JOINT SECURITY & ROBUSTNESS ENHANCEMENT FOR QUANTIZATION EMBEDDING

Min Wu

Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742, U.S.A.

ABSTRACT

This paper studies joint security and robustness enhancement of quantization based data embedding for multimedia authentication applications. We present analysis showing that through a look-up table (LUT) of non-trivial run that maps quantized multimedia features randomly to binary data, the detection error probability can be considerably smaller than the traditional quantization embedding. We quantify the security strength of LUT embedding and enhance its robustness through distortion compensation. Introducing a joint security and capacity measure, we show that the proposed distortion compensated LUT embedding provides joint enhancement of security and robustness over the traditional quantization embedding.

1. INTRODUCTION

Data hiding in multimedia signals has been an active research area in recent years. One potential application is to use the embedded data to verify whether or not a multimedia host signal has been tampered [1]. The data embedding mechanism for these authentication applications should be secure enough to prevent an adversary from forging the embedded data at his/her will [2]. In addition, semi-fragileness is often preferred to allow for distinguishing content changes versus non-content changes. Robustness against moderate compression is desirable as the multimedia data with authentication watermark embedded in may inevitably go through lossy compression, such as in the emerging application of building trustworthy digital cameras [3]. In this paper, we focus on jointly enhancing the robustness and security of core embedding mechanisms that can be used as building blocks for authentication.

While spread spectrum techniques has been widely used to embed a small number of bits robustly in multimedia signals [4], quantization based embedding is more popular for such high-rate data hiding applications as authentication. A popular technique, often known as odd-even embedding [5] or dithered modulation [6], is to choose a quantization step size q and round a feature, which can be a sample or a coefficient of the host signal, to the closest even multiples of q to embed a "0" and to odd multiples to embed a "1". Motivated by Costa's information theoretical result [7], distortion compensation has been proposed to be incorporated into quantization-based embedding [6][8], where the quantization embedding result is combined linearly with the host signal to form a watermarked signal. Using the optimal compensation factor that is a function of watermark-to-noise ratio (WNR), distortion compensated version of odd-even embedding can reach higher payload than the odd-even embedding alone.

This research was supported in part by research grants from U.S. National Science Foundation CCR-0133704 (CAREER), and Minta Martin Foundation. The author can be reached at minwu@eng.umd.edu.

One of the main problems of quantization based embedding is security. An adversary who knows the embedding algorithm can change the embedded data at his/her will, which presents concerns of counterfeiting attacks on authentication [2]. There are three directions to alleviate this security problem. The first way is to encrypt the data to be embedded using a secure cipher such as AES and RSA. The second approach is to provide security to feature extraction, such as deriving features through projecting a set of samples/coefficients along a direction specified by a key [9]. The third approach is to add security to the embedding mechanism itself to make it difficult for an adversary to embed a specific bit at his/her will. Since the first and second approaches involve multiple samples or coefficients, they cannot always allow the localization of tampered regions to fine scale, which is a desirable feature for authentication [1][3]. In this paper, we concentrate on the third approach. More specifically, we propose new enhancement strategies for quantization based embedding, which leads to joint improvement of security and robustness. It can also be combined with other two approaches to further enhance the security strength.

Our proposed approach is built on top of a general embedding technique known as look-up table (LUT) embedding. A pixel-domain LUT embedding scheme was proposed by Yeung et al. [1], and was extended to quantization based embedding in a transform domain in our earlier work [3]. The proprietary look-up table can be generated from a cryptographic key and add security to embedding. With the same quantization step size, the LUT embedding generally introduces larger distortion than the traditional odd-even embedding, making it less popular. In this paper, however, we present analysis showing that at the same WNR, the probability of detection error for LUT embedding can be smaller than the odd-even embedding. We further quantify the security strength of LUT embedding and analyze the effect of distortion compensation on it. As will be seen, our proposed distortion compensated LUT embedding provides joint enhancement of security and robustness over the traditional quantization embedding.

The paper is organized as the follows. We begin with a general formulation of LUT embedding in Section 2. The security and robustness of LUT embedding are analyzed in Section 3 and Section 4, respectively. We then propose and analyze distortion compensated LUT embedding in Section 5 and demonstrate its capability of joint enhancement of security and robustness.

2. LOOK-UP TABLE (LUT) EMBEDDING

We focus on quantization based embedding in scalar features and use uniform quantizers in this paper. A proprietary look-up table $T(\cdot)$ is generated beforehand. The table maps every possible quantized feature value randomly to "1" or "0" with a con-

straint that the runs of "1" and "0" are limited in length. To embed a "1" in a feature, the feature is unchanged if the entry of the table corresponding to that feature is also a "1". If the entry of the table is a "0", then the feature is changed to its nearest neighboring values for which the entry is "1". The embedding of a "0" is similar. For example, we consider a uniform quantizer with quantization step size $q = 10$ and a look-up table $\{..., T(7q) = 1, T(8q) = 0, T(9q) = 0, T(10q) = 1, ...\}$. To embed a "1" to a coefficient "84", we round it to the nearest multiples of 10 such that the multiple is mapped to "1" by the LUT. In this case, we found that "70" satisfies this requirement and use "70" as the watermarked pixel value. Similarly, to embed a "0" in this pixel, we round it to "80".

This embedding process can be abstracted into the following formula, where X_0 is the original feature, Y is the marked one, b is a bit to be embedded in, and $Quant(\cdot)$ is the quantization operation:

$$Y = \begin{cases} Quant(X_0) & \text{if } T(Quant(X_0)) = b\,, \\ X_0 + \delta & \text{otherwise}\,. \end{cases} \quad (1)$$

Here, $\delta \triangleq \min_{|d|}\{d = Quant(x) - X_0 \text{ s.t. } T(Quant(x)) = b\}$. The extraction of the embedded data is simply by looking up the table: $\hat{b} = T(Quant(Y))$, where $\hat{b}$ is the extracted bit.

3. QUANTIFYING SECURITY OF LUT EMBEDDING

During the LUT embedding of Eq. 1, when $T(Quant(X_0))$ does not match the bit to be embedded (b), we need to find a nearby entry in LUT that is mapped to b. As such, the run of "1" and "0" entries of an LUT need to be constrained to avoid excessive modification on the feature. We denote the maximum allowable run of "1" and "0" as r. To analyze security as a function of r, we start with the case of $r = 1$, which leads to only two possible tables. One table has "1" for odd entries and "0" for even entries, and the other has "0" for odd and "1" for even. This is essentially the odd-even embedding [5] or the dithered modulation embedding [6]. Since there is little uncertainty in the table, unauthorized persons can easily manipulate the embedded data, and/or change some feature values while retaining the embedded values. As we discussed earlier in this paper, the traditional quantization embedding, or equivalently the choice of $r = 1$, is not appropriate for authentication applications if no other security measures, such as a careful design of what data to embed, are taken.

When r is greater than 1, the number of LUTs satisfying the run constraint can be computed through a recursive relation. For example, a binary LUT can be constructed by equiprobably initializing each of the first two entries to 0 or 1, and generating the remaining entries with maximum allowable run of 2. Let the number of k-entry LUTs that satisfy the above conditions be F_k. We can show that F_k equals to twice the *Fibonacci* series: $F_{k+1} = F_k + F_{k-1}$ for $k \geq 2$, and $F_0 = 2, F_1 = 2, F_2 = 4$. For a binary LUT with length 256 and maximum run of 2, the total number of such LUTs is on the order of 10^{53}, which is a significant increase from only 2 possible tables for run 1.

We further quantify the uncertainty of LUT embedding by identifying the generation process of binary LUT as a $2r$-state Markov chain illustrated in Fig. 1. We can show that the stationary probability of both $0^{(i)}$ and $1^{(i)}$ states is

$$\pi(0^{(i)}) = \pi(1^{(i)}) = \frac{2^{r-i-1}}{2^r - 1} \quad (2)$$

for $i = 1, ..., r$, and the entropy rate of the stationary process $\{Z_1, Z_2, ...\}$ in unit of bit is

$$\lim_{n\to\infty} \frac{H(Z_1, ..., Z_n)}{n} = \lim_{n\to\infty} H(Z_n|Z_{n-1}) = 1 - \frac{1}{2^r - 1}. \quad (3)$$

For example, in the case of maximum allowable run $r = 2$, the LUT generation process is a 4-state Markov chain with an entropy rate of $2/3$ bit. In contrast, the entropy rate with maximum run of 1 (or equivalently, the odd-even embedding) is 0 bit. This indicates that the uncertainty of LUT has increased significantly with a slight increase of the maximum allowable run.

It is important to note that the security quantified in this section concerns how much uncertainty (against an adversary's guess) a basic embedding mechanism can offer to each individual feature. Zooming into an LUT embedding mechanism that is already sufficiently secure at the individual feature level, another security aspect addresses how feasible it is for an adversary to derive the LUT from a number of watermarked features [2]. Such a threat can be alleviated by introducing location dependency so that effectively different LUTs are used for different features. Interested readers can refer to [2] for details.

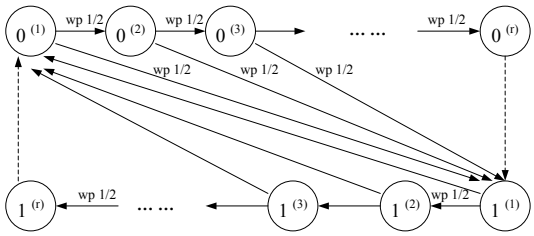


Fig. 1. A Markov chain model for LUT table generation, where the transition probability is $1/2$ for solid arrow lines and 1 for dash arrow lines.

4. ROBUSTNESS ANALYSIS ON LUT EMBEDDING

Though bringing higher security, the increase in the allowable run r will inevitably lead to larger embedding distortion when a feature value of the host signal is not mapped by LUT to the bit to be embedded. In this section, we analyze the mean squared distortion introduced by LUT embedding and the probability of detection error under additive white Gaussian noise (AWGN).

Using the stationary state probability of the Markov Chain model in Fig. 1, we can show that the overall mean squared distortion incurred by LUT embedding with binary LUT and maximum allowable run $r = 2$ is $q^2/2$, where q is the quantization step size [10]. This is larger than the MSE distortion of $q^2/3$ by the odd-even embedding (or equivalently, LUT embedding with run 1). However, with larger allowable run in LUT, stronger noise that drags a watermarked feature out of the enforced interval does not necessarily lead to errors in detection. Therefore, the probability of detection error can be reduced.

To quantify the robustness in terms of the probability of detection error, we assume that the watermarked feature is at $k'q$ and that the additive noise follows i.i.d. Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with zero mean and variance σ^2. The probability of detection error under Gaussian noise can be approximated by [10]

$$P_e^{(r=2)} \approx \frac{4}{3} \cdot \mathcal{Q}(q/2\sigma) = \frac{4}{3} \cdot \mathcal{Q}(\sqrt{\gamma/2}) \quad (4)$$

where the Q-function $\mathcal{Q}(x)$ is the tail probability of a Gaussian random variable $\mathcal{N}(0, 1)$. The watermark-to-noise ratio (WNR)

γ is defined as as the ratio of MSE distortion introduced by watermark embedding to that by additional noise, and we have $\gamma = q^2/2\sigma^2$ for the LUT embedding with maximum allowable run $r = 2$. We have compared with the simulation result for maximum allowable run $r = 2$ and found that the analytic approximation and simulation conform very well [10]. In contrast, for LUT with maximum run of 1, detection error occurs as soon as the noise is strong enough to drag the watermarked feature to the quantization intervals next to the $k'q$ interval. The probability of detection errors for this embedding is

$$P_e^{(r=1)} \approx 2\times[\mathcal{Q}(\sqrt{3\gamma}/2)-\mathcal{Q}(\sqrt{3\gamma}\cdot 3/2)+\mathcal{Q}(\sqrt{3\gamma}\cdot 5/2)]. \quad (5)$$

Using a total of 500,000 simulation points at each WNR ranging from -6dB to +10dB, we compare the probability of detection error vs. WNR for maximum allowable run r of 1, 2, 3, and infinity, respectively. As can be seen from Fig. 2, P_e of maximum run of 2 (solid line) is significantly smaller than run of 1 (dot line) for up to $4dB$-advantage at low and medium WNR, and is slightly higher at high WNR. In addition, the further increase of LUT's run (dot-dash line and dash line) gives only a small amount of reduction of P_e at low WNR and much larger P_e at medium and high WNR. This indicates that LUT embedding with maximum allowable run of 2 can potentially provide higher robustness as well as higher security than the commonly used quantization embedding with equivalent run 1. In the next section, we explore techniques that further improve the robustness and capacity of LUT embedding.

5. DISTORTION COMPENSATED LUT EMBEDDING

Motivated by Costa's information theoretical result [7], distortion compensation has been proposed and incorporated into quantization-based embedding [6][8], where the LUT enforced feature is combined linearly with the original feature value to form a watermarked feature. Using an optimal scaling factor that is a function of WNR, distortion compensated version of odd-even embedding provides higher capacity than without compensation [6]. The basic idea behind such improvement is to render more separation between the watermarked feature values while keeping the MSE distortion introduced by the embedding process unchanged. In this section, we propose to apply distortion compensation to LUT embedding and study the impact of distortion compensation on the reliability of LUT embedding.

Linear Distortion Compensation Let X_0 be the original unmarked feature, X_1 the output from LUT embedding alone (with maximum allowable LUT run $r = 2$), and Y the finally watermarked feature after distortion compensation. We use a quantization step size of q/α to produce X_1 in the LUT embedding step, where $\alpha \in (0, 1]$ is also used as a weighting factor in distortion compensation: $Y = \alpha X_1 + (1-\alpha)X_0$. The overall mean squared distortion introduced by this distortion compensated embedding remains the same as in the non-compensated version that uses a quantization step size of q. One criterion for selecting of α is to maximize the following "SNR":

$$SNR^{(r=2)} = \frac{2\cdot(q/\alpha)^2}{(1-\alpha)^2\frac{(q/\alpha)^2}{2}+\sigma_n^2}. \quad (6)$$

Here the "signal" power in the numerator is the mean squared distance between two neighboring, perfectly enforced feature values that represent "1" and "0", respectively; the "noise" power in the denominator is the mean squared deviation away from a perfectly enforced feature, where the deviation is introduced by both distortion compensation and additional noise of variance σ_n^2. The α value that maximizes the above SNR can be found as $\alpha_{opt}^{(r=2)} = 1\,/\left(1+\frac{1}{WNR}\right)$.

Robustness and Capacity We quantify the robustness of different embedding settings through their embedding capacities at a wide range of WNRs. For simplicity, the channel between embedding and detection is modelled as a binary symmetric channel (BSC) with cross-over probability being the probability of detection error P_e. We compare the BSC embedding capacity of five cases in Fig. 3, namely, maximum allowable run of 2 with and without distortion compensation, constant run of 1 (traditional odd-even embedding) with and without compensation, and maximum allowable run of infinity (i.e. no run constraint) with compensation. From the cross marked line to the dash line, we see that the embedding capacity when maximum allowable run is 2 increases significantly for up to $4dB$-advantage in WNR after applying distortion compensation. We also observe that when keeping all other conditions identical and only varying the maximum allowable run of LUT, the increase in allowable run gives higher embedding capacity in low WNR when no compensation is used (dot line and cross marks), and a moderately smaller capacity when distortion compensation is applied (solid line, dash line, and circles).

The difference in capacity, or equivalently in the probability of detection error P_e, for different embedding settings is also reflected in our proof-of-concept experiments with the 512×512 Lenna image. One bit is embedded in each pixel through LUT embedding with run constraint $r = 2$ and linear distortion compensation. The embedded raw data forms a 512×512 pattern shown in Fig. 5(a). We have also implemented a embedding scheme using the same LUT but without compensation, as well as the popular odd-even embedding with and without compensation. The base quantization step q is 3 and the PSNRs of watermarked images are about $42dB$. We then add white Gaussian noise to watermarked images and tailor its strength to give a WNR of $0dB$ in all tests. The detection errors on uncoded data of 512×512 bits are visualized in Fig. 5, from which we can see an improvement of distortion compensation on reducing the raw bit error rate by 10% (Fig. 5(d)). Channel coding can be applied to provide reliable communication at targeted WNRs, as demonstrated in [10].

Joint Security and Capacity Measure The above comparison, however, concerns mainly the robustness/ capacity and does not include information about security. To take into account both capacity and security issues, we define a joint measure $J(H, C)$ as a function of the entropy rate H of the embedding mapping and the embedding capacity C. One simple choice of $J(\cdot,\cdot)$ is a linear combination of the entropy rate and the embedding capacity under binary symmetric channel (BSC) assumption for additive noise. That is,

$$J = \omega H_{LUT} + (1-\omega)\cdot C_{LUT}, \quad (7)$$

where H_{LUT} is the entropy rate of LUT table given by Eq. 3, C_{LUT} is the BSC embedding capacity, and $\omega \in [0, 1]$ is a weighting factor to provide a desirable emphasis on security and robustness issues. We plot this joint measure at $0dB$ WNR for maximum LUT run of 1 and 2, respectively, with different weight ω and different compensation settings. We see from Fig. 4 that distortion

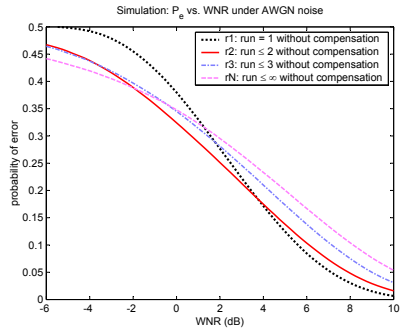


Fig. 2 Detection error probability under AWGN noise for LUT embedding with different maximum allowable LUT runs.

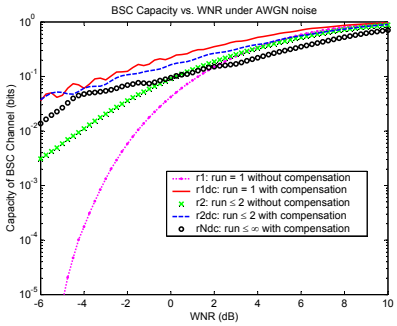


Fig. 3 BSC embedding capacity under different maximum allowable LUT runs and different compensation settings.

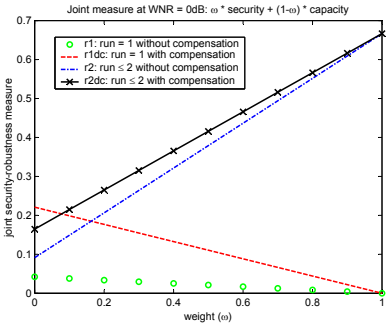
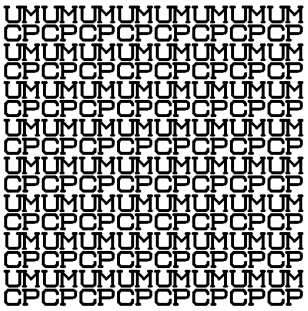
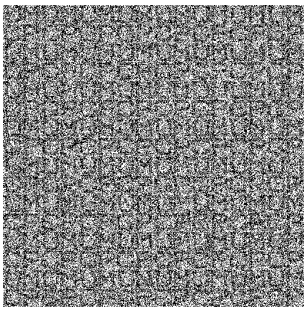


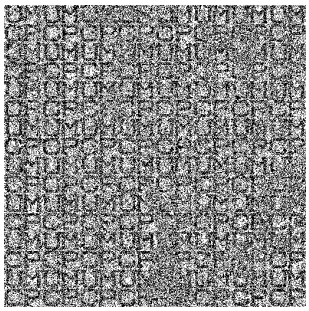
Fig. 4 Linear joint security and capacity measure of LUT embedding as a function of weight ω at a WNR of $0dB$.



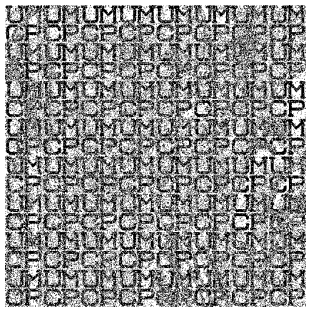
(a) originally embedded pattern



(b) run=1, no compensation, extraction error 38.8%



(c) run$\leq$2, no compensation, extraction error 35.1%



(d) run$\leq$2, with compensation, extraction error 26.0%

Fig. 5. Visualization of raw error patterns under WNR of 0dB by LUT embedding with different settings.

compensated embedding with run constraint of 2 (cross marked line) gives the highest J until the weight ω goes below 0.15 and security is not much concerned, where the joint measure for the traditional odd-even embedding with distortion compensation (dash line) becomes higher. This suggests that as long as some level of security is desired, by slightly increasing the allowable LUT run from 1 to 2 and by applying distortion compensation, we can provide joint improvement of security and robustness to quantization based embedding.

6. CONCLUSIONS

In summary, this paper studies the joint enhancement of security and robustness for quantization based data embedding. We start with a general embedding approach that employs a look-up table to map quantized multimedia features to binary data. We quantify the security strength of LUT embedding in terms of entropy rate and have shown that the security is improved significantly with a slight increase of the allowable LUT run from 1 to 2. We present analysis showing that LUT embedding with larger run constraints can have smaller probability of detection error for up to $4dB$-advantage in WNR. We then explore distortion compensation on LUT embedding to further enhance its robustness and provide an additional advantage of up to $4dB$ in WNR. Through a joint security and capacity measure, we have shown that our proposed distortion compensated LUT embedding with maximum allowable run of 2 offers joint enhancement of security and robustness over the traditional quantization embedding that has an equivalent run of 1. This joint enhancement makes the proposed embedding scheme an attractive building block for authentication applications.

7. REFERENCES

[1] M. M. Yeung and F. Mintzer: "An Invisible Watermarking Technique for Image Verification", *ICIP'97*.

[2] M. Holliman and N. Memon, "Counterfeiting Attacks on Oblivious Blockwise Independent Invisible Watermarking Schemes", *IEEE Trans. on Image Proc.*, vol.9, no.3, 2000.

[3] M. Wu and B. Liu: "Watermarking for Image Authentication", *ICIP'98*.

[4] I. Cox, J. Kilian, T. Leighton, and T. Shamoon: "Secure Spread Spectrum Watermarking for Multimedia", *IEEE Transaction on Image Processing*, vol.6, no.12, pp.1673-1687, 1997.

[5] M. Wu and B. Liu: "Data Hiding in Image and Video: Part-I – Fundamental Issues and Solutions", accepted by *IEEE Trans. on Image Processing*, Nov. 2002, to appear.

[6] B. Chen and G.W. Wornell: "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding", *IEEE Trans. on Info. Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.

[7] M.H.M. Costa: "Writing on Dirty Paper", *IEEE Trans. on Info. Theory*, vol. IT-29, no. 3, May 1983.

[8] P. Moulin and J.A. O'Sullivan: "Information-Theoretic Analysis of Information Hiding", preprint, Sept. 1999, revised Dec. 2001, http://www.ifp. uiuc.edu/ ∼moulin/ paper.html.

[9] M. D. Swanson, B. Zhu, and A. H. Tewfik: "Robust Data Hiding for Images", *Proc. of IEEE DSP Workshop*, 1996.

[10] M. Wu: "Joint Security and Robustness Enhancement for Quantization Based Embedding," accepted by *IEEE Trans. on Circuits and Systems for Video Technology*, May 2003, to appear.