A MULTIPLE DESCRIPTION CODING SCHEME BASED ON THE CHINESE REMAINDER THEOREM

Çağatay Candan

Multimedia Communications Laboratory
Georgia Institute of Technology, Atlanta, GA 30332
candan@ece.gatech.edu

ABSTRACT

A novel multiple description coding scheme based on the Chinese remainder theorem is presented. The main advantages of the scheme are the uniform diffusion of information to multiple packets, the flexible control mechanism over the representation redundancy, and the possibility of having a detail preserving post-processing routine to reduce the reconstruction error. We believe that the described Chinese Remainder Theorem based approach can be useful for other coding problems such as coding with side information and information hiding.

1. INTRODUCTION

Reliable and efficient transfer of information from one place to another is the common goal of communication and signal processing. Different coding schemes have been designed to address various user requirements (lossless, almost lossless, lossy, minimum delay etc.) and system constraints (phone-line, internet, wireless channel etc). In this paper, we present a new approach based on a well known number theoretical result which can be useful for three different coding problems: multiple description coding (MDC), source coding with side information, and information hiding. The presentation given here discusses the problem mainly from the MDC point of view. The connections with other coding regimes are discussed at the end of the paper.

Multiple description coding aims at efficiently representing the input signal with a number of streams (or packets) such that the reconstructed signal quality increases with the number of streams successfully received. MDC is aimed to lessen the effects of unreliable media such as the internet which may fail at packet delivery.

By definition, a MDC scheme should be able to reconstruct a reasonable approximation of the input signal with the partial information on the transmitted signal. A method proposed by Servetto et. al. can be described as follows [1]: The input signal is quantized with a set of quantizers. The output of the nth quantizer forms the nth packet. Quantizers of [1] are designed such that the reconstruction resolution improves with the reception of each packet (description). A simple but illustrative example can be given as follows: Assume that the data to be transmitted takes integer values and let's say that the input signal x takes the value of 4 at time n, i.e. $x[n] = 4$. If we use two *central quantizers* with quantization step sizes of 3 and 7, the quantized signal values become $x_{Q3}[n] = 3$ and $x_{Q7}[n] = 7$. If the receiver receives only one of the transmitted packets, the reconstruction point is the value of the quantized signal that is received. If both packets are successfully delivered, the receiver can achieve a better resolution on $x[n]$, since the reconstruction point possibilities after the reception of the packet carrying x_{Q3} is $\{2, 3, 4\}$, the possibilities due to the other packet is $\{4, 5, 6, 7, 8, 9, 10\}$; the intersection of these sets gives the reconstruction point at 4.

In this paper, we propose an alternative approach. Instead of sending the quantized signal values, we propose to transmit the quantization errors. For this purpose, we use the mid-thread quantizers instead of central-quantizers. A mid-thread quantizer quantizes the input value to the quotient of the input and step size division. The quantization error of this quantizer becomes the remainder of the division. Let $x[n] = 10$, then the quantization error of mid-thread quantizers with step sizes 3 and 7 can be calculated using modulo arithmetic, $(x[n])_3 = 1$ and $(x[n])_7 = 3$ where we represent the modulo N equivalent of $(\cdot)$ with $(\cdot)_N$. Similar to the above approach, we send the modulo reduction results as the information packets. This time, the possibility sets formed by the reception of packets are infinite sets, i.e. $\{1, 4, 7, 10, ...\}$ and $\{3, 10, 17, ...\}$ for the first and second packet respectively. The intersection of these two sets gives us another infinite set $\{10, 31, 52, ...\}$. Unless we have a-priori information about the source of the signal, we cannot single out the input from the final possibility set. If the receiver knows that $x[n]$ is the output of a source with 16 possible outcomes (i.e. an image with 16 colors), we can then uniquely construct $x[n] = 10$ from the information packets that are received. The simplified discussion given in this paragraph is essentially the theorem known as the Chinese Remainder Theorem in number theory.

2. SCHEME

In this section, we focus on multiple description coding of images. For this presentation we assume that the images are represented with 8-bits. We start with a brief discussion of the Chinese Remainder Theorem (CRT):

CRT: If $P_1, P_2, ..., P_n$ are integers with no common factors (pairwise relatively prime) then the integers between 0 and $(-1 + \prod_{i=1}^{n} P_i)$ and the set of tuples-of-integers formed by reduction of x in modulo $P_1, P_2, ..., P_n$ are in one-to-one relation. (That is, for every tuple of integers, there exists one and only one integer in the interval $[0, \prod_{i=1}^{n} P_i - 1]$.)

The mapping between tuples and scalars for the two prime case ($n = 2$) can be written as follows:

$$x = (\ (x)_A B (B^{-1})_A + (x)_B A (A^{-1})_B\ )_{AB} \qquad (1)$$

As before $(\cdot)_A$ denotes equivalent of $(\cdot)$ in modulo A. If A and B are relatively prime, the inverses required in (1) are guaranteed to exist. Readers may notice the resemblance of equation (1) to an interpolation relation (say sinc interpolation) or to the well known practical method of partial fraction expansion of ratios of polynomials.

The interested reader may consult [2] for other applications of the CRT. Two important applications are the NIM arithmetic system for efficient fixed-point DSP algorithm implementations and Good's method for fast calculation of the DFT for composite block lengths.

We describe the proposed scheme through an example. The first step of the scheme is the partitioning of the image. For the sake of simplicity we choose to partition the image into 2 by 2 blocks. After the partitioning operation, the two most significant bits of each pixel is extracted and the CRT is applied to the remaining 6 bits. Below, we present a block of data that is processed accordingly:

$$\begin{bmatrix} 110 & 140 \\ 45 & 190 \end{bmatrix} = \begin{bmatrix} 64+46 & 128+12 \\ 0+45 & 128+62 \end{bmatrix} \equiv$$

$$\left\{ \begin{bmatrix} 64 & 128 \\ 0 & 128 \end{bmatrix}_{SB}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}_2, \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}_5, \begin{bmatrix} 4 & 5 \\ 3 & 6 \end{bmatrix}_7 \right\}$$

Going over the above example, the upper left pixel of the block has the value of 110 and the two most significant bits of this pixel is 01 which contributes the value of 64 to the accumulative value of 110. The contribution of the remaining six bits becomes $110 - 64 = 46$. The two most significant bits of every pixel in the partition is similarly collected in the matrix with the (SB)subscript. The other matrices on the second line correspond to modulo reduction of the less significant bit contribution. For the upper left pixel with less significant bit contribution value 46, the modulo reduction in moduli $\{2, 5, 7\}$ is applied to get $(46)_2 = 0$ and $(46)_5 = 1$ and $(46)_7 = 4$ and these residual values are placed in upper left hand part of the matrices with the subscripts matching the modulo value.

As can be seen from the discussion on decoding, we have chosen to separate the significant bits from the others to reduce the complexity of decoding when few packets are received.

The contribution of the least significant 6 bits ranges from 0 to 63. We have the perfect construction of the less significant bits from the residual information, since the chosen moduli values $\{2, 5, 7\}$ have the product of 70, which is greater 64 implying the existence of perfect construction according to the CRT. At the next step, we construct the packets to be transmitted.

To make the relative importance of the packets identical to each other, which is important if the media treats all packets identically, we place the same number of elements from SB and moduli matrices into every packet. If we aim to packetize the input into 4 packets, the following can be one of the several ways of diffusing the data.

$$\left\{ \begin{bmatrix} (064)_{SB} & (2)_5 \\ (1)_2 & (6)_7 \end{bmatrix}_{P_1}, \begin{bmatrix} (4)_7 & (0)_2 \\ (000)_{SB} & (2)_5 \end{bmatrix}_{P_2}, \right.$$
$$\left. \begin{bmatrix} (1)_5 & (128)_{SB} \\ (3)_7 & (0)_2 \end{bmatrix}_{P_3}, \begin{bmatrix} (0)_2 & (5)_7 \\ (0)_5 & (128)_{SB} \end{bmatrix}_{P_4} \right\}$$

From the above representation, it can seen that every packet carries information about the most significant bits of a pixel in the block and some partial information about the less significant bits of the other pixels in the block. For example packet 1 (P_1) has the MSB bits of the upper left pixel of the image block, other values in packet 1 are the modulo reductions of the least significant 6 bits of other pixels in the block.

It is clear that if all packets are successfully received, we can reverse the process described to re-assemble the matrices described and apply the inverse mapping available through the CRT to reconstruct the original signal.

If all of the packets are not received, the decoding procedure becomes non-trivial. As expected from the discussion at the introduction, the decoding reduces to selection of best representative from a set of possibilities. For example, after the reception of the first packet (P_1), the possibilities for pixels are:

$$\begin{aligned} \text{UL} &= \{64\} + \{0, 1, 2, \ldots, 63\} \text{ (64 cases)} \\ \text{LL} &= \{0, 64, 128, 192\} + \{1, 3, 5, \ldots, 63\} \text{ (128 cases)} \\ \text{UR} &= \{0, 64, 128, 192\} + \{2, 7, 12, \ldots, 62\} \text{ (52 cases)} \\ \text{LR} &= \{0, 64, 128, 192\} + \{6, 13, 20, \ldots, 62\} \text{ (36 cases)} \end{aligned}$$

The UL,LL,UR,LR denote upper/lower and left/right.

A decoding strategy has to be established to differentiate between the possibilities. If the neighboring blocks of data are received with higher reliability (less uncertainty), the information from the neighboring regions can be used to make a selection. A post-processing method based on this idea is developed in the next section. It is interesting to note that the decoding problem is reduced to picking the right answer out of some possibilities as in a multiple choice test.

As soon as we receive the second packet (P_2), the choices can be updated as follows:

$$\begin{aligned} \text{UL} &= \{64\} + \{4, 11, 18, \ldots, 60\} \text{ (9 cases)} \\ \text{LL} &= \{0\} + \{1, 3, 5, \ldots, 63\} \text{ (32 cases)} \\ \text{UR} &= \{0, 64, 128, 192\} + \{2, 12, 22, \ldots, 62\} \text{ (28 cases)} \\ \text{LR} &= \{0, 64, 128, 192\} + \{27, 62\} \text{ (8 cases)} \end{aligned}$$

It is clear that with the reception of more packets, the uncertainty about the right choice gets smaller and the decoding problem gets easier. This procedure continues until all packets are received. If one or more packets is lost, a choice from the set of possibilities has to be made. In the next section, we propose a decision recipe and give some computer experimentation results to determine the quality loss in partial reception cases.

3. COMPUTER EXPERIMENTS

We present a simple decoding method which can be used progressively. We continue using the same example presented in the previous section.

If only one description for a block is received, we choose the reconstruction point as the median value of possibility set.

If two or three packets are received, the selection is made such that the sum of the pair-wise distance between the pixels of the block is the minimum one among the choices or in other words the selection is made such that the resultant partition is the smoothest one among the choices.

If all four packets are received, the CRT is applied to get the perfect construction.

The results of the decoding with the mentioned decoding algorithm can be seen in Figure 1. In this simulation, we assume that the lost packets are of the same type throughout the image. For example, the image in the upper left hand corner of the Figure 1 is reconstructed from the information contained in Packet 1, we assume that for this image none of the packets numbered 2, 3, and 4 are received throughout the image. A more realistic situation is the independent reception of blocks of the image. We believe that the simulation we present illustrates the worst case condition.

Fig. 1. Upper left, lower left, upper right and lower right images are reconstructed with the reception of packets of type {1},{1,2},{1,2,3} and {1,2,3,4} respectively. The corresponding PSNR values are given in Table 1.

Figure 2 represents the reconstructed images after the application of a simple post processing operation. The post-processing operation can be considered as external help to make a better selection out of the possibility set. The effect of the post-processing is clearly visible on the reconstructed image with 3 packets. In this case, the 3x3 median filter is applied to the reconstructed image and the median filter output is stored. The post-processed image in Figure 2 is formed by re-mapping the pixel values of the reconstructed image to the value in their possibility set which is closest to the stored median filter output. With median filtering operation the general trend of a block is estimated from its neighbors and with the re-mapping operation the lost detail due to smoothing is restored back. It can be observed that due to this combined effect smooth regions of the image, such as shoulder part or the background have zero error after post-processing.

Fig. 2. Reconstruction results after the application of post-processing.

To understand the effect of post processing on quality, we give the histogram of the error magnitude ($Error_{kn} = |R_{kn} - O_{kn}|$) before and after post-processing in Figure 3. The top point of two stacked bars (black bar + white bar) shows the number of pixels in error at a given error magnitude level before post-processing. The black part of the bar shows the pixels in error at a particular error magnitude after the post-processing. The overall area of the white region gives the number of corrected pixels by post processing. Different from many schemes, the error magnitude does not take arbitrary values, but a pre-determined multiple of moduli, (either multiple of 2, 5, 7 or MSB errors of $\{64, 128, 192\}$). We reiterate that the transmitted signal is the quantization error and the reconstruction error is due to the errors in the estimation of the quotient.

The redundancy of the representation is also an important factor. One feature of this scheme is its explicit redundancy value. The representation discussed in the previous section has the redundancy of $log_2(70/64)$ bits per pixel which implies an addition transfer of 0.13 bits/pixel. Changing the moduli values to $\{3, 7, 8\}$, we get a more redundant scheme with the redundancy of 1.4 bits/pixel. The reconstruction with the more redundant scheme is easier than the former one due to the faster elimination of choices. The PSNR values of the

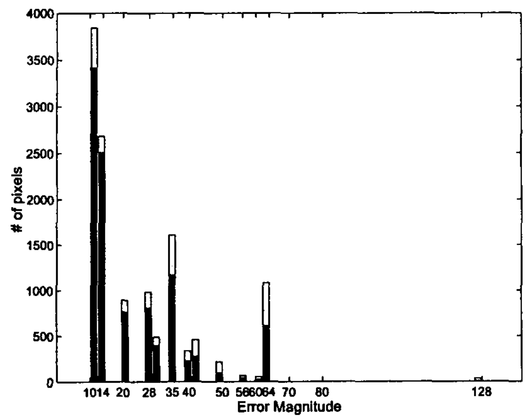

Fig. 3. Histogram of error before and after post processing.

reconstructed images for the more redundant scheme can be seen in Table 1.

Method/# of Packets	1	2	3	4
$\{2,5,7\}$ scheme	19.07	19.75	25.59	∞
$\{2,5,7\}$ + post-proc.	21.87	21.90	27.55	∞
$\{3,7,8\}$ scheme	19.06	22.05	26.77	∞
$\{3,7,8\}$ + post-proc.	21.80	24.55	29.05	∞

Table 1. PSNR values of the reconstructed images for two different schemes.

4. CONCLUSIONS

In this section, we elaborate on the method proposed, discuss some advantages and disadvantages of the CRT based approach and point out some connections with the other problems.

The application of the CRT to the multiple description coding problem can be seen as a natural extension of the CRT. The CRT establishes a mapping between scalars and vectors and this type of mapping from a single unit of information to the multiple units of information is the basic idea of MDC.

In many multiple description coding schemes, the important part of the data such as the DC value of the block is repeatedly sent in different packets. In the CRT based approach, all information units have an identical importance level making the method more suitable for the lossy media.

Another interesting fact of the CRT based scheme is the systematic reduction of the uncertainty on the perfect construction at each successful transmission. For the partial information reconstruction case, the decoder of the CRT based approach has to make a decision out of a possibility set. Even though this decision process may seem arbitrary at first sight, it should be noted that decoding of the all MDC algorithms is based on the estimation of the missing information. For example, in [3] the variance of the transform coefficients in the missing packets is estimated from the received ones. Different from the other schemes, the packets of the proposed CRT framework contains only the detail information and the general trend is left to be guessed, which is in many cases easier to estimate. This effect can be seen from Figure 2 where the smooth regions are perfectly constructed with only 3 packets; but the unpredictable information contained in the 4th packet causes visible reconstruction errors around the edges.

If the reconstruction output of the proposed scheme is cascaded with a post-processing filter followed by a re-mapping operation as described before, some reductions in the reconstruction error can be accomplished. Different from the most post-processing methods, with this method the detail information can be restored back after the initial smoothing stage of post-processing.

The uniform distribution of the information to the packets and the controlled redundancy of the representation are among the other advantages of the scheme.

The main disadvantage of the presentation given here is its space domain definition. We believe that the CRT based ideas can also be applied in transform domain and a carefully designed decoding method may lead to good partial reconstruction results in transform domain also. We hope that this paper would be a precursor for such an activity.

The Chinese Remainder Theorem is also applicable to the problem of source coding with side information. This coding problem assumes that the receiver has access to some side information about the transmitted signal. The coding problem for this application is the design of the most compact representation of the signal that is not yet available to the receiver. [4] proposes to use cosets of linear error-correction codes to systematically partition the space. It can be seen that the CRT based approach can be used in a similar fashion to partition the space.

Another application area can be information hiding. The QIN modulation method, as described in [5], can be seen as a method of partitioning the coding domain into equivalence classes with the condition that elements (codewords) of an equivalence class (codebook for source coding) are sufficiently close that the source coding distortion (quantization error) is small. The hidden information with this scheme lies in the choice made in the selection of the quantizer or the equivalence class. We believe that an application of the CRT based ideas can also be interesting for this problem.

5. REFERENCES

[1] S. Servetto, V. Vaishampayan, and N. Sloane, "Multiple description lattice vector quantization," *Data Compression Conference*, pp. 13–22, 1999.

[2] J. H. McClellan and C. M. Rader, *Number theory in digital signal processing*. Prentice Hall, 1979.

[3] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Process.*, pp. 351–366, 2001.

[4] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *Data Compression Conference*, pp. 158–167, 1999.

[5] B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Information Theory*, vol. 47, pp. 1423–1443, 2001.