New Properties of AMBTC

Kai-Kuang Ma 1, Member, IEEE
Sarah A. Rajala 2, Senior Member, IEEE

1 Institute of Microelectronics
National University of Singapore
Block 750E, Chai Chee Rd, #07-03/04
Chai Chee Industrial Park
Singapore 1646

2 Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7911

EDICS Category: 1.1 Image Coding

Abstract — Eight new properties of absolute moment block truncation coding (AMBTC) are presented with proof. The main purposes of this paper are two-fold: (1) provide fundamental insights into the AMBTC algorithm and (2) show that AMBTC is the optimum choice among the 1-bit moment preserving quantizers for achieving the minimum mean square error and the minimum computations. Coding performance simulated in the full-band and subband environments further demonstrates its consistency with the theoretical performance bounds.

Corresponding Address:

Dr. Sarah A. Rajala, Professor
Department of Electrical and Computer Engineering
North Carolina State University
Campus Box 7911
Raleigh, NC 27695-7911
Telephone: (919) 515-5114
Telefax: (919) 515-5523
Email: sar@eos.ncsu.edu
1 Introduction

A moment preserving quantizer (MPQ) compresses a set of discrete samples subject to preserving certain moments of these data. Since no transformation technique is involved, it is essentially a time-domain coding process. In addition, since the quantization process is based on a block of pixels, the MPQ actually performs vector quantization. In general, a family of moment preserving quantizers can be formed by preserving various sample moments in each quantizer. A special class of moment preserving quantizer, called absolute moment block truncation coding (AMBTC) [3], is further investigated here through the derivation and proof of eight new properties.

Compared with certain well-known coding methods (such as discrete cosine transform and vector quantization), the AMBTC algorithm is extremely simple computationally while providing competitive image quality at high bit rates [3]. However, the image quality severely degrades at low bit rates. Recently, Ma and Rajala [1, 2] incorporated time-domain AMBTC with frequency-domain subband coding and a new bit allocation algorithm to yield subband AMBTC (or SAMBTC). The SAMBTC system produces high quality imagery at low bit rates for both monochrome and color images recorded in YIQ and $L^*u^*v^*$ color spaces [1]. In addition, SAMBTC results in a system that is suitable for real-time implementation and many applications [1]. Since the AMBTC algorithm is the fundamental building block of the SAMBTC system, it is important to understand its behavior. In this paper, eight new properties of the AMBTC are presented with proof.

This paper is organized as follows. Since a special relationship exists between the AMBTC and two-level (or 1-bit) minimum mean-square-error quantizer (MMSEQ),
both algorithms are briefly described in Section 2. Section 3 shows two properties, the mean and the variance of the AMBTC quantization error. Section 4 shows four properties describing the relationship among the AMBTC’s input, output, and the resulting quantization error. Section 5 proves there is a relationship between the AMBTC and the MMSEQ and indicates that the AMBTC is the optimum choice in the class of moment preserving quantizers subject to achieving the minimum mean square errors and minimum computations. To demonstrate this, two monochrome digital images, LENA and HOUSE (with $256 \times 256$, 8 bpp each), are used for simulation. The coding performance of the AMBTC and the MMSEQ in the full-band and subband environments are documented in Sections 6 and 7, respectively. The results are consistent with the derivations presented in Section 5. The last section is a conclusion.

2 AMBTC and MMSEQ

Given an image, it is first divided into nonoverlapping blocks. The AMBTC algorithm is then independently applied to each block such that the gray value of each pixel within the block is quantized into either level a or level b (called quantization levels). Since there are only two levels, this type of quantizer is called two-level (or 1-bit) quantizer. Assume the number of pixels which are quantized into a and b within a block is $(m - q)$ and $(q)$, respectively; where $m$ is the total number of pixels in that block. The AMBTC algorithm requires that the block (or local) mean $\eta$ and the first absolute central moment $\alpha = \frac{1}{m} \sum_{i=1}^{m} |x_i - \eta| = \frac{1}{m} \sum_{i=1}^{m} |y_i - \eta|$ are preserved; where $x_i$ and $y_i$ are the gray value of input and output pixels, respectively. That is,
at the quantizer's output,

\[ m\eta = \sum_{i=1}^{m} y_i = qb + (m - q)a \]  \hspace{1cm} (2.1)

\[ m\alpha = \sum_{i=1}^{m} |y_i - \eta| = q(b - \eta) + (m - q)(\eta - a). \]  \hspace{1cm} (2.2)

Furthermore, a more efficient way of computing \( \alpha \) was derived as [3, Equation (8)]

\[ \alpha = \frac{2}{m} \left[ \left( \frac{1}{m} \sum_{f_{x_i \geq \eta}} x_i \right) - \eta q \right] . \]  \hspace{1cm} (2.3)

Using Equations (2.1), (2.2), and (2.3), the quantization output levels were derived in

\[ a = \eta - \frac{m\alpha}{2(m-q)} = \frac{1}{m-q} \left( \sum_{f_{x_i \leq \eta}} x_i \right) = \frac{1}{m-q} \left( \sum_{i=1}^{m-q} x_{ai} \right) \]  \hspace{1cm} (2.4)

\[ b = \eta + \frac{m\alpha}{2q} = \frac{1}{q} \left( \sum_{f_{x_i > \eta}} x_i \right) = \frac{1}{q} \left( \sum_{i=1}^{q} x_{bi} \right) \]  \hspace{1cm} (2.5)

where \( x_{ai} \in \{x_i \mid x_i \text{ quantized to level } a\} \) with \( m-q \) elements in the set and \( x_{bi} \in \{x_i \mid x_i \text{ quantized to level } b\} \) with \( q \) elements in the set. The quantities \( m-q \) and \( q \) equal to the number of 0's and 1's in the bit plane, respectively. The quantization levels \( a \) and \( b \) are assumed to be 8 bits each.

Another special type of 1-bit quantizer is the MMSEQ. The two-level MMSEQ does not require any of the moments to be preserved, but optimally (in the mean-square-error sense) quantizes a block of samples into two quantization levels. The two-level MMSEQ has the same algorithm structure as the AMBTC except for using an optimum block threshold \( T \) (instead of using block mean \( \eta \) as required in the
AMBTC algorithm) for the generation of the two quantization levels as follows. The quantization output levels for the MMSEQ are:

\[
\tilde{a} = \frac{1}{m - p} \left( \sum_{x_i \leq T} x_i \right) \quad \text{and} \quad \tilde{b} = \frac{1}{p} \left( \sum_{x_i > T} x_i \right)
\]  

(2.6)

where \((m - p)\) and \((p)\) are the number of 0's and 1's in the bit plane, respectively. In general, the threshold \(T\) is found through an iterated search such that the minimum mean square error for each block of pixels is achieved.

To prove the properties, the following model of the AMBTC algorithm is used:

\[
y = x + e
\]  

(2.7)

where \(x\) is the input image (uncompressed), \(y\) is the output image (compressed) from the AMBTC, and \(e\) is the resulting quantization error. For numerical comparison, the mean-square error (MSE) and peak-to-peak signal-to-noise ratio (PSNR) are used. They are defined as:

\[
MSE \triangleq \frac{1}{m} \sum_{i=1}^{m} [y_i - x_i]^2 \quad \text{per pel}
\]  

(2.8)

where \(y_i \in \{a, b\}\) in AMBTC or \(y_i \in \{\tilde{a}, \tilde{b}\}\) in MMSEQ and

\[
PSNR \triangleq 10 \log_{10} \left( \frac{(255)^2}{MSE} \right) \quad \text{dB.}
\]  

(2.9)
3 AMBTC Quantization Error

Two properties of the quantization error, $E[e]$ and $\sigma_e^2$, are derived in this section.

**Property 1** The quantization (or truncation) error of AMBTC has zero mean, i.e. $E[e] = 0$.

**Proof:** By definition, the total quantization error for a given block is equal to the summation of all pixel values at the output, subtracted by the total pixel values at the input. In addition, by using the Equation (2.1), the proof can be shown as follows.

$$E[e] = \frac{1}{m} \sum_{i=1}^{m} e_i$$

$$= \frac{1}{m} \left( \left[ (m - q) a + (q) b \right] - m\eta \right)$$

$$= \left[ \left( \frac{m-q}{m} \right) a + \left( \frac{q}{m} \right) b \right] - \eta$$

$$= 0$$

Alternatively,

$$E[y] = E[x + e] \implies E[e] = E[y] - E[x] = \eta - \eta = 0.$$  \(\square\)

**Comments:** Preserving the block mean ($E[y] = E[x]$) implies that the quantizer output is an unbiased estimate of the input.
Property 2  The variance of the quantization error is equal to the input variance subtracted by the output variance, i.e. \( \sigma_e^2 = \sigma_x^2 - \sigma_y^2 \).

Proof:  To prove the first part of the statement,

\[
\]

\[\Rightarrow \eta^2 + \sigma_x^2 = \eta^2 + \sigma_y^2 - 2E[ye] + \sigma_e^2\]

By Property 3, thus \( \sigma_e^2 = \sigma_x^2 - \sigma_y^2 \).

Comments:  Since \( \sigma_e^2 \geq 0 \), this property implies that the average power of the quantization output is reduced by the average power of the quantization noise. By Property 1, \( \sigma_e^2 = E[e^2] \) is the MSE or noise power. An alternate method for determining the \( \sigma_e^2 \) is stated as Property 4.

4  The Relationship Among the Input, Output, and Quantization Error

Four properties are developed in this section; the quantities \( E[ye] \), \( E[xe] \), and \( E[xy] \) as well as the relationship between the input and output variances.

Property 3  The quantization error is orthogonal to the quantizer output, i.e. \( E[ye] = 0 \).

Proof:

\[
E[ye] = \frac{1}{m} \sum_{i=1}^{m} y_i e_i = \left( \frac{1}{m-q} \right) \sum_{i=1}^{m-q} y_i(y_i - x_{ai}) + \left( \frac{1}{q} \right) \sum_{i=1}^{q} y_i(y_i - x_{bi})
\]

For inputs \( x_{ai} \) and \( x_{bi} \), the quantization outputs are a and b, respectively. Now it will be shown that each summation term in the above equation is zero. For the first
summation term,

\[ \sum_{i=1}^{m-q} y_i(y_i - x_{ai}) = \sum_{i=1}^{m-q} a(a - x_{ai}) = a \left[ \frac{\sum_{i=1}^{m-q} 1}{m-q} \right] - \sum_{i=1}^{m-q} x_{ai} \]

\[ = a \left[ (m-q)a - \left( \sum_{i=1}^{m-q} x_{ai} \right) \right] = 0. \quad \text{(by Equation (2.4))} \]

Similarly, for the second summation term,

\[ \sum_{i=1}^{q} y_i(y_i - x_{bi}) = \sum_{i=1}^{q} b(b - x_{bi}) = b \left[ \frac{\sum_{i=1}^{q} 1}{q} \right] - \sum_{i=1}^{q} x_{bi} \]

\[ = b \left[ qb - \left( \sum_{i=1}^{q} x_{bi} \right) \right] = 0. \quad \text{(by Equation (2.5))} \]

Hence, \( E[ye] = 0 \). That is, the quantization error and the quantizer output are orthogonal. \( \square \)

**Property 4** The signal and quantization noise are nonpositively correlated at the output of the AMBTC quantizer, i.e. \( E[xe] = -\sigma_e^2 \leq 0 \).

**Proof:** By Property 3,

\[ 0 = E[ye] = E[(x + e)e] = E[xe] + E[e^2] = E[xe] + \sigma_e^2 \]

Hence, \( \sigma_e^2 = -E[xe] \). Alternatively,

\[ E\left[y^2\right] = E\left[(x + e)^2\right] = E\left[x^2\right] + 2E[xe] + E\left[e^2\right] \]
Knowing that AMBTC preserves the first moment and using Property 2,

\[
\sigma_y^2 + \eta^2 = \sigma_x^2 + \eta^2 + 2E[xe] + \sigma_e^2
\]

\[
E[xe] = \frac{1}{2} \left[ \sigma_y^2 - \sigma_x^2 - \sigma_e^2 \right] = -\sigma_e^2 \leq 0
\]

The equality holds when all the pixels within the block have the same value. \( \square \)

Comments: The nonpositive correlation between the random variables \( x \) and \( e \) matches the characteristics of the AMBTC model. Since \( y = x + e \), \( y \) preserves the moment of \( x \), and due to the moment conservation property, an increase in the quantization error \( e \) causes a decrease in the input signal \( x \), and vice versa.

Property 5  The correlation of the input and the output is equal to the second moment of the output, i.e. \( E[xy] = E[y^2] \).

Proof:

\[
E[y^2] = E[(x + e)y] = E[xy] + E[ye]
\]

by Property 3, thus \( E[xy] = E[y^2] \). \( \square \)

Before proving the next property, it is necessary to introduce a family of means called the exponential mean \([4]\) or power mean \([5]\). It is defined as follows.

Definition  The number

\[
C_\nu = \left( \frac{x_1^\nu + x_2^\nu + \ldots + x_n^\nu}{n} \right)^{\frac{1}{\nu}} \quad \text{where} \quad \nu \neq 0.
\]
is called the exponential mean or power mean of order $\nu$ of the numbers $x_1, x_2, \ldots, x_n$. In particular, the first-order power mean

$$C_1 = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

is the arithmetic mean of the numbers $x_1, x_2, \ldots, x_n$, and the second-order power mean

$$C_2 = \left(\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}\right)^{\frac{1}{2}}$$

is called the quadratic mean or root mean square mean of the numbers $x_1, x_2, \ldots, x_n$. It is worth mentioning two other power means: the harmonic mean, $C_{-1}$ and geometric mean, $C_0$. Note, the parameter $\nu$ is not allowed to be zero in the definition. However, when $\nu = 0$, $C_0$ can be shown to be the geometric mean using L'Hôpital's Rule [5]. There is an important relationship among these means, stated as theorems in [5] and [4] for various ranges of power $\nu$. The theorems can be combined together and rephrased as follows.

**Theorem**  If $p$ and $q$ are any positive number and $p < q$, then

$$C_{-1} \leq C_0 \leq C_p \leq C_q$$

where the equality holds if and only if $x_1 = x_2 = \ldots = x_n$.

**Proof:** See [5] and [4].

Of particular interest is the relationship; harmonic mean $\leq$ geometric mean $\leq$ arithmetic mean $\leq$ quadratic mean. Both [5] and [4] provide tedious proofs for the general case. Here, a different approach is used to prove the relationship between the
arithmetic mean and quadratic mean and is stated as a lemma below. This lemma will be used in Property 6.

Lemma  

The arithmetic mean is equal to or less than the quadratic mean. That is,

\[
\frac{x_1 + x_2 + \cdots + x_n}{n} \leq \left( \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} \right)^{\frac{1}{2}} \tag{4.1}
\]

Proof: For a positive integer \( n \) and \( n \geq 2 \) (i.e., considering two terms at least) where \( i, j = 1, \ldots, n \). It is true that

\[
\left( \frac{n - 1}{2} \right) (x_1 - x_2 - \cdots - x_n)^2 = \left( \frac{n - 1}{2} \right) \left( \sum_{i=1}^{n} x_i^2 - 2 \sum_{i \neq j} x_i x_j \right) \geq 0,
\]

Hence,

\[
\left( \frac{n - 1}{2} \right) \sum_{i=1}^{n} x_i^2 \geq (n - 1) \sum_{i \neq j} x_i x_j \geq \sum_{i \neq j} x_i x_j \tag{4.2}
\]

Assuming the lemma is true and taking square of both sides of the inequality, Equation (4.1):

\[
\left( \sum_{i=1}^{n} x_i \right)^2 \leq n \sum_{i=1}^{n} x_i^2
\]

\[
\sum_{i=1}^{n} x_i^2 + 2 \sum_{i \neq j} x_i x_j \leq n \sum_{i=1}^{n} x_i^2
\]

\[
\left( \frac{n - 1}{2} \right) \sum_{i=1}^{n} x_i^2 \geq \sum_{i \neq j} x_i x_j \equiv \text{Equation (4.2)}
\]

Hence, the arithmetic mean is less than or equal to quadratic mean. Note that the equality holds when all \( x_i \) have the same value. \( \Box \)
Property 6 The variance of the output is less than or equal to the variance of the input, that is, $\sigma^2_x \geq \sigma^2_y$. (Note that the difference of the two is equal to $\sigma^2_e$, by Property 2)

Proof: There are two ways to prove it. Since $0 \leq \sigma^2_e$ and $\sigma^2_e = \sigma^2_x - \sigma^2_y$, $\sigma^2_x$ must be greater than or equal to $\sigma^2_y$. The proof is complete.

Alternatively, using the fact that the first moment is preserved at the output of AMBTC then,

$$\sigma^2_y = \frac{1}{m} \left[ (m-q)(a-\eta)^2 + q(b-\eta)^2 \right] = \frac{1}{m} \left[ \sum_{i=1}^{m-q} (a-\eta)^2 + \sum_{i=1}^{q} (b-\eta)^2 \right]$$

(4.3)

The input variance is

$$\sigma^2_x = \frac{1}{m} \sum_{i=1}^{m} (x_i - \eta)^2 = \frac{1}{m} \left[ \sum_{i=1}^{m-q} (x_{ai} - \eta)^2 + \sum_{i=1}^{q} (x_{bi} - \eta)^2 \right]$$

(4.4)

In order to prove that $\sigma^2_x - \sigma^2_y \geq 0$, consider the difference between the first summation term in Equations (4.3) and (4.4):

$$\left[ \sum_{i=1}^{m-q} (x_{ai} - \eta)^2 \right] - \left[ \sum_{i=1}^{m-q} (a-\eta)^2 \right]$$

$$= \sum_{i=1}^{m-q} \left[ x_{ai}^2 + \eta^2 - 2\eta x_{ai} - a^2 - \eta^2 + 2a\eta \right]$$

$$= \sum_{i=1}^{m-q} x_{ai}^2 - 2\eta \sum_{i=1}^{m-q} x_{ai} - (m-q)a^2 + 2a \eta (m-q)$$
Using Equation (2.4), the second term is canceled by the fourth term and the above equation can be simplified to

\[
\sum_{i=1}^{m-q} x_{ai}^2 - (m - q) \left[ \frac{1}{m - q} \sum_{i=1}^{m-q} x_{ai} \right]^2 = \sum_{i=1}^{m-q} x_{ai}^2 - \frac{1}{m - q} \left( \sum_{i=1}^{m-q} x_{ai} \right)^2 \geq 0
\]

That is,

\[
\left[ \left( \frac{1}{m - q} \sum_{i=1}^{m-q} x_{ai}^2 \right)^{\frac{1}{2}} \right] \geq \left( \frac{1}{m - q} \sum_{i=1}^{m-q} x_{ai} \right)
\]

Here the left-hand term is the quadratic mean and the right-hand term is the arithmetic mean. From the lemma, it is known that this difference must be greater than or equal to zero. Similarly, it can be shown for the second summation term in Equations (4.4) and (4.3) that

\[
\sum_{i=1}^{q} (x_{bi} - \eta)^2 \geq \sum_{i=1}^{q} (b - \eta)^2
\]

Hence,

\[
\sigma_x^2 \geq \sigma_y^2
\]

The equality holds if all the pixels within the block have the same value. In such a case, there is no quantization error for that block, that is, \( \sigma_e^2 = 0 \) and \( \sigma_y^2 = \sigma_x^2 \). □
5 Relationship Between AMBTC and Minimum Mean-Square-Error Quantizer

Property 7 The AMBTC is optimum in the minimum mean-square error sense under the constraint of preserving the first absolute central moment. Furthermore, if the AMBTC is equivalent to the two-level MMSEQ, then the bit plane has an equal number of 0's and 1's [6].

Proof: Using the Lagrange multiplier optimization technique and the MSE as the fidelity criterion for measuring distortion, the performance index $J_{MSE}(a, b)$ can be defined as

$$
J_{MSE}(a, b) = \sum_{x_i \leq \eta} (x_i - a)^2 + \sum_{x_i > \eta} (x_i - b)^2
$$

$$
= \sum_{x_i \leq \eta} x_i^2 - 2a \left( \sum_{x_i \leq \eta} x_i \right) + a^2 \left( \sum_{x_i \leq \eta} 1 \right)
$$

$$
+ \sum_{x_i > \eta} x_i^2 - 2b \left( \sum_{x_i > \eta} x_i \right) + b^2 \left( \sum_{x_i > \eta} 1 \right) \quad (5.1)
$$

First, it is proven that the quantization levels $a$ and $b$ of AMBTC do yield minimum mean square error (MMSE) under the constraint of preserving the first absolute moment. Let $\lambda$ and $\mathcal{H}(\cdot)$ be the Lagrange multiplier and the Lagrangian (or Hamiltonian), respectively. The optimization problem can then be formulated as:
\[ \mathcal{H}(a, b, \lambda) = J_{MSE}(a, b) + \lambda \left[ \left( \sum_{i=1}^{m} |x_i - \eta| \right) - m \alpha \right] \]

\[ = J_{MSE}(a, b) + \lambda [q(b - \eta) + (m - q)(\eta - a) - m \alpha] \quad (5.2) \]

where \( \alpha \) is the first absolute central moment and the last equality is obtained by using Equation (2.2). To find the stationary point, the necessary conditions are: (* denotes the stationary point or extremum)

\[ \frac{\partial \mathcal{H}(\cdot)}{\partial a} \bigg|_{a=a^*} = 0; \quad \frac{\partial \mathcal{H}(\cdot)}{\partial b} \bigg|_{b=b^*} = 0; \quad \frac{\partial \mathcal{H}(\cdot)}{\partial \lambda} = 0 \quad (5.3) \]

Combining Equations (5.3a) and (2.1), the Lagrange multiplier \( \lambda \) is derived as follows:

\[ 0 = \frac{\partial \mathcal{H}(\cdot)}{\partial a} \bigg|_{a=a^*} = -2 \left( \sum_{x_i \leq \eta} x_i \right) + 2a^* \left( \sum_{x_i \leq \eta} 1 \right) - \lambda (m - q) \quad (5.4) \]

\[ \Rightarrow \lambda = -\frac{2}{m - q} \left( \sum_{x_i \leq \eta} x_i \right) + 2a^* \]

\[ = -\frac{2}{m - q} \left( \sum_{x_i \leq \eta} x_i \right) + \left( \frac{2m}{m - q} \right) \eta - \left( \frac{2q}{m - q} \right) b^* \quad (5.5) \]

Combining Equations (5.3b) and (5.5), the optimum quantized level \( b^* \) is derived as follows:

\[ 0 = \frac{\partial \mathcal{H}(\cdot)}{\partial b} \bigg|_{b=b^*} = 2 \left( \sum_{x_i > \eta} x_i \right) - 2b^* \left( \sum_{x_i \geq \eta} 1 \right) + \lambda q \quad (5.6) \]
\[ b^* = \left( \frac{1}{2} q \right) \left[ 2 \left( \sum_{x_i > \eta} x_i \right) + \lambda q \right] = \frac{1}{q} \left( \sum_{x_i > \eta} x_i \right) + \frac{\lambda}{2} \]

\[ = \frac{1}{q} \left( \sum_{x_i > \eta} x_i \right) - \frac{1}{m-q} \left( \sum_{x_i \leq \eta} x_i \right) + \left( \frac{m}{m-q} \right) \eta - \left( \frac{q}{m-q} \right) b^* \]

Hence,

\[ \left( \frac{m}{m-q} \right) b^* = \frac{1}{q} \left( \sum_{x_i > \eta} x_i \right) - \frac{1}{m-q} \left( \sum_{x_i \leq \eta} x_i \right) + \left( \frac{m}{m-q} \right) \eta \quad (5.7) \]

\[ b^* = \frac{m-q}{mq} \left( \sum_{x_i > \eta} x_i \right) - \frac{1}{m} \left( \sum_{x_i \leq \eta} x_i \right) + \eta \]

\[ = \frac{m-q}{mq} \left( \sum_{x_i > \eta} x_i \right) - \frac{1}{m} \left( \sum_{x_i \leq \eta} x_i \right) + \frac{1}{m} \left( \sum_{x_i > \eta} x_i \right) + \frac{1}{m} \left( \sum_{x_i \leq \eta} x_i \right) \]

\[ = \left( \frac{m-q}{mq} + \frac{q}{mq} \right) \left( \sum_{x_i > \eta} x_i \right) \]

\[ = \frac{1}{q} \left( \sum_{x_i > \eta} x_i \right) \]

\[ = b \quad (5.8) \]

Similarly, \( a^* \) can be found by using Equations (2.1) and (5.8):

\[ a^* = \frac{m \eta - q b^*}{m-q} \]

\[ = \frac{1}{m-q} \left( m \eta - \sum_{x_i > \eta} x_i \right) \quad (5.9) \]
From Equations (5.8) and (5.10), the quantization levels of AMBTC is the stationary point, i.e., \((a, b) = (a^*, b^*)\).

The second order derivatives, Hessian or curvature matrix, is required to determine whether the stationary point \((a, b)\) is a minimum or a maximum point. The sufficient condition is:

\[
H = \begin{bmatrix}
\frac{\partial^2 \mathcal{H}(\cdot)}{\partial a^2} & \frac{\partial^2 \mathcal{H}(\cdot)}{\partial a \partial b} \\
\frac{\partial^2 \mathcal{H}(\cdot)}{\partial b \partial a} & \frac{\partial^2 \mathcal{H}(\cdot)}{\partial b^2}
\end{bmatrix} = \begin{bmatrix}
(m - q) & 0 \\
0 & q
\end{bmatrix}
\]  

(5.11)

The Hessian matrix \(H\) is positive semi-definite due to the fact that all the major minors are greater than or equal to zero; that is, \((m - q) \geq 0\) and \(|H| = (m - q)q \geq 0\). The equality holds when all the pixels of the current block have the same gray levels (i.e. \(q = 0\) and \(H\) is positive definite). Therefore, the stationary point \((a, b)\) is a minimum point. Hence, the AMBTC does provide the minimum mean-square error under the constraint of preserving the first absolute central moment.

On the other hand, the two-level MMSEQ quantization levels \(\bar{a}\) and \(\bar{b}\) defined in Equation (2.6) yield the minimum mean square error without constraint. By com-
paring Equations (2.6), (2.4), and (2.5), the AMBTC is obviously identical to the two-level MMSEQ when \( T = \eta \) (thus \( p = q \)).

Finally, it is true that for a given pair of quantization levels \( \bar{a} \) and \( \bar{b} \) for the 1-bit quantizer, the optimum threshold for that block is \( (\bar{a} + \bar{b})/2 \), in the minimum mean-square error sense [7]. Hence,

\[
T = \frac{\bar{a} + \bar{b}}{2}
\]

Comparing the coefficients of Equations (5.12) and (2.1), it can be seen that the AMBTC converges to the two-level MMSEQ, i.e., if \( T = \eta, \bar{a} = a, \bar{b} = b \), then \( \frac{m-q}{m} = \frac{q}{m} = \frac{1}{2} \). Thus, the bit plane has an equal number of 0’s and 1’s. □

**Property 8** If the bit plane has an equal number of 0’s and 1’s, then the output standard deviation \( \sigma_y \) is equal to the first absolute central moment \( \alpha \) (i.e. \( \sigma_y = \alpha \)) and the quantizer outputs \( a = \eta - \alpha \) and \( b = \eta + \alpha \). In addition, the skewness \( \gamma \) is zero.

**Proof:** For AMBTC, \( \alpha \) is conserved; that is,

\[
\alpha \triangleq \alpha_x = \frac{1}{m} \sum_{i=1}^{m} |x_i - \eta| = \frac{1}{m} \sum_{i=1}^{m} |y_i - \eta|
\]

\[
= \alpha_y = \left( \frac{q}{m} \right) (b - \eta) + \left( \frac{m-q}{m} \right) (\eta - a)
\]

where \( \alpha_x \) and \( \alpha_y \) are the first absolute central moment of input and output, respectively. When the bit plane contains an equal number of zeros and ones, that is, \( \left( \frac{m-q}{m} \right) = \left( \frac{q}{m} \right) = \frac{1}{2} \), then \( \alpha_y = (b - a)/2 \) and \( \alpha_y^2 = (a^2 + b^2 - 2ab)/4 \). On
the other hand,

\[
\sigma_y^2 \triangleq \left( \frac{1}{m} \sum_{i=1}^{m} y_i^2 \right) - \eta^2
= \frac{1}{m} \left[ (m-q)a^2 + qb^2 \right] - \eta^2
= \frac{a^2 + b^2}{2} - \left( \frac{a + b}{2} \right)^2
= \frac{1}{4} (a^2 + b^2 - 2ab) = \alpha^2 = \alpha^2
\]

(5.13)

Hence, \( \alpha = \sigma_y \). From Equations (2.4) and (2.5), \( a \) and \( b \) can now be shown to be
\( a = \eta - \alpha \) and \( b = \eta + \alpha \).

To prove the second part of the property, use the definition of the skewness.

\[
\gamma_y \triangleq \frac{E[(y - \eta)^3]}{\sigma_y^3} = \frac{\bar{y}^3 - 3 \eta \bar{y}^2 + 2 \eta^3}{\sigma_y^3}
= \frac{1}{\sigma_y^3} \left[ \frac{a^3 + b^3}{2} - 3 \left( \frac{a + b}{2} \right) \left( \frac{a^2 + b^2}{2} \right) + 2 \left( \frac{a + b}{2} \right)^3 \right] = 0.
\]

6 Performance Comparison in the Full-Band Case

It has been shown that both AMBTC and MMSEQ have identical formulae for generating the output levels for an 1-bit quantizer. However, the AMBTC and the MMSEQ do differ in the calculation of the quantization threshold. AMBTC uses the block mean as the threshold while MMSEQ requires an iterated search for the optimum threshold on each block. Delp and Mitchell [8] suggest a simple method for
finding the threshold by testing every pixel value of the current block and picking the one which results the smallest MSE value as the threshold. This exhaustive search is inefficient and may not be practical when the window size is large. Hui [9] proposes the following algorithm to iteratively search for the threshold, and the required computations are significantly reduced.

Let \( x_{\text{min}} \) and \( x_{\text{max}} \) be the minimum and maximum gray-level values of the pixels in the block and be initially assigned to \( \bar{a} \) and \( \bar{b} \), individually. The corresponding optimum threshold \( T \) is calculated using Equation (5.12) as the initial threshold for an iterated search process. With this initial threshold \( T \), the new quantization levels \( \bar{a} \) and \( \bar{b} \) are derived using Equation (2.6) and are then compared with quantization levels of the previous stage to see if they are the same. If convergence fails, a new threshold is calculated using Equation (5.12) again based on new quantization levels. The search procedure is iteratively repeated until either threshold \( T \) or quantization levels \( \bar{a} \) and \( \bar{b} \) are unchanged. As a result, the minimum mean square error is achieved.

Before presenting the simulation results, we would like to point out that Hui's algorithm (as described above) is the same as the well-known LBG [10] algorithm for designing vector quantizers (VQ). Given a number of classes \( N \) in VQ, the LBG algorithm iteratively searches the optimum partition for a set of (training) samples and the representation level for each class such that the minimum MSE quantization distortion is achieved. The search process ends when the current distortion is fairly close to the distortion achieved in the previous iteration. The relationship between the Hui and LBG algorithms is as follows. Since there are two quantization levels in the MMSEQ, the number of VQ classes is set to two (i.e., \( N = 2 \)). For a given set of quantization levels \( \bar{a} \) and \( \bar{b} \), the optimum threshold is the mid-point between these two levels [7] (i.e., Equation (5.12)). On the other hand, for a fixed partition
T, the optimum representation level for each class is the centroid of the class. Notice that Equation (2.6) computes the centroid of each class in the MMSEQ. In addition, a fixed quantization level set \((\tilde{a}, \tilde{b})\) results in a fixed distortion \(D\). Therefore, either quantity can be used to terminate the search process (Hui [9] uses quantization levels and LBG [10] uses distortion).

Power-of-two window sizes are used in the simulation of AMBTC and MMSEQ for both LENA and HOUSE images. The results are documented in Table 6.1 for LENA and Table 6.2 for HOUSE. Given a fixed window size, the required number of iterations of the MMSEQ for each block vary. Generally speaking, the flat areas tend to have a smaller number of iterations compared with the busy areas. For the MMSEQ, the total number of iterations is the sum of the iterations of individual blocks (column 3 in Tables 6.1 and 6.2). For the various window sizes shown in Tables 6.1 and 6.2, the average number of iterations \(\text{per block}\) is equal to the total number of iterations (column 3) divided by the total number of windows (column 2), the results

<table>
<thead>
<tr>
<th>Window</th>
<th># Blocks</th>
<th>MMSEQ</th>
<th>AMBTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Ave.(/block)</td>
<td>Total</td>
</tr>
<tr>
<td>2 × 2</td>
<td>16384</td>
<td>31675</td>
<td>1.9333</td>
</tr>
<tr>
<td>4 × 4</td>
<td>4096</td>
<td>8797</td>
<td>2.1477</td>
</tr>
<tr>
<td>8 × 8</td>
<td>1024</td>
<td>2866</td>
<td>2.7988</td>
</tr>
<tr>
<td>16 × 16</td>
<td>256</td>
<td>1014</td>
<td>3.9609</td>
</tr>
<tr>
<td>32 × 32</td>
<td>64</td>
<td>294</td>
<td>4.5938</td>
</tr>
<tr>
<td>64 × 64</td>
<td>16</td>
<td>96</td>
<td>6.0000</td>
</tr>
<tr>
<td>128 × 128</td>
<td>4</td>
<td>29</td>
<td>7.2500</td>
</tr>
<tr>
<td>256 × 256</td>
<td>1</td>
<td>7</td>
<td>7.0000</td>
</tr>
</tbody>
</table>

Table 6.1: Iteration comparisons of block threshold searching between MMSEQ and AMBTC using LENA
<table>
<thead>
<tr>
<th>Window</th>
<th># Blocks</th>
<th>MMSEQ Total</th>
<th>Ave./block</th>
<th>AMBTC Total</th>
<th>Ave./block</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 × 2</td>
<td>16384</td>
<td>30019</td>
<td>1.8322</td>
<td>16384</td>
<td>1</td>
</tr>
<tr>
<td>4 × 4</td>
<td>4096</td>
<td>8813</td>
<td>2.1516</td>
<td>4096</td>
<td>1</td>
</tr>
<tr>
<td>8 × 8</td>
<td>1024</td>
<td>2989</td>
<td>2.9189</td>
<td>1024</td>
<td>1</td>
</tr>
<tr>
<td>16 × 16</td>
<td>256</td>
<td>1061</td>
<td>4.1445</td>
<td>256</td>
<td>1</td>
</tr>
<tr>
<td>32 × 32</td>
<td>64</td>
<td>361</td>
<td>5.6406</td>
<td>64</td>
<td>1</td>
</tr>
<tr>
<td>64 × 64</td>
<td>16</td>
<td>109</td>
<td>6.8125</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>128 × 128</td>
<td>4</td>
<td>25</td>
<td>6.2500</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>256 × 256</td>
<td>1</td>
<td>7</td>
<td>7.0000</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.2: Iteration comparisons of block threshold searching between MMSEQ and AMBTC using HOUSE

are list in the column 4. On the other hand, AMBTC requires only one iteration to quantize each block, i.e., no iterated search. Therefore, the AMBTC’s computational load is about one half to one seventh of the MMSEQ’s (compare columns 4 and 6 on both tables).

Note that both the AMBTC and the MMSEQ are time-domain (i.e., full-band) approaches. The coding performance in the full-band case has been evaluated for both AMBTC and MMSEQ using LENA and HOUSE. The numerical and perceptual performance results are documented as follows. For the numerical comparison, the MSE and PSNR are calculated for several window sizes and shown in the Tables 6.3 and 6.4 for LENA and HOUSE, respectively. For the perceptual comparison, the reconstructed images from both algorithm have been compared side by side. Their perceptual differences at each window size is consistently almost unnoticeable. To show the quality of the decoded images for both AMBTC and MMSEQ, a set of compressed images using 8 × 8 and 32 × 32 windows are presented in Figures 6.1 and

22
6.2 for LENA and in Figures 6.3 and 6.4 for HOUSE.

<table>
<thead>
<tr>
<th>LENA Image</th>
<th>AMBTC</th>
<th>MMSEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>Bit Rate, b (bpp)</td>
<td>MSE (/pel)</td>
</tr>
<tr>
<td>2 x 2</td>
<td>5.0</td>
<td>14.021</td>
</tr>
<tr>
<td>4 x 4</td>
<td>2.0</td>
<td>63.941</td>
</tr>
<tr>
<td>8 x 8</td>
<td>1.250</td>
<td>134.931</td>
</tr>
<tr>
<td>16 x 16</td>
<td>1.06250</td>
<td>229.467</td>
</tr>
<tr>
<td>32 x 32</td>
<td>1.015625</td>
<td>374.059</td>
</tr>
</tbody>
</table>

Table 6.3: The numerical performance comparisons of full-band AMBTC and MMSEQ for various bit rates using LENA.

<table>
<thead>
<tr>
<th>HOUSE Image</th>
<th>AMBTC</th>
<th>MMSEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>Bit Rate, b (bpp)</td>
<td>MSE (/pel)</td>
</tr>
<tr>
<td>2 x 2</td>
<td>5.0</td>
<td>20.947</td>
</tr>
<tr>
<td>4 x 4</td>
<td>2.0</td>
<td>90.528</td>
</tr>
<tr>
<td>8 x 8</td>
<td>1.250</td>
<td>168.017</td>
</tr>
<tr>
<td>16 x 16</td>
<td>1.06250</td>
<td>251.807</td>
</tr>
<tr>
<td>32 x 32</td>
<td>1.015625</td>
<td>409.759</td>
</tr>
</tbody>
</table>

Table 6.4: The numerical performance comparisons of full-band AMBTC and MMSEQ for various bit rates using HOUSE.

7 Performance Comparison in the Subband Case

Multi-dimensional subband coding theory was first introduced by Vetterli [11]. Later, Woods and O’Neil [12] utilized subband coding technique to compress monochrome digital images. Since then subband coding has been well recognized as a novel approach for compressing images and video [13]. Ma and Rajala [1, 2] recently introduced a new subband coding method which incorporates time-domain absolute
Figure 6.1: Full-band coded LENA images using $8 \times 8$ window (i.e., 1.25 bpp): (a) AMBTC and (b) MMSEQ.
Figure 6.2: Full-band coded LENA images using 32 × 32 window (i.e., 1.015625 bpp): (a) AMBTC and (b) MMSEQ.
Figure 6.3: Full-band coded HOUSE images using $8 \times 8$ window (i.e., 1.25 bpp): (a) AMBTC and (b) MMSEQ.
Figure 6.4: (a) and (b) Full-band coded HOUSE images using $32 \times 32$ window (i.e., 1.015625 bpp): (a) AMBTC and (b) MMSEQ.
moment block truncation coding along with a newly developed subband dynamic bit allocation algorithm [1] to yield subband AMBTC (or SAMBTC). It was shown that SAMBTC completely eliminates the blocking artifacts which AMBTC suffers at low bit rates. In summary, SAMBTC achieves two major objectives: (1) high quality monochrome and color imagery at low bit rates and (2) flexibility and practicality for real-time application and implementation [1]. Similarly, the MMSEQ algorithm can be substituted for the AMBTC algorithm in SAMBTC yielding the SMMSEQ. The overall image quality of SAMBTC and SMMSEQ is significantly superior to the one of AMBTC and MMSEQ at the same bit rate, respectively.

Again, it has been shown in Section 5 that both AMBTC and MMSEQ have identical formulae for generating the output levels for an 1-bit quantizer except for the difference in calculating the quantization thresholds. The performance of AMBTC and MMSEQ in the time domain (i.e., the full-band case) has been compared in Section 6. The MSE and PSNR between AMBTC and MMSEQ is quite small, and the difference of the image quality is almost unnoticeable. In this section, the performance of AMBTC and MMSEQ in the subband representation is compared. For numerical comparison, the MSE and PSNR of the SAMBTC and SMMSEQ are calculated for several bit rates and are shown in the Table 7.1 for LENA and Table 7.2 HOUSE. For perceptual comparison, the reconstructed images have been compared side by side. The perceptual difference in the reconstructed images is even smaller than their full-band counterparts. To illustrate, the decoded images at bit rates 1.25 bpp and 1.015625 bpp (rows 3 and 5 of Tables 7.1 and 7.2) are shown in Figures 7.1 and 7.2 for LENA and in Figures 7.3 and 7.4 for HOUSE. Recall that the MMSEQ requires an iterated search for each block's threshold. The required computations for MMSEQ is about two to seven times that for AMBTC at the same bit rate (see Tables 6.1 and
Thus, AMBTC is a practical alternative to MMSEQ for achieving the minimum mean-square-error, especially in the subband environment. Hence, AMBTC is the optimum choice among the 1-bit moment preserving quantizers.

<table>
<thead>
<tr>
<th>LENA image</th>
<th>SAMBTC</th>
<th>SMMSEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit Rate, (bpp)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B$</td>
<td>$b = B/16$</td>
<td>$MSE$ (/pel)</td>
</tr>
<tr>
<td>80</td>
<td>5.0</td>
<td>16.053375</td>
</tr>
<tr>
<td>32</td>
<td>2.0</td>
<td>50.582260</td>
</tr>
<tr>
<td>20</td>
<td>1.250</td>
<td>71.741592</td>
</tr>
<tr>
<td>17</td>
<td>1.0625</td>
<td>81.768188</td>
</tr>
<tr>
<td>16.25</td>
<td>1.015625</td>
<td>87.719543</td>
</tr>
</tbody>
</table>

Table 7.1: Comparison between SAMBTC and SMMSEQ at various bit rates using LENA. Note that the $B$ is the total bit rate for 16 equally divided subbands where $b$ is the average bit rate per subband.

<table>
<thead>
<tr>
<th>HOUSE image</th>
<th>SAMBTC</th>
<th>SMMSEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit Rate, (bpp)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B$</td>
<td>$b = B/16$</td>
<td>$MSE$ (/pel)</td>
</tr>
<tr>
<td>80</td>
<td>5.0</td>
<td>23.559525</td>
</tr>
<tr>
<td>32</td>
<td>2.0</td>
<td>82.994003</td>
</tr>
<tr>
<td>20</td>
<td>1.250</td>
<td>130.872787</td>
</tr>
<tr>
<td>17</td>
<td>1.0625</td>
<td>148.455826</td>
</tr>
</tbody>
</table>

Table 7.2: Comparison between SAMBTC and SMMSEQ at various bit rates using HOUSE. Note that the $B$ is the total bit rate for 16 equally divided subbands where $b$ is the average bit rate per subband.
Figure 7.1: Subband coded LENA images at 1.25 bpp: (a) SAMBTC and (b) SMMSEQ.
Figure 7.2: (a) SAMBTC and (b) SMMSEQ. Subband coded LENA images at 1.015625 bpp: (a) SAMBTC and (b) SMMSEQ.
Figure 7.3: Subband coded HOUSE images at 1.25 bpp: (a) SAMBTC and (b) SMMSEQ.
Figure 7.4: Subband coded HOUSE images at 1.015625 bpp: (a) SAMBTC and (b) SMMSEQ.
8 Conclusion

Eight new properties of AMBTC [3] are presented in this paper with proof and coding performance simulation. These provide fundamental insights into the algorithm and demonstrate that AMBTC is the optimum two-level moment preserving quantizer for achieving the minimum mean-square-error and the least computations. Finally, it was shown in [1] that SAMBTC is significantly superior to AMBTC in compressing both monochrome images and color images recorded in different color spaces, this means that SAMBTC is the optimum choice of subband two-level moment preserving quantizers.
Bibliography


Biography

Kai-Kuang Ma (S-'80, M-'84) received the B.E. degree in electronic engineering from Chung Yuan Christian College of Science and Engineering, Taiwan, Republic of China, the M.S. degree in electrical engineering from Duke University, Durham, North Carolina, and the Ph.D. degree in electrical engineering at North Carolina State University, Raleigh, North Carolina. He had seven years of industrial experience with IBM and worked on various product developments including T1, VLSI, and multirate digital signal processing and speech processing for multi-media applications and real-time implementation. He is now a Member of Technical Staff in the Institute of Microelectronics at National University of Singapore. His research interests are in the areas of color image/video compression, simulated annealing optimization, and artificial neural networks with applications in signal processing.

Dr. Ma is a member of Sigma Xi and Eta Kappa Nu.

Sarah A. Rajala received the B.S.E.E. degree from Michigan Technological University, Houghton, in 1974, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1977 and 1979, respectively. Since 1979, she has been a member of the Electrical and Computer Engineering Department at North Carolina State University, Raleigh, where she currently holds the position of Professor. Professor Rajala has been an active researcher in the areas of target acquisition and tracking, motion estimation, color image processing, and time-varying image analysis with application to image coding.