Secured Data Hiding based on Compression Function and Quantization

Ajit Danti
Dept. of Computer Applications,
Jawaharlal Nehru National College of Engg.
Shimoga

G.R.Manjula
Dept. of Computer Science and Engg.
Jawaharlal Nehru National College of Engg.
Shimoga

ABSTRACT

Data hiding is the process of secretly embedding information inside a data source without changing its perceptual quality. In this paper, Quantization Index Modulation and the compression function of μ-Law standards for quantization are used. The proposed method transforms the host signal into the logarithmic domain using the μ-Law compression function. Then, the transformed data is quantized uniformly and the result is transformed back to the original domain using the inverse function. The scalar and the vector methods along with a secret key for data hiding will make the method more secure and efficient. The experimental results demonstrate the robustness of the proposed approach.

General Terms

Data hiding, Security, Information embedding techniques.

Keywords

Quantization, Data hiding, Information security, Digital water marking.

1. INTRODUCTION

The advancement of the internet has resulted in many new opportunities for the creation and delivery of contents in the digital form. The recent growth of network multimedia system has increased the need for the security of digital media. This is important for the protection and enforcement of intellectual property rights. Techniques are needed to prevent the copying, forgery and unauthorized distribution of data. Without such methods, placing data on a public networks puts them at a risk of theft and alteration. Information hiding methods such as cryptography, steganography and watermarking will result in reduction of theft and alteration of data.

Watermarking is a technique of hiding the secure information into the data without introducing a noticeable change. These systems embed one signal, called an "embedded signal" or a "watermark", within another signal, called a "host signal". The embedding must be done such that the embedded signal causes no degradation to its host. The science of concealing information is known as "Steganography" and the technology of "Digital Watermarking" has taken the root from it. Digital watermarking and information embedding systems have a number of multimedia applications and has been widely studied during the recent years for the purpose of copyright protection, authentication, fingerprinting, copy protection, secret communication and broadcast monitoring.

Quantization in the image processing, is a lossy compression technique achieved by compressing a range of values to a single quantum value. Quantization refers to the process of approximating the continuous set of values in the image with a finite set of values[2]

Among many watermarking schemes, the class of host-interference rejecting embedding methods called Quantization Index Modulation (QIM) has grabbed the attention of researches due to its good rate of distortion-robustness tradeoffs[1]. QIM is a new class of host-interference rejecting methods. In QIM, the watermark data is embedded by quantizing the host signal featured by a set of quantizers, each of which associated with a different message. The QIM methods provably better rate-distortion-robustness performance than methods such as spread spectrum methods and generalized low bit modulation methods.

Due to the perceptual advantages of logarithmic quantization based methods, the compression function of μ-Law standards for quantization is used[2]. The μ-law algorithm is a companding algorithm, primarily used in the digital telecommunication systems of North America and Japan[14]. Companding algorithms reduce the dynamic range of an audio signal. In analog systems, this can increase the signal-to-noise ratio (SNR) achieved during transmission, and in the digital domain, it can reduce the quantization error (hence increasing signal to quantization noise ratio). These SNR increases can be traded instead for reduced bandwidth for equivalent SNR. μ-Law encoding effectively reduced the dynamic range of the signal, thereby increasing the coding efficiency while biasing the signal in a way that results in a signal-to-distortion ratio that is greater than that obtained by linear encoding for a given number of bits.

The paper is organized as follows: In Section 2, Logarithmic Quantization Index Modulation method is presented. In Section 3, the implementation issues are presented. Performance analysis and comparison results with other works are reported in Section 4. Finally, concluding remarks are made in Section 5.

2. LOGARITHMIC QUANTIZATION & INDEX MODULATION (LQIM) METHOD

Brain Chen and Gregory [1] examined the various information embedding problem and proposed a new class of host interference rejecting embedding methods namely Quantization Index Modulation that performs provable better than the methods like least bit modulation (LBM) and spread spectrum methods.

Comesana and Gonzalez [5] presented a novel quantization-based watermarking technique robust to scaling attacks, with both differential and non-differential versions. They developed the quantization-based watermarking approach in the logarithmic domain. The host signals are transformed using the logarithmic function. Then, quantization is performed using uniform step sizes on transformed signal in order to embed data within host signal. If the original signal is zero, the logarithm of zero is undefined The function used has not any freedom of controlling the distortion in fixed step sizes and the distortion cannot be minimized according to the host signal distribution.

Barni, Bartolini, Rosa and Piva [6] derived a new watermark detection algorithm which is optimum under the Neyman–Pearson criterion. The new algorithm permits to minimize the missed detection probability subject to a given constraint on the maximum allowable false detection rate. The new algorithm, which relies on Bayes statistical detection theory and on a new approach to the modeling of DFT coefficients, permits to improve significantly the performance of correlation-based decoders in terms of watermark robustness.

Cox and Paul [7] have discussed the susceptibility of watermarking algorithms to tampering. A series of attacks that are all independent of the underlying algorithms used for watermarking are summarized. Collision attacks in the watermarking schemes are discusses. "Unrestricted-key watermarking schemes has been discussed in which the attacker knows how to detect a watermark, but despite the knowledge the attacker cannot remove or alter the watermark.

Jonathan Pinel, Laurent Girin, Cleo Baras and Mathieu Parvaix [8] have designed a high-capacity watermarking technique for audio signals. This technique is suitable for many uncompressed audio signals, more particularly for 16-bit Pulse Coded Modulation (PCM) signals as widely used in audio-CD and wav formats. The technique is based on the application of the Quantization Index Modulation (QIM) technique on the MDCT (Modified Discrete Cosine Transform) coefficients of the signal. The underlying basic principle is that, if those coefficients can be significantly modified by quantization in audio compression schemes such as MPEG MP3/AAC without quality impairments, they can also be modified to embed watermark codes.

Yargic[9] have presented a novel data hiding methods that embed secret data during mixed excitation linear prediction (MELP) coding of the speech signal[6] .The secret data bits are hidden by using quantization index modulation (QIM) which is carried out in the multistage vector quantization (MSVQ) of line spectral frequencies (LSF) parameters. According to the steganalysis results, methods that apply QIM on the third and fourth indexes of MELP-MSVQ assure both lower distortion and better steganographic imperceptibility.

Ramachandran, Chou and Ortega [11] proposed a robust method of imperceptible audio data-hiding technique that. A new and improved audio data-hiding technique was derived a method of data hiding was developed that represents the codebook as a tree structure and varies the height of the tree based on perceptual constraints given by the audio signal. The data hiding methods used uniform quantization. The uniform quantization is optimum when the host signal is uniformly distributed. For the hosts with non-uniform distributions, a set of optimum quantizer levels exists, by use of which, quantization introduces minimum distortion to the host signal

Coax, Kilian, Leighton, Shamoon [12] presented a paper which describes a secure algorithm for digital watermarking. The approach works on the additive spread spectrum methods and inserts the data in the most significant spectral components of the data. It has been developed for two principle applications: to identify the owner and to ensure strong resilience to multiple documents or collision attacks.

Zoran Peric, Aleksandar Mosic and Stefan Panic[13] presented a novel coding algorithm based on loss compression using scalar quantization switching technique. The algorithm of switching is performed by the estimating input variance and further coding with Nonuniform Switched Scalar Compandor (NSSC). An accurate estimation of the input signal variance is needed when finding the best compressor function for a compandor implementation. It enables quantizers to be adapted to the maximal amplitudes of input signals. The main contribution of the model is reaching the loss compression through reaching the higher quality of the signal-to-quantization noise ratio (SQNR) in a wide range of signal volumes (variances) with respect to the necessary robustness over a broad range of input variances, and applying possibility for VOIP applications and an effective coding of signals that likewise speech signals follow Gaussian distribution and have the time varying characteristics

A new watermarking technique called Logarithmic Quantization Index Modulation [4] has been developed and published by Kalantari and Seyed Mohammad Ahadi. The host signal features are transformed using a logarithmic function and then quantized uniformly regarding the watermark data. The watermarked data is obtained by applying inverse transform to the quantized data. Euclidean distance decoder is used to extract the watermark data.

The rationale behind the logarithmic quantization is that since signal's amplitudes are more concentrated around zero, more step sizes should be devoted to quantizing smaller amplitudes

and less should be associated to the larger amplitudes. This also leads to a more uniform signal-to-quantization error ratio for different amplitudes.
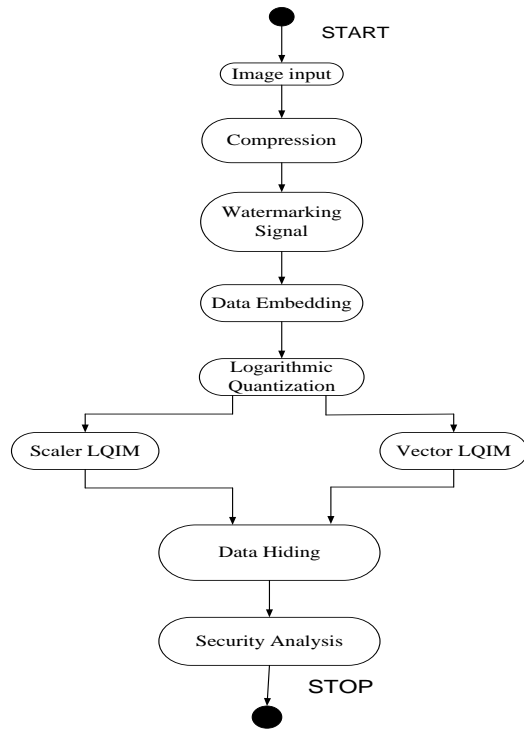

Figure 1: Block Diagram for LQIM method

In order to perform logarithmic quantization, the host signal must be transformed using the following compression function.

$$F(X) = \text{sgn}(x) \frac{\ln(1 + \mu \mid x \mid)}{\ln(1 + \mu)} \text{ -1} \leq \text{x} \leq \text{1}$$

1

The quantized data is then expanded, in order to obtain the watermarked signal. In order to extract the watermark data; the Euclidean distance decoder is used. Euclidean distance decoder can be implemented in the original or the transformed domain.

3. IMPLEMENTATION

The logarithmic quantization index modulation can be implemented using scalar and vector methods. Thus the main application can be further divided into vector logarithmic quantization methods and scalar quantization methods.

3.1 Scalar LQIM

In order to perform logarithmic quantization, the host signal must be transformed using the following compression function:

$$c = \frac{\ln(1 + \mu \frac{\mid x \mid}{X_s})}{\ln(1 + \mu)}, \ \mu > 0, X_s > 0$$

2

where μ is a parameter defining the compression level and X_s is the parameter that scales the host signal. The best value X_s is the value which spreads most of the host signal samples into the range [0, 1]. The values that are larger than one can be converted to the logarithmic domain and be used for embedding. The function used in μ-law does not perform well for these values. There exists only a small number of such values, and thus, there is no major problem with them. Since the host signal samples (or features) may be inappropriately large, X_s should be carefully selected according to the histogram in a way that the host signal samples spread well into the range [0, 1].
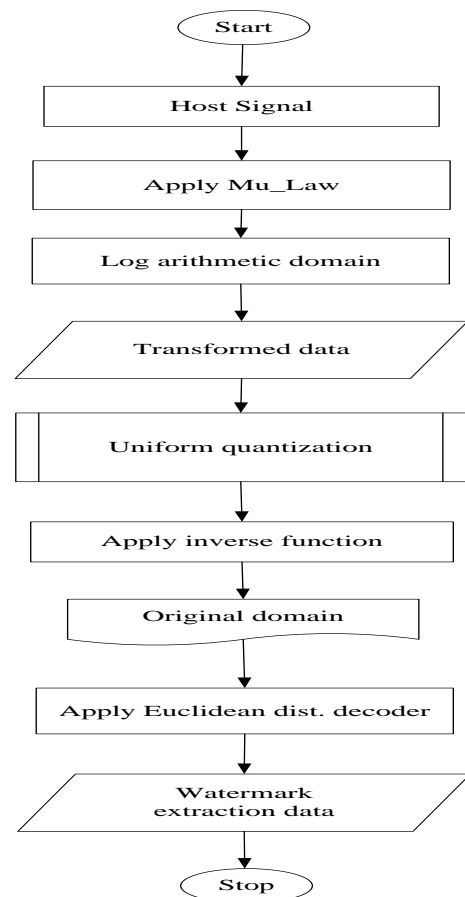

Figure 2: Block Diagram for scalar LQIM

3.2 Vector LQIM

This method can be generalized to N–dimensional space. For this, the magnitude of an N-dimensional vector is quantized. The surface of each sphere is the reconstruction area instead of the reconstruction point in scalar LQIM. Each vector, in order to lie on the surface of a sphere or a hyper sphere, is moved according to the watermark data either upward or downward on the line connecting origin to it. The described method can be implemented as simple as scalar LQIM. The goal is to embed one bit of message into the whole vector. For this, the normalized magnitude is calculated Watermark data can also be extracted using minimum distance decoder.
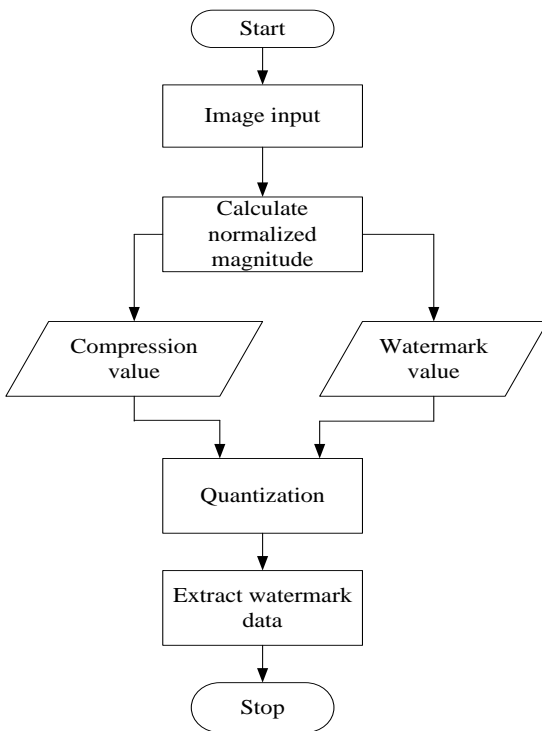

Figure 3:Flowchart for vector LQIM

4. PERFORMANCE ANALYSIS

Our purpose is to show that LQIM provides less distortion and more robustness The reported results in Table 1 refer to Entropy and Mean Squared Error (MSE).

Entropy: Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Low entropy images, such as those containing a lot of black sky, have a very little contrast and large number of pixels with the same or similar pixel value. An image that is perfectly flat will have entropy of zero. Image entropy is calculated with the formula

$$Entropy = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad \textbf{3}$$

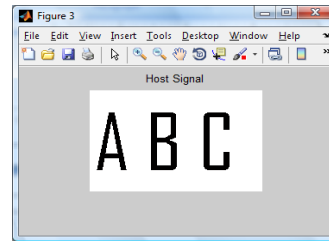where p_i is the probability of occurrence of pixel values i and n is the maximum pixel value.

Mean Squared Error (MSE): Mean squared error is the measure used to quantify the difference between the initial and distorted or noisy image. MSE measures the average of the squares of the "errors". The error is the amount by which the value implied by the estimator differs from the quantify to be estimated. MSE can be calculated using the equation(4)

$$MSE = \sum_{Y=1}^{M} \sum_{X=1}^{N} [I(x, y)\ I(x, y)]^2 / MN \qquad \textbf{4}$$

where I(x,y) is the original image, I'(x,y) is the approximated version and M,N are the dimensions of the images

(a): ESI1.bmp (b): ESI2.bmp (c): ESI3.bmp

(d)ABCD.bmp

(e):GOOD.bmp

Figure 4 (a) to (e) shows sample cover images

(a)Baboon128_color   (b): Lena.bmp

Figure 5 shows the secret images

Table 1 shows the Entropy and MSE between the cover images and secret image(Data hiding using scalar quantization).

Cover Image	Secret Image	Entropy difference	MSE
ESI1	baboon128_color	0.0000000	0.568367
ESI1	lena	0.000000	0.563878
ESI2	baboon128_color	0.0000000	0.560151
ESI2	lena	0.0000000	0.5600012
ESI3	baboon128_color	0.0000000	0.574879
GOOD	baboon128_color	0.0000000	0.808543
ABCD	baboon128_color	0.0000000	0.869972
ABCD	lena	0.0000000	0.872315

Table 2 shows the Entropy and MSE between the cover images and secret image(Data hiding using Wavelet based fusion method).

Cover Image	Secret Image	Entropy	MSE
ESI1	baboon128_color	1.1042	0.00076
ESI1	lena	1.1000	0.00073
GOOD	baboon128_color	0.7404	0.00009
ABCD	baboon128_color	1.1155	0,00076
ABCD	lena	0.8064	0.00008

5. CONCLUSIONS

The experimental result shows that the proposed method provides perceptual advantages that lead to stronger watermark insertion. This method does not have the drawbacks of a previous logarithmic quantization-based method proposed in [5]. By the above we infer that using LQIM technique data can be hidden more effectively which results in less distortion and also the technique provides more robustness.

5.1 Future enhancement

- The existing work can be modified to design a spherical Logarithmic Quantization, which is a vector quantization method for efficiently digitizing analog signals at a high dynamic range and with very low distortion while preserving the original waveform as closely as possible. The technique of SLQ is universally applicable and achieves an efficient digital representation of waveforms with high long term as well as high segmental signal-to-noise ratios.
- Another feasibility of the enhancement of the current work is to design rational Dither Modulation using logarithmic quantization with optimum parameter. In this future work, a new **logarithmic quantization** for Rational Dither **Modulation** (RDM) can be presented.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Chen and G. W. Wornell, "Quantization index modulation methods: 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.

[2] A.M.Kondaz"Digital Speech: Coding for Low Bit Rate Communication System", 2nd edition Wiley Publishers.

[3] Rafael C Gonzalez and Richard E Woods, Digital Image Processing. 3rd Edition Pearson Education Hall.

[4] N.K.Kalantari, Seyed Mohammad Ahadi "A Logarithmic Quantization Index Modulation for Perceptually Better Data Hiding" IEEE Trans on Image Processing, Vol 19, no 6, June 2010.

[5] P.Comesana and F.Perez Gonzalez "On a watermarking scheme in the logarithmic domain and its perceptual advantages", presented at the IEEE Int. Conf. Image Processing, San Antonio, TX, Sep. 2007

[6] M.Barti, F.Bartolini, A DE Rosa and A.Piva, "A new decoder for the optimum recovery of non-additive watermarks," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673-1687, Dec. 1997.

[7] I.J.Coax and Jean Paul , "Some general methods for tampering with watermarks", IEEE J.Sel.Areas Commun, vol 16, no4, pp587-593, May 1998.

[8] Jonathan Pinel, Laurent Girin, Cléo Baras and Mathieu Parvaix "A high-capacity watermarking technique for audio signals based on MDCT-domain quantization" Proceedings of 20th International Congress on Acoustics, ICA, 23–27 August 2010

[9] A.U. Yargıc "Hidden data transmission in mixed excitation linear prediction coded speech using quantization index modulation" Published in IET Information Security, Received on 3rd June 2009, Revised on 3rd May 2010

[10] B. Chen, Design and Analysis of Digital Watermarking, Information Embedding and Data hiding systems. PhD Thesis, MIT Cambridge, MA, June 2000.

[11] J.Chou, K.Ramachandran and A.Ortega, "Next generation techniques for robust and imperceptible audio data hiding," Proc IEEE Int. Conf. Acoustics, Speech. Signal processing (ICASSP), May 2001, vol. 3, pp. 1349-1352.

[12] I. J. Cox, J. Killian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing, vol. 6, pp. 1673–1687, Dec. 1997.

[13] Zoran Peric, Aleksandar Mosic And Stefan Panic "Coding Algorithm Based on Loss Compression using Scalar Quantization Switching Technique and Logarithmic Companding" Journal Of Information Science And Engineering 26, 967-976, 2010

[14] T. Villmann and S. Haase "Mathematical Aspects of Divergence Based Vector Quantization Using Frechet-Derivatives" Machine Learning Reports, Research group on Computational Intelligence,, 2010

[15] William Staling "Cryptography and Network Security" 2nd edition Pearson Education Hall.