Abstract: - In this paper we present a framework for estimating corrupted coefficients in multiple description coding (MDC) of images. Descriptions are generated by partitioning the domain of wavelet coefficients into sets whose elements are maximally separated from each other. Consequently an edge-adaptive estimation algorithm is proposed to explore the local texture statistical characteristics for estimating the lost coefficients. The edge adaptation property of our algorithm significantly increases the intra-subband dependencies employed in reconstruction. Extensive experimental results show that impressive improvements on both objective and subjective measures can be achieved.

Key-Words: - Multiple description coding, error concealment, covariance method, domain partitioning based MDC, edge-adaptive algorithm, mutual information.

1 Introduction

Recently, a promising approach, multiple description coding (MDC), has emerged to enhance the resilient coding and transmission for image and video signals. Typically, a multiple description (MD) coder is a source coding technique that provides a graceful quality degradation in the presence of packet losses during transmission. A conventional media coder generates only one stream, while a MD coder encodes a media source into two or more bitstreams. These bitstreams, also called descriptions, are generated in such a way that each description can be independently decoded to produce a signal of basic quality. The decoder can reconstruct a signal of improved quality with more than one descriptions received. In order to achieve this goal, each description should contain a sufficient amount of information about the original source. That is, a certain amount of correlation has to be adopted to help reconstruct signals from each description.

A number of MDC methods for image and video coding have been proposed during the past decade [1]. Vaishampayan proposed a simple and practical MDC scheme, known as multiple description scalar quantizers (MDSQ), to generate two sub-streams by producing two indices for each quantization level [2]. The index assignment is designed to be equivalent to a fine quantizer when both indices are received, but a coarse quantizer when only one index is received. More complicated quantizers, such as trellis-coded vector quantizers and entropy coders were later designed to improve the coding efficiency [3], [4], [5]. An alternative way of designing MDC algorithms uses correlation-inducing transforms [6], [7]. Wang proposed applying a pair-wise correlating transform (PCT) to each pair of uncorrelated variables obtained from the Karhunen-Loeve transform (KLT) [6]. Goyal et al. further generalized Wang’s work to any number of variables, and coined the term generalized multiple description coding (GMDC) [7]. GMDC was later applied to image coding with correlating transforms [8]. As compared to the MDSQ method, PCT/GMDC has been shown to perform well at low rates [9]. Other popular MDC schemes introduce redundancy in different ways: through frame expansion [8] or unequal forward error correction [10].

Recently, Bajic [11] proposed a novel wavelet multiple description coding scheme named as domain-based MDC that is created by partitioning the wavelet transform coefficients. The method is bandwidth-efficient in the sense that no extra redundancy is introduced. Only the inherent spatial redundancy in the transform domain of the signal is employed to make the bitstream robust and enable lost descriptions to be estimated from the received ones. However, this MDC system only focuses on finding the best way to partition the transform coefficients but ignores efficiently adopting the inherent correlation within the signals to reconstruct the corrupted coefficients at the decoder.
In this paper, we concentrate on improving the reconstruction of lost descriptions based on domain-based MD coder. Instead of only depending on the immediate available neighbors of the corrupted samples for reconstruction at the receiver, we explore all the edge orientation correlations within a local window so that we can estimate the corrupted descriptions using the covariance method adaptively corresponding to the local texture statistical characteristics. As a result the lost coefficients can be more accurately reconstructed. In this sense, our reconstruction algorithm is edge-adaptive to the local texture. Hence the proposed scheme is named as edge-adaptive MD coder (EA-MDC).

2 Domain partitioning based MDC

Successful estimation of missing signal coefficients relies on the autocorrelation of the signal. The autocorrelation is typically a decreasing function of distance. It is expected to get the best results if the missing coefficient is estimated from its immediate neighbors. Moreover, the distortion will be a non-increasing function of the size of the available neighbors surrounding a missing signal coefficient for typical estimation algorithms [11]. Hence, Bajic proposed an optimal way to partition the coefficients in transform domain by maximizing the minimum distance within the same partition [12], i.e. the minimum intra-partition distance. Each partition is formed as a description.

An example, given in Fig. 1, illustrates a two-level wavelet decomposition of a 16×16 image, generating \( P = 4 \) descriptions. Numbers 0 through 3 indicate the partition to which the corresponding subband coefficient belongs. In this case, the minimum intra-partition distance is maximized [12] so that more neighbors from other descriptions are expected within the same subband. Assume the subbands are ordered in a zig-zag manner from the lowest frequency band to the highest frequency band, and indexed as \( k = 0, 1, ..., 6 \). Given the partitioning scheme \( p(x) \) for the lowest frequency band, the partition for each higher frequency band is obtained by applying modulo-shifted partitioning function, \( p^{(k)}(x) = [p(x) + k] \mod P \), to the partitioning function for the lowest frequency band.

It is apparent from Fig. 1 that each description contains the same number of coefficients from each tree and also the same number of coefficients from each subband. This makes the descriptions “equally important”. After partitioning, each description is encoded independently from other descriptions.

3 Edge-adaptive MD coder

Reconstructing a lost description from the received ones is naturally related to error concealment techniques. In this section, we present a MD coder scheme that improves the error concealment process for better reconstruction quality. First we generate quantized coefficients with a conventional encoder. Then the optimal domain partitioning is applied to generate the descriptions. They are encoded independently and transmitted through different channels. At the decoder side, the lost samples are reconstructed by an edge-adaptive (EA) error concealment algorithm according to their local texture statistical characteristics.

3.1 Edge-adaptive error concealment algorithm

This section addresses the problem of estimating a lost coefficient \( X_0 \) from its available neighbors. The available neighbors could be either coefficients belonging to the successfully received descriptions, or previously recovered coefficients. To facilitate the discussion, the set \( L_x = \{X_1, X_2, ..., X_L\} \), \( L \geq 1 \), also named as local window, is used to denote the available neighbors of \( X_0 \). The atomic problem in recovery is to obtain the optimal estimation of \( X_0 \) in the sense of maximizing \( p(X_0 \mid L_x) \).

Although the wavelet coefficients for natural images can not be modeled by a stationary Gaussian process it can be viewed as locally stationary. Hence, the conditional probability distribution function \( p(X_0 \mid L_x) \) is mostly characterized by the second-order statistics (covariance). On the other hand, the \( L \)th-order Markov property holds for image source, i.e. \( p(X_0 \mid X_1, X_2, ...) = \)
The Maximum A Posterior (MAP) estimation $p(X_0 \mid L_x)$ comes down to the linear minimum mean square error (MMSE) estimation problem of minimizing $E[(X_0 - \hat{X}_0)^2]$, where $\hat{X}_0$ is the linear estimation of $X_0$ based on the set of its immediate neighbors $N_x = \{X_1, X_2, \ldots, X_N\}$

$$
\hat{X}_0 = \sum_{k=1}^N w_k X_k .
$$

According to the classical Wiener filtering theory [13], the optimal weights, $\bar{w} = [w_1, w_2, \ldots, w_N]^T$ can be determined by the covariance of the Gaussian process

$$
\bar{w} = (R_{xx})^{-1} \tilde{r}_x ,
$$

where $\tilde{r}_x = [r_1, r_2, \ldots, r_N]^T$, $r_k = \text{cov}(X_k X_l)$, $1 \leq k \leq N$, and $R_{xx} = [r_{ij}]$, $R_{ij} = \text{cov}(X_k X_l)$, $1 \leq k, l \leq N$. The vector $\bar{w}$ in (2) is actually the projection of $X_0$ onto the subspace spanned by $N_x$ in the least-square sense. Under the assumption of a locally stationary process, we can estimate $R_{xx}$ based on the covariance [13] as,

$$
\tilde{R}_{xx} = C^T C , \quad \tilde{r}_x = C^T \tilde{I} .
$$

where $\tilde{I} = [X_1, X_2, \ldots, X_L]^T$ is an $L \times 1$ vector composed of all the available neighbors inside the local window $L_x$. $C$ is an $L \times N$ matrix whose $l$th row contains the $N$ immediate neighbors used to estimate $X_l$, and is given by

$$
C = \begin{bmatrix}
X_{1,1} & \cdots & X_{1,N} \\
\vdots & \ddots & \vdots \\
X_{L,1} & \cdots & X_{L,N}
\end{bmatrix} .
$$

By combining (2) and (3), the optimal weight $\bar{w}$ is obtained by

$$
\bar{w} = (C^T C)^{-1} (C^T \tilde{I}) .
$$

Image intensity field experiences a sharp transition across the edge orientation and is almost homogeneous along the edge orientation. These geometric constraints of edges hold in the wavelet domain, especially for high resolution subbands. The above covariance-based estimation can be viewed as an edge-adaptive error concealment method that is capable of tuning $\bar{w}$ along an arbitrarily-oriented edge across the local window. Therefore, important edge features can be well estimated. On the other hand, this method is also enhanced with the way to generate descriptions by DP-MDC. When only one description is lost, all the immediate neighbors of lost coefficients are available at the advantage of optimal partitioning. Hence, more accurate local texture information can be retrieved from the local window.

**Fig. 2.** Definition of local window $L_x$ and the set of neighbors $N_x$ under the 4-description case described as in Fig. 1. The white points are coefficients of the successfully received descriptions, the gray points represent the corrupted coefficients from the lost description. The center dark point is the one to be estimated currently.

Intuitively the local neighbors can be divided into two classes: edge samples (close to edges) and non-edge samples (away from edges, i.e. samples in smooth areas). For the non-edge samples, the matrix $C^T C$ is often singular or close to singular, thus the solution of $\bar{w}$ is not unique but lies in the hyperplane $\sum_{k=1}^N w_k = 1$, which is approximately isotropic for any direction. In this case the solution of (5) cannot be determined or may cause ill-posed estimation. Hence the simple bilinear interpolation instead of EA method is applied under this situation.

The elements of $L_x$ and $N_x$ include the available neighbors of $X_0$, which are the coefficients from the received descriptions and the corrupted coefficients that have been previously reconstructed. However, some neighbors of $X_0$ might not be available. Hence, it is necessary to check the validity of the elements in $N_x$ and $L_x$. An element is said to be valid if it is either successfully received or previously reconstructed. As shown in Fig. 2, the valid $L_x$ includes all the coefficients inside the $(2T+1) \times (2T+1)$ square window except for $X_0$ and the coefficients that
have not been reconstructed yet. The valid $N_X$ includes the immediate eight coefficients of the successfully received descriptions. Accordingly, the related rows and elements for invalid coefficients will be removed from $C$ and $\bar{l}$.

### 3.2 Evaluation for EA algorithm

Although wavelet transform nearly decorrelates images and can be viewed as an approximate KLT, significant dependencies still exist between wavelet coefficients [14]. Especially strong dependencies in the form of spatial clusters exist between wavelet coefficients inside each subband, which is well known as intra-band or intra-scale correlation [15].

Liu et al. [16] uses mutual information as defined in (6) to measure the dependencies within subbands:

$$\begin{align*}
I(X,Y) &= \int \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \, dx \, dy \\
&= E_{xy}\left[ \log \frac{p(x,y)}{p(x)p(y)} \right].
\end{align*}$$

Let $I(X;N_X)$ denote the mutual information between $X$ and its immediate neighbors $N_X$, thus $N_X$ can provide information to $X$ through a many-to-one mapping function $T = f(X;N_X)$, in the sense that

$$I(X;N_X) = I(X;T).$$

However, $I(X;N_X)$ is only the upper bound of intra-scale dependency one can use to estimate $X$. A function $T$ can only approach $I(X;N_X)$. Hence a well-designed function $T$ will reach a tight bound so that $I(X;N_X) \rightarrow I(X;T)$. In other words, the intra-band correlation can be better adopted by a well-designed mapping function. The maximum likelihood (ML) algorithm is a natural way to optimize this problem. The EA algorithm is actually a Least-Square estimation approach, which is equivalent to the special case of ML with equal prior probability under the assumption that the local texture is stationary [17]. Therefore the EA algorithm is expected to hit a tighter bound of $I(X;T)$ compared to simple bilinear interpolation.

We test the images of Lena and Barbara ($512 \times 512$) in order to evaluate the EA algorithm. The images are decomposed with four levels. The wavelet coefficients are partitioned into 4 descriptions without quantization as in Fig. 1. It is assumed that one of the descriptions is completely lost and all the other descriptions are successfully received.

### Table 1. Mutual information for high-pass bands of Lena.

<table>
<thead>
<tr>
<th>Level</th>
<th>BI</th>
<th>EA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1808</td>
<td>0.3364</td>
</tr>
<tr>
<td>2</td>
<td>0.2652</td>
<td>0.2993</td>
</tr>
<tr>
<td>3</td>
<td>0.2191</td>
<td>0.2354</td>
</tr>
<tr>
<td>4</td>
<td>0.0971</td>
<td>0.1436</td>
</tr>
</tbody>
</table>

### Table 2. Mutual information for high-pass bands of Barbara.

<table>
<thead>
<tr>
<th>Level</th>
<th>BI</th>
<th>EA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1553</td>
<td>0.2702</td>
</tr>
<tr>
<td>2</td>
<td>0.2076</td>
<td>0.4115</td>
</tr>
<tr>
<td>3</td>
<td>0.2478</td>
<td>0.5235</td>
</tr>
<tr>
<td>4</td>
<td>0.2574</td>
<td>0.5255</td>
</tr>
</tbody>
</table>

We calculate the mutual information as in [18] between the original coefficients and the reconstructed ones using bilinear interpolation (BI) and EA algorithm, respectively. Since the LL band is dominated by low frequency coefficients, the EA method does not show apparent improvement over BI. Hence only the high-pass subbands are considered here. Experimental results are listed in Tables 1 and 2, where the decomposition level index is in the ascending order from coarser to finer resolutions. Apparently, the EA algorithm achieves larger mutual information compared to simple linear interpolation. A better reconstruction is expected when more intra-scale dependencies are employed to estimate the lost coefficients.

### 3.3 Fast implementation for EA algorithm

The drawback with the EA algorithm is its huge computational complexity. The main bottleneck lies in the computation of the covariance matrix $\hat{R}_{xx}$ by (3). To facilitate the application of the EA algorithm, we provide a low-complexity approach to efficiently speed up the calculation without performance sacrifice.

Since the local windows are overlapped each other, partial row elements are shared by the matrix $C$ as defined in (4). This information should be considered while executing the EA algorithm instead of repeating the burdensome calculation.

Fig. 3 illustrates the overlapped local windows for a 4-description case. Where we use indices 0 through 4 to denote the current, top, left-top, right-top and left corrupted coefficients, respectively. The local windows of coefficients 1 to 4 share elements with coefficient 0, only the 4 points at the right-bottom corner will introduce unknown information for coefficient 0. Each element inside a local
window is related to one row of $C$. We rewrite (4) as $C = [C_0, ..., C_m, ..., C_L]$, where $C_m$ denotes the $m$th row of $C$. Thus $R_{X} = C^T C = \sum_m C_m^T C_m$. If the matrices, $C_m^T C_m$, related to the elements inside the overlapped area are pre-stored, we can spare a lot of multiplications. As can be seen from Fig. 3, only few elements at the right-bottom corner are involved in the multiplications.

Let $L_{X_0}$ denotes the local windows for the corrupted coefficient $X_0$ to be reconstructed currently, $R_{X_0} = C_L^T C_0$ be the covariance matrix for estimating $X_0$, and $L_{X_j}$, $j = 1, 2, ..., M$ denote the local windows used for neighboring corrupted coefficients that have been reconstructed previously, where $M$ is the number of local windows that overlap with $L_{X_0}$. $C_{i,j}$, $i = 1, 2, ..., L$ denote the $i$th row of $C$ for $L_{X_j}$. The idea of reducing complexity of EA algorithm is to retrieve matrices $R_{i,j} = C_{i,j}^T C_{i,j}$ for $L_{X_j}$ that have been previously calculated and stored in the memory, so that $R_{X_0}$ can be obtained by

$$R_{X_0} = \sum_j \sum_i R_{i,j} + \sum_m C_{0,m}^T C_{0,m}. \quad (8)$$

where $C_{0,m}$ denotes the rows of $C_0$ not available from the memory, which introduces multiplications in computing $R_{X_0}$. The fast implementation is summarized as follows. A vector $M_R$ is used to represent the validity of each row in $C_0$.

**Fast implementation for computing $R_{X_0}$**

- Set $R_{X_0} = [0]_{N \times N}$, $M_R = 0$.
- For $j = 1$ to $M$ do
  - Find the overlapping area between $L_{X_0}$ and $L_{X_j}$. Each element in the overlapping area is mapped to a corresponding entry $m$ of $M_R$.
  - If $M_R(m) = 0$, set $M_R(m) = 1$, retrieve the $R_{i,j}$ related to that element, $R_{X_0} = R_{X_0} + R_{i,j}$.
- For $m = 1$ to $M$ do
  - If $M_R(m) = 0$, $R_{X_0} = R_{X_0} + C_{0,m}^T C_{0,m}$.

**4 Simulation results**

We compare our algorithm against the DP-MDC [11] for robust image transmission. The MD image coder is based on the wavelet codec from [19] for both methods. The wavelet coefficients are partitioned into the descriptions as described in Section 2 for each subband. Each description is independently quantized and encoded using a context-based adaptive arithmetic coder. Four level decompositions are applied to the images. For both methods, the LL band is duplicated in every description.

In DP-MDC, missing coefficients are bilinearly interpolated from available immediate neighbors for the high-pass subbands, whereas the EA algorithm is used in our MDC scheme.

We carried out a set of experiments on the 512 × 512 Lena and Barbara images. In accordance with [11], the bitrates of the conventional single description coding are fixed at 0.21 bpp and 0.4 bpp for Lena and Barbara, respectively. We encode the images into 16 descriptions. In each experiment we fix the number of lost descriptions and completely remove a set of descriptions according to a given description loss ratio.
Fig. 4. Average PSNRs of different MDC algorithms for: (a) Lena, (b) Barbara

Fig. 5. Sample image reconstruction for Lena with descriptions loss ratio of 25%: (a) image reconstructed by DP-MDC, PSNR = 26.81 dB; (b) image reconstructed by EA-MDC, PSNR = 28.00 dB.

Fig. 6. Sample image reconstruction for Barbara with descriptions loss ratio of 25%: (a) image reconstructed by DP-MDC, PSNR = 24.30 dB; (b) image reconstructed by EA-MDC, PSNR = 25.97 dB.
Table 3 includes the average PSNR comparison results of Lena between DP-MDC and our algorithm. The PSNR performance improvement over DP-MDC ranges from 0.6 to 1.2 dB under different description loss ratios. Similar experimental result for Barbara is compared in Table 4. Because of the rich textures in Barbara, higher improvement, ranging from 0.8 to 1.6 dB is achieved by using our EA algorithm.

Table 3. Performance comparison in PSNR (dB) for Lena.

<table>
<thead>
<tr>
<th>Loss ratio (%)</th>
<th>0</th>
<th>6.25</th>
<th>12.5</th>
<th>18.75</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP-MDC</td>
<td>32.36</td>
<td>30.18</td>
<td>28.50</td>
<td>27.47</td>
<td>26.74</td>
</tr>
<tr>
<td>EA-MDC</td>
<td>32.36</td>
<td>30.75</td>
<td>29.39</td>
<td>28.52</td>
<td>27.91</td>
</tr>
</tbody>
</table>

Table 4. Performance comparison in PSNR (dB) for Barbara.

<table>
<thead>
<tr>
<th>Loss ratio (%)</th>
<th>0</th>
<th>6.25</th>
<th>12.5</th>
<th>18.75</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP-MDC</td>
<td>28.78</td>
<td>27.16</td>
<td>25.90</td>
<td>25.03</td>
<td>24.25</td>
</tr>
<tr>
<td>EA-MDC</td>
<td>28.78</td>
<td>27.99</td>
<td>27.18</td>
<td>26.52</td>
<td>25.85</td>
</tr>
</tbody>
</table>

Average PSNR results are also plotted in Fig. 4 that clearly shows remarkable gains of EA-MDC over DP-MDC, especially for images with a lot of high frequency components. It appears that the visual quality is also improved. Figs. 5 and 6 respectively compare the Lena and Barbara images reconstructed by both methods with 25% description lost. Noticeable improvement can be found around the edges as shown in Figs. 7 and 8, for instance, the hair, the eyes and the rim of the hat in Lena, and the table cloth, the right arm and the trousers texture in Barbara.

5 Conclusion

This paper introduces a scheme for improving the state-of-art domain-based MD coder. The proposed EA-MDC successfully retrieves the correlations along all the orientations between the corrupted coefficient and its available neighbors inside a local window by using the covariance method. Hence the weights of linear estimation are well manipulated to emphasize the directions along the important high-frequency components. With the ease of adapting estimation to the dominant edge orientations, EA-MDC can effectively reconstruct the lost descriptions by taking full advantages of the intra-scale dependency within the wavelet subbands. The simulation results show that improvement can be obtained in terms of average PSNR of up to 1.6 dB.

References:


