ESTIMATION OF QUANTIZATION STEP SIZE AGAINST AMPLITUDE MODIFICATION ATTACK IN SCALAR QUANTIZATION-BASED AUDIO WATERMARKING

Siho Kim and Keunsung Bae

School of Electronic and Electrical Engineering, Kyungpook National University
1370, Sankyuk-Dong, Buk-Gu, 702-701, Daegu, Korea
Tel: +82-53-950-5527, Fax: +82-53-950-8827
E-mail: si5@mir.knu.ac.kr, ksbae@mir.knu.ac.kr

ABSTRACT

Scalar quantization-based watermarking schemes are very vulnerable to amplitude modification attack. To overcome this problem, we propose a novel and robust algorithm that estimates the modified quantization step size by searching *QE* (quantization error) function. For efficient searching of *QE* curve, we analyze the peak curve of *QE* analytically and derive the equations to determine an appropriate search interval. The search interval can be determined from the mean and variance of an audio signal regardless of its probability density function shape. The experimental results demonstrate that the proposed algorithm provides both exact estimation of the modified quantization step size and good watermark detection performance under AWGN attack.

1. INTRODUCTION

Recently, scalar quantization watermarking schemes [1,2] based on the Costa's dirty paper coding [3] are drawing more attention due to its good performance. They do not need original host signals for watermark detection and host signals do not affect the performance of watermark detection, while the blind SS (spread-spectrum) and echo hiding schemes suffer significantly from the host signal interference. However scalar quantization watermarking schemes are known to be very vulnerable to amplitude modification attack.

So far, there were several researches to solve the performance degradation due to amplitude scaling. Eggers et al. [4] embedded a pilot sequence into a cover signal to estimate the possible amplitude modification in extractor. This method has a shortcoming that a large number of samples for a pilot signal encroach on the embedding space for watermark message. Lee et al. [5] proposed the estimation scheme of a scale factor using EM (Expectation Maximization) algorithm. However, EM algorithm needs a large number of samples for accurate estimation of a scale factor and then it can cause the impractical complexity. On the other hand, Perez-Gonzalez proposed Rational Dither Modulation (RDM) which embeds a watermark signal on the domain irrespective of amplitude scaling [6]. This method outperforms the original DM under amplitude scaling attack but its performance becomes poorer than the original DM in case of no amplitude modification attack. In addition, Oostveen et al. [7] proposed a quantization step size proportional to the intensity of image signal using Weber's law to avoid the effect of amplitude scaling. As a similar approach, Li et al. [8] proposed to use an adaptive step size that is determined using Watson's perceptual model. However, it is difficult to apply these schemes to an audio signal because they exploit the perceptual property of human for image signal.

In this paper, we propose a novel and robust algorithm that can estimate the modified quantization step size under amplitude modification attack. It searches the quantization step size that minimizes the quantization error of the received audio signal. In that case it is very important to determine an appropriate search interval. In our previous research [9,10] we determined the search interval numerically as well as analytically with the assumption of Gaussian distribution for an audio signal. In this paper, we analyze the quantization error function for the audio signal regardless of its pdf (probability density function) shape using Gaussian mixture model, and derive the formula to determine an appropriate search interval analytically. We also discuss the search range of the *QE* curve and allowed range of scaling attack to apply our estimation algorithm correctly.

2. SCALAR QUANTIZATION-BASED WATERMARKING SCHEME

In quantization-based watermarking such as QIM [1] or SCS [2], the binary watermark message $d \in \{0,1\}$ is embedded into a host signal using a dithered scalar quantizer $Q_{\Delta,d}$, which is defined by

$$Q_{\Delta,d}(x) \equiv Q_{\Delta}\left(x+\frac{\Delta}{2}\cdot d\right)-\frac{\Delta}{2}\cdot d \Bigg|_{Q_{\Delta}(x)\equiv\left\lfloor \frac{x}{\Delta}+0.5\right\rfloor\cdot\Delta}, \quad (1)$$

where $Q_\Delta(\cdot)$ is a uniform scalar quantizer and Δ is a quantization step size. Then the watermarked signal s is obtained from the host signal x as follows:

$$s = x + \alpha \cdot [Q_{\Delta_e,d}(x) - x], \tag{3}$$

where α and Δ_e are the embedding parameters. If the amplitude modification attack that scales the watermarked signal by a scaling factor g, and AWGN attack v is assumed, then the received signal r can be written by

$$r = g \cdot (s + v). \tag{4}$$

Then the quantization step size of the received signal changes to $\Delta_d = g \cdot \Delta_e$. Thus it is necessary to estimate the scale factor g or modified quantization step size Δ_d for exact watermark detection.

3. ESTIMATION OF QUANTIZATION STEP SIZE

In [9,10], we defined the quantization error function as

$$QE(\Delta) \equiv E\left[\left(r - Q_\Delta(r)\right)^2\right]. \tag{5}$$

In addition, we normalize the QE function with $\Delta^2/12$ and define it as a normalized QE ($\equiv QE_N$) as follows:

$$QE_N(\Delta) \equiv 1 - \frac{12}{\Delta^2} \cdot QE(\Delta). \tag{6}$$

If we denote the quantization step size of a received audio signal as Δ_q, QE function has minimum value when the quantization step size Δ of a uniform scalar quantizer $Q_\Delta(r)$ is equal to Δ_q. Therefore the estimated quantization step size $\hat{\Delta}_d$ can be obtained from the equation

$$\hat{\Delta}_d = \arg\min_{\Delta} \left(QE_N(\Delta)\right). \tag{7}$$

In this estimation algorithm, it is very important to determine an appropriate search interval satisfying both detection performance and computational complexity at the same time.

3.1. Analysis of the peak curve of QE

To find an appropriate search interval, we investigated the curve of QE function. If we denote the pdf of a host signal as $f(x)$, then QE function can be rewritten as

$$QE(\Delta) = \sum_{k=-\infty}^{\infty} \left[\Delta_q \cdot f(k \cdot \Delta_q) \cdot e_\Delta^2(k)\right]. \tag{8}$$

Here $e_\Delta(k)$ denotes the error caused by quantizing a host signal, whose quantization step size is Δ_q, with Δ. In general, any probability density function can be represented by Gaussian mixture as shown in

$$f(x) = \sum_{i=1}^{L} \eta_i \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}, \tag{9}$$

where σ_i and μ_i are the standard deviation and mean of the i^{th} mixture component, respectively, and η_i is the existence probability of i^{th} mixture. When the quantization step size Δ is in the neighborhood of Δ_q, the quantization error $e_\Delta(k)$ can be represented as a linear equation with slope $\delta(=\Delta - \Delta_q)$ if we assume that there are no samples over maximum value K as shown figure 1. As a result, $e_\Delta(k)$ is expressed as

$$e_\Delta(k) = \delta \cdot k \,\big|_{-K \le k \le K} \tag{10}$$

From (9) and (10), QE function given in (8) can be rewritten by

$$QE(\Delta) = \int \sum_{i=1}^{L} \eta_i \frac{\Delta_q}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(\Delta_q k - \mu_i)^2}{2\sigma_i^2}} \cdot (\delta \cdot k)^2 dk. \tag{11}$$

Equation (11) is valid for the range of (12) and (13).

$$|\delta_i \cdot K_i| \le \frac{\Delta}{2}\Big|_{\Delta = \Delta_q + \delta_i}, \tag{12}$$

$$\delta \le \min(|\delta_i|), \tag{13}$$

where $K_i \equiv \max(|K_i^{\max}|, |K_i^{\min}|)$, and

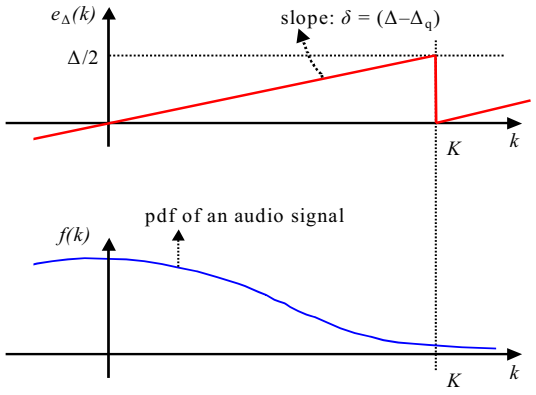

Fig. 1 Error function and pdf of a host audio signal

$K_i^{\max} \equiv (\mu_i + m \cdot \sigma_i)/\Delta_q$, $K_i^{\min} \equiv (\mu_i - m \cdot \sigma_i)/\Delta_q$.

By expanding (11), we can derive

$$QE(\Delta) = \frac{(\Delta - \Delta_q)^2}{\Delta_q^2} \sum_{i=1}^{L} \eta_i \left(C_\sigma(m) \cdot \sigma_i^2 + C_\mu(m) \cdot \mu_i^2 \right) \quad (14)$$

$$where, \begin{cases} C_\sigma(m) = -\sqrt{\frac{2}{\pi}} \cdot m \cdot e^{-\frac{m^2}{2}} + erf\left(\frac{m}{\sqrt{2}}\right) \\ C_\mu(m) = erf\left(\frac{m}{\sqrt{2}}\right) \end{cases}$$

If m is greater than 3, $\left(-\sqrt{\frac{2}{\pi}} \cdot m \cdot e^{-\frac{m^2}{2}} + erf\left(\frac{m}{\sqrt{2}}\right) \right) \cong 1$ and $erf\left(\frac{m}{\sqrt{2}}\right) \cong 1$. Then QE is approximated as follows:

$$QE(\Delta) \cong \frac{(\Delta - \Delta_q)^2}{\Delta_q^2} \sum_{i=1}^{L} \eta_i \cdot (\sigma_i^2 + \mu_i^2) \Big|_{m \geq 3} . \quad (15)$$

Here $\sigma_i^2 + \mu_i^2$ represents the power of i^{th} mixture, thus $\sum_{i=1}^{L} \eta_l \cdot (\sigma_i^2 + \mu_i^2)$ is the total power of a host audio signal. Finally, the QE can be expressed by

$$QE(\Delta) \cong \frac{(\Delta - \Delta_q)^2}{\Delta_q^2} (\sigma_s^2 + \mu_s^2), \quad (16)$$

where σ_s and μ_s are the standard deviation and mean of a host audio signal, respectively. From the equation, we see that the QE curve in the neighborhood of the peak can be represented by a second-order equation, which is not related to the shape of probability density function of the host signal in a limited range.

3.2. Determination of Search Interval

To find the minimum search interval, which guarantees to detect the value that exceeds $\beta \cdot QE_N(\Delta_q)$, we should solve the following equation:

$$QE_N(\Delta_q + \delta_w) = \beta \cdot QE_N(\Delta_q) \Big|_{0<\beta<1}, \quad (17)$$

where δ_w is half of the peak width. Hence $2 \cdot |\delta_w|$ becomes a minimum search interval. By substituting (16) into (17) using equation (6) and solving it, we can obtain

$$\delta_w = \Delta_q^2 \cdot \sqrt{\frac{1-\beta}{12 \cdot (\sigma_s^2 + \mu_s^2)}} \Big|_{m \geq 3} . \quad (18)$$

In (18), β must satisfy the following condition, where $\beta_{\min}$ is derived from equation (12), (13) and (18).

$$\beta_{\min} < \beta < 1$$
$$\beta_{\min} \cong 1 - \frac{3 \cdot \left(\sigma_s^2 + \mu_s^2\right)}{\left(|\mu_d| + m \cdot \sigma_d\right)^2} \Bigg|_{m \geq 3,\ |\mu_d| + m \cdot \sigma_d >> \Delta_q} \quad (19)$$

where μ_d and σ_d are the mean and standard deviation of dominant mixture that makes an important role to construct the shape of QE curve. Therefore it is required to estimate the parameters of Gaussian mixture for exact $\beta_{\min}$. However we could find out that 0.95 of $\beta_{\min}$ is suitable for a normal audio signal through the experiments.

If we assume that the range of an amplitude scaling attack is $g_{\min} \leq g \leq g_{\max}$, then the searching range Δ should satisfy the following conditions:

$$k_1 \Delta_e \leq \Delta \leq k_2 \Delta_e \quad (20)$$

$$\frac{1}{2} \cdot g_{\max} < k_1 < g_{\min} \text{ and } k_2 > g_{\max} \quad (21)$$

where Δ_e is the quantization step size used in embedding process of watermark data.

3.3. Fine Searching

The searching process with the search interval determined from (18) provides the coarse position of the quantization step size. Therefore, to find an exact quantization step size, we need further search process in the neighborhood of the coarse peak by dense interval. It is done using the following formula, which is obtained from (16) with the maximum peak $(\Delta_1, QE_N(\Delta_1))$ and the second maximum peak $(\Delta_2, QE_N(\Delta_2))$ detected from the coarse searching.

$$\hat{\Delta}_d = \frac{1}{2} \left(\Delta_1 + \Delta_2 + \frac{QE_N(\Delta_1) - QE_N(\Delta_2)}{\alpha(\Delta_1 - \Delta_2)} \right) \Bigg|_{\alpha = \frac{12 \cdot (\sigma_s^2 + \mu_s^2)}{\Delta_q^4}} \quad (22)$$

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to validate our proposed algorithm, we applied it to an audio watermarking system, which embeds the watermark signal in wavelet filter bank domain with 32 subbands. An audio signal is normalized to a pre-defined power before watermark embedding and extracting. We use 93 seconds long audio signal that is made up of various genre of short music signals. It has a 44.1kHz sampling frequency with 16bit quantization. We set β to 0.95.

Figure 2 shows the original QE curve and the QE curve calculated from (16). We can find that two curves are almost same in the neighborhood of Δ=154.9, which is the

exact quantization step size of the received audio signal. Figure 3 shows the trace of the estimated quantization step size $\hat{\Delta}_d$ on a frame basis under AWGN attack. Although its performance decreases as AWGN attack increases, it can be seen that a comparatively exact quantization step size is estimated in most frames. Table 1 shows the BER of watermark detection for the methods with three different quantization step sizes in the presence of AWGN. The 'Embedding' represents the one that uses the quantization step size used in embedding the watermark data, and 'Exact' denotes the one that uses the modified quantization step size for amplitude scaling. Finally 'Estimated' denotes the quantization step size estimated from our algorithm. It is shown that the proposed method provides almost same detection performance as compared with the method 'Exact' although BER increases as WNR (Watermark to Noise Ratio) decreases.

5. CONCLUSION

In this paper, we analyze the quantization error in the quantization-based watermarking and express it as a second-order equation. And then, we derive the formula to determine an appropriate search interval for robust and efficient estimation of the modified quantization step size from the *QE* equation. The experimental results demonstrated that the shape of *QE* curve from the derived *QE* function is nearly similar to that of the real audio signal. It also has been shown that the proposed method gives almost same detection performance as compared with the method using the exact quantization step size used for amplitude scaling under AWGN attack.

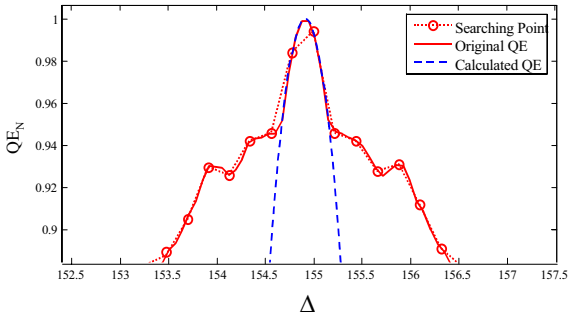

Fig. 2 Comparison of the *QE* curve

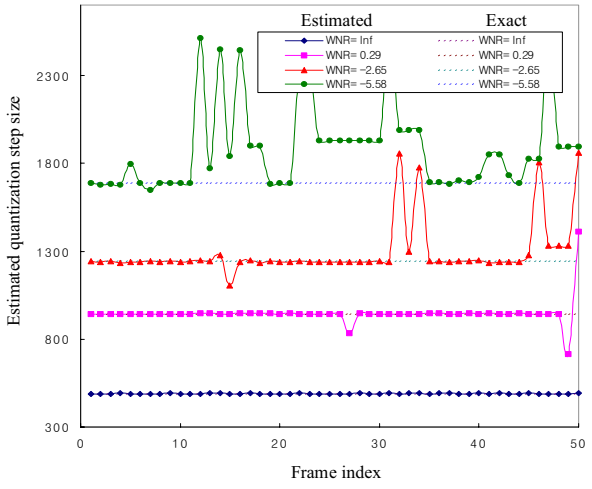

Fig. 3 Trace of the estimated $\hat{\Delta}_d$ in AWGN attack

Table 1. BER with respect to WNR in AWGN attack (g=1.2) [%]

WNR [dB]	Embedding	Exact	Estimated
-11.03	46.8	46.2	46.1
-8.11	44.8	43.7	43.5
-5.17	41.7	39.7	39.5
-2.24	37.1	33.5	33.4
0.71	31.3	24.9	24.8
3.67	24.9	14.7	14.7
6.64	19.8	5.7	5.8
9.62	17.1	1.00	1.0

6. REFERENCES

[1] B. Chen and G. W. Wornell, "Quantization Index Modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory*, vol. 47, no. 4, 2001.

[2] J. J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding," *IEEE Trans. on Signal Processing*, vol. 51, no. 4, pp. 1003-1019, April 2003.

[3] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory*, vol.29, no.7, pp. 439-441, 1983.

[4] J. J. Eggers, R. Bauml, and B. Girod, "Estimation of amplitude modifications before SCS watermark detection," in *Proc. SPIE: Multimedia Systems and Applications IV*, vol. 4675, pp. 387-398, 2002.

[5] K. Lee, D. Kim, T. Kim, and K. Moon, "EM estimation of scale factor for quantization-based audio watermarking," in *Proc. International Workshop on Digital Watermarking*, pp. 335-346, 2003.

[6] F. Perez-Gonzalez, M. Barni, A. Abrardo, and C. Mosquera, "Rational dither modulation: a novel data-hiding method robust to value-metric scaling attacks," in *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, pp.139-142, 2004.

[7] J. C. Oostveen, A. A. C. Kalker, and M. Staring, "Adaptive quantization watermarking," in *Proc. SPIE: Security, Steganography, and Watermarking of Multimedia Contents VI*, vol. 5306, pp. 37-39, 2004.

[8] Qiao Li, and I. J. Cox, "Using perceptual models to improve fidelity and provide invariance to volumetric scaling for quantization index modulation watermarking," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. II-1-4, March 2005.

[9] S. Kim, and K. Bae, "Robust estimation of amplitude modifications for scalar Costa scheme based audio watermarking," in *Proc. International Workshop on Digital Watermarking 2004*, Seoul, pp.121-135, 2004.

[10] S. Kim and K. Bae, "Analysis of optimal searching interval for estimation of amplitude modifications in quantization-based audio watermark detection," in *Proc. 8th International Symposium on Signal Processing and Its Applications*, pp.107-110, 2005.