VECTOR QUANTIZATION INDEX MODULATION WATERMARKING USING CONCENTRIC HYPERSPHERICAL CODEBOOKS

Nima Khademi Kalantari and Seyed Mohammad Ahadi

Electrical Engineering Department, Amirkabir University of Technology, 424 Hafez Avenue, Tehran 15914, Iran
nimakhademi@aut.ac.ir and sma@aut.ac.ir

ABSTRACT

In this paper, a digital watermarking system based on vector quantization is presented. Each vector containing N samples is mapped on the surface of the hyperspheres each of which are associated with a message to embed the digital watermark. We called this method Vector Quantization Index Modulation (VQIM) since it is conventional QIM in the N-dimensional space. The performance of the method and its comparison to orthogonal code-based watermarking is investigated. Furthermore, we implemented the VQIM method on a real audio watermarking system and adopted it with the human auditory system. The experimental results show the robustness of this scheme against common attacks in audio watermarking such as MP3 compression, lowpass filtering, resampling etc.

Index Terms— Vector Quantization Index Modulation, hyperspherical codebooks, decoder optimization

1. INTRODUCTION

Digital watermarking is the art of embedding information (digital watermark) in multimedia data in a way that does not interfere with the normal usage of it. There exist several applications such as copyright protection, content authentication, copy protection etc. for digital watermarking that have made it one of the hottest research topics in recent years.

Three fundamental properties should be considered in designing every watermarking system. These are robustness against attacks, imperceptibility of the watermark embedding and watermark embedding data rate. Besides, there exist other issues such as security, computational complexity, etc. The importance of each of these requirements is dependent on the application in which the watermarking system is going to be used. For example, in fingerprinting, robustness is very crucial while in real-time applications the system should have low computational complexity.

Many watermarking schemes have already been proposed. Among all, much attention is paid to the class of Quantization Index Modulation (QIM) algorithms [2], because of its good rate-distortion-robustness trade-offs and low complexity of encoder and decoder. In this method, the host data is quantized according to the watermark data with a special quantizer chosen from two or more available ones. Data embedding in this class can be done by quantizing single samples (scalar quantization) or by vector quantization. Using single sample quantization results in Dither Modulation (DM) [2] scheme or Scalar Costa Scheme (SCS) [1]. The higher dimensional quantizers, according to rate-distortion theory, improve quantization performance due to better sphere-packing properties. In the vector quantization-based approach, the major problem is the design of the codebooks for quantizing the host data to embed the watermark data. Several methods have been proposed in which vector quantization is used [3], [4], [5], and [6]. Lattice-based codebook methods are investigated in [3] and [4]. In [5] and [6] the concept of equi-energetic codewords, i.e. all codewords lying on the surface of a sphere, is used to divide space into hyperconic regions centered in the origin.

The goal of this paper is to propose a system with low complexity, minimum distortion and optimal decoder. Vector quantization is a time consuming process. By using concentric hyperspherical codebooks the problem of vector quantization was reduced to scalar quantization and therefore low complexity was achieved. Also the distortion was minimized by selecting the radii of the hyperspheres in a way that the proportion of the radii of two consecutive hyperspheres be Δ. Furthermore, the performance of the decoder was optimized by considering the statistical analysis of the detection process.

The remainder of this paper is organized as follows. In Section 2 the proposed method is described. Simulation results and comparisons are presented in Section 3. Section 4 explains the implementation of the proposed scheme on a real audio watermarking system. Simulation results are depicted in Section 5 and Section 6 concludes the paper.

2. VECTOR QUANTIZATION INDEX MODULATION

Quantization Index Modulation was proposed by Chen and Wornell [2] in which host data is quantized with quantizers that are labeled with message bits to embed the watermark data. Encoding in scalar QIM with uniform quantizers is performed as follows.

$$y_n(x_n, m_n) = Q(x_n; m_n) \tag{1}$$

where $Q(\cdot)$ is the quantization function, x_n the host signal, y_n the watermarked signal and $m_n \in \{0,1\}$ (for two quantizers set) is the message bit to be embedded. The decoder extracts the bit from the received signal according to minimum Euclidean distance rule

$$\hat{m}_n = \arg \min_{l \in \{0,1\}} \left\| z_n - Q(z_n; l) \right\| . \tag{2}$$

For extending the scalar QIM to Vector QIM (VQIM) consider host signal $C = \{c_1, c_2, ..., c_N\}$ in N dimensional space. The goal is to embed one bit in the whole of this vector. Information embedding for $N = 3$ in VQIM is depicted in Fig.1. As we can see from Fig.1 the surface of each sphere is our reconstruction area instead of the reconstruction point in scalar QIM. Each vector, in order to lie on the surface of a sphere or a hypersphere, according to the watermark data is moved either upward or downward on the line connecting origin to it. This can be obtained by the following equations.

$$R = \sqrt{\frac{c_1^2 + c_2^2 + ... + c_N^2}{N}} , \tag{3}$$

$$R'(R, m) = Q(R; m) , \tag{4}$$

$$C_w = \alpha \cdot C \quad , \quad \alpha = \frac{R'}{R} . \tag{5}$$

2.1. Distortion minimization

As seen in Fig.1, the radii of the spheres for minimizing the distortion are not uniformly distributed. We chose radius of each hyperspher with the proportion of Δ^n in respect to the radius of the basic sphere[1](R_b), where parameter n is the sphere's index. For example, the spheres with even index were used to embed one and those with odd index were used to embed zero. This was done for two reasons. First, the variance of α was minimized and therefore the distortion was reduced. Second, with this radii arrangement, the system performance is not much dependent on host data variance. The simulation results proved our claim.

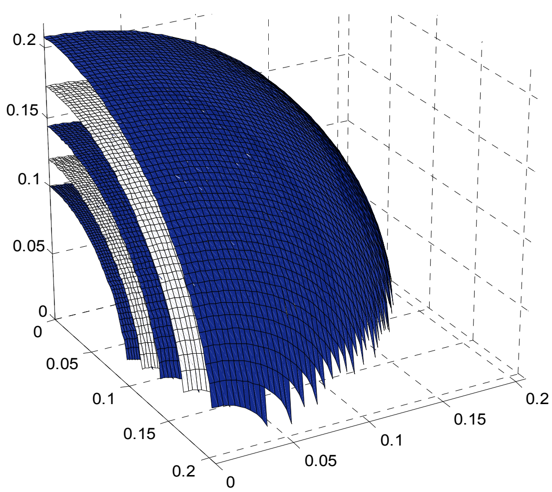

Figure 1. Reconstruction areas in VQIM method. For better illustration only one eighth of each sphere is shown. The innermost sphere is the basic sphere. The surfaces with the same color belong to the same quantizer set.

2.1. Decoding optimization

We used the statistical analysis of the decoding process under AWGN channel to optimize the detection. Suppose the host signal has Gaussian distribution with zero mean and σ_c^2 variance, $N(0, \sigma_c^2)$. We know from the probability theory that if x is a random variable with $N(0,1)$ distribution, x^2 has χ^2 distribution with one degree of freedom and that the summation of N random variables with χ_1^2 distribution has χ_N^2 distribution. If we assume that Δ is near 1 (it is a correct assumption since we do not want to introduce much distortion for imperceptibility of the watermark embedding) then we have

$$\frac{R'^2}{\sigma_c^2} \sim \chi_N'^2 . \tag{6}$$

Consequently, the received signal is added with zero mean and σ_n^2 variance Gaussian noise $N(0, \sigma_n^2)$. Therefore, we have

$$R'' = \sqrt{\frac{(c_1 + n_1)^2 + ... + (c_N + n_N)^2}{N}} . \tag{7}$$

Detection is carried out as:

$$m_n = \begin{cases} 1 & R'' \in one\_regions \\ 0 & otherwise \end{cases} . \tag{8}$$

Our goal is to design zero and one regions to have minimum error. Correct detection is happened when R'' is in the region where watermark data is embedded, i.e.

$$\hat{m} = m \quad \text{If} \quad T' < Z = \frac{R''}{R'} < T \text{ [2]}. \tag{9}$$

Similar to (6) we have

$$\frac{R''^2}{\sigma_c^2 + \sigma_n^2} \sim \chi_N''^2 . \tag{10}$$

Again from the probability theory (by assuming large noise variance and therefore R''^2 and R'^2 to be independent) we know

$$F = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_n^2} \cdot \frac{R''^2}{R'^2} \sim f_{N,N} , \tag{11}$$

where $f_{N,N}$ is F-distribution with N and N degrees of freedom.

$$g_F(f) = \frac{\Gamma(N)}{\Gamma(\frac{N}{2})\Gamma(\frac{N}{2})} f^{\frac{N}{2}-1} (1+f)^{-N} , f > 0 . \tag{12}$$

Also Z has the following distribution

[1] All vectors inside this sphere are not quantized

[2] For the sake of simplicity, the correct region higher and lower than this area is not considered

$$Z \sim \sqrt{\frac{\sigma_c^2 + \sigma_n^2}{\sigma_c^2}} 2z \cdot g_Z(z^2) . \qquad (13)$$

According to the detection theory, it is favorable for us to have $P\{z < T'\} = P\{z > T\}$. The F-distribution has the following property

$$p\left\{ f < \frac{1}{T} \right\} = P\{f > T\}. \qquad (14)$$

It can also be easily shown that for the distribution function in (13) the equality in (14) is right. However, in order to save space, the proof is not included here. Thus, $T' = \frac{1}{T}$ in (9). Note that this result is obtained by assuming low noise variance in (13). Experimentally we found that the property in (14) is also correct when the noise variance is small. It is obvious that by using noise variance estimation and applying it in the threshold, better results could be obtained. Finally we must find the relationship between T and Δ. It is obvious that the higher bound of a zero region must be equal to the lower bound of the next region. Thus we have

$$R_b \cdot \Delta^n . \frac{1}{T} = R_b \cdot \Delta^{n-1} . T \qquad (15)$$

If we simplify the above equation and cross out R_b and Δ^{n-1}, then T can be obtained as

$$T = \sqrt{\Delta} . \qquad (16)$$

3. SIMULATION ON ARTIFICIAL SIGNAL

For evaluating the system performance and its comparison with previous vector quantization methods, VQIM scheme described in section 2 was simulated. For simulation, both the host signal C and the noise attack were generated according to the WGN model. The results were obtained in different Watermark to Noise Ratios (WNR), defined as

$$WNR = 10 \log\left(\frac{E[\|C_w - C\|^2]}{E[\|R - C_w\|^2]}\right) \qquad (17)$$

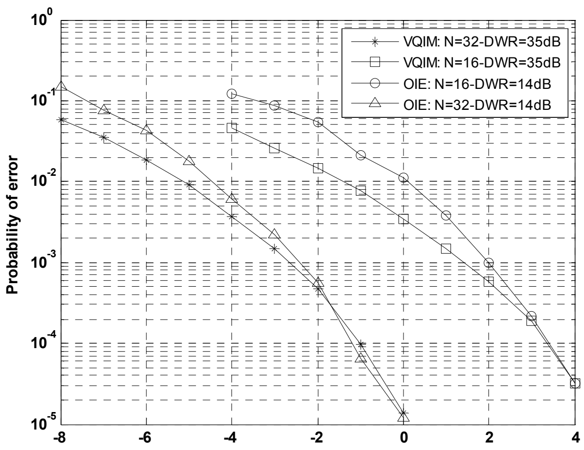

Figure 2. VQIM and OIE performance

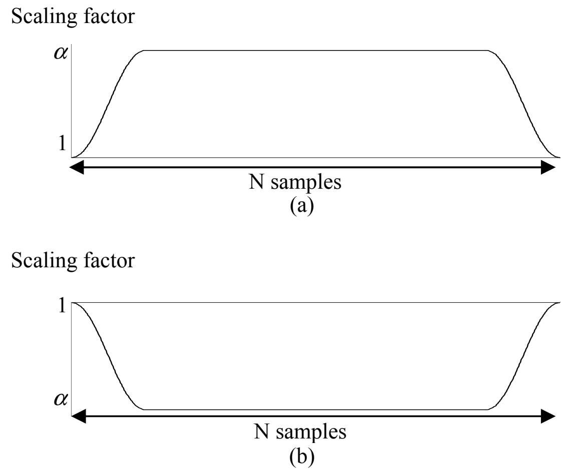

Figure 3. The progressive scaling curves (a) upward movement (b) downward movement.

where R is the received vector. The Document to Watermark Ratio (DWR) which is defined below was fixed for each plot.

$$DWR = 10 \log\left(\frac{E[\|C\|^2]}{E[\|C_w - C\|^2]}\right) . \qquad (18)$$

Fig.2 shows the bit error probability as a function of WNR. The DWR was fixed at 35dB and the system was simulated for N=32 and N=16, i.e. 1/32 bit per sample and 1/16 bit per sample respectively. We compared our scheme with the algorithm of Abrardo and Barni [5] (Orthogonal Informed Embedding (OIE)) in which they split the space into hyperconic regions by using equi-energetic orthogonal codes and quantizing the host data by these regions to embed the watermark bitstream.

As seen in Fig.2, our scheme has better performance in comparison to the OIE scheme. In fact, since our decoder was not optimized for high watermark strength (large Δ), the results in low DWR were not better than the OIE method.

4. IMPLEMENTATION ON REAL AUDIO WATERMARKING SYSTEM

In this section, the implementation of our scheme on audio files will be discussed. As described in previous sections, each vector was quantized and then samples in each vector were multiplied by α to move upward or downward on the surface of the desired hypersphere. Since α is different for adjacent vectors, this will cause discontinuities in vector boundaries. In the audio files, this will result in annoying audible sounds. For solving the discontinuities problem, we adopted a progressive scaling scheme as used in [7], where, at the boundaries, the scaling factor is 1 and then increases or decreases to α. Fig. 3 shows the curve used for smoothing both the upward and downward movements.

Furthermore, in order to embed data in audio signals, Δ should be near 1 to ensure inaudibility which results in bad performance. The most important frequency content

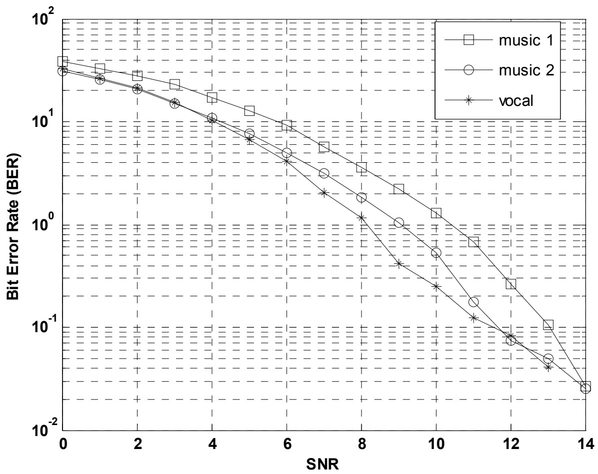

Figure 4. BER(%) vs. SNR(dB) in Additive White Gaussian Noise.

in an audio signal is its low frequency part. Therefore, we first filtered out the high frequency components and then embedded the watermark data in the audio signal consisting of low frequency components. The low frequency audio was subtracted from the original audio and the remainder was added, without any modifications, at the end of the encoding process to the watermarked low frequency part.

5. SIMULATION ON REAL SIGNALS

The simulation was carried out for the evaluation of the proposed scheme's performance on three mono audio signals, each with the length of about 20s, sampled at 44.1 kHz and quantized with 16 bits. Two of them were music clips including both high and low frequency components and one of them was a vocal clip, mainly composed of low frequency components.

Each bit of the watermark signal was embedded within a vector with the length of 256 samples. Therefore, data rate was about 150 bits per second. Also, sharp lowpass filter with the cut off frequency fixed at 2 kHz was used for filtering the audio signal. Radius of the basic sphere (R_b) and step size (Δ) were selected equal to 0.058 and 1.15, respectively. The Bit Error Rates (BER) of detection after adding White Gaussian Noise for different Signal to Noise Ratios (SNR) are depicted in Fig.4. BER for each SNR was found after 50 attempts and then averaging all the results.

Table 1. Robustness of the proposed scheme against common attacks

Attacks	BER(%)		
	Music1	Music2	vocal
MP3 128Kbps	0	0	0
MP3 96Kbps	0	0	0
MP3 64Kbps	4.1	0.6	0
MP3 32Kbps	12.3	3.9	1.5
Requantization 16 to 8	0	0	0
Lowpass 3KHz	0	0	0
Resampling 44/22/44	0	0	0
Resampling 44/11/44	0	0	0
Resampling 44/6/44	0	0	0

Table 1 shows the results of detection after applying some common attacks to the watermarked audio signals. As seen in this table, the proposed algorithm has good robustness against MP3 compression, even in high compression rates, lowpass filtering and resampling to different sampling rates.

6. CONCLUSION

In this paper, a vector quantization-based watermarking scheme has been presented. As described, this method has low complexity because of its simplicity and changing the vector quantization problem to one dimensional quantization. Also, by using nonlinear quantization step sizes, distortion was minimized. Furthermore, decoding process was optimized by using statistical analysis. Although the decoder is not optimized for all DWRs, we can always work near the optimal point, and therefore, our decoder is suboptimal. In comparison to OIE, as demonstrated in Fig.2, our method has better performance. The problem of sensitivity of QIM approaches to gain can be solved by applying Rational Dither Modulation (RDM) [8] to our scheme which can be investigated in future.

7. REFERENCES

[1] J. J. Eggers, R. Bäuml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 1003–1019, Apr. 2003.

[2] B. Chen and G. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[3] J. J. Eggers, J. K. Su, B. Girod. A blind watermarking scheme based on structured codebooks. In *Secure Images and Image Authentication, Proc. IEE Colloquium*, pp. 4/1–4/6, London, UK, April 2000.

[4] R. Fischer, R. Tzschoppe, and R. B¨auml, "Lattice costa schemes using subspace projection for digital watermarking," *European Trans. Telecommunications*, vol. 15, no. 4, pp. 51–362, Aug. 2004.

[5] A. Abrardo and M. Barni, "Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding," *IEEE Trans. Signal Processing*, vol. 53, No. 2, pp. 824-833, 2005.

[6] M. L. Miller, G. J. Doerr, and I. J. Cox, "Applying informed coding and embedding to design a robust, high capacity, watermark," *IEEE Trans. Image Process.*, vol. 13, no. 6, pp. 792–807, Jun. 2004.

[7] W. Lie, and L. Chang, "Robust and High-Quality Time-Domain Audio Watermarking Based on Low-Frequency Amplitude Modification," *IEEE Transactions on Multimedia*, vol. 8, no.1, pp. 46-59, Feb, 2006

[8] F. Pérez-González, C. Mosquera, M. Barni, and A. Abrardo. "Rational Dither Modulation: a high-rate data-hiding method robust to gain attacks," *IEEE Trans. on Signal Processing*, 53(10):3960-3975, October 2005. Third supplement on secure media.