一般化メディアンカット量子化とその応用

Generalized Median Cut Quantization and its Applications

貴家仁志 † 小林弘幸 †† 岩橋政宏 †††
† 首都大学東京システムデザイン学部
† 東京都立産業技術高等専門学校 ものづくり工学科
††† 長岡技術科学大学電気系

Hitoshi KIYA† Hiroyuki KOBAYASHI†† Masahiro IWAHASHI†††
†Tokyo Metropolitan University
††Tokyo Metropolitan College of Industrial Technology
†††Nagaoka University of Technology

アブストラクト 本報告では，従来のメディアンカット量子化を拡張し，変換後の階調数が 2 のべき乗に限定されず，かつ荷重メディアン値を用いる一般化メディアンカット量子化を提案する．提案する量子化は，画像を任意の階調数に変換でき，演算時間も比較的高速である．さらに，一般の画像において Lloyd-Max 量子化とほぼ同程度の量子化誤差となることを確認している．また，本量子化のロッシー画像符号化への応用例についても報告する．

1 はじめに

近年，8 ビットを超えるビット深度の画像を取り扱うデバイスなどが普及し，高解像度・高階調である高品質な画像を取り扱う機会が増加している．このような画像に対して，標準階調を前提にした各種機器との互換性やトランスコーディングなどの目的で，ビット深度を変換する目的のための量子化という操作が改めて再注目されている [1–4].

代表的な量子化に四捨五入 (丸め量子化) や切り捨てがある．これらの量子化処理は，処理が単純で高速な処理が可能であるが，信号値の分布に偏りがあった場合に，一般に大きな量子化誤差を発生してしまう．特に，ヒストグラムに疎性が存在する場合には，量子化後の代表値の個数が減少してしまい，大きな量子化誤差を発生させてしまう．

このような信号値の偏りを考慮した量子化法に，Lloyd-Max 量子化がある [5,6]．この方法は，ある条件の下で量子化の平均自乗誤差を最小にすることができるが，繰り返し計算を必要とし，高速な量子化処理の適用には適さない．本稿では，このような背景から，画像信号のような大きな信号値の偏りを持ち，かつヒストグラムの疎性が存在する信号に対して有効なメディアンカット量子化に着目する [7]．従来のメディアンカット量子化は，主にビット数削減を目的としており，分割を繰り返して使用されることが多かった．本稿では，2 のべき乗の制約を解除し，分割を繰り返さない一般化メディアンカット量子化を提案する．提案法では，さらに重み付きメディアン値を使用する．

次に，提案法のロッシー画像符号化への応用とその有効性を示す．高ダイナミックレンジ (HDR) 画像や高ビット深度画像の取り扱いでは，しばしば量子化操作が求められる．提案される一般化メディアンカット量子化は，その特徴から細かいレート制御が可能であり，ヒストグラムが疎性を持つ場合においても，比較的効率的な圧縮を可能とする．JPEG2000，JPEG-LS，JPEG-LS ニアロスレスなどとの比較において，特に高画質な圧縮領域において，提案法の有効性が確認される．

2 量子化の準備

2.1 代表的量子化

本稿では，画像信号に対する量子化について検討する．そのため，対象とする入出力の信号値はすべて整数値となる．画像信号に対する量子化とは，入力画像の近接する濃度値を統合し，クラス分けを行うマッピング作業と考えることができる．今，N ビットの画像を M ビット $(N > M)$ に量子化する場合を例とすると，入力画像 $P(x, y)$ に対する量子化は，クラス分けのマッピングテーブル $Q(k)$, $(k = 0, 1, \cdots, 2^N - 1)$ を用いて，次式のように表現される．

$$I(x, y) = Q(P(x, y)), \qquad \text{for all } x, y \tag{1}$$

ここで，$I(x, y)$ は分類されたクラスの番号を示すインデックス値に相当し，その値域は 0〜$2^M - 1$ となる．本稿で

は，この量子化された画像 $I(x,y)$ をインデックス画像と呼ぶ．

インデックス画像は，各インデックスに対応する代表値テーブル $R(n)$, $(n=0,1,\cdots,2^M-1)$ を用いて，次式のように N ビットの濃度値を持つ画像 $P'(x,y)$ に変換される．

$$P'(x,y)=R(I(x,y)), \qquad \text{for all } x,y \tag{2}$$

通常の量子化では $N>M$ であり，インデックス値で表現できる階調数は，入力画像に比べて低減することになる．すなわち，入力画像におけるいくつかの異なる濃度値は，量子化後の出力画像において同じ濃度値に変換される．この操作によって発生する出力画像 $P'(x,y)$ と入力画像 $P(x,y)$ の誤差が量子化誤差である．

量子化誤差の発生の仕方は，量子化のアルゴリズムに依存する．代表的な量子化アルゴリズムとしては以下のようなものがある．

1. 下位ビット切り捨て
2. 丸め処理
3. メディアンカット
4. Lloyd-Max

前者二つの量子化は，各クラスに対応する階調の幅を均等に量子化していく均等量子化である．画像に依存せず高速に処理が可能であるが，一般に量子化誤差が大きくなってしまう．後者二つの量子化は，各クラスに対応する階調の幅が異なる不均等量子化である．これらは，画像の濃度値の頻度に従ったクラス分けを行うため，前者に比べて量子化誤差の発生が抑えられる．

2.2 メディアンカット量子化

ここでは，不均等量子化のうちメディアンカット量子化について説明する．メディアンカット量子化は，画像における全画素の濃度値を大きさ順に並べ，その数列の中央値 (メディアン値) を持ってクラス分けを行うアルゴリズムである．この時，メディアン値 m は次式で計算される．

$$m=MED\{P(x,y), \quad \text{for all } x,y\} \tag{3}$$

ここで，MED{} は {} 内の数列を大きさ順に並べた際の中央値を取り出す関数である．中央値で分割しているため，二つのクラスにはほぼ同数の濃度値が存在することになる．このようなクラスの分割を 2^M-1 回繰り返し行うことで，最終的に全画素を 2^M 個のクラスに分割することができる．2 回目以降の分割におけるクラスの選択については自由度があり，クラス内の分散や濃度値の種類などを指標とすることが多い．

クラス分け終了後，原画像の各画素値に対して，対応するクラスのインデックス値を割り振ることで，M ビットのインデックス画像 $I(x,y)$ が得られる．この時，逆量子化に使用する代表値テーブル $R(n)$ の値は，各クラスに対応する画素値の範囲から一つを選択すればよく，その選択にも自由度が存在する．

3 提案法

3.1 一般化メディアンカット

本稿では，前述のメディアンカット量子化を二つの点で拡張する．一つ目の拡張は，画素値の頻度数ではなく画素値の種類でクラス分けを行う点である．この拡張は，荷重メディアン値を用いたメディアンカット量子化と考えることができる．二つ目の拡張は，分割作業を複数回実行するのではなく，一回の作業で L 回に分割する点である．これら二つの拡張を施した量子化を本稿では一般化メディアンカット量子化と呼ぶ．

拡張 A: 画素値の種類による分割

提案法ではインデックス値を有効に利用するために，ヒストグラムの頻度が零でない濃度値に着目し，その濃度値の数が各クラスごとに均等になるように拡張する．この操作は，画像内に存在する濃度値の頻度を すべて 1 としたメディアンカット量子化に相当し，荷重メディアン値を用いたある特別なメディアンカット量子化と解釈することもできる．

荷重メディアン値 m' は，画素値 $P(x,y)$ の位置に対応した重み $w(x,y)$ を用いて，次式のように計算される．

$$m'=MED\{w(x,y)\diamond P(x,y), \quad \text{for all } x,y\} \tag{4}$$

ここで，$w(x,y)\diamond P(x,y)$ は，次式のように $P(x,y)$ の出現回数を $w(x,y)$ とする操作を示す．

$$w(x,y)\diamond P(x,y)=\underbrace{P(x,y),P(x,y),\cdots,P(x,y)}_{w(x,y)\text{ 個}} \tag{5}$$

本稿では，この重み $w(x,y)$ を

$$w(x,y)=\frac{1}{H(P(x,y))}, \qquad \text{for all } x,y \tag{6}$$

と選ぶ．ここで，$H(k)$ は濃度値 k における画像のヒストグラム頻度を示す．荷重メディアンの重みが頻度の逆数になっているため，画像内に存在する濃度値に対する頻度はすべて 1 となる．すなわち，この重み付きメディアン値 m' を用いた量子化は二つのクラスに分割され，二つのクラスに存在する非零の濃度値の種類は原則として同数となる．

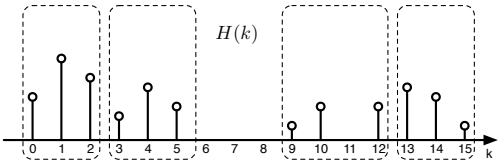
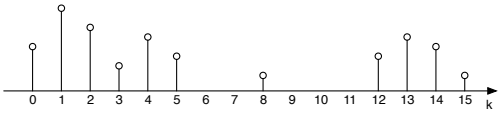


図 1: 提案法によるクラス分割例 ($N = 4$, $L = 4$)



(a) ヒストグラム $H(k)$

0	1	2	3	4	5	8	12	13	14	15
0	1	2	3	4	5	6	7	8	9	10

Ne = 11, s = 1, Nh = 3

(b) 有効濃度値の位置 m_i

0	0	1	1	2	2	-	-	3	-	-	-	4	5	6	7
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

(c) 量子化マッピングテーブル $Q(k)$

1	2	4	8	12	13	14	15
0	1	2	3	4	5	6	7

(d) 代表値テーブル $R(n)$

図 2: 提案法の例示 ($N = 4, L = 8$)

拡張 B: クラスの一括分割

従来のメディアンカット量子化では，ヒストグラムの分割作業を繰り返し行うことで，複数個のクラス分割を行っていた．この再帰的作業により分割数が多い場合には，計算時間が長くなってしまう．

一方，提案法では N ビットの画像から L 階調の画像へ一括量子化を行う．この処理によって，L 個のクラスには非零の濃度値の種類は原則として同数になる．図 2 に提案法の分割例を例示する．

上述の提案法の特徴によって，提案法は多くの疎性を持つ画像に対して，効果的に量子化することが可能となる．すなわち，L 階調への量子化において，L 個の代表値の個数の確保が原則保証されることになる．

3.2 提案法の実行手順

N ビットの画像 $P(x, y)$ に対して，L 階調に変換する量子化を行い，インデックス画像 $I(x, y)$ および代表値テーブル $R(n)$ を作成する手順を説明する．提案法ではまず，式 (2) の代表値テーブル $R(n)$ と，式 (1) の量子化マッピングテーブル $Q(k)$ を設計する．その後，この $Q(k)$ を用いて，インデックス画像 $I(x, y)$ を作成する．

(A) ヒストグラムと有効濃度値の個数 N_e の算出

まず画像からヒストグラムを算出し，有効濃度値の数 N_e を算出する．

(A-1) 画像 $P(x, y)$ からヒストグラム $H(k), k = 0, 1, \cdots, 2^N - 1$ を作成する (図 2(a) 参照).

(A-2) $H(k)$ から，有効濃度値 ($H(k) \neq 0$) の個数 N_e ($N_e \leq 2^N$) を算出する．

(A-3) 有効濃度値の位置 ($H(k) \neq 0$ となる k) を $m_0, m_1, \cdots, m_{N_e-1}$ とする (図 2(b) 参照).

(B) 量子化マッピングテーブル $Q(k)$ および代表値テーブル $R(n)$ の決定

有効濃度値の個数 N_e の値によって，手順を二つの場合に分類する．$N_e \leq L$ の場合には，以下のアルゴリズム (B1) を実行する．この条件では，適切な量子化を施すことによって，各クラスに割り当てられる有効濃度値は高々 1 個となり，量子化誤差の発生を回避できることに注意する．

(B1-1) $k = m_i$, $(i = 0, 1, \cdots, N_e - 1)$ の位置の量子化テーブル $Q(m_i)$ を設定する．

$$Q(m_i) = i, \qquad (i = 0, 1, \cdots, N_e - 1) \tag{7}$$

(B1-2) 代表値テーブル $R(n)$, $(n = 0, 1, \cdots, L - 1)$ を設定する．先頭の N_e 個には有効濃度値を，後半の $L - N_e - 1$ 個には，後述の簡略表現のために最終値と同じ値を入れておく．

$$R(n) = \begin{cases} m_n, & (n = 0, 1, \cdots N_e - 1) \\ m_{N_e-1}, & (n = N_e, \cdots, L - 1) \end{cases} \tag{8}$$

一方，$N_e > L$ の場合には，以下のアルゴリズム (B2) を実行する．この条件では，一般に量子化誤差が発生する．本アルゴリズムでは，L 個の各クラスに原則等しい個数の有効濃度値を割り振る．

(B2-1) N_e を L で割り，小数点以下を切り捨てた値を s とする．

$$s = \lfloor N_e / L \rfloor \tag{9}$$

ここで，$\lfloor s \rfloor$ は s を超えない最大の整数である．各クラスに s 個の有効濃度値を割り振った場合，

$$N_h = N_e - Ls \tag{10}$$

個の有効濃度値の割り振りが残る．そのため，N_h 個のクラスには $s+1$ 個の有効濃度値を割り振る必要がある．どのクラスに $s+1$ 個の有効濃度値を割り振るかについては，自由度があるが，以下では説明を簡単にするため，前半の N_h 個のクラスに $s+1$ 個の有効濃度値を配布する場合を例に示す．

(B2-2) $k = m_i, (i = 0, 1, \cdots, N_e - 1)$ の位置の量子化テーブル $Q(m_i)$ を設定する (図 2(c) 参照)．

$$Q(m_i) = \begin{cases} \left\lfloor \dfrac{i}{s+1} \right\rfloor, & (i < N_h(s+1)) \\ \left\lfloor \dfrac{i - N_h(s+1)}{s} \right\rfloor + N_h, & (i \geq N_h(s+1)) \end{cases} \quad (11)$$

(B2-3) 代表値テーブル $R(n)$ を設定する．各クラスの代表値の決定についても，自由度があるが，輝度値の重心を例にすると以下のように定式化できる (図 2(d) 参照)．

$$R(n) = \begin{cases} \dfrac{\sum_{i=(s+1)n}^{(s+1)n+s} H(m_i) \times m_i}{\sum_{i=(s+1)n}^{(s+1)n+s} H(m_i)} & (n < N_h) \\ \dfrac{\sum_{i=sn+N_h}^{sn+Nh+s-1} H(m_i) \times m_i}{\sum_{i=sn+N_h}^{sn+N_h+s-1} H(m_i)} & (n \geq N_h) \end{cases} \quad (12)$$

(C) 量子化テーブル $Q(k)$ による量子化，代表値テーブル $R(n)$ による逆量子化

量子化および逆量子化は，(B) で作成した $Q(k)$ および $R(n)$ を用いて，式 (1) および式 (2) で実現できる．ただし，$Q(k)$ の値域および $R(n)$ の定義域は，2^M ではなく設定した階調数 L になっている．

3.3 代表値テーブルの表現

提案法において N ビット画像を復元するためには，インデックス画像だけでなく，L 個の要素を持つ代表値テーブル $R(n)$ を用意する必要がある．ただし，代表値テーブル $R(n)$ は単調増加数列となっており，差分を保存するなどの工夫により，効率的に符号化することが可能である．今，代表値テーブル $R(n)$ から以下のように差分テーブル $R_d(n)$ を作成する．

(a) Cafe
$(4,992 \times 3,328, N = 16,$
$N_e = 64,486)$

(b) CT image
$(512 \times 512, N = 12,$
$N_e = 1,853)$

図 3: 使用した画像

$$R_d(n) = \begin{cases} R(0), & (n = 0) \\ R(n) - R(n-1), & (n = 1, \cdots, L-1) \end{cases} \quad (13)$$

$R(n)$ は単調増加であるため，$R_d(n)$ は値の小さい正の整数値になっている．符号化等ではこの $R_d(n)$ を保持すればよい．$R_d(n)$ からは以下の式で，$R(n)$ を復元できる．

$$R(n) = \begin{cases} R_d(0), & (n = 0) \\ R_d(n) + R(n-1), & (n = 1, 2, \cdots, L-1) \end{cases} \quad (14)$$

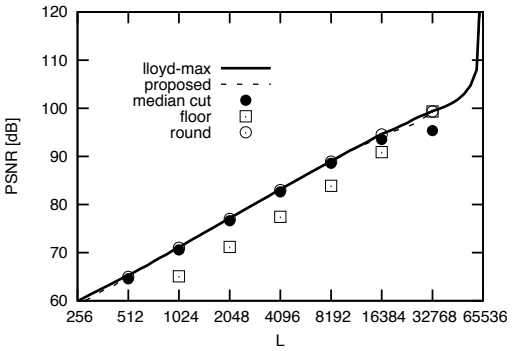
4 評価と応用例

4.1 比較

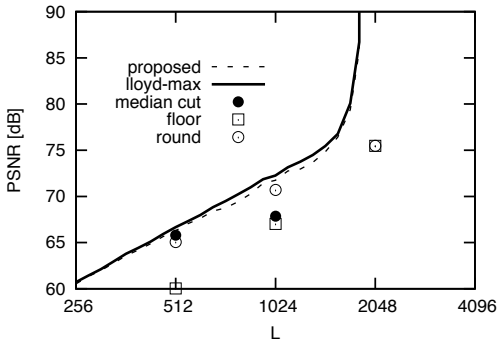
提案法の有効性をシミュレーションで確認する．使用する画像は，図 3 に示す 16 ビット (最大 65,536 階調) グレースケール画像 Cafe と 12 ビット (最大 4,096 階調) グレースケール画像 CT image を利用した．前者の Cafe は，サイズが $4,992 \times 3,328$ で，有効濃度値の数 N_e が 64,486 階調である疎性の低い画像である．一方，CT image はサイズが 512×512 で，有効濃度値の数 N_e が 1,853 階調である疎性が高い画像である．

図 4 は，これらの画像に対して量子化を行い，量子化後の階調数 L と逆量子化画像の原画像に対する PSNR の関係を示したものである．量子化は，提案法 (proposed)，Lloyd-Max およびメディアンカット (median cut)，ビット単位の切り捨て (floor) および丸め (round) を行った．後者二つの均等量子化と異なり，前者三つの量子化では，有効濃度値以上の L に対して，ロスレスで量子化可能になっている．提案法は，Lloyd-Max に及ばないものの，ほぼ同等の歪み特性が得られている．疎性が低い画像 Cafe の場合，提案法は丸め処理とあまり PSNR で差が見られないが，疎性が高い画像 CT image の場合，丸め処理に比べ高い PSNR が得られている．

図 5 は量子化にかかった時間を示したものである．シミュレーションは Core2 duo 2.8[GHz] の CPU を持つ PC 上で動作する MATLAB で実行した．ただし，MATLAB に標準で用意されている lloyds 関数は，再帰の度に量子化を繰り返すため非常に動作が遅い．そのため，ここで



(a) 画像 Cafe に対する結果



(b) 画像 CT image に対する結果

図 4: 逆量子化画像の原画像に対する PSNR(L:量子化後の階調数)

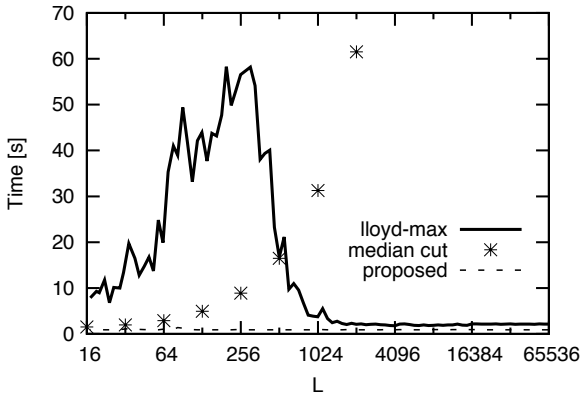


図 5: 量子化の実行時間 (L:量子化後の階調数)

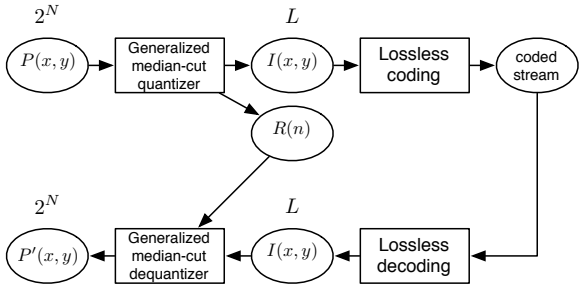


図 6: ロッシー符号化システム

は [2] で使用されているヒストグラムを用いたアルゴリズムに修正している．Lloyd-max 量子化はクラスの組み合わせのパターン数が増大する 256 付近でもっとも時間がかかっている．また，メディアンカット量子化は分割を繰り返し実行するため階調数に比例して時間が増加している．一方，提案法は L に依存せず，ほぼ一定時間で処理が終了していることがわかる．

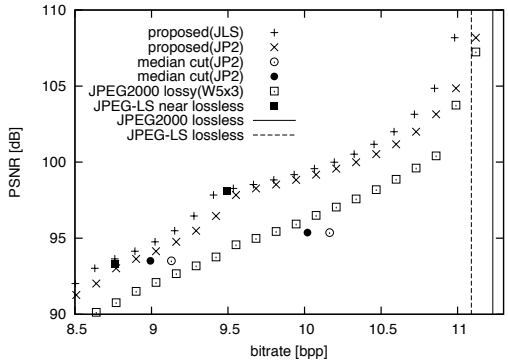
4.2 ロッシー符号化

提案法の応用例として，提案する量子化とロスレス符号化を組み合わせたロッシー符号化について説明する．図 6 にシステムの構成図を示す．符号化側では，ロスレス符号化器の前処理として，一般化メディアンカット量子化器を配置している．復号側では，ロスレス復号されたインデックス画像が一般化メディアンカット逆量子化され，画像 $P'(x,y)$ が復元される．復号画像の歪み特性は，一般化メディアンカット量子化器の部分で制御される．

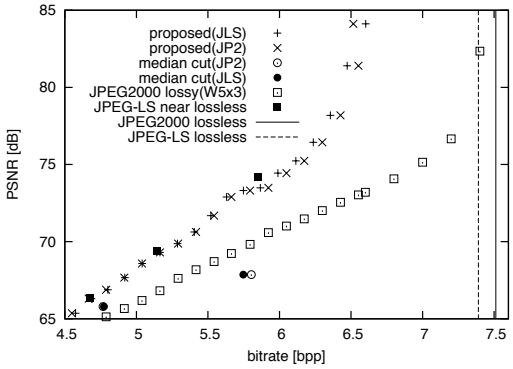
また，JPEG2000 符号化では，疎なヒストグラムを持つ画像の場合，ヒストグラムパッキングすることで，ビットレートを低減できることが知られている [8,9]．提案法で作成されるインデックス画像 $I(x,y)$ は，0〜$L-1$ までのインデックス値のみで構成されており，そのヒストグラムはパッキングされた状態になっている．すなわち，量子化によるインデックス化はビットレート低減効果も期待される．

なお，提案法ではテーブル値のデータ量も考慮する必要がある．提案法では，テーブル値のデータは式 (13) によりテーブル $R(n)$ の差分値 $R_d(n)$ に変換されている．そのため，値自体は非常に小さく分散も低いため，圧縮がしやすいデータ構造になっている．今回は，$R_d(n)$ をバイナリで書き出したファイルを bzip2 アルゴリズム [10] で圧縮した．例えば，Cafe において，$L = 64,486$ とした場合，バイナリデータ自体は 128,972[bytes] であるが，bzip2 圧縮すると 241 [bytes] となった．これは，符号化後の JPEG2000 ファイルの 0.001% に相当する大きさである．

図 7 は符号化画像のビットレートと歪みの関係を示したものである．図 3 の二つの画像に対して，提案法によるロッシー符号化，メディアンフィルタによるロッシー符号化，整数フィルタ (W5x3) による JPEG2000 lossy 符号化，JPEG-LS ニアロスレス符号化を行った．提案法のロスレス符号化は，JPEG2000 と JPEG-LS の双方について実行している．図中の実線と破線はそれぞれ JPEG2000 および JPEG-LS でロスレス符号化した時のビットレートを示している．



(a) Cafe



(b) CT image

図 7: 復号画像のレート歪み曲線

提案法は，JPEG2000 の lossy 符号化に比べ，高い PSNR を実現しており，JPEG-LS のニアロスレスモードとほぼ同等の PSNR を示している．また，JPEG-LS のニアロスレスモードに比べ，細かいレート歪み制御が行えていることがわかる．文献 [4] の方法も，提案法と同様にヒストグラムのパッキングを行いニアロスレスを実現するものであるが，JPEG-LS のニアロスレスと同様に粗いレート制御しか行えない．また，図 7(b) に示されるように，元々疎性が高い画像に対しては，提案するニアロスレス符号化の持つヒストグラムパッキング効果が大きく，ロスレスに比べて 1[bpp] 弱のレート低減効果が見られる．

5 おわりに

本報告では，従来のメディアンカット量子化を拡張し，量子化後の階調数を任意に設定可能な一般化メディアンカット量子化法を提案した．提案する量子化法は，Lloyd-Max 法に比べ高速に動作し，ほぼ同程度の量子化誤差が実現できる．また，提案法とロスレス符号化を組み合わせることで，ニアロスレス符号化が実現できることを示した．このニアロスレス符号化器は，従来の JPEG-LS 等のニアロスレス符号化に比べて，歪み誤差を細かく制御することができる．

参考文献

[1] 伊藤健, 板東幸浩, 高村誠之, 上倉一人，八島由幸: “ビット深度変換処理を用いた高ビット深度画像の符号化に関する基礎検討,” 2009 信学総大, S-5，2009.

[2] Takeshi Ito, Yukihiro Bandoh, Takamura Seishi, Hirohisa Jozawa: “A Coding Method for High Bit-Depth Images Based on Optimized Bit-Depth Transform”, 17th International Conference on Image Processing, pp.3141–pp.3144 2010.

[3] M. Aguzzi and M. Albanesi, “A novel approach to sparse histogram image lossless compression using JPEG2000,” Electronic letters on computer vision and image, 2006.

[4] E. Nasr-Esfahani, S. Samavi, N. Karimi, and S. Shiran, “Near lossless image compression by local packing of histogram,” IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1197—1200, 2008.

[5] Lloyd, S.P., “Least Squares Quantization in PCM,” IEEE Transactions on Information Theory, Vol. IT-28, pp. 129–137, March, 1982.

[6] Max, J., “Quantizing for Minimum Distortion,” IRE Transactions on Information Theory, Vol. IT-6, pp. 7—12, March, 1960.

[7] P. Heckbert, “Color image quantization for frame buffer display,” Computer Graphics, vol.16, no.3, July 1982.

[8] A. J. Pinho, “An Online Preprocessing Technique for Improving the Lossless Compression of Images with Sparse Histograms,” Signal Processing Letters, IEEE, vol. 9, no. 1, pp. 5—7, 2002.

[9] P. J. S. G. Ferreira and A. J. Pinho, “Why Does Histogram Packing Improve Lossless Compression Rates?,” IEEE Signal Processing Letters, vol. 9, no. 8, pp. 259—261, Aug. 2002.

[10] bzip2 web page: http://www.bzip.org/ .