=(9, 3) [bit],  $Pr=0.2$

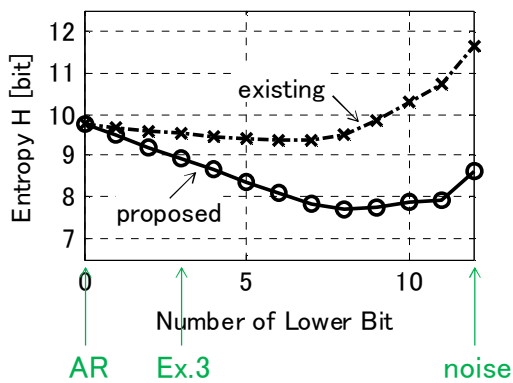
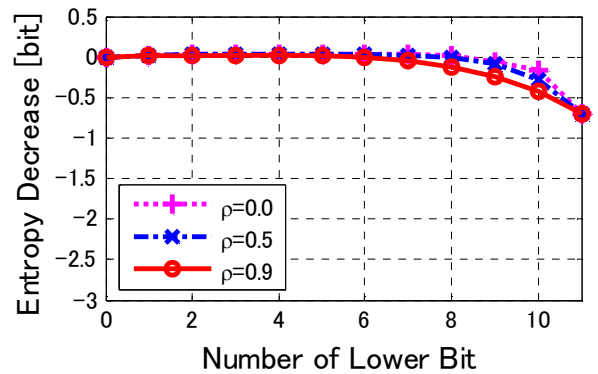
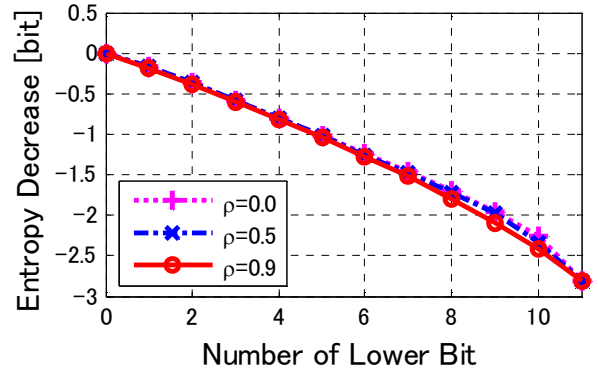


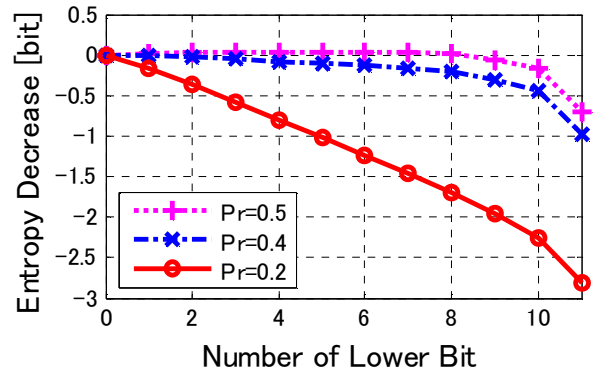
図 9 下位ビットがスパースな混合 AR モデル に対する効率評価の結果



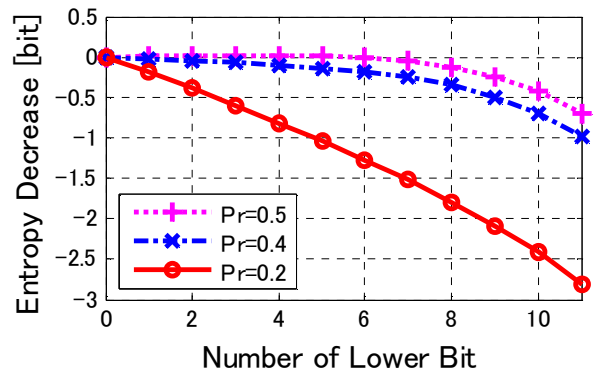
(a)  $Pr=0.5$



(b)  $Pr=0.2$



(c)  $\rho=0.0$



(d)  $\rho=0.9$

図 10 下位ビットがスパースな混合 AR モデル に対する類別符号化の効率

## まとめ

本報告では、上位ビットが AR(1)モデルであり、下位ビットが白色雑音である、混合 AR モデルを提案した。このモデルに対しては、ロスレス符号化のための DPCM が、符号化効率を低下させる場合があることを示した。また、上位には DPCM を適用するが、下位には適用しない、ビットプレーン類別符号化を提案した。実験により、下位ビットが 8 [bit] 以上の場合、0.12 [bpp] 以上のエントロピ低減効果が得られることがわかった。更に、下位ビットがスパースな場合は、下位のビット数に拘わらず、類別符号化の効果があることがわかった。

## 参考文献

- [1] Y. Bandoh, Guoping Qiu, M. Okuda, Scott Daly, Til Aach, Oscar Au, "Recent advances in high dynamic range imaging technology", IEEE International conference image processing (ICIP '10), Sept. 2010.
- [2] K. Hirakawa, Patrick J. Wolfe, "Optimal exposure control for high dynamic range imaging", IEEE International conference image processing (ICIP '10), Sept. 2010.
- [3] Oscar Au, C.H. Liu, C.H. Cheng, Ka Yue Yip, "Two-level optimized tone mapping for high dynamic range images", IEEE International conference image processing (ICIP '10), Sept. 2010.
- [4] Guoping Qiu, Yujie Mei, Kin-Man Lam, Min Qiu, "Tone mapping HDR images using optimization: a general framework", IEEE International conference image processing (ICIP '10), Sept. 2010.
- [5] G. Ward and M. Simmons, "Subband encoding of high dynamic range imagery", Proc. Symposium on applied perception in graphics and visualization, pp.83-90, 2004.
- [6] S. Liu, W. S. Kim, A. Vetro, "Bit-depth scalable coding for high dynamic range video", SPIE conference on visual communications and image processing, TR2007, 078, Jan. 2008.
- [7] M. Winken, D. Marpe, H. Schwarz, T. Wiegand, "Bit-depth scalable video coding", IEEE International conference image processing (ICIP '07), vol.1, pp.5-8, Sept. 2007.
- [8] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, M. Wien, Joint Draft ITU-T Rec. H.264 ISO/IEC 14496-10, "Scalable video coding", July, 2007.
- [9] S. Park and K. R. Rao, "Bit-depth scalable video coding based on H.264/AVC", IEICE Trans. Fundamentals of Electronics, Commun. Computer Sciences, Vol. E91-A, No.6, pp.1541-1544, 2008.
- [10] T. Ito, Y. Bandoh, S. Takamura, H. Jozawa, "A coding method for high bit-depth images based on optimized bit-depth transform" IEEE International conference image processing (ICIP '10), Sept. 2010.
- [11] H. Kikuchi, W. Otake, M. Iwahashi, "Bit rate reduction of enhancement layer in bit-depth scalable coding", International symposium intelligent signal processing, communications and systems (ISPACS), pp.264-267, Dec. 2009.

## 付録 A AR(1)モデルの分散

サンプル数が  $N$  である信号  $x(n)$  の  $z$  変換が、式(1)で定義されるとき、この信号の分散は、

$$\sigma_x^2 = E_x - m_x^2 \quad (\text{A.1})$$

但し、

$$\begin{cases} E_x = \frac{1}{N} \cdot \frac{1}{2\pi} \int_0^{2\pi} |X(e^{j\omega})|^2 d\omega \\ m_x = \frac{1}{N} \cdot X(1) \end{cases} \quad (\text{A.2})$$

として計算される。ここで、 $E_x$  および  $m_x$  はそれぞれエネルギー密度および平均であり、 $x(n)$  からは、

$$E_x = \frac{1}{N} \cdot \sum_{n=0}^{N-1} x^2(n) \quad , \quad m_x = \frac{1}{N} \cdot \sum_{n=0}^{N-1} x(n) \quad (\text{A.3})$$

と計算される。また、一般に、

$$\frac{1}{\pi} \int_0^\pi \frac{1}{a - b \cos \omega} d\omega = \frac{1}{\sqrt{a^2 - b^2}}, \quad a > |b| \quad (\text{A.4})$$

が成り立つ。このため、式(4)のモデルについては、

$$\frac{1}{\pi} \int_0^\pi |X_A(e^{j\omega})|^2 d\omega = \frac{1}{\sqrt{\lambda^2 - 2^2}} \quad (\text{A.5})$$

となる。従って、式(A.2)より、

$$E_A = \frac{1}{N} \cdot \frac{1}{\sqrt{\lambda^2 - 2^2}}, \quad m_A = \frac{1}{N} \cdot \frac{1}{\sqrt{\lambda - 2}} \quad (\text{A.6})$$

となり、式(A.1)より、

$$\sigma_A^2 = \frac{1}{N} \cdot \frac{1}{\sqrt{(\lambda+2)(\lambda-2)}} \left( 1 - \frac{1}{N} \cdot \sqrt{\frac{\lambda+2}{\lambda-2}} \right) \quad (\text{A.7})$$

として式(6)が得られる。

## 付録 B エントロピと分散の関係

例えば、 $y$  の値が  $y \in [-\Delta/2, \Delta/2]$  の範囲で一様に分布する場合は、

$$H_y = - \int_{-\Delta/2}^{+\Delta/2} \frac{1}{\Delta} \log_2 \frac{1}{\Delta} dy = \log_2 \Delta \quad (\text{B.1})$$

$$\sigma_y^2 = \int_{-\Delta/2}^{+\Delta/2} \frac{1}{\Delta} \cdot y^2 dy = \frac{\Delta^2}{12} \quad (\text{B.2})$$

より、

$$H_y = \frac{1}{2} \log_2 \Delta^2 = \frac{1}{2} \log_2 12 \sigma_y^2 \quad (\text{B.3})$$

となり、 $\gamma_y=12$  となる。同様に計算することで、Gaussian 分布、Laplacian 分布、Gamma 分布では、 $\gamma_y$  はそれぞれ、 $2\pi e$ 、 $2e^2$ 、 $(4/3)\pi e^{(1-C)}$  となる。ここで、 $e$  は自然対数、 $C=0.5772\cdots$  はオイラー定数である。