A TRANSCODING ROBUST DATA HIDING METHOD FOR IMAGE COMMUNICATION APPLICATIONS

Çağatay Candan

Department of Electrical Engineering
Middle East Technical University (METU), Ankara Turkey
ccandan@metu.edu.tr

ABSTRACT

We present a data embedding method for image communication applications. Our goal is to implement novel multimedia applications such as multi-language captions, interactive programming and title specific features over the existing image communication channel. To this aim, we present a data embedding method for JPEG images which has the desired degree of robustness to transcoding or bitrate adjustments that may take place in the communication channel. The described system is designed for JPEG images but can be extended to MPEG systems.

1. INTRODUCTION

Data hiding is the art of embedding bits into a multimedia signal with little or no perceived distortion. Security applications such as the copyright protection, signal authentication, fingerprinting are the most well known applications of data hiding. A hiding based solution to the security problems could not be established until now in spite of some significant achievements, [1, 2]. The research focus of data hiding is currently shifting from security applications towards the application specific designs with less stringent attack requirements. The considerations such as hiding capacity, embedding distortion and implementation complexity are being explicitly introduced as the central figure of merits, [3]. In this paper, we present an algorithm that is designed to establish novel multimedia applications over the existing image communication channel.

The conventional systems use MPEG compression and standard communication protocols for video broadcasting. Offering new applications within the existing compression and communication framework is highly desirable from service providers' viewpoint. We propose a data hiding based method to implement such applications. The data of a multimedia application can be embedded into the compressed stream and then be transmitted through the conventional channel. The receiver can extract the hidden bits, while doing the standard picture decoding operations and can implement an action based on the extracted application bits and the content. The multimedia applications such as multi-language captions, commentaries, special

features and a variety of title specific applications can be implemented using the described data hiding solution. The only cost of the implemented applications is the additional distortion on the transmitted content signal due to embedding operation. In [7], we have presented an algorithm which minimizes the embedding distortion, in this paper we present a hiding algorithm robust to the bitrate changes of the delivered stream.

The described data hiding method is given for JPEG images, but can be easily adapted to I frames of MPEG systems. The distinguishing feature of the method is its transcoding robustness at high embedding payloads. We note that transcoding operation is inevitable in many communication scenarios and perhaps it is the most critical attack for communication applications. For example, the content signal can be compressed further at an intermediate link (say at a satellite link), or can be compressed to meet the specifications of a client device. A good data hiding method for the communication applications should be able to deliver the application at different network conditions. The described method can embed thousands of bits in a TV resolution JPEG image with very little distortion. The robustness of the method is controlled by a parameter. The cost of increased robustness is the increased distortion on the content signal.

In data hiding literature, robustness to a given attack is established in two ways. The first approach is the usage of redundancy or the usage of diversity techniques at embedding, [8]. The second approach is based on finding a robust domain, if possible an invariant one, to the given attack. A simple example is the utilization of the DFT magnitude invariancy to the circular shift attack. The idea of invariant domains has been successfully applied to joint rotation-scale-translation invariancy in [9]. The invariancy to more general geometric attacks has also been studied in [10].

We use the invariancy approach to establish the transcoding robustness. It is clear that the redundancy approach is not suitable for the high capacity communications applications. Invariancy approach is not easy to implement since a domain totally invariant to transcoding is not possible to define. Different from geometric attacks, transcoding is not an invertible operation or in other words there is some inevitable loss of information after

transcoding. The robustness to transcoding is an ambitious goal for high capacity embedding applications. The image coders are built to discard the redundant bits and our goal of transcoding robust embedding is just the opposite.

In literature there are many robust data hiding methods for JPEG attack. Most of these methods are designed for security applications and are not applicable to communications application area. In particular, these methods either lack the high capacity requirement or the efficient decoding mechanism requirement. For example methods whose decoding is based on matched filtering are not applicable for communication applications. The matched filter decoders evaluate the inner product of the received signal with the candidate signals to check whether a particular mark is embedded or not. Therefore, matched filtering is not possible to implement when there are too many candidates to check. For communication applications embedding only 8 bytes of information may result in an infeasible number of inner product evaluations (can be up to 2^{64} if the hidden bits are independent from each other).

In the next section, we give a short description of JPEG system, illustrate its inefficiencies and propose a robust embedding rule.

2. METHOD

JPEG Compression: JPEG is a transform domain quantization algorithm. The 8x8 blocks of the original image are individually transformed to the DCT domain and non-uniformly quantized with a pre-determined quantization table. The quantization table entries are selected to minimize the overall perceived distortion. The quality factor parameter (QF) of the JPEG system controls the quantization and therefore the bitrate of the compressed image. The quality factor is a real number between 0 and 100 and is used to generate a multiplication factor which scales the entries of the default quantization step sizes. The relation between QF and the multiplication factor is defined as:

$$\text{factor} = \begin{cases} 5000/\text{QF} & \text{QF} < 50 \\ 200 - 2\text{QF} & \text{QF} \geq 50 \end{cases}$$

The entries of the default quantization table (quantization table for QF=50) are scaled and rounded to the nearest integer, $\tilde{S}_{jk} =$ round($S_{jk} \times$ factor/100). For example when QF is changed from 50 to 25, the step-sizes are doubled, less number of bits are needed to express non-zero quantized coefficients.

The success of JPEG transform is based on the removal of the signal dependency in a typical block, i.e. removal of intra-block redundancy. It is known that as the correlation of data in a given block increases, the DCT transform approaches the optimal KL Transform, [11]. The inter-block dependencies are almost totally ignored in the JPEG system. The only effort to reduce the inter-block redundancy is the differential coding of the DC coefficients of the consecutive blocks.

In Figure 1, we give an illustration of inter-block redundancy. The image on the upper left corner is formed by the collection of DC coefficients from every block of Lena Image. The image on its right is the collection of first DCT coefficient (the coefficient on the right of DC coefficient at every individual block). Other images are the collection of similar low frequency coefficients. We call the collection of the same indexed DCT coefficients as the DCT channels, [12]. Apart from the differential encoding of the DC coefficient, JPEG system does not take any measure to reduce the redundancy in DCT channels. Recently, Tran has introduced a context based arithmetic coder to reduce the inter-block redundancy, [13]. To illustrate the amount of residual redundancy, we have estimated the correlation in the DCT channels shown in Figure 1. The matrix entries below indicate the correlation (calculate row-wise) in the DCT channels of Figure 1.

Fig. 1. DCT Channels: Collection of the same indexed DCT coefficients. The DC channel and eight other low frequency channels are illustrated for Lena image.

$$\text{correlation coef.} = \begin{bmatrix} 0.89 & 0.45 & 0.49 \\ 0.50 & 0.55 & 0.52 \\ 0.60 & 0.55 & 0.53 \end{bmatrix}$$

Even though the correlation of 0.89 in the DC channel is virtually eliminated by the application of differential coding, the correlation of 0.5 still remains in other channels.

Method: A domain whose coefficients can be altered without significant distortion is needed to define a successful hiding method. As illustrated low frequency DCT channels can be used for this purpose. And to sustain the high transcoding robustness, we give an invariancy relation as described below.

First we note that transcoding is nothing but the requantization of DCT coefficients with a coarser quantizer. It is easy to see that the relative order of the coefficients that are quantized is

preserved after requantization (since quantization, rounding to a nearest multiple of step-size, is a monotonic function). That is, if the coefficient A is greater than or equal to B, the same order relation is preserved when A and B are quantized:

$$A \geq B \Leftrightarrow \text{round}(\frac{A}{\text{step}}) \times \text{step} \geq \text{round}(\frac{B}{\text{step}}) \times \text{step}$$

We propose the following embedding rule: The coefficients in a DCT channel is partitioned into 3x3 blocks. The coefficients in each 3x3 block are sorted in the increasing order. We call the sorted coefficients as the list and denote them with $x_s(k)$ where $x_s(k+1) \geq x_s(k) \geq x_s(k-1)$. Note that the order of coefficients in the list does not after quantization. We define an embedding rule based on the locations of the coefficients in the list.

The embedding rule compares the mean and median of the middle third portion of the list: $\{x_s(4), x_s(5), x_s(6)\}$. If the bit to be hidden is a "1", the coefficients in the list are changed such that the mean of the middle third coefficients, $1/3(x_s(4) + x_s(5) + x_s(6))$, is greater than the median, $x_s(5)$; after hiding. If the bit is a "0", the mean after hiding should be below the median.

The hidden bit extraction process is the application of a simple check on the received coefficients. At the decoder, DCT channels are formed and partitioned into 3x3 blocks and sorted in the increasing order, $x_s(k)$, $x_s(k+1) \geq x_s(k) \geq x_s(k-1)$. The hidden bit is extracted by comparing mean and median of the middle third of the list.

$$\text{hidden bit} = \begin{cases} 1 & 1/3(x_s(4) + x_s(5) + x_s(6)) > x_s(5) \\ 0 & 1/3(x_s(4) + x_s(5) + x_s(6)) < x_s(5) \\ e & 1/3(x_s(4) + x_s(5) + x_s(6)) = x_s(5) \end{cases}$$

Since the relative ordering of coefficients are preserved, we expect to have same mean-median relation after the application of quantization.

If the calculated mean and median values of are different from each other, the decoded bit is guaranteed to be the bit embedded at the encoder (assuming that there are no other coefficient modifying attacks in the channel). If the mean and median values are the same, we can not decode that particular hidden bit. We call such decoding ambiguities as erasures, following the error-correction tradition, [14].

Next, we introduce a parameter to improve the robustness of the described system. Embedding rule is modified to the following where Δ is a free variable:

$$\text{hidden bit} = 1 \rightarrow 1/3(\tilde{x}_s(4) + \tilde{x}_s(5) + \tilde{x}_s(6)) > \tilde{x}_s(5) + \Delta$$
$$\text{hidden bit} = 0 \rightarrow 1/3(\tilde{x}_s(4) + \tilde{x}_s(5) + \tilde{x}_s(6)) < \tilde{x}_s(5) + \Delta$$

In the above relation the parameter Δ is the guaranteed minimum value between the mean and median after hiding. The earlier design has Δ as zero. The simulation results given in the next section illustrate the effectiveness of the Δ parameter on reduction of erasures.

Selection of the DCT Channel: The described algorithm can be used to embed bits into any DCT channel. In order to minimize the embedding distortion we have chosen to use the DCT channel with the smallest quantization step-size. The channel with the smallest quantization step size (using the default quantization table) is the one formed by the collection of coefficients from the first row and third column of each block. Depending on the application requirements, we can use more than one DCT channels for embedding.

3. COMPUTER EXPERIMENTS

We present two computer experiments. The first experiment is on the selection of DCT channels. A 512x512 Lena image at the QF=75 is selected as the content image. For the given image, DCT channels have the size of 64x64 and 411 bits can be embeddedd in each channel. The following table shows the embedding distortion of the Lena image when 411 randomly generated bits are embedded in different low frequency DCT channels (Δ is selected as 0.33).

$$\text{After Hiding PSNR} = \begin{bmatrix} 36.94 & 44.99 & 48.03 \\ 45.7 & 47.26 & 46.64 \\ 46.57 & 46.37 & 44.12 \end{bmatrix}$$

The results of the experiment confirm that the lowest distortion occurs for the 1-3 channel as anticipated.

The second experiment is on the effectiveness of Delta parameter. 411 randomly generated bits are embedded at four different values of Δ, $\Delta = \{0, 0.33, 0.66, 1\}$. The composite image is then transcoded to different QF levels. The Table 1 illustrates the reduction in percentage of erasure bits at different levels. The introduced embedding distortion on the Lena image for the increasing value of Δ is $\{49.83, 47.45, 45.31, 43.38\}$ dB.

The choice of $\Delta = 0.33$ seems to be a reasonable choice having little embedding distortion and high robustness at almost all cases except the harsh quantization case of QF=25. In the extensions section, we propose another method in addition to the increase in Δ parameter to reduce the number of erasures.

4. EXTENSIONS

On Choice of Δ: We present a bound for Δ guaranteeing a desired robustness level. Assume that hidden bit to be embedded is a "1", that is $x_4 + x_6 > 2x_5 + 3\Delta$. Assume that the carrier is compressed after hiding and at the the new compression level the step-size values are $1 + p$ times the original ones ($p > 0$). Under these conditions, the coefficients x_k ($k = \{4, 5, 6\}$) will then be mapped to $\langle \frac{x_k}{1+p} \rangle$. $\langle \cdot \rangle$ denotes the rounding operation. For small enough p, $\frac{x_k}{1+p}$ can be approximated with $x_k(1 - p)$.

QF	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
$\Delta = 0$	41.2	31.7	30.8	29.7	31.5	10.8	8.8	7.7	11.5	1.3	0	0.2	5.4	0	0	0	0
$\Delta = 0.33$	34.2	11.1	9.5	7.4	9.0	1.1	0	0	2.2	0	0	0	0	0	0	0	0
$\Delta = 0.66$	12.2	2.2	0.2	0	1.3	0.9	0	0	0	0	0	0	0	0	0	0	0
$\Delta = 1.00$	3.8	0.2	0	0	1.36	0	0	0	0	0	0	0	0	0	0	0	0

Table 1. The percentage of erasures at varying transcoding levels. The robustness parameter is denoted by Δ. The compression levels are denoted by the quality factor QF.

A bound on Δ can be given as follows:

$$
\begin{aligned}
\langle \frac{x_4}{1+p} \rangle + \langle \frac{x_6}{1+p} \rangle &\approx \langle x_4(1-p) \rangle + \langle x_6(1-p) \rangle \\
&\geq x_4 + x_6 - [\langle x_4 p \rangle + \langle x_6 p \rangle] \\
&\geq 2x_5 + 3\Delta - [\langle x_4 p \rangle + \langle x_6 p \rangle] \\
&\geq 2\langle \frac{x_5}{1+p} \rangle + 3\Delta - [\langle x_4 p \rangle + \langle x_6 p \rangle]
\end{aligned}
$$

From the last relation, we can deduce that if $3\Delta - [\langle x_4 p \rangle + \langle x_6 p \rangle] > 0$ is satisfied, the mean and median values after compression are different from each other.

On Erasure Coding: It is possible to decrease the number of erasures by an increase in Δ at the expense of embedding distortion. Similarly the error correction codes can be used to reduce the number erasures at the expense of hiding capacity. An error correction code with the minimum distance d can correct $d-1$ erasures, [14, page 227]. The combination of an (n,k) error-correction code with sufficiently high minimum distance (enough to correct erasures on the 10% of the received bits) with $\Delta = 0.33$ embedding level produce zero decoding error for almost all transcoding cases in Table 1.

5. CONCLUSIONS

In this paper, we have presented a data hiding algorithm designed for the image communication channel. The proposed algorithm can be immediately put into use to establish novel application over the communication channel within the existing compression and communication framework. The hiding method has adjustable robustness to the transcoding operations at the expense of embedding distortion. Some experiments on the selection of parameters and an analytical bound on the robustness parameter are presented.

The presented method is based on a trade-off between embedding distortion and transcoding robustness. The results of the paper illustrates one of the reachable conditions in the resolution of this trade-off. It can be possible that a superior algorithm with lower distortion at the same level robustness may exist. We believe that the communication applications has many practically and academically interesting trade-offs to pursue. Finally we invite the readers to the following demo site for an interactive demo of the presented method:

http://www.eee.metu.edu.tr/˜ccandan/demos/

6. REFERENCES

[1] M. Barni and F. Bartolini, "Data hiding for fighting piracy," *IEEE Signal Processing Magazine*, pp. 28–39, 2004.

[2] "Special issue on enabling security technologies for digital rights management," *Proceddings of IEEE*, June 2004.

[3] "What is the future for watermarking ?," *IEEE Signal Processing Magazine*, vol. 20-5, pp. 55–59, 2003.

[4] Ç. Candan and N. Jayant, "A new interpretation of data hiding capacity," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2001.

[5] Ç. Candan and N. Jayant, "A minimum distortion data hiding technique for compressed images," *International Workshop on Multimedia Signal Processing*, 2002.

[6] Ç. Candan and N. Jayant, "A Secure Authentication Method for JPEG Images with Tampering Localization ," *Submitted to Trans. Signal Proc.*, 2004.

[7] Ç. Candan and N. Jayant, "A Minimum Distortion Data Hiding Scheme For JPEG Images ," *in preparation* .

[8] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, pp. 1673–1687, 1997.

[9] J. J. K. O'Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Process.*, pp. 303–317, 1998.

[10] M. Alghoniemy and A. H. Tewfik, "Geometric invariance in image watermarking," *IEEE Trans. Image Process.*, vol. 13, pp. 145–153, 2004.

[11] A. Jain, "A sinusoidal family of unitary transforms," *IEEE Trans. Pattern Analysis, and Machine Intelligence*, pp. 356–365, 1979.

[12] P. Moulin and M. K. Mıhçak, "A framework for evaluating the data-hiding capacity of image sources," *IEEE Trans. Image Process.*, vol. 11, pp. 1029–1042, 2002.

[13] C. Tu and T. D. Tran, "Context based entropy coding of block transform coefficients for image compression," *IEEE Trans. Image Process.*, vol. 11, pp. 1271–1283, 2002.

[14] S. Wicker, *Error control systems for digital communication and storage*. Prentice Hall, 1995.

[15] G. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, pp. 30–44, 1991.

[16] P. Wong and O. Au, "A capacity estimation technique for JPEG-to-JPEG image watermarking," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 13, pp. 746–752, 2003.