

# H.264 の量子化特性に関する一検討

## A Study on Quantization Characteristics of H.264 Encoding

蝶野 慶一 仙田 裕三 宮本 義弘  
Kei-ichi Chono Yuzo Senda Yoshihiro Miyamoto

NEC メディア情報研究所  
Media and Information Research Laboratories, NEC Corporation

### 1. あらまし

H.264[1]は、MPEG-2/4 後継の最新動画画像符号化国際標準である。当初の ITU-T による H.264(H.26L)の規格化は、QCIF や CIF などの低解像度な映像を 32~256kbps の低ビットレートで高品質に伝送することを目的に 1998 年に開始した。そして、その後の MPEG との共同作業でビットレート・解像度が拡張され、最終的に 64kbps の QCIF から 50Mbps の HDTV を含む規格となった [2]。

しかしながら、前記の拡張によって追加された高ビットレートにおけるエンコーダの量子化特性は、十分に検討されていない。このため H.264 開発リファレンス(Joint Model)には、H.26L 開発当初の低ビットレート符号化で有効な切捨て重視の量子化が実装されたままである [3]。

そこで、本稿は H.264 符号化でのビットレート毎の量子化丸め位置と量子化特性の関係を解析する。以下、2節で Joint Model の量子化、3節で量子化丸め位置と量子化特性の関係、4節で数値実験、5節でまとめを述べる。

### 2. Joint Model の量子化

H.264 は、4x4 画素のブロックを符号化の最小構成単位とする。H.264 の変換は、DCT を整数値化した変換であり、予測誤差信号を  $D = \{d_{00} \dots d_{33}\}$ 、変換係数を  $C$  とすると、式(1)で定義される。

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix} \begin{pmatrix} d_{00} & d_{01} & d_{02} & d_{03} \\ d_{10} & d_{11} & d_{12} & d_{13} \\ d_{20} & d_{21} & d_{22} & d_{23} \\ d_{30} & d_{31} & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{pmatrix} \quad (1)$$

H.264 の量子化は、変換係数のゲイン補正パラメータを  $Q(\text{rem}, i, j)$ 、量子化パラメータを  $qp$ 、量子化の出力を  $CQ$  とすると、式(2)で定義される。

$$CQ = (C(i, j) \times Q(qp \% 6, i, j) + f \times 2^{(15+qp/6)}) / 2^{(15+qp/6)} \quad (2)$$

ここで式(2)の  $f$  は、量子化の丸め位置を決定するパラメータである。Joint Model(JM)ではフレーム内予測ブロック(intra)で  $|f|=1/3$ 、フレーム間予測ブロック(inter)で  $|f|=1/6$  である。

式(2)より H.264 は、ブロックの予測モードによらず、線形量子化を用いる(説明を省略するが、Intra16x16MB の DC 係数、Chroma の DC 係数の量子化も線形量子化である [3])。

### 3. 量子化丸め位置と量子化特性の関係

#### 3.1. 線形量子化器の入出力特性

2節で説明した式(2)の JM の量子化を、より一般的な線形量子化に置き換えて、量子化特性を考える。量子化ステップサイズを  $q$ 、量子化器の入力を  $x$ 、量子化丸め位置パラメータを  $f$ 、逆量子化器の出力を  $y$  とすると、線形量子化の入出力特性は図1のようになる。

量子化の入出力特性は正負対称であるから、正の領域だ

けを数式で表現すると、範囲  $[X_k, X_{k+1})$  を  $y_k$  に量子化する線形量子化の入出力は以下となる。

$$X_k = \begin{cases} 0 & k=0 \\ (k-f)q & k \geq 1 \end{cases} \quad (3)$$

$$y_k = kq \quad (4)$$

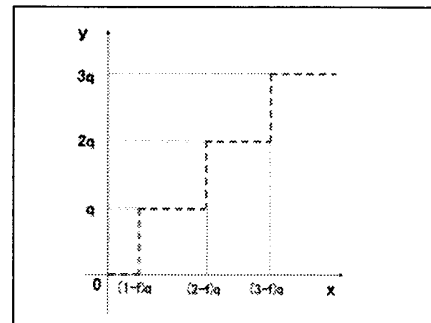


図1 線形量子化器の入出力特性

#### 3.2. 量子化ノイズとエントロピーの関係

予測誤差信号の変換係数  $x$  が式(5)の分散  $\sigma^2$  のラプラス分布であると仮定し、量子化の入出力特性は正負対称なので、正の領域を用いて量子化の平均符号化ノイズとエントロピーの関係を考える。

$$P(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}}{\sigma}|x|} \quad (5)$$

範囲  $[X_k, X_{k+1})$  が  $y_k$  に量子化されるとすると、平均符号化ノイズ  $N$  とエントロピー  $H$  は次式となる。

$$N = 2 \sum_{k=0}^{\infty} \int_{X_k}^{X_{k+1}} (x - y_k)^2 P(x) dx \quad (6)$$

$$H = -P_0 \log_2 P_0 + \sum_{k=1}^{\infty} P_k (1 - \log_2 P_k) \quad (7)$$

ここで  $(1 - \log_2 P_k)$  の 1 は正負符号のビットである。また、 $P_k$  は量子化係数がレベル  $\pm k$  となる確率であり、

$$P_k = 2 \int_{X_k}^{X_{k+1}} P(x) dx \quad (8)$$

と表される。

以上の数式から、線形量子化器の平均ノイズ  $N_l$  とエントロピー  $H_l$  を導出すると、

$$N_l = \sigma^2 \{1 - bc((1-2f)a^2 + \sqrt{2}a)/(1-b)\} \quad (9)$$

$$H_l = -P_0 \log_2 P_0 + bc + \sqrt{2}abc(1/(1-b) - f) \log_2 e - bc \log_2(1-b) \quad (10)$$

となる。ここで  $a = q/\sigma$ 、 $b = e^{-\sqrt{2}a}$ 、 $c = e^{\sqrt{2}af}$ 、 $P_0 = 1 - bc$  である。

最も誤差の小さい四捨五入の  $f=1/2$ 、JM に記載される  $f=1/3$  および  $f=1/6$  それぞれの量子化特性を図2に示す。図2の横軸は1係数あたりのエントロピー、縦軸は量子化にされた予測誤差信号の SNR である。SNR は以下の式(11)で求めており、量子化予測誤差信号の伝送による SNR 改善量となる。

$$SNR = 10 \log_{10} \frac{\sigma^2}{N_f} \quad [dB] \quad (11)$$

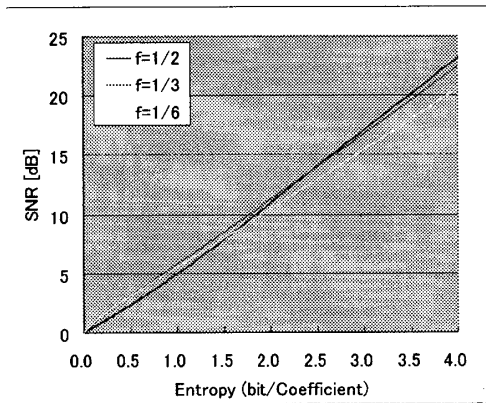


図2 エントロピーと符号化ノイズ

図2のグラフより理論的には、量子化丸め位置パラメータ  $f$  を、1変換係数あたりのエントロピーが

- 0.5bit 以下で  $f=1/6$
- 0.5~2.5bit で  $f=1/3$
- 2.5bit 以上で  $f=1/2$

とすることで最も高い符号化効率を得られる。

JM は2節で述べたように、ビットレートに依存することなく intra ブロックで  $f=1/3$ 、inter ブロックで  $f=1/6$  とする。通常、intra ブロックの予測性能は inter ブロックよりも低いいため、intra ブロックには inter ブロックよりもビットを多く割り当てる。このことを考慮すると、Level1(QCIF, 15fps, 64kbps)では、1係数あたりの平均ビット数は 0.11bit で、JM の上記のパラメータは理に合う。しかし、Level3(SDTV, 30fps, 10Mbps)では、1係数あたりの平均ビット数は 0.64bit であり、特に inter ブロックでは  $f=1/6$  の符号化効率は最適ではないと考えられる。

エントロピー(ビットレート)毎に最適な量子化丸め位置 ( $f=1/2, 1/3, 1/6$ ) を4節の数値実験で確認する。

#### 4. 実験結果

量子化丸め位置と量子化特性の関係を正確にみるために、孤立係数打ち切り[3]をオフとし、ピクチャタイプではなく MB タイプで量子化丸め位置を制御するように改良した JM80 を実験に利用した。

CIF 解像度の4つのシーケンスを Ropt=OFF、固定量子化 (QP=0, 6, , , 42, 48)、CABAC で、 $f=1/2, 1/3, 1/6$  の3通りの量子化丸め位置で符号化した結果の一例を図3, 4にそれぞれ示す(すべてのシーケンスで同じ傾向であった)。PSNR は、符号化誤差の2乗平均を  $e^2$ 、原画像に含まれる A/D 変換誤差の2乗平均を  $1/12$  として、式(12)とする。

$$PSNR = 10 \log_{10} \frac{255^2}{e^2 + 1/12} \quad [dB] \quad (12)$$

実験結果より、量子化丸め位置パラメータ  $f$  を

- intra ブロックでは、約 2.2bit/pixel 以上の高レートで  $f=1/2$ 、その他のレートで  $f=1/3$

- inter ブロックでは、約 2.2bit/pixel 以上の高レートで  $f=1/2$ 、約 0.6~2.2bit/pixel のレートで  $f=1/3$ 、約 0.6bit/pixel 以下の低レートで  $f=1/6$
- とすることで最も高い符号化効率を得られる。この結果は、3節で求めた理論上のレート歪特性を最適にする量子化丸め位置パラメータにおおまかに一致する。

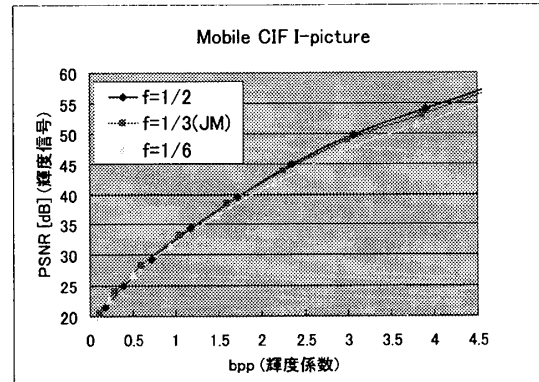


図3 Mobile I ピクチャ量子化特性

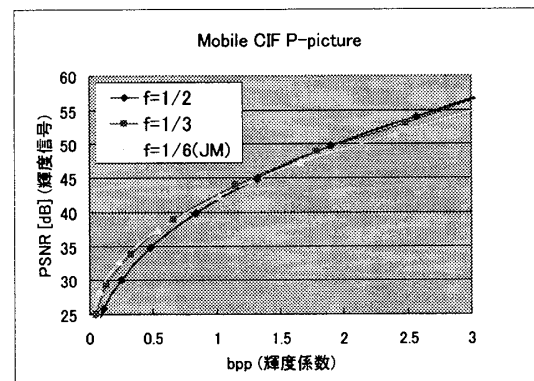


図4 Mobile P ピクチャ量子化特性

#### 5. まとめ

本稿は、H.264(JM)の量子化を述べ、ビットレート毎の量子化丸め位置と量子化特性の関係を解析した。解析の結果、JM の量子化丸め位置パラメータ  $f$  をそのまま利用するのではなく、

- intra ブロックでは、約 2.2bit/pixel 以上の高レートで  $f=1/2$ 、その他のレートで  $f=1/3$
- inter ブロックでは、約 2.2bit/pixel 以上の高レートで  $f=1/2$ 、約 0.6~2.2bit/pixel のレートで  $f=1/3$ 、約 0.6bit/pixel 以下の低ビットレートで  $f=1/6$

とすることで、より高い符号化効率を得られることを示した。2.5bit/pixel を上回る高ビットレートで映像を符号化するには、提案する量子化丸め位置を採用することで JM の符号化効率を約 2[dB]改善できる。

#### 参考文献

- [1] ISO/IEC 14496-10 "Advanced Video Codec"(AVC) & ITU-T Rec. H.264, 2003.
- [2] Thomas Wiegand, "Overview of the H.264/AVC Video Coding Standard", in IEEE Transaction on Circuits and Systems for Video Technology, July 2003.
- [3] Gary Sullivan, "Joint Model Reference Encoding Methods and Decoding Error Concealment Methods", JVT-J049, March 2004.