LOCALLY ADAPTIVE INTRASUBBAND INTERPOLATION OF LOST LOW FREQUENCY COEFFICIENTS IN WAVELET CODED IMAGES

Joost Rombaut, Aleksandra Pižurica and Wilfried Philips

Ghent University, Dept. TELIN, St-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

ABSTRACT

This paper presents a novel passive error concealment method for wavelet coded images. The proposed method is a locally adaptive directional interpolation approach, where the interpolation weights are estimated based on the available local context. For each lost low frequency coefficient, we estimate the optimal interpolation weights based on the errors that would arise by horizontally and vertically interpolating the available neighbors of the lost coefficient. Compared to older methods of similar complexity, the proposed scheme estimates the lost coefficients much better: on average, the PSNR is increased with up to 0.6 dB. The results also indicate improvements over the best available state-of-the-art techniques. The reconstructed images also look better. As our method is fast and of low complexity, it is widely usable.

***Index Terms*—** Image reconstruction, image communication, error concealment, wavelet coding.

1. INTRODUCTION

In lossy packet networks such as the Internet, information often gets lost due to, e.g., network congestion. This data loss is particularly annoying for compressed data, as the loss of a single bit can make the rest of the data stream unusable. These problems are typically solved by protecting the data (e.g., forward error correction) or by resending lost packets. A good overview of the corresponding *Active Error Concealment* techniques, is given in [1]. In certain applications, the packet retransmission is not an option, either because it is too slow (e.g., for real time video) or because there is no return channel (e.g., broadcasting). In these cases, *Passive Error Concealment* [1] is essential.

Passive error concealment exploits the redundancy in the image: lost data is *estimated* from its correctly received *neighboring* data. Therefore, neighboring data must be spread over different packets. This is called dispersive packetization. Examples of packetization techniques are parity based slicing [2] or a packetization based on the partitioning of the $\mathbb{Z}^2$ lattice [3]. We use the packetization strategy of [3], but any dispersive packetization strategy will work with our reconstruction algorithm.

In this paper, we focus on wavelet based image coding. We compress images by dispersively spreading neighboring wavelet coefficients over different packets, and by coding these packets independently from each other with the coder of [3]. These packets are then sent over a lossy packet network. If a packet gets lost during the transmission, the missing data are typically replaced by zeros, which results in annoying black holes in the received image. As the low frequency coefficients contain most of the energy of the signal, their loss has the greatest impact on the quality of the received content. Although the wavelet transform tends to decorrelate the signal, there are substantial spatial dependencies between the coefficients, especially in the low-pass subband. These spatial dependencies can be used for the estimation of a lost coefficient.

Existing passive error concealment techniques for wavelet coded images vary in complexity and speed. A very simple and straightforward technique is bilinear interpolation, where a lost coefficient is replaced by the mean value of its four adjacent neighbors. The efficiency of this method was demonstrated in [3]. In [4], a lost low frequency coefficient is interpolated by fitting a cubic interpolative surface to the known coefficients. Correct edge placement is achieved by adapting the interpolation grid in horizontal and/or vertical direction according to the high frequency content. This method gives better results than the bilinear interpolation, but is also more complex and slower which may be less suited for low-end video clients such as portable devices with only a small processing capacity. Furthermore, the method is only tested on uncompressed images and not in a realistic compression scenario. In [5], the low frequency subband is repaired by a Maximum A Posteriori (MAP) approach, using a Markov random field prior in each subband. The potential functions are adapted locally by estimating the edge characteristics based on the evolution of the coefficients across scales. This technique gives better results than the bilinear interpolation, but it requires much more computational effort. Other notable works about passive error concealment for wavelet coded images include [6] and [7]. However, these methods recover complete image blocks for block based image coders such as JPEG and JPEG2000, while our method reconstructs lost wavelet coefficients which are much more scattered thanks to the dispersive packetization.

It was shown in [3] that bilinear interpolation performs well with dispersive packetization schemes. This method is very fast and gives good results in smooth areas where there is much correlation between neighboring coefficients. The problems arise near edges where the coefficients are rapidly changing. In this paper, we aim at making the interpolation scheme locally adaptive without excessive increase in complexity. We propose a novel interpolation technique that has a complexity similar to that of bilinear interpolation but yielding a reconstruction quality that is even higher than the quality achieved by the adaptive MAP approach of [5].

In the next section we describe the proposed interpolation method. Results and findings are in Section 3, and in Section 4 we draw the conclusions and give some remarks about further work.

2. THE PROPOSED RECONSTRUCTION ALGORITHM

Most wavelet coders use a bi-orthogonal wavelet transform [8]. Our reconstruction method is developed for this type of transformation, but can also be extended to other types of wavelet transforms.

A. Pižurica is a postdoctoral research fellow of the FWO, Flanders, Belgium. **Corresponding author:** Joost Rombaut. Tel: +32 9 264 95 30, Fax: +32 9 264 42 95, E-mail: Joost.Rombaut@telin.ugent.be, Web: http://telin.ugent.be/∼jorombau/.

In the remainder, we use the following notation: LL^n denotes the low pass subband (the scaling coefficients) at the decomposition level n; the wavelet coefficients are organized into the subbands LH^ℓ, HL^ℓ and HH^ℓ, which denote respectively horizontal, vertical and diagonal details at the decomposition level ℓ where $\ell \in \{1, \ldots, n\}$.

The missing high frequency coefficients are difficult to estimate because of the sparse and decorrelated representation. However, errors in the high frequency content are less visible than errors in the low frequency content and hence simple interpolation techniques are effective in this case. We estimate lost LH^n and HL^n coefficients by a one dimensional linear interpolation as in [3, 4], and set lost HH^n coefficients to zero.

In the remainder of this section, we describe our reconstruction method for lost LL^n coefficients. For clarity, we omit the index n, which denotes the scale. The subscripts denote the spatial position, e.g., $LL_{i,j}$ denotes the scaling coefficient at spatial position (i, j).

2.1. Detection of the local correlation

The proposed method adaptively estimates the local correlation in the horizontal and vertical direction, and adapts the interpolation weights accordingly. A lost coefficient $LL_{i,j}$ is estimated by an adaptive weighted averaging in two directions: vertically (using the upper and lower coefficients $LL_{i-1,j}$ and $LL_{i+1,j}$), and horizontally (using the left and right coefficients $LL_{i,j-1}$ and $LL_{i,j+1}$).

We denote the local horizontal interpolation by $\widehat{LL}^H_{i,j} = (LL_{i,j-1} + LL_{i,j+1})/2$, and the vertical interpolation by $\widehat{LL}^V_{i,j} = (LL_{i-1,j} + LL_{i+1,j})/2$. The proposed interpolator is

$$\widehat{LL}_{i,j} = \alpha^V_{i,j} \widehat{LL}^V_{i,j} + \alpha^H_{i,j} \widehat{LL}^H_{i,j}. \quad (1)$$

The weighting factors for the vertical and horizontal direction, $\alpha^V_{i,j}$ and $\alpha^H_{i,j}$, are estimated locally at each spatial position (i, j). We relate these local interpolation weights to a measure of the local correlation in the corresponding directions. We estimate the local correlation from the errors that would arise by horizontally and vertically interpolating the neighbors of the lost coefficient.

We define the horizontal interpolation error measure as

$$E^H_{i,j} = (LL_{i-1,j} - \widehat{LL}^H_{i-1,j})^2 + (LL_{i+1,j} - \widehat{LL}^H_{i+1,j})^2, \quad (2)$$

and the vertical interpolation error measure as

$$E^V_{i,j} = (LL_{i,j-1} - \widehat{LL}^V_{i,j-1})^2 + (LL_{i,j+1} - \widehat{LL}^V_{i,j+1})^2. \quad (3)$$

In case where $E^H_{i,j} > E^V_{i,j}$, it is likely that vertical correlation is dominant at position (i, j); the opposite is true for $E^H_{i,j} < E^V_{i,j}$.

2.2. Defining the optimal interpolation weights

In this section, we experimentally determine the optimal (in the mean squared error sense) relationship between the interpolation weights and the interpolation error measures $E^H_{i,j}$ and $E^V_{i,j}$. Our experiments showed that $\alpha^H_{i,j}$ and $\alpha^V_{i,j}$ do not depend on the exact values of $E^H_{i,j}$ and $E^V_{i,j}$, but only on the ratio $E^H_{i,j}/E^V_{i,j}$. We define the error ratio

$$R_{i,j} = \frac{E^H_{i,j}}{E^V_{i,j}} \quad (4)$$

as a measure of the local correlation direction. For all the low frequency coefficients of 146 different images, we calculated this error ratio $R_{i,j}$. We then quantized the obtained range of error ratio values into 20 intervals. Next, we grouped the low frequency coefficients for which the error ratio was within the same interval. For each of these groups of coefficients, we jointly optimized the interpolation weights $\alpha^H_{i,j}$ and $\alpha^V_{i,j}$ with the least squares method. The resulting optimal values of $\alpha^H_{i,j}$ and $\alpha^V_{i,j}$ in function of $R_{i,j}$ are given in Fig. 1.

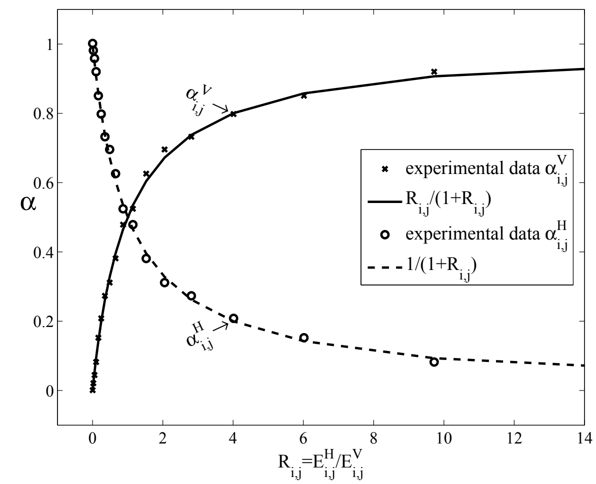

Fig. 1. Optimal $\alpha^H_{i,j}$ and $\alpha^V_{i,j}$ values and our approximation functions.

Based on the experimental data from Fig. 1, we propose the following model for $\alpha^H_{i,j}$ and $\alpha^V_{i,j}$:

$$\alpha^H_{i,j} = \frac{1}{1 + R_{i,j}}, \quad (5)$$

$$\alpha^V_{i,j} = \frac{R_{i,j}}{1 + R_{i,j}}. \quad (6)$$

This model fits the data very well and it yields an accurate estimation of the optimal interpolation weights from the error ratio $R_{i,j}$. By substituting Eq. (4) in Eq. (5) and (6), we obtain respectively:

$$\alpha^H_{i,j} = \frac{E^V_{i,j}}{E^H_{i,j} + E^V_{i,j}}, \quad (7)$$

$$\alpha^V_{i,j} = \frac{E^H_{i,j}}{E^H_{i,j} + E^V_{i,j}}. \quad (8)$$

Note that, if $E^H_{i,j} \gg E^V_{i,j}$, then $\alpha^V_{i,j} \approx 1$ and $\alpha^H_{i,j} \approx 0$, such that the lost coefficient $LL_{i,j}$ is reconstructed by vertical interpolation. Vice versa, if $E^H_{i,j} \ll E^V_{i,j}$, then $\alpha^V_{i,j} \approx 0$ and $\alpha^H_{i,j} \approx 1$, such that the lost coefficient $LL_{i,j}$ is reconstructed by horizontal interpolation.

If $E^H_{i,j} = E^V_{i,j}$, there is no preferential interpolation direction and equations (7) and (8) yield in this case $\alpha^H_{i,j} = 1/2$ and $\alpha^V_{i,j} = 1/2$, which is equivalent to bilinear interpolation. Note that, independent of $E^H_{i,j}$ and $E^V_{i,j}$, $\alpha^H_{i,j} + \alpha^V_{i,j} = 1$ always holds[1].

2.3. Iterative extension

In the proposed method, a lost coefficient is interpolated from its four most adjacent neighbors. Therefore, if a neighbor of a lost coefficient is also missing, this has an impact on its reconstruction.

[1] If $E^H_{i,j} = E^V_{i,j} = 0$, then we choose $\alpha^H_{i,j} = \alpha^V_{i,j} = 1/2$.

Table 1. Average PSNR [dB] for: Bilinear Interpolation (BI), the adaptive MAP approach of [5], and the proposed method, for n lost packets.

average PSNR for *Lena*			
n	BI	MAP [5]	Proposed
0	32.18	32.18	32.18
1	28.73	29.05	29.13
2	26.72	27.16	27.23
3	25.24	25.74	25.80
4	24.04	24.59	24.61

average PSNR for *Peppers*			
n	BI	MAP [5]	Proposed
0	31.66	31.66	31.66
1	27.89	28.05	28.33
2	25.75	25.94	26.27
3	24.20	24.40	24.72
4	22.93	23.15	23.44

average PSNR for *Couple*			
n	BI	MAP [5]	Proposed
0	33.33	33.33	33.33
1	29.91	29.98	30.12
2	27.94	28.01	28.20
3	26.52	26.57	26.80
4	25.34	25.41	25.68

Although the loss of adjacent coefficients is avoided as much as possible by the dispersive packetization [3], it is still occasionally possible to lose adjacent coefficients, especially for high packet loss rates.

In the proposed method, the interpolation weights are also estimated from the surrounding coefficients. Loss of any of these adjacent coefficients decreases the reliability of the estimated interpolation weights. A solution for this problem is iteratively recalculating the lost coefficients. Of course, this quality improvement comes at the expense of a higher computational cost, as more iterations need to be performed.

3. RESULTS

We tested the proposed interpolation method in an experiment similar to the experiment of [5]. For each of our three test images (*Lena*, *Peppers*, and *Couple*), we spread the coefficients over 16 packets using the dispersive packetization scheme of [3]. Each packet was then coded independently of the other packets by using the coder of [3]. This makes each packet independently decodable which is important in case of packet loss. The *Lena* (512×512) and *Peppers* (512×512) images were encoded at respectively 0.208 and 0.207 bits per pixel (bpp) with four levels of wavelet decomposition. The *Couple* (256×256) image was encoded at 0.840 bpp, also with four levels of wavelet decomposition. For each image this gives an average packet size of 430 Bytes which is suitable for Internet transmission without fragmentation. In all tests we used the Daubechies 9/7 bi-orthogonal wavelet filtering.

We then simulated the transmission of each image (i.e., the 16 packets of compressed coefficients) over a lossy packet network by simulating the loss of *every combination* of n packets for n=1,...,4. Each damaged image was repaired with the bilinear interpolation [3], with the adaptive MAP approach of [5], and with the proposed locally adaptive method. For each n, we calculated the average PSNR of the reconstructed images for each reconstruction method. The results of this experiment are given in Table 1.

If no packets are lost, the PSNR of the reconstructed image is equal to the PSNR of the originally broadcasted compressed image. For the *Lena* and the *Peppers* images which are compressed at respectively 0.208 and 0.207 bpp, the PSNR of the compressed image is respectively 32.18 dB and 31.66 dB. The PSNR of the compressed *Couple* image (0.840 bpp) is 33.33 dB.

From the results in Table 1, we notice that our proposed method outperforms the bilinear interpolation with 0.2 up to 0.4 dB for low packet loss rates and with 0.3 up to 0.6 dB for high packet loss rates. For *Lena*, the average PSNR of the images reconstructed by the proposed method is similar to the average PSNR of the images reconstructed by the adaptive MAP approach of [5]. For the *Peppers* and *Couple* images, our proposed method outperforms the adaptive MAP approach of [5] with 0.1 up to 0.3 dB for low packet loss rates and with 0.3 dB for high packet loss rates.

It is interesting to note that although our proposed method yields a similar or a higher PSNR than the adaptive MAP approach of [5], it is of a much lower complexity. As explained in [5], the adaptive MAP error concealment method requires 1540 additions and 1456 multiplications for each lost coefficient. Our proposed method requires only 14 additions and 13 multiplications for each lost coefficient, which is a reduction of a factor 100 compared to the method of [5]. For comparison, the bilinear interpolation requires 3 additions and 1 multiplication for each lost coefficient.

Table 2. Average PSNR [dB] for the adaptive MAP approach of [5] and the proposed method for the *Peppers*-image, for n lost packets. The proposed method is iterated twice.

n	MAP [5]	Proposed	Gain
0	31.66	31.66	
1	28.05	28.33	0.28
2	25.94	26.31	0.37
3	24.40	24.81	0.41
4	23.15	23.58	0.43

Furthermore, it is important to note that while the proposed method outperforms the adaptive MAP approach of [5], we only use a simple one dimensional filtering for the *high* frequency coefficients. In [5], the *high* frequency coefficients are reconstructed with a more complex adaptive MAP method.

In the previous experiment, each lost coefficient was reconstructed only once, such that the loss of adjacent coefficients decreased the reconstruction quality. Table 2 shows the results for the *Peppers*-image where our reconstruction method has been iterated twice. This technique is also used in the adaptive MAP approach of [5], where the method is always iterated twice to cope with the loss of adjacent coefficients. For small packet loss rates, these iterations have little or no impact on the reconstruction quality of the proposed method. This is due to the dispersive packetization: lost coefficients are as far apart from each other as possible, so that there is little or no interference between their reconstruction. For higher packet loss rates ($n > 1$), the iterative method brings a gain in PSNR. For $n = 4$, there is an increase of 0.14 dB compared to the non-iterative approach. This also means that for $n = 4$, our iterative method outperforms the adaptive MAP approach of [5] with 0.43 dB for the $Peppers$ image. Note that two iterations of our proposed method need exactly twice as much additions and multiplications as our non-iterative approach, which is still negligible compared to the number of operations of the adaptive MAP approach of [5].

We also illustrate the visual results on two images. Fig. 2 (a) is the *Lena*-image compressed at 0.208 bpp. Fig. 2 (b) is the *Lena*-image after the loss of packet 12 (i.e., 6.25% of the coefficients lost). Fig. 2 (c), (d) and (e) are the images after reconstruction with respectively bilinear interpolation, the adaptive MAP approach of [5] and the proposed reconstruction method. As indicated by the PSNR, the quality improvement is obvious.

In Fig. 3 we illustrate our iterative approach (Table 2) for high packet loss rates. Fig. 3 (a) is the *Peppers*-image compressed at

Fig. 2. **(a)** *Lena* compressed at 0.208 bpp (PSNR = 32.18 dB). **(b)** *Lena* after loss of packet 12. **(c)** Bilinear interpolation (PSNR = 28.29 dB). **(d)** MAP approach of [5] (PSNR = 28.54 dB). **(e)** The proposed method (PSNR = 28.77 dB).

0.207 bpp. Fig. 3 (b) is the *Peppers* image with 4 lost packets (25% of the coefficients). Without reconstruction, it is difficult to see its content. Fig. 3 (c) is the reconstruction with the adaptive MAP approach of [5]. In this example with a lot of adjacent low frequency coefficients lost, we repaired the image with the iterative version of our reconstruction method with 2 iterations. The result is shown in Fig. 3 (d). We obtained an increase in PSNR of 0.42 dB.

4. CONCLUSION

In this paper, we presented a novel locally adaptive interpolation method for lost low frequency wavelet coefficients in image and video communication. Our method estimates the optimal interpolation weights from neighboring coefficients using novel error measures for horizontal and vertical interpolation, calculated from the neighbors of the lost coefficient.

The experiments on different images have demonstrated that the proposed method outperforms the bilinear interpolation for up to 0.4 dB in case of 6.25% of the coefficient lost, and up to 0.6 dB in case of 25% of the coefficient lost. Compared to a more complex locally adaptive MAP reconstruction method, our method performs up to 0.3 dB better. An iterative version of the proposed method increases its PSNR gain even more. While yielding a very good reconstruction quality, our method is of very low complexity.

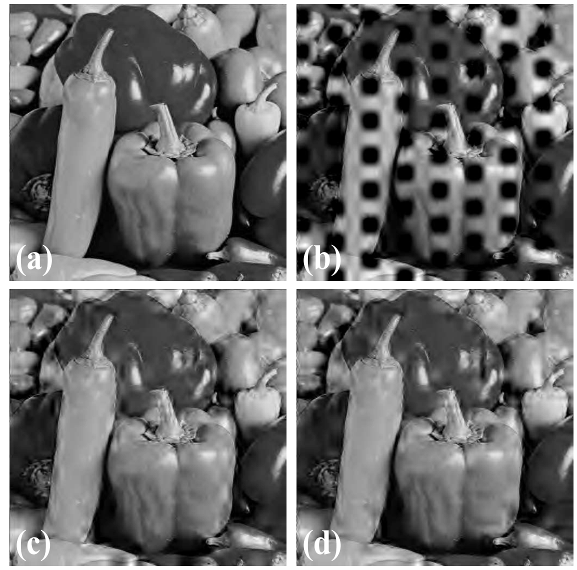

Fig. 3. **(a)** *Peppers* compressed at 0.207 bpp (PSNR = 31.66 dB). **(b)** *Peppers* after loss of packets 5, 6, 9 and 10. **(c)** Iterative reconstruction with the adaptive MAP approach of [5] (PSNR = 22.85 dB). **(d)** Iterative reconstruction with our proposed reconstruction method (PSNR = 23.27 dB).

5. REFERENCES

[1] Y. Wang and Q. Zhu, "Error control and concealment for video communication: A review," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.

[2] M. Stoufs, J. Barbarien, F. Verdicchio, A. Munteanu, J. Cornelis, and P. Schelkens, "Error protection and concealment of motion vectors in MCTF-based video coding," in *Proc. Optics East*, Philadelphia, PA, USA, Oct 2004, vol. 5607, pp. 71–80.

[3] I.V. Bajić and J.W. Woods, "Domain-based multiple description coding of images and video," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1211–1225, October 2003.

[4] S.S. Hemami and R.M. Gray, "Subband-coded image reconstruction for lossy packet networks," *IEEE Transactions on Image Processing*, vol. 6, no. 4, pp. 523–539, April 1997.

[5] I.V. Bajić, "Adaptive MAP error concealment for dispersively packetized wavelet-coded images," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1226–1235, May 2006.

[6] S.D. Rane, J. Remus, and G. Sapiro, "Wavelet-domain reconstruction of lost blocks in wireless image transmission and packet-switched networks," in *Proc. ICIP2002*, Rochester, NY, USA, September 2002, vol. 1, pp. 309–312.

[7] O.G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising—part II: Adaptive algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 555–571, March 2006.

[8] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, December 1993.