Quantization Watermarking in Three-Dimensional Wavelet Transform Domain

Mohsen Ashourian
1- Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Isfahan, Iran.
Email: ashourian@iaumajlesi.ac.ir

Received: October 2019 Revised: January 2020 Accepted: February 2020

ABSTRACT:
Quantization watermarking is a technique for embedding hidden copyright information based on dithered quantization. This non-blind scheme is only practical for watermarking applications, where the original signal is available to the detector as for a fingerprinting purpose. The goal of this paper is to analyze the quantization watermarking in the three-dimensional wavelet transform. We consider the nonlinear effect of dithered quantization in the time-domain representation of the filter bank. We derive a compact and general form for distortion in the host video due to the encoding and embedding process. The formulation has the capacity to be simplified and optimized for different filter banks and dither signals. We provide some supporting experiments for the three-dimensional wavelet analysis of video signal.

KEYWORDS: Quantization Watermarking, Wavelet Transform, Dithering.

1. INTRODUCTION

Digital watermarking is an approach for copyright protection by embedding copyright information directly in the host data. In quantization watermarking, the signature information is used as a dither signal in the process of quantizing the host signal [1,2]. This method is especially useful for those applications where compression of the host signal with embedding the secret information is jointly implemented. Further extension and improved types of quantization watermarking has been reported recently [3, 4].

Using theoretical results of dithered quantization [2], Eggers and Girod derived a mathematical analysis of quantization watermarking based on probability density function (PDF) of the host and signature signals [1]. This paper extends their results for quantization watermarking of wavelet coefficients of the host signal. Since quantization is a non-linear process, the time-domain framework introduced by Nayebi et. al. [5,6] is selected in order to analyse quantization watermarking in the wavelet domain. We derive a statistical form of this formulation that could be used to analyse the distortion in the host signal due to embedding the signature information and encoding the subbands [6,7].

The following section is organized as follows. In Section 2, we provide a brief explanation on dithered quantization, we explain the modification of time-domain formulation of filter bank for the quantization watermarking effect and its simplification in Section 3. Section 4 provides a sample analysis for video signal and Section 5 concludes the paper.

2. DITHERED QUANTIZATION

Fig. 1 shows the non-subtractive dithering scheme. The host signal x(n) is the main input to quantizer, and the signature signal d(n) is the added dither.

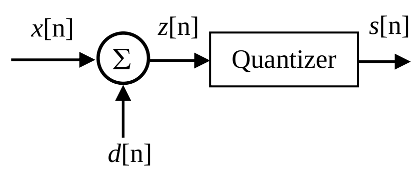

Fig. 1. Non-Subtractive dithered quantization

The quantization error in the uniform scalar quantizer is only limited to (-Δ/2 , Δ/2), where Δ is the quantization step size. The aim of dithering in the classical signal coding system is to improve perceptual quality of the reconstructed signal by changing the quantization error spectrum [2]. The characteristic function of the input z(n) to the quantizer can be written by:

$$M_z(ju) = M_x(ju)M_d(ju) \quad (1)$$

Where, M_x and M_d are the characteristic functions of the input signal and the dither respectively. It can be proved [1] that the characteristic function of the quantization noise e(n) is represented in terms of the input signal and the dither signal characteristic functions [1].

$$M_e(ju) = \sum_{b=-\infty}^{\infty} M_z(j\,\frac{2\pi b}{\Delta}) \cdot Sinc(\frac{\Delta}{2}(u + 2\pi b/\Delta)) \quad (2)$$

By substituting $M_z(u)$ from Equ. (2) by Equ.(1), we have

$$M_e(ju) = \sum_{b=-\infty}^{\infty} M_x(j\,\frac{2\pi b}{\Delta}) \cdot M_d(j\,\frac{2\pi b}{\Delta}) \cdot Sinc(\frac{\Delta}{2}(u + 2\pi b/\Delta)) \quad (3)$$

Using the characteristic function, the energy of quantization noise can be computed by

$$E[e^2] = -\frac{d^2}{du^2} M_e(u)\Big|_{u=0} = \frac{\Delta^2}{12} + \sum_{\substack{b=-\infty \\ b\neq 0}}^{\infty} \frac{(-1)^b}{2(\pi b/\Delta)^2} M_x(\frac{j2\pi b}{\Delta}).\, M_d(\frac{j2\pi b}{\Delta}) \quad (4)$$

The cross-correlation of the quantization error and the dither signal is

$$E[ed] = \sum_{\substack{b=-\infty \\ b\neq 0}}^{\infty} \frac{(-1)^b}{2\pi b/\Delta} M_x(\frac{j2\pi b}{\Delta}).\, Im\{\, M_d^1(\frac{j2\pi b}{\Delta})\} \quad (5)$$

Where

$$M_d^1(\frac{j2\pi b}{\Delta}) = \frac{d}{du}\{M_d(\frac{j2\pi b}{\Delta}\}\Big|_{u=0} \quad (6)$$

In most cases, the signal PDF and its characteristic function are even functions, therefore the summation in Equ. (5) can be written as

$$E[ed] = \sum_{b=1}^{\infty} \frac{(-1)^b}{\pi b/\Delta} M_x(\frac{j2\pi b}{\Delta}).\, Im\{\, M_d^1(\frac{j2\pi b}{\Delta})\} \quad (7)$$

If we want to check only the existence or absence of the watermark signal, we can use a correlation detector. In this case the output of the correlation detector for the two case, of existence of the watermark (U_1) and its absence (U_0) could be derived by [1]:

$$U_1 = \frac{E[(e(n)d(n)]}{\sigma_d^2} + 1 \quad (8)$$

$$U_0 = \frac{E[(e(n)d(n)]}{\sigma_d^2} \quad (9)$$

Which implies that the absolute value of the normalized cross-correlation between the quantization error and the dither signal ($E[ed]/\sigma_d^2$) should be as small as possible. At the same time, to reduce the perceptual distortion, we should minimize the quantization error.

3. DITHERED QUANTIZATION

3.1. Number General Formulation

In this section we derive a formulation that shows the effect of the added watermark in the wavelet transform domain. The distortion could be calculated based on reconstruction error in the corresponding filter bank.

Fig. 2 shows the block diagram of the conventional filter bank structure. Fig. 3 shows the filter bank with dithered quantizers. As depicted in Fig. 3, we model the effect of quantization as additive, but not necessarily un-correlated signals, $F_i(m)$. The dither signal in each channel is represented by $D_i(m)$. The output signal $\hat{x}(n)$ is synthesized from the quantized subband signals. In order to analyze and design the subband coder, we only consider a uniform M band filter bank with filters of length N and the overall delay of Δ samples, such that L=N / M is an integer. This result can be easily extended to non-uniform or multidimensional subbands. The relationship between input and output of the system, in the time-domain can be expressed as

$$\hat{x}(Mm) = s^T(A^T x_I(Mm) + F_q(m) + D(m)) \quad (10)$$

Where input and output vectors of length **M** and **I** are

$$\hat{x}_M(n) = [\hat{x}(n+M-1), \hat{x}(n+M-2), \ldots, \hat{x}(n)]^T \quad (11)$$

$$x_I(n) = [x(n), x(n-1), \ldots, x(n-I+1)]^T \quad (12)$$

The parameter I is equal to ***2N-M***, since the analysis and synthesis filter together create 2N delay, and the output is calculated in M points, therefore we need ***2N-M+1*** points of input. Note that, in Equ. (10) because of mixing the two terms of quantization and dithering, before up-samplers, the time index of the input and the output signal changed from ***m*** to ***mM***. Finally, **A** is a block Toeplitz matrix of size **I** × **N** defined as

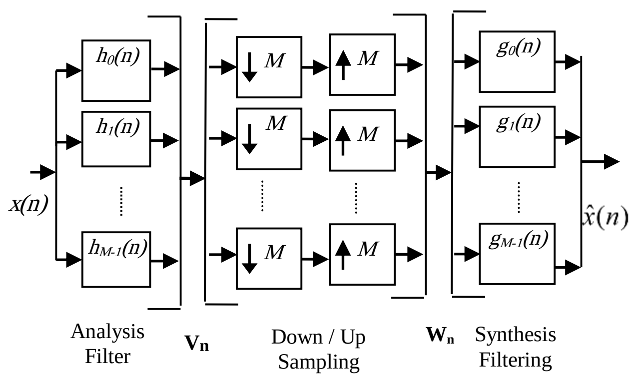

Fig. 2. Block Diagram of a Basic Filter Bank.

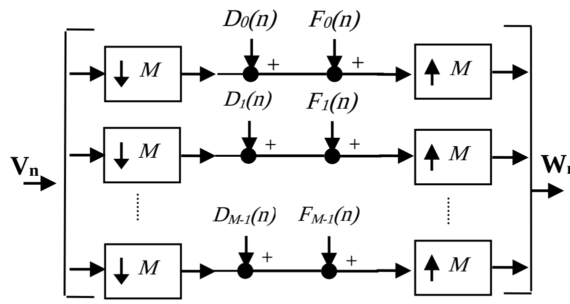

Fig. 3. Down/ Up sampling part of the subband coder with including the dithered quantizer.

$$A(n) = \begin{bmatrix} [P^t] & 0 & .. & 0 \\ 0 & [P^T] & & \vdots \\ \vdots & 0 & \ddots & \vdots \\ & : & \ddots & 0 \\ 0 & 0 & .. & [P^T] \end{bmatrix} \quad (13)$$

Where **P** is an **M × N** matrix whose **i**th row is comprised of the coefficients of the **i**th analysis filter, and **O** is an **M × M** zero matrix. The matrix ***s*** consists of the synthesis filter coefficients

$$s = \begin{bmatrix} g_0(0) & g_1(0) & \cdots & g_{M-1}(0) \\ g_0(1) & g_1(1) & \cdots & g_{M-1}(1) \\ \vdots & : & \ddots & \vdots \\ g_0(N-1) & g_1(N-1) & .. & g_{M-1}(N-1) \end{bmatrix} \quad (14)$$

and $g_i(j)$ denotes the **j** coefficient of the **i** synthesis filter. Finally, the vector *D(m)* represents the dither signal, and $\boldsymbol{F_q}(\boldsymbol{m})$ the quantization noise:

$$F_q(m) = [q^T(m),\ q^T(m-1), \dots, q^T(m-L+1)]^T \quad (15)$$

$$D(m) = [d^T(m),\ d^T(m-1), \dots, d^T(m-L+1)]^T \quad (16)$$

Where

$$d(m) = [d_0(m),\ d_1(m),\ \dots\ ,\ d_{M-1}(m)]^T \quad (17)$$

$$q(m) = [q_0(m),\ q_1(m), \dots, q_{M-1}(m)]^T \quad (18)$$

are the dither and the quantization signal at time **m**.

Assuming that the filter bank is perfect reconstruction, the relationship between input and output of the filter bank can be written as

$$\hat{x}(Mm) = b^T x_I(Mm) \quad (19)$$

Where the matrix **b** denotes one period of the ideal impulse response of the system. Therefore, using (10) and (20), one period of the output error can be written as

$$e = \hat{x}(Mm) - x(Mm - \Delta)$$
$$= (As - b)^T x_I(Mm) + s^T F_q(m) + s^T D(m) \quad (20)$$

We use the mean square error of the output as a criterion for minimization assuming that the input signal can be modelled as zero-mean, wide sense stationary (WSS) sources

$$\bar{\sigma}_e^2 = \frac{1}{M} Trace[E\{ee^T\}]$$
(21)

Where "Trace" denotes the sum of main matrix diagonal elements. Using (21), it can be shown that

$$\begin{aligned} \bar{\sigma}_e^2 = {} & \frac{1}{M} Trace[(As-b)^T R_{xx}(As-b)] + \\ & \frac{1}{M} Trace[s^T R_{qq} s] + \frac{2}{M} Trace[(As-b)^T R_{xq} s] + \\ & \frac{1}{M} Trace[s^t R_{dd} s] + \frac{2}{M} Trace[(As-b)^T R_{xd} s] + \\ & \frac{2}{M} Trace[s^T R_{qd} s] \end{aligned} \quad (22)$$

$\mathbf{R_{xx}}$, $\mathbf{R_{qq}}$ and $\mathbf{R_{dd}}$ are the input, the quantization noise and the dither signal autocorrelation matrices respectively, while $\mathbf{R_{xq}}$,$\mathbf{R_{xd}}$ and $\mathbf{R_{qd}}$ represents the cross-correlations between these signals.

Equ. (22) is the basic formulation that shows the distortion effect due to dither signal and quantization error. The optimization should be undertaken with addition of a Lagrangian cost with constraint of a fixed total bit-rate R_T

$$\sum_{k=0}^{M-1} R_k = R_T \quad (23)$$

Where R_k denotes the bit-rate in subband k . In the encoding process, two situations are of particular interest.

- Subband signals are split into blocks and the number of bits used for a given block depends on the dynamic range of these signals.
- Entropy coding is performed in each subband. In this case the optimization is carried out under the constraint of a given entropy budget H_T .

For the uniform scalar quantization, which we use in quantization watermarking, R_k and Δ_k are related as $R_k = log_2(d_k/\Delta_k)$, where d_k is the dynamic range of the signal and Δ_k is the quantization bin.

In order to maximize the correct watermark detection probability, for each subband quantizer, we should minimize the normalized absolute value of cross-correlation between the quantization noise and the watermark $\left|E[e_{q_i}.d_i]\right|/\sigma^2_{d_i}$, for each channel [1]. At the same time, we should minimize the total reconstruction error $\bar{\sigma}_e^2$ to reduce visible distortion.

3.2. Simplification of Formulation

In the design of the wavelet-based quantization watermarking scheme, it sounds reasonable to simplify the design by selecting a perfect reconstruction filter bank (***As*=*b***) and optimize the system performance by a proper selection of quantizers and the variance of the added watermark signal to each subband. By selecting perfect reconstruction filters, we can simplify Equ. (22)

$$\bar{\sigma}_e^2 = \frac{1}{M} Trace\,[s^T R_{qq} s] + \frac{1}{M} Trace\,[s^t R_{dd} s] + \frac{2}{M} Trace\,[s^T R_{qd} s] \quad (24)$$

In the second step, we analyse the quantization error and the dither signal.

Since in watermarking, we prefer to maintain the signal quality, we consider only high bit-rate quantizer. At high bit-rates, the quantization process can be approximated by an additive and uncorrelated white noise. This means the cross-correlation term R_{qd} in Equ. (24) is zero; therefore, Equ. (24) is simplified to

$$\bar{\sigma}_e^2 = \frac{1}{M} Trace[s^T R_{qq} s] + \frac{1}{M} Trace[s^t R_{dd} s \rightleftarrows] \quad (25)$$

At high bit-rates, we can consider the quantization noise as a memoryless signal; therefore, R_{qq} is a diagonal matrix. The elements on the main diagonal are the variances of the quantization noise of subbands and can be calculated using Equ. (2)

$$R_{qq}(i,i) = E[e^2_{q_i}] = \frac{\Delta_i^2}{12} + \sum_{b=1}^{+\infty} \frac{(-1)^b}{(\pi b/\Delta_i)^2} M_{x_i}(\frac{j2\pi b}{\Delta_i}).M_{d_i}(\frac{j2\pi b}{\Delta_i}) \quad (26)$$

Where, M_{d_i} is the characteristic function of the dither signal in the i^{th} subband.

The dither signal is usually selected to be a memoryless signal with bipolar, Gaussian or uniform distributions. In these cases, the diagonal element of R_{dd} is equal to the dither variance in each channel.

3.2.1. High-Bit Rate Quantizer + Paraunitary Filter Bank

For the of the paraunitary filter bank, the subband signals are orthogonal and we have

$$\sum_{i=0}^{L-1} Q_i Q_{i+j} = \begin{cases} I & j = 0 \\ 0 & j \neq 0 \end{cases} \quad (27)$$

We can use Equ.(27) and expand the two terms in Equ. (25).

$$\bar{\sigma}_e^2 = \frac{1}{M}\sum_{i=1}^{M}\{\sigma^2_{q_i} + \sigma^2_{d_i}\}\sum_{j=1}^{N} g_i^2(j) \quad (28)$$

4. RESULTS AND DISCUSSION

In order to verify the analytical results, we need a large number of experiments on different configurations of filter banks, bit-allocations and different input data-sets and dithering.

Fig. 4 shows the structure of 3-D subband decomposition used for both video signals. A three-dimensional subband coder uses a unique approach for encoding intra-frame and inter-frame redundancy in a video sequence. The video signal passed through a 3-D filter bank, and then different subbands are encoded based on their visual importance [9]. The terms HP and LP refer to high-pass filtering and low-pass filtering, where the subscripts t, h, and v refer to temporal, horizontal, and vertical filtering respectively. The selected subband framework consists of 11 spatio-temporal frequency bands. The temporal frequency decomposition is restricted to only two subbands due to potential delay problems in a practical implementation and reducing dependency in coding consecutive frames. The image frames are filtered temporally using the two-tap Harr basis functions [9]. Temporal decomposition is followed by horizontal-spatial filtering and vertical-spatial filtering using 9/7 biorthogonal filters [5].

Selection of optimum quantizer for different subbands based on their statistical characteristics and visual importance is the key factor for developing subband coder. Fig. 5 shows the frequency map of the 11 video subbands which can be classified as below:

1. Band 1, the low temporal and spatial frequency band, is a blurred version of the original video frame. It has much higher energy compared to other subbands and has the most visual importance. However, while all the subbands histogram follow well a generalized Gaussian distribution, this subband does not follow any fixed distribution [10, 11].
2. Bands 2-7, the low temporal and high spatial frequency bands, include information of texture and sharpness of video signal in the spatial domain. Depends on amount of these information in scene, the energy of these bands could be higher or lower. Among these bands, bands 4 and 7 have much lower due to two times highpass filtering (vertical and horizontal).
3. Band 8, the high temporal and low spatial frequency band, has higher average energy compared to other high temporal bands, and it shows the major changes in consecutive video frames.
4. Bands 9-11, the high temporal and high spatial frequency bands (Bands 9-11) have low energy, but high variation in time. They represent sharp and fast objects movements in the video scene.

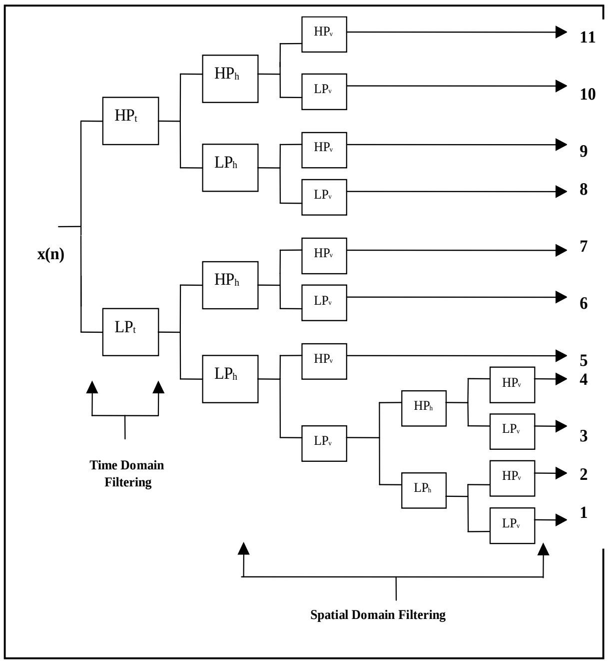

Fig. 4. Three-Dimensional Filter Bank Structure.

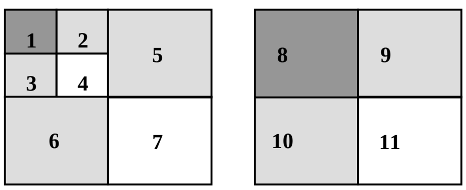

Fig. 5. Three-Dimensional Filter Bank Frequency Map

The amplitude distribution of Bands 1 and 8 does not follow any fixed probability distribution function (PDF) [9].

For 3-D wavelet decomposition of video, we should implement one level of wavelet transform in time domain and later implement spatial domain of wavelet transform on each temporal subbands. (Figs. 7-8)

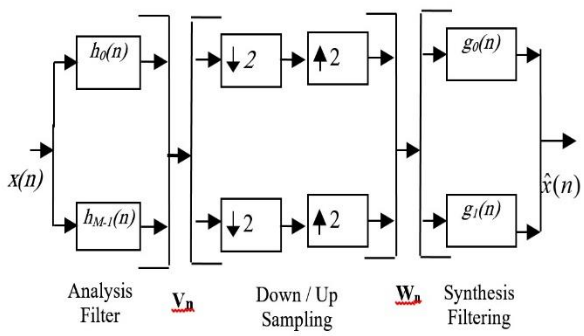

Fig. 6. Time domain Filter Bank Structure

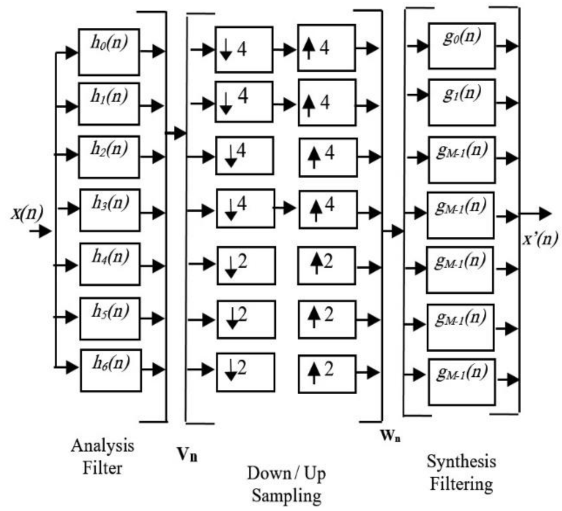

Fig. 7. Spatial Domain Filter Bank Structure

5. CONCLUSION

In this paper, we derived a general formulation for the reconstruction error and watermark detection in quantization watermarking in the filter bank domain. We did not assume any constraint on the type of filter bank and quantizers. We also simplified the mathematical formulation for the high and low bit rates. Some experimental results for embedding data in a video subband coder are reported.

REFERENCES

[1] J. Eggers and B. Girod. "**Quantization Effects on Digital Watermarks**". *Signal Processing*, Vol.81, No.2, pp. 239–263, 2001.

[2] R. M. Gray and T. G. Stockham. "**Dithered Quantizers**". *IEEE Trans. on Information Theory*, Vol. 39, No.5, pp. 805–812, 1993.

[3] Fan, Di, Ying Wang, and Chunwei Zhu. **"A blind watermarking algorithm based on adaptive quantization in Contourlet domain."**, *Multimedia Tools and Applications* 78.7, pp. 8981-8995, 2019.

[4] Sridhar, B. **"Cross-Layered Embedding of Watermark on Image for High Authentication."** *Pattern Recognition and Image Analysis* 29.1, pp.194-199, 2019.

[5] K. Nayebi , T.P.I. Barnwell and M.J.T. Smith. "**Time -Domain Filter Bank Analysis: A New Design Theory**". *IEEE Transaction. on Signal Processing*, Vol.44, No. 6, pp. 1412–1429, 1992.

[6] I. Sodagar, T.P.I Barnwell and M.J.T. Smith. "**On the Statistically Optimum FIR Filter Bank Design**". *IEEE Int. Conf. on Acoustic, Speech and Signal Proceeding*, pp. 227–230, 1994.

[7] M. Ashourian, Z. Mohdyusof. "**Scalar quantization Error Analysis for Image Subband Coding using Dithering**". *SPIE Int. Conf. on Visual Comm. and Image Proc*., Vol. 3, pp. 1434–1442, Australia 2000.

[8] Ashourian, Mohsen, and Yo-Sung Ho. "**Analysis of Quantization Watermarking in the wavelet**

transform domain." *Seventh International Symposium on Signal Processing and Its Applications*, 2003. Proceedings.. Vol. 2. IEEE, 2003.

[9] K.A. Birney, and T.R. Fischer. "**On the Modeling of DCT and Subband Image data for Compression**". *IEEE Transactions on Image Processing*, Vol.4, No.2, pp. 186–193, 1995.

[10] Atta, R. "**Optimal bit Allocation for Subband Video Coding**." *IET image processing* 4.5 (2010): 365-373.

[11] N, Mohammad Ali, Ch. Vorakulpipat, and H. Gamboa Rosales. "**Three-Dimensional (3D) Watermarking**." *Digital Watermarking. Springer*, Singapore, pp. 81-99, 2018.