Passive error concealment for wavelet-coded I-frames with an inhomogeneous Gauss-Markov Random Field Model

Joost Rombaut, Aleksandra Pižurica, *Member, IEEE,* and Wilfried Philips, *Member, IEEE,*

***Abstract*— In video communication over lossy packet networks (e.g., the Internet) packet loss errors can severely damage the transmitted video. The damaged video can largely be repaired with passive error concealment, where neighboring information is used to estimate missing information. We address the problem of passive error concealment for wavelet coded data with dispersive packetization. The reported techniques of this kind have many problems and usually fail in the reconstruction of high frequency content. This paper presents a novel locally adaptive error concealment method for subband coded I-frames based on an inhomogeneous Gaussian Markov Random Field model. We estimate the parameters of this model from a local context of each lost coefficient, and we interpolate the lost coefficients accordingly. The results demonstrate a significant improvement over the reported related methods both in terms of objective performance measures and visually. The biggest improvement of the proposed method compared to the state-of-the-art in the field is the correct reconstruction of high frequency information such as textures and edges.**

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

I. Introduction

In lossy packet networks such as the Internet, packets can be dropped, e.g., in case of network congestion. Even over highly protected fiber-based ATM networks, error bursts or processing delays can cause loss of data packets. This data loss is particularly annoying for compressed data, as the loss of a single bit can in some cases make the rest of the data stream unusable. These problems are typically solved by protecting the data (e.g., forward error correction, error control coding) or by implementing a protocol for resending lost packets. A good overview of the corresponding *Active Error Concealment* techniques, is given in [1]. However, even with data protection, if the error rate is too high (exceeding the error correction capacity), the receiver may not be able to decode the correctly received information. Also, in certain applications, 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* (also known as *post processing error concealment*) is essential. In video applications, passive error concealment [1] exploits the remaining redundancy within and between frames: missing pixel values (or coefficients, or blocks) are *estimated* from the correctly received *neighboring* pixels (or coefficients, or blocks). An implicit assumption by passive error concealment schemes is that the probability of a simultaneous loss of two or more neighboring pixels (or coefficients, or blocks) is low. This maximal availability of correctly received neighbors is ensured by spreading spatially adjacent pixels (or coefficients, or blocks) over different packets. This is called dispersive packetization and is also known as multiple description coding [2]. Our method requires a dispersive packetization scheme, but it is not restricted to any particular scheme.

We focus on error concealment for wavelet-coded intra coded frames (I-frames). In particular, these I-frames are the reference frames in traditional Motion Compensated Prediction (MCP) coding, or the lowest frequency band in Motion Compensated Temporal Filtering (MCTF) coding which is mainly used in wavelet-based scalable video coding. In this work, we do not address the concealment for lost motion vectors nor for missing samples from the predictive frames (P-frames) in the MCP approach or for missing samples from the high temporal frequency bands in the MCTF approach. The lost motion vectors can be estimated using motion vector reconstruction methods from Discrete Cosine Transform (DCT) based video coding, e.g., vector median filters [3], [4]. A missing sample from a P-frame can be estimated as the median of its available spatial neighbors as shown in [2]. The optimization of this concealment scheme for P-frames will be the topic of further research. As the reconstruction of wavelet-coded I-frames is in essence equal to the reconstruction of wavelet-coded images, we will use the terms *image* and *I-frame* interchangeably in the rest of this paper, and we will also compare our method to the existing passive error concealment methods for wavelet-coded images. To evaluate our error concealment scheme for wavelet-coded I-frames, 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 [2]. Then we simulate the loss of coded packets. In reality, if a packet gets lost during the transmission, the missing data are typically replaced by zeros, which results in annoying black

Manuscript received ...; revised .... Current version published .... The associate editor coordinating the review of this manuscript and approving it for publication was ....

J. Rombaut and W. Philips are with the Department for Telecommunications and Information Processing (UGent-TELIN-IPI-IBBT), Ghent University, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium (e-mail: joost.rombaut@telin.ugent.be).

A. Pižurica is with the Department for Telecommunications and Information Processing (UGent-TELIN-IPI-IBBT), Ghent University, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium, and also with the FWO, Flanders, Belgium.

Digital Object Identifier ...

holes in the incompletely received images. In our simulations, we recover the underlying image by an adaptive interpolation of the lost coefficients.

A. Related Work

Compared to more common block-based approaches, relatively few passive concealment methods were reported for wavelet coded images and video [2], [5]–[7]. For the reconstruction of lost low frequency coefficients (the scaling coefficients), the existing methods are mainly traditional error concealment algorithms from the image domain which are slightly adapted to be able to work in the wavelet domain. Bilinear interpolation (the lost coefficient is replaced by the average of its four adjacent neighbors) has proved very efficient despite its simplicity [2]. The results of this simple interpolation are quite good in smooth areas but artifacts arise near edges and other discontinuities. In [5], a bicubic interpolation method was proposed. Correct edge placement is achieved by adapting the interpolation grid in horizontal and/or vertical direction according to the high frequency content. The method in [5] was only tested on uncompressed images and not in a realistic compression scenario. In [6], a Maximum A Posteriori (MAP) approach was applied using a Markov random field prior. This technique gives slightly better results than bilinear interpolation, but it requires much more computational effort. In our previous work [7], the lost low frequency coefficients are reconstructed with a locally adaptive interpolation scheme. For each lost coefficients, the interpolation weights are estimated based on the interpolation errors that arise from interpolating nearest correctly received neighbors in both horizontal and vertical directions. While this method of [7] interpolates the lost data from the immediate horizontal and vertical neighbors like the proposed method, the concept behind determining the weights is totally different. In our earlier method [7], we derive a closed form expression under the side constraint that the sum of the horizontal and vertical interpolation weights equals 1. In the proposed method, we employ the GMRF prior model and all the underlying expressions are different. For the reconstruction of lost high frequency coefficients (the wavelet coefficients), even fewer methods were proposed. In [2], [5], [7], lost coefficients from the LH and HL-subbands are reconstructed by a one dimensional linear interpolation in the direction where there has been only low-pass filtering in the wavelet transform. The lost coefficients in the HH-subbands are set to zero as errors in these subbands are less visible. In [6], lost high frequency coefficients are reconstructed with a simplified version of the MAP approach that is used for the lost low frequency coefficients.

Other recent approaches are the block-based techniques of [8] and [9]. The approach of [8] is very different from ours because it recovers complete blocks of wavelet coefficients (for block-based wavelet coders such as JPEG2000) simultaneously, while we consider the case of (more or less) isolated lost coefficients because of the dispersive packetization of coefficients. Traditional error concealment algorithms such as [9] also recover complete pixel blocks for block-based image coders such as JPEG and JPEG2000, and are also not applicable in a non-block-based coding approach.

B. The Proposed GMRF Approach

In this paper, we propose a reconstruction scheme for both low frequency and high frequency coefficients based on a Gauss-Markov Random Field (GMRF) model. Markov Random Field (MRF) models [10] have already proved to be very effective for image processing, as these models are ideal to model global (joint) image statistics in terms of local spatial interactions. In a typical gray-level image the intensity, i.e., the brightness level of each pixel, is very similar to the intensities of surrounding pixels. Therefore, the concept of a *neighborhood*, which is usually a local surrounding, is central for MRFs. MRF models have been successfully applied in different areas of image processing, such as: texture modeling and classification [11]–[13], image restoration [14], [15], segmentation of noisy and textured images [16]–[20], deinterlacing [21], and reconstruction of lost pixel blocks in DCT image coding [22]. Although the wavelet transform tends to decorrelate the image, neighboring coefficients are still statistically dependent to some degree [23]. Significant wavelet coefficients appear in spatial clusters. The scaling coefficients (a low pass filtered, subsampled version of the original image) closely resembles the original image. Therefore, an MRF model is also very suitable for wavelet-based image processing such as wavelet-based image denoising [24], [25], and reconstruction of lost coefficients in wavelet based image coding [6]. In image processing, commonly used MRF models are automodels [26]. An important class, the *auto-normal* models (also known as *Gaussian Markov Random Field models*) [12], [27], [28] are often used, e.g., for segmentation [16], [17], [19] and texture analysis [12]. In most of these applications, the MRF model is homogeneous which means that the clique potential does not depend on the position in the image. We develop an *inhomogeneous* GMRF model with local parameter estimation and apply it to modeling statistical dependencies among the wavelet coefficients within a given subband. Based on this model, we develop a passive error concealment method.

The clique potential (i.e., the function that controls the local statistical properties) of a GMRF model is of quadratic form, which tends to smoothen the processed image. As noted in [22], this type of potential function may result in blurred estimates when reconstructing lost pixel blocks. To address this issue, the authors of [22] use (a modified version of) the *Huber* function [29], which does not penalize the edges as heavily as a quadratic function, and thus preserves edges and other structural information much better. The drawback of using the piecewise Huber function is that the optimal solution, i.e., the minimization of the energy function, can only be reached by iteratively recalculating the lost pixels. In [22], this is done with the iterative conditional modes algorithm, resulting in slower processing. In a related MRF-based approach for wavelet coded images [6], the clique potential is also a Huber function. There, the minimization of the energy function is based on the golden search procedure, which also needs a lot of operations for each lost coefficient. In [6], the clique potentials for each lost coefficient depend on the corresponding high-frequency coefficients. While this approach allows a nice adaption of the clique potentials to local edge structures, it also

poses some problems. Firstly, it hinders parallel processing of the subbands (because of possibly missing high-frequency coefficients need to be reconstructed first). Secondly, in some important coding strategies like zero-tree based coders [30], the high frequency coefficients necessary for the reconstruction of lost low frequency coefficients according to this scheme are also lost since they are encoded in the same zero-tree.

The main contribution of this paper is a new error concealment method for wavelet coefficients based on an original inhomogeneous GMRF model. The GMRF model is attractive because of its computational simplicity. Its quadratic clique potentials yield a fast calculation of the lost data without iterative recalculation as is the case for more complex piecewise clique potentials such as in the method of [6]. However, the homogeneous GMRF model tends to smooth out the discontinuities. Therefore we introduce an original inhomogeneous model, which keeps the simplicity of the ordinary GMRF but enables an efficient local adaptation. This is done through local estimation of the parameters of the GMRF model. Our approach offers several improvements over the competing recent schemes like the MRF-based scheme [6] and our previous method [7]. Firstly, the reconstruction quality in textured areas is much better (up to 2.5 dB in PSNR and with obvious visual differences). Secondly, the proposed method is much faster (e.g., the speed improvement over [6] is a factor of 10), and suitable for a parallel implementation which can yield a further speed-up. This is interesting especially when we consider the reconstruction of I-frames in video communication. In this case, the communication delay should be kept as small as possible. The proposed method is also less restrictive in terms of the coders than other related methods like [6]: contrasting to these approaches, our method can be used with, e.g., tree-like coders as in [30], wavelet layered PCM coders as in [2], the quadtree-limited coder as in [31], ...

In the next section, we introduce the Gauss-Markov Random Field (GMRF) model. In Section III, we explain the proposed interpolation method. Results and findings are in Section IV, and conclusions in Section V.

II. The Gauss-Markov Random Field (GMRF) model

In this section, we briefly introduce the GMRF model. For more details, we refer to [12], [19]. We use the following notation:

s, t	$s = (s_1, s_2)$, $t = (t_1, t_2)$, the coordinates of grid points on a two dimensional (2D) lattice.
Ω	$\{s : 0 \leq s_1 \leq M-1, 0 \leq s_2 \leq N-1\}$, a 2D lattice of $M \times N$ grid points.
$\mathbf{X}$	Random field on Ω, represented as a vector by row-wise scan ordering: $\mathbf{X} = \{X_0, \ldots, X_{M\times N-1}\}$.
X_s	Random variable at site s.
$\mathbf{x}$	A configuration of $\mathbf{X}$, corresponding to a given realization of the image field. The notation $\mathbf{X} = \mathbf{x}$ will be used to abbreviate the joint event $(X_0 = x_0, \ldots, X_{M\times N-1} = x_{M\times N-1})$.
η	Set of coordinate *offsets* defining a neighborhood.
$\eta^\star$	Set of *anti-causal* offsets defining a neighborhood.
η_s	Neighborhood of s; $\eta_s = \{s + r : r \in \eta\} = \{s \pm r : r \in \eta\}$.

$\mathbf{X}$ exhibits the Markov property which means that the conditional probability of the pixel value X_s given all other pixel values in the image reduces to a function of neighboring pixel values only:

$$P(X_s = x_s | \mathbf{X}_{S\setminus s} = \mathbf{x}_{S\setminus s}) = P(X_s = x_s | \mathbf{X}_{\eta_s} = \mathbf{x}_{\eta_s}). \quad (1)$$

The neighborhood η_s usually consists of the spatially adjacent neighbors up to a certain distance, i.e., the order of the neighborhood. For the first-order case, the set of coordinate offsets defining the neighborhood is $\eta = \{(1,0), (-1,0), (0,1), (0,-1)\}$, and the neighboring elements are $\eta_s = \{s + r : r \in \eta\}$. A visualization of the nth order neighborhoods is given in Fig. 1, where each nth order neighbor of the element s is marked with the label n.

5	4	3	4	5
4	2	1	2	4
3	1	s	1	3
4	2	1	2	4
5	4	3	4	5

Fig. 1. The elements of the nth order neighborhood set are marked with n [13].

The joint probability of an MRF can be written as a Gibbs distribution for the neighborhood η [10]:

$$P(\mathbf{X} = \mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{H(\mathbf{x})}{T}\right) \quad (2)$$

where Z is a partition constant, T is called the temperature, and $H(\mathbf{x})$ is the energy function. In case $\mathbf{X}$ is an MRF, the energy function is a sum of *clique* potentials:

$$H(\mathbf{x}) = \sum_{C\in\mathcal{C}} V_C(\mathbf{x}), \quad (3)$$

with $\mathcal{C}$ the set of all possible cliques. A clique is a set of pixels that are all neighbors of one another. Typically in image processing, cliques consisting of two sites (pair-site cliques) are used because of their simple form and the relatively low computational cost. Single-site cliques may be additionally used if the prior probabilities of different labels are known. When using higher order neighborhoods, the number of clique types grows rapidly with the neighborhood order, which can significantly increase the computational complexity. Some researchers use higher-order neighborhoods with pair-site cliques only (for a motivation see [14]). In the proposed method, we use the first-order neighborhood consisting of single-site and pair-site cliques. For estimating the parameters of our inhomogeneous model, we will employ larger neighborhoods, but these should not be confused with the first-order neighborhood from which we interpolate the missing data. Commonly used MRF models in image processing are *auto-models* [10], where the energy function is of the form:

$$H(\mathbf{x}) = \sum_{s\in S} x_s g_s(x_s) + \sum_{s\in S}\sum_{r\in\eta} \theta_r^s x_s x_{s+r}, \quad (4)$$

in which $g_s(x_s)$ are arbitrary functions and θ_r^s are constants reflecting the pair-site interactions. When the distribution is *homogeneous*, we have $\theta_r^s = \theta_r$ regardless of the spatial position s. This is typically the case for texture processing.

An important class of auto-models are *auto-normal* models, also called Gaussian Markov Random Field (GMRF) models

[10], [12], [19]. These models are defined for continuous MRFs. Their energy function (4) is of the form:

$$H(\mathbf{x}) = \sum_{s\in S} \frac{x_s^2}{2\sigma^2} + \sum_{s\in S}\sum_{r\in\eta} \frac{\theta_r^s x_s x_{s+r}}{2\sigma^2}. \quad (5)$$

The conditional probability of a GMRF is given by:

$$p(X_s = x_s|\mathbf{X}_{\eta_s} = \mathbf{x}_{\eta_s}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{\left[x_s - \sum_{r\in\eta} \theta_r^s x_{s+r}\right]^2}{2\sigma^2} \right\} \quad (6)$$

with $\theta_r^s = \theta_{-r}^s, \forall r \in \eta$, as been proved in [12]. The two dimensional, homogeneous, bilateral, GMRF model with symmetrical neighborhood is defined as [12], [19]:

$$X_s = \sum_{r\in\eta} \theta_r X_{s+r} + N_s, \quad (7)$$

where N_s is a sample of zero mean Gaussian noise, with autocorrelation given by [10, p. 18–19], [12, p. 85], [19, p. 252]:

$$E[N_s N_{s+r}] = \begin{cases} \sigma^2 & \text{if } r = (0,0), \\ -\theta_r\sigma^2 & \text{if } r \in \eta, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

From (8), it follows that $\forall r \in \eta : \theta_r = \theta_{-r}$. As such, we can rewrite (7) as

$$X_s = \sum_{r\in\eta^\star} \theta_r (X_{s+r} + X_{s-r}) + N_s, \quad (9)$$

where $\eta^\star$ is defined as the set of anti-causal offsets. For the first-order case, the set of anti-causal offsets defining the neighborhood is $\eta^\star = \{(1,0),(0,1)\}$, and the coordinates of the neighboring elements are $\eta_s = \{s \pm r : r \in \eta^\star\}$.

The expected value of a coefficient x_s in an homogeneous GMRF model ($\forall s : \theta_r^s = \theta_r$) is the value with highest probability (see [12, p. 85] and [19, p. 255]). To find this value, the derivative of (6) with respect to x_s is set to zero. The expected value is then:

$$\widehat{x}_s = \sum_{r\in\eta} \theta_r x_{s+r}. \quad (10)$$

and since $\forall r \in \eta : \theta_r = \theta_{-r}$:

$$\widehat{x}_s = \sum_{r\in\eta^\star} \theta_r (x_{s+r} + x_{s-r}). \quad (11)$$

The optimal (in the Mean Squared Error sense) interpolation weights $\widehat{\boldsymbol{\theta}} = [\widehat{\theta}_r, r \in \eta^\star]$ are estimated by the *linear least squares* method [12], [19]:

$$\widehat{\boldsymbol{\theta}} = [\widehat{\theta}_r, r \in \eta^\star] = \left[\sum_{t\in\Omega'} \mathbf{q}_t^T \mathbf{q}_t\right]^{-1} \left[\sum_{t\in\Omega'} \mathbf{q}_t^T x_t\right], \quad (12)$$

where

$$\mathbf{q}_t = [x_{t+r} + x_{t-r}, r \in \eta^\star], \quad (13)$$

and Ω' is the pool of samples from which the weighting factors $\widehat{\theta}$ are calculated.

III. The proposed reconstruction algorithm

The new method proposed in this paper is meant for the reconstruction of wavelet coded I-frames in broadcasted video streams. In this paper, we will use a critically sampled bi-orthogonal wavelet transform, which is used in most of the wavelet coders [30], but the proposed method can be applied to other types of wavelet transforms [32]–[35], and other multiresolution codecs (such as those based on bandelets, contourlets, etc.). We use the following notation: LL^J denotes the low-pass subband (the scaling coefficients) at the decomposition level J; the wavelet coefficients are organized into the subbands LH^j, HL^j and HH^j, which denote respectively horizontal, vertical and diagonal details at the decomposition level j where $j \in \{1,\ldots,J\}$).

In this section, we describe our reconstruction method for both the scaling coefficients and for the lost wavelet coefficients. As we reconstruct each subband independently, we will represent a scaling/wavelet coefficient as X_s, thereby — for clarity — we will omit the subband type and the scaling indices J or j. As previously introduced, the subscript s will denote the spatial position.

In our reconstruction algorithm we use a first order Gaussian Markov process for the reconstruction of lost coefficients. As each coefficient depends on its first order neighborhood, $\eta^\star = \{(1,0),(0,1)\}$, a lost coefficient is interpolated partially from its vertical neighbors, partially from its horizontal neighbors. However, in contrast to the homogeneous GMRF model introduced in the previous section where the weighting factors θ_r were calculated over the global image/subband and are the same for all interpolated pixels/coefficients, we now estimate the weighting factors θ_r^s locally for each lost coefficient. In our approach, the clique potentials are different for horizontal and for vertical cliques (θ_r^s depends on r), and as they now depend on the spatial position of the central pixel (θ_r^s depends on s), our model is an *inhomogeneous* GMRF model. The energy function and the conditional probability of an inhomogeneous GMRF model are given in (5) and (6) respectively. The expected value of a lost coefficient according to the proposed GMRF model is:

$$\begin{aligned} \widehat{x}_s &= \sum_{r\in\eta^\star} \theta_r^s (x_{s+r} + x_{s-r}) \\ &= \theta_{(1,0)}^s (x_{(s_1+1,s_2)} + x_{(s_1-1,s_2)}) + \theta_{(0,1)}^s (x_{(s_1,s_2+1)} + x_{(s_1,s_2-1)}) \end{aligned} \quad (14)$$

We estimate the weighting factors as:

$$\widehat{\boldsymbol{\theta}^s} = \begin{bmatrix} \widehat{\theta}_{(1,0)}^s \\ \widehat{\theta}_{(0,1)}^s \end{bmatrix} = \left[\sum_{t\in\Omega_s} \mathbf{q}_t^T \mathbf{q}_t\right]^{-1} \left[\sum_{t\in\Omega_s} \mathbf{q}_t^T x_t\right], \quad (15)$$

where $\mathbf{q}_t$ is defined as in (13):

$$\mathbf{q}_t = [x_{(t_1+1,t_2)} + x_{(t_1-1,t_2)}, x_{(t_1,t_2+1)} + x_{(t_1,t_2-1)}]. \quad (16)$$

In contrast to (12), the interpolation weights $\widehat{\boldsymbol{\theta}^s}$ are now calculated from a pool Ω_s of selected samples, instead of one big pool Ω' containing *all* samples as is the case in the homogeneous GMRF model. The choice of Ω_s is important for the quality of the reconstruction. If Ω_s is too big (too many neighbors are involved in the estimation of $\boldsymbol{\theta}^s$), spatially

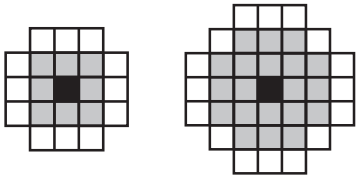
varying statistics will be averaged out which is detrimental in edge regions. On the other hand, if Ω_s is too small, $\widehat{\boldsymbol{\theta}^s}$ cannot capture textural properties. We experimentally optimized the size of Ω_s for the loss of low-frequency and high-frequency coefficients which is given in Section III-A and Section III-B respectively.

In some cases, (15) may not be stable. This is the case if the determinant of $\sum_{t\in\Omega_s} \mathbf{q}_t^T \mathbf{q}_t$ is zero or close to zero. If the equation matrix is ill-conditioned, this may lead to unpredictable (and probably incorrect) results for $\widehat{\boldsymbol{\theta}^s}$. While this instability problem can be solved by using an orthogonal decomposition (like the singular value decomposition) to solve the linear least squares problem, these methods are relatively slow compared to the normal equations method of (15), which is why we have chosen the latter. Furthermore, the probability that (15) is unstable is very low and instability mainly occurs in flat areas. In flat areas, lost coefficients can be easily reconstructed with a bilinear interpolation in case of a low-frequency coefficient, with a horizontal or vertical interpolation in case of an LH or HL coefficient respectively, or by replacing it by zero in case of an HH coefficient. An efficient and simple solution to find these unstable cases is to check if the absolute value of the determinant of $\sum_{t\in\Omega_s} \mathbf{q}_t^T \mathbf{q}_t$ is bigger than a certain threshold. In practice, 10^{-10} is a useful threshold.

A. The reconstruction of lost low frequency coefficients

When the coefficient x_s is lost, it cannot be used to estimate the optimal interpolation weights $\widehat{\boldsymbol{\theta}^s}$. Also, its first order neighbors cannot be part of Ω_s, since they require x_s in the calculation of $\mathbf{q}_t$ in (16). However, as the first order neighbors generally have the highest correlation with the lost coefficient, it would be best to include them in Ω_s. To be able to do this, we calculate a rough estimate of x_s with a bilinear interpolation scheme as we did in [7] from the four nearest coefficients. If it is impossible to calculate the bilinear interpolation, the average of the available first order neighbors is used.

Secondly, the lost coefficient is re-estimated with the proposed GMRF interpolation scheme of (14) and the weighting factors are estimated by (15). For the reconstruction of a lost low frequency coefficient, we experimentally found that the optimal Ω_s (Fig. 2 (a)) consists of the first and the second order neighbors, and of the bilinear interpolation of the lost coefficient x_s itself. Note that if any of the necessary neighboring coefficients used in the calculation of $\widehat{\boldsymbol{\theta}^s}$ are lost, their bilinearly interpolated value (or average of the available first order neighbors) is used. By taking more coefficients (from third or higher order neighborhoods) into account, the reconstruction quality decreases. This is so because there is little correlation between a low frequency coefficient and its nth order neighbors for $n > 2$. In particular, this is true for wavelet decompositions with bigger decomposition depth.



(a) (b)

Fig. 2. Visualization of the optimal Ω_s for the reconstruction of **(a)** low-frequency coefficients, and of **(b)** high-frequency coefficients. The lost coefficient x_s is marked as a black square. The set Ω_s consist of all grey coefficients and the black center coefficient. The coefficients used in the calculation of $\{\boldsymbol{q}_t : t \in \Omega_s\}$ are all the white, grey and black coefficients.

B. The reconstruction of lost high frequency coefficients

Like in the case of lost low frequency coefficients, the reconstruction of lost high frequency coefficients benefits from taking into account first order neighbors in Ω_s. Since x_s, which is then necessary in the calculation of $\mathbf{q}_t$ in (16), is missing, we calculate a initial estimate of it. This initial estimate is:

- for the HL-subbands, a vertical linear interpolation since there is generally much more vertical than horizontal correlation in this subband. This is due to the horizontal high pass filtering in the creation of the HL-subbands.
- similarly for the LH-subbands, a horizontal linear interpolation.
- for the HH-subbands, a replacement of the lost coefficient by zero, since a large majority of the coefficients in these subbands are zero or nearly zero. This is especially true after compression. Furthermore, replacing a large coefficient by zero results in a loss of detail that is visually much less disturbing than the artifacts that arise from falsely creating a large coefficient. We also experimentally examined some other simple initialization approaches, like bilinear interpolation, median value filtering, horizontal and vertical linear interpolation, ... These other approaches brought no improvement as compared to initializing the missing coefficients by zero, so this simple initialization is reasonable.

Note that these initial estimates correspond to the reconstruction schemes used in [2], [5], [7].

Secondly, the lost high frequency coefficient is re-estimated with the proposed GMRF interpolation scheme of (14) and the weighting factors are estimated by (15). For the reconstruction of the high frequency coefficients, we experimentally found that the optimal Ω_s (Fig. 2 (b)) consists of the first, the second, the third and the fourth order neighbors, and of the initial estimate of the lost coefficient x_s itself. Note that if any of the neighboring coefficients is also lost, it is also replaced by its initial estimate. By increasing the number of elements in Ω_s, the reconstruction quality may increase a little bit more, but this is too small compared to the invested processing time.

C. Summary of the proposed algorithm

In a real time communication application with packet loss, the receiver knows which coefficients are lost, since both the sender and the receiver know which packetization is used. In our passive error concealment method, subbands are processed

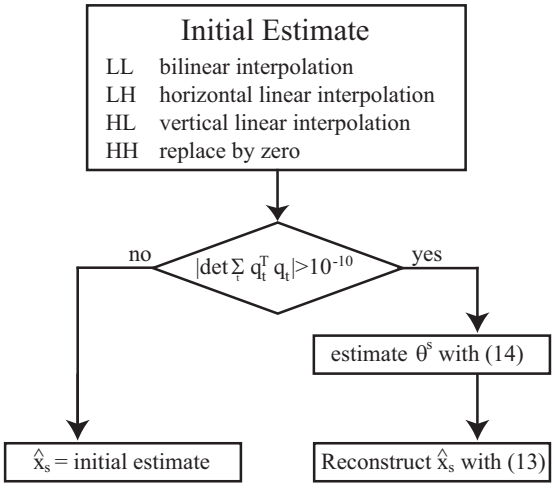


Fig. 3. Summary of the proposed algorithm for the concealment of a lost coefficient X_s.

independently which facilitates parallel processing. We can summarize our algorithm for the reconstruction of a lost coefficient x_s as follows (Fig. 3):

- Step 1: Calculate an initial estimate of x_s. For LL coefficients: bilinear interpolation; for HL coefficients: vertical linear interpolation; for LH coefficients: horizontal linear interpolation; for HH coefficients: replacement by zero.
- Step 2: Check the stability of (15). If it is unstable, stop. Otherwise, go to step 3.
- Step 3: Estimate $\boldsymbol{\theta^s}$ from the neighborhood Ω_s using eq. (15). For an LL coefficient, Ω_s consists of 9 coefficients; for an high frequency coefficient, Ω_s consists of 21 coefficients (see Fig. 2).
- Step 4: Estimate $\widehat{x}_s$ from its first order neighbors using eq. (14).

D. Complexity analysis of the proposed algorithm

We define the number of coefficients in Ω_s as N_{Ω_s}. The number of operations for the calculation of one lost coefficient depends only on N_{Ω_s}. Note that for the calculation of a lost LL-coefficient, $N_{\Omega_s} = 9$. For the calculation of a high frequency coefficient, $N_{\Omega_s} = 21$. The calculation of a lost coefficient consists of three main steps, as explained above. These numbers of operations necessary for each step are explained in detail in the appendix.

- An initial estimate. For the LL-subband this is a bilinear interpolation with 3 additions (A) and 1 multiplication (M). For the LH and HL-subbands, this is a linear interpolation with 1A and 1M. For the HH-subband this is a replacement by zero which requires no operations.
- The calculation of $\widehat{\boldsymbol{\theta^s}}$ which requires $(7N_{\Omega_s} - 2)$A and $(5N_{\Omega_s} + 9)$M.
- The calculation of $\widehat{x}_s$ which requires 3A and 2M.

This requires for the reconstruction of a lost LL-coefficient: 67A and 57M; for the reconstruction of a lost LH or HL-coefficient: 149A and 117M; and for the reconstruction of a lost HH-coefficient: 148A and 116M.

For comparison, the bilinear interpolation scheme requires 3A and 1M for each lost low frequency coefficient, and 1A and 1M for each lost LH or HL-coefficient. As given in [6], the adaptive MAP error concealment method requires 1540A and 1456M for each lost coefficient. The adaptive method of [7] requires 14A and 13M for each lost low frequency coefficient, and 1A and 1M for each lost LH or HL-coefficient. As such, the proposed method is much faster than the method of [6] but slower than the bilinear interpolation as being used in [2] and the adaptive method of [7]. However, the increase in complexity as compared to bilinear interpolation as in [2] is justified by the quality of the reconstructed images, as will be shown in Section IV. In Section IV-C, we propose some techniques to speed up the computation process without sacrificing too much of the reconstruction quality.

IV. Results

A. General case

Typical packet loss rates in the Internet for communication without packet retransmission techniques such as the User Datagram Protocol (UDP), are in the range of 2% to 10% [1]. For communication with packet retransmission techniques such as the Transmission Control Protocol (TCP), each packet will eventually reach the receiver as lost packets are retransmitted by the sender. As explained in the introduction in Sect. I, these packet retransmissions introduce a communication delay which is intolerable in time-sensitive applications such as two-way video communication.

As data packet loss is typically bursty in nature, the instantaneous packet loss rate can be much higher than the average loss rate of 2% to 10%. We tested the proposed interpolation method in a realistic scenario with both low and high packet loss rates. In this experiment, we simulated the transmission of 416 test images/I-frames of 512×512 pixels over a lossy packet network. The wavelet coefficients (calculated with the Daubechies 9/7 bi-orthogonal wavelet with four levels[1] of wavelet decomposition) of each I-frame were each time stored in 16 packets using the dispersive packetization strategy of [2].

After packetization, each packet was coded independently of the other packets at a bit rate of about 0.8 bpp by using the coder of [2]. The independent coding guarantees independent decodability which is important for error recovery. We then simulated the loss of every combination of p packets for $1 \leq p \geq 4$. For $p = 1$, 2, 3 and 4, there are respectively 16, 120, 560 and 1820 possible combinations of lost packets. The lost low frequency coefficients were repaired using four reconstruction methods: bilinear interpolation as in [2], the adaptive MAP approach of [6], the adaptive method of [7] and the proposed GMRF method. For each p, we calculated the average PSNR of all reconstructed I-frame for each reconstruction method. The results of this experiment are given in Table I.

[1] We use four levels of wavelet decomposition as this provides the ideal trade off between compression ratio (which is close to optimal for four or more levels) and reconstruction quality (reconstruction of low frequency coefficients gives optimal results for four or fewer levels).

TABLE I

AVERAGE PSNR [DB] OF *all* THE 416 TEST I-FRAMES FOR: BILINEAR INTERPOLATION (BI), THE ADAPTIVE MAP APPROACH [6], THE ADAPTIVE METHOD OF [7] AND THE PROPOSED GMRF METHOD, FOR *all* POSSIBLE COMBINATIONS OF p LOST PACKETS

p	BI	MAP [6]	Method of [7]	Proposed GMRF
0	40.01	40.01	40.01	40.01
1	33.85	33.90	34.37	34.68
2	31.40	31.38	31.93	32.26
3	29.73	29.67	30.24	30.53
4	28.37	28.33	28.88	29.10

If no packets are lost, the PSNRs of the reconstructed I-frames are equal to the PSNRs of the broadcast compressed I-frames. On average the PSNR of our test I-frames, compressed at 0.8 bpp, is 40.01 dB. The results in Table I show that our proposed method outperforms bilinear interpolation on average by 0.85 dB for low packet loss rates and by 0.75 dB for high packet loss rates. The average improvement over the adaptive MAP approach of [6] is 0.8 dB for both low and high packet loss rates. The average improvement over the method of [7] is 0.3 dB for low packet loss rates and 0.2 dB for high packet loss rates.

For all possible combinations of three lost packets of all 416 test I-frames, we calculated the difference between the PSNRs of the I-frames reconstructed by the proposed method and the PSNRs of the I-frames reconstructed by the bilinear interpolation method. In Fig. 4 (a) we plot the frequency distribution of this PSNR difference. For only 4.25% of the test frames, the PSNR of the reconstruction with the proposed method is lower than the PSNR of the reconstruction with the bilinear interpolation method. In only 1.10% of the cases the PSNR drop is bigger than 0.5 dB. In 95.8% of the cases, the proposed method performs better, or more specifically: in 34% of the cases there is a PSNR gain between 0.0 and 0.5 dB, in 31% of the cases the gain is between 0.5 and 1.0 dB, in 17% of the cases the gain is between 1.0 and 1.5 dB, in 13% of the cases the gain from our method is over 1.5 dB. The frequency distribution of the PSNR difference for one, two, and four lost packets is similar to the one for three packets.

In Fig. 4 we also plot the frequency distribution of the difference between the PSNRs of the I-frames reconstructed by the proposed method and respectively the PSNRs of the I-frames reconstructed by the adaptive MAP approach of [6] (Fig. 4 (b)) and the adaptive method of [7] (Fig. 4 (c)). The results for the adaptive MAP approach of [6] are similar to the results for the bilinear interpolation: for 7.7% of the I-frames, the reconstruction with the proposed method performs worse than the reconstruction with the approach of [6]. In only 2.0% of the cases the PSNR drop is bigger than 0.5 dB. In 92.3% of the cases, the proposed method performs better, or more specifically: in 30% of the cases there is a PSNR gain between 0.0 and 0.5 dB, in 26% of the cases the gain is between 0.5 and 1.0 dB, in 18% of the cases the gain is between 1.0 and 1.5 dB, in 18% of the cases the gain from our method is over 1.5 dB.

The frequency distribution for the PSNR difference between the proposed method and the method of [7] is different from the previous two: here, in 14.0% of the cases, the reconstruction with the proposed method performs worse than the reconstruction with the adaptive method of [7]. However, again in only 1.71% of the cases the PSNR drop is bigger than 0.5 dB. In 86.0% of the cases, the proposed method performs better, or more specifically: in 64% of the cases there is a PSNR gain between 0.0 and 0.5 dB, in 18% of the cases the gain is between 0.5 and 1.0 dB, in 3% of the cases the gain from our method is over 1.0 dB.

B. Specific Images

In Table II, we show the results from a similar experiment on some well known images such as *Lena* (512×512 image), *Couple* (256×256 image), *Peppers* (512×512 image) all compressed at 0.8 bpp. For the fourth image *Barbara*, an image with a lot of high frequency information due to the edges and textures on the tablecloth and clothes, we show the results for three bitrates: 0.2 bpp, 0.8 bpp and 2.0 bpp.

If no packets are lost, the PSNR of the reconstructed image is equal to the PSNR of the broadcast compressed image. For *Lena*, *Couple*, *Peppers* (all compressed at 0.8 bpp), the PSNR of the compressed image is respectively 38.29 dB, 33.33 dB and 36.69 dB. The PSNR of *Barbara* compressed at 0.2, 0.8 and 2.0 bpp is respectively 25.32 dB, 32.81 dB and 40.54 dB.

The results in Table II agree with the results of Table I: the proposed method outperforms the bilinear interpolation and the methods of [6] and [7]. In particular, the results for *Barbara*, compressed at three different bitrates, clearly show the strength of the proposed method. At low bitrates (e.g., 0.2 bpp), a lot of the high frequency information is discarded during the compression. In that case, our method outperforms bilinear interpolation and the approach of [6] by 0.3 to 0.6 dB, and the improvement over [7] is between 0.25 and 0.4 dB. This quality improvement is mainly due to the improved estimation of the lost low-frequency and mid-frequency information. For high bitrates (e.g., 2.0 bpp), much more high frequency information is still present after compression. In this case, our method outperforms bilinear interpolation and the approach of [6] by 1.6 to 2.5 dB, and the gain over [7] is between 1.3 and 2.1 dB. This quality improvement is due to not only the improved estimation of the lost low frequency information, but also because of the improved estimation of the lost high frequency information.

Fig. 5 and 6 show visual results for two images. Fig. 5 compares the performance of bilinear interpolation, the methods of [6], [7] and the proposed method for the *Lena*-image compressed at 0.8 bpp with one packet lost (i.e., 6.25% of the coefficients lost). This example clearly illustrates the

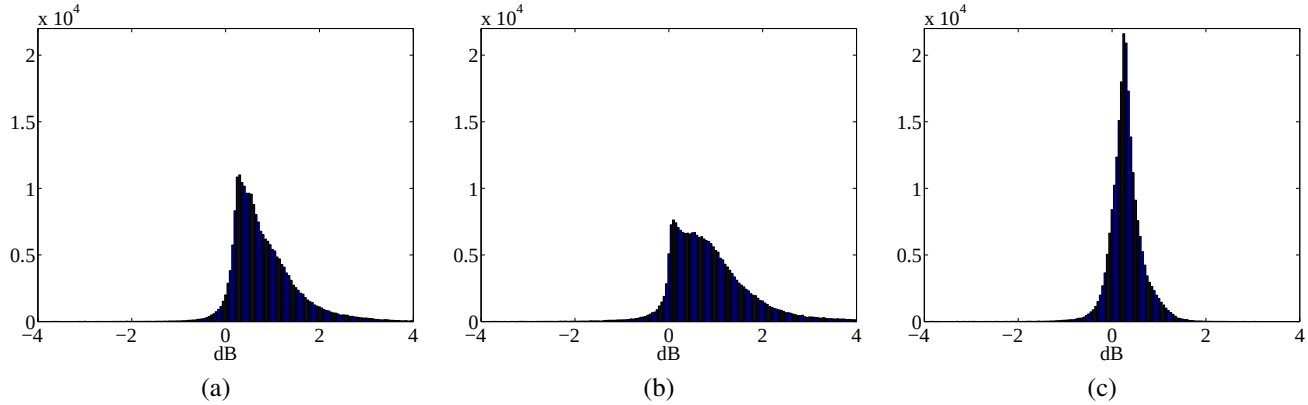


Fig. 4. Frequency distribution of the PSNR difference between the PSNRs of the I-frames reconstructed by the proposed method and the PSNRs of the I-frames reconstructed by respectively: **(a)** the bilinear interpolation, **(b)** the adaptive MAP approach of [6], **(c)** the adaptive method of [7].

TABLE II

AVERAGE PSNR [DB] FOR: BILINEAR INTERPOLATION (BI), THE ADAPTIVE MAP APPROACH [6], THE ADAPTIVE METHOD OF [7] AND THE PROPOSED GMRF METHOD, FOR *all* POSSIBLE COMBINATIONS OF p LOST PACKETS.

Lena (0.8 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	38.29	38.29	38.29	38.29
1	30.52	30.94	30.93	31.85
2	27.75	28.18	28.19	29.05
3	25.96	26.37	26.41	27.13
4	24.57	24.99	25.04	25.61

Couple (0.8 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	33.33	33.33	33.33	33.33
1	30.27	29.94	30.50	30.55
2	28.38	27.94	28.66	28.71
3	26.95	26.49	27.26	27.30
4	25.74	25.33	26.11	26.09

Peppers (0.8 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	36.69	36.69	36.69	36.69
1	29.64	29.62	30.14	30.45
2	26.87	26.77	27.39	27.70
3	25.04	24.90	25.54	25.82
4	23.61	23.47	24.09	24.33

Barbara (0.2 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	25.32	25.32	25.32	25.32
1	24.22	24.28	24.34	24.58
2	23.29	23.37	23.46	23.83
3	22.45	22.54	22.66	23.07
4	21.68	21.78	21.91	22.30

Barbara (0.8 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	32.81	32.81	32.81	32.81
1	28.07	28.19	28.36	29.66
2	25.74	25.84	26.05	27.47
3	24.12	24.19	24.44	25.72
4	22.85	22.89	23.16	24.22

Barbara (2.0 bpp)

p	BI	MAP [6]	Method of [7]	Proposed
0	40.54	40.54	40.54	40.54
1	29.38	29.48	29.76	31.90
2	26.39	26.45	26.76	28.65
3	24.52	24.55	24.87	26.44
4	23.11	23.13	23.44	24.70

benefits of our GMRF method for the reconstruction of the low frequency content. Near the strongest edges we are able to do a correct interpolation, where the bilinear interpolation and the approach of [6] fail. This explains the PSNR gain of more than 1.5 dB of the proposed method over the bilinear interpolation and the PSNR gain of 1.2 dB over [6]. The gain of 0.5 dB of the proposed method over the method of [7] can be explained by the further improved reconstruction of the low frequency content but also by the improved reconstruction of the high frequency content.

Fig. 6 shows visual results for the *Barbara*-image compressed at 2.0 bpp with two packets lost (i.e., with 12.5% of the coefficients lost). The proposed method has a PSNR gain of 2.2 dB over the bilinear interpolation, of 2.1 dB over the adaptive MAP approach of [6], and of 1.7 dB over the method of [7]. Again, this example clearly shows the correct interpolation of low frequency content by the proposed method. Moreover, the proposed method correctly interpolates the lost high frequency coefficients where all the reference methods fail. This is visible in enlarged details in Fig. 7. These enlarged image parts clearly show that the proposed method correctly interpolates the missing data on the trousers of Barbara, while the other methods cause a striping effect which is undesirable. Similar effects can be observed on the tablecloth.

C. Fast version

In Section III-D we counted how many operations that are necessary for the reconstruction of one coefficient. If we extrapolate this, the required number of operations (additions and multiplications) for the reconstruction of a 512×512 I-frame with four levels of wavelet decomposition and with 1 of 16 packets lost is: $22 \cdot 10^3$ for the bilinear interpolation, $49 \cdot 10^6$ for the adaptive MAP approach [6], $23 \cdot 10^3$ for the

(a) (b) (c) (d) (e) (f)

Fig. 5. **(a)** *Lena* compressed at 0.8 bpp (PSNR = 38.29 dB). **(b)** *Lena* after loss of packet 4. **(c)** Reconstruction with the bilinear interpolation (PSNR = 30.59 dB). **(d)** Reconstruction with the adaptive MAP approach of [6] (PSNR = 30.95 dB). **(e)** Reconstruction with the adaptive method of [7] (PSNR = 31.63 dB). **(f)** Reconstruction with the proposed method (PSNR = 32.14 dB).

adaptive method of [7], and $4.3 \cdot 10^6$ for the proposed method.

For time-critical applications, we can also consider a fast version of the proposed method. In this fast version, we reconstruct the coefficients of LH^1 and HL^1 with a one dimensional linear interpolation instead of the proposed GMRF method, and we also do not process the HH^j subbands with the proposed method but we replace lost HH^j coefficients by zero. In this way, we can reduce the number of operations to $0.7 \cdot 10^6$, which is a speed increase of a factor 5.9, without a big quality reduction in most cases.

In Table III, we show the results for the fast version of the proposed method for the same experiment as in Table I. We can conclude that, in general, the fast version of our method gives only a small loss in quality, compared to our *normal*

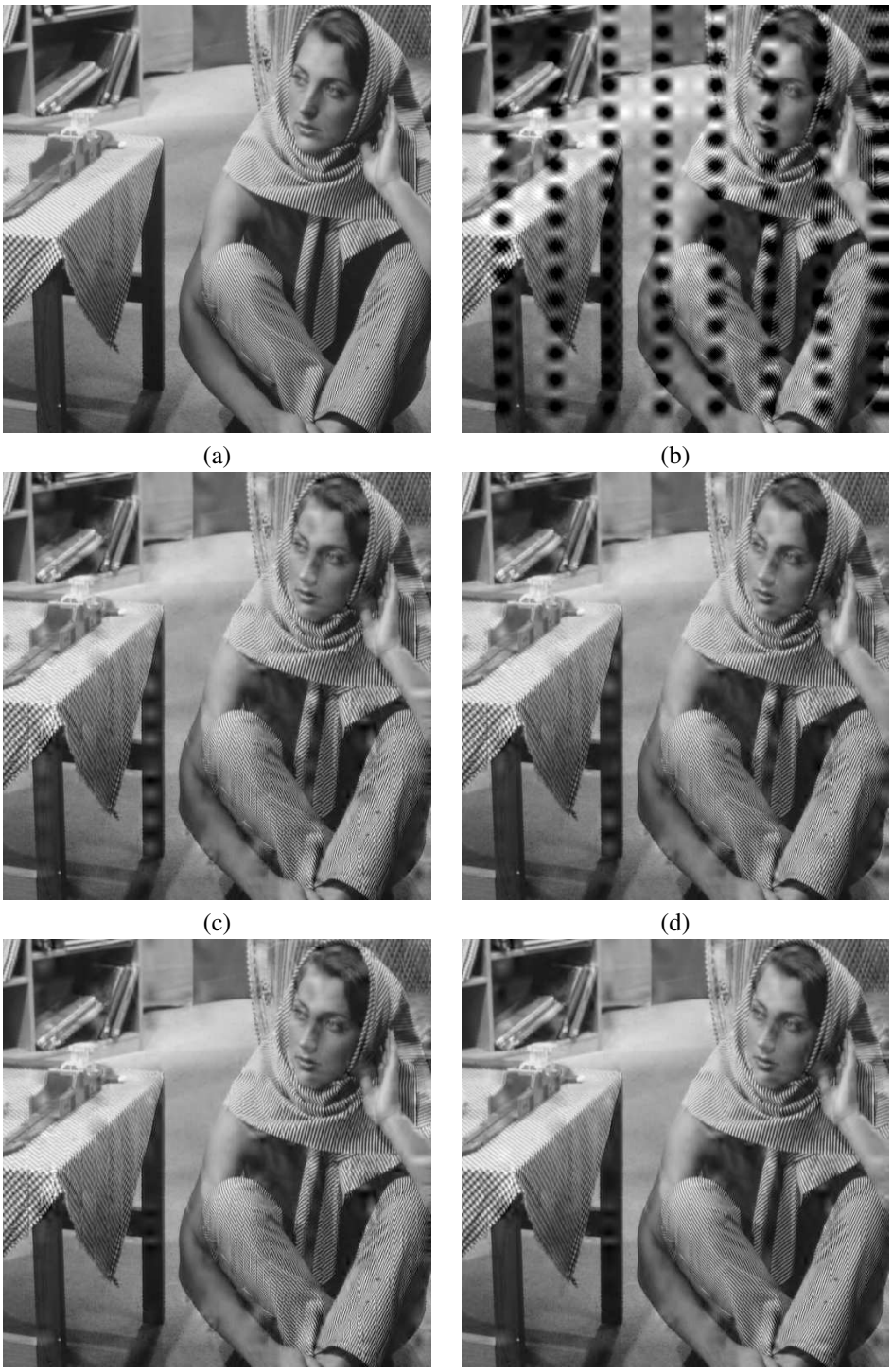


Fig. 6. **(a)** *Barbara* compressed at 2.0 bpp (PSNR = 40.54 dB). **(b)** *Barbara* after loss of packets 4 and 12. **(c)** Reconstruction with the bilinear interpolation (PSNR = 26.08 dB). **(d)** Reconstruction with the adaptive MAP approach of [6] (PSNR = 26.22 dB). **(e)** Reconstruction with the adaptive method of [7] (PSNR = 26.60 dB). **(f)** Reconstruction with the proposed method (PSNR = 28.27 dB).

proposed method. However, for some specific images, the resulting performance loss is higher than shown in Table III.

For *Lena*, *Couple* and *Peppers* the quality drop of the fast version (Table IV) is low (on average about 0.01 dB). This is because these images have few significant coefficients in the HH^j subbands and in the subbands HL^1 and LH^1. For *Barbara*, the fast version results in a much bigger quality drop because this image has much more high frequency information (fine textures) than the other three images. Note that our fast version still yields a much better quality than the three reference methods.

The quality drop resulting from the simplification in the fast version of the method increases with an increasing bitrate: if *Barbara* is encoded at a low bitrate, a lot of the high

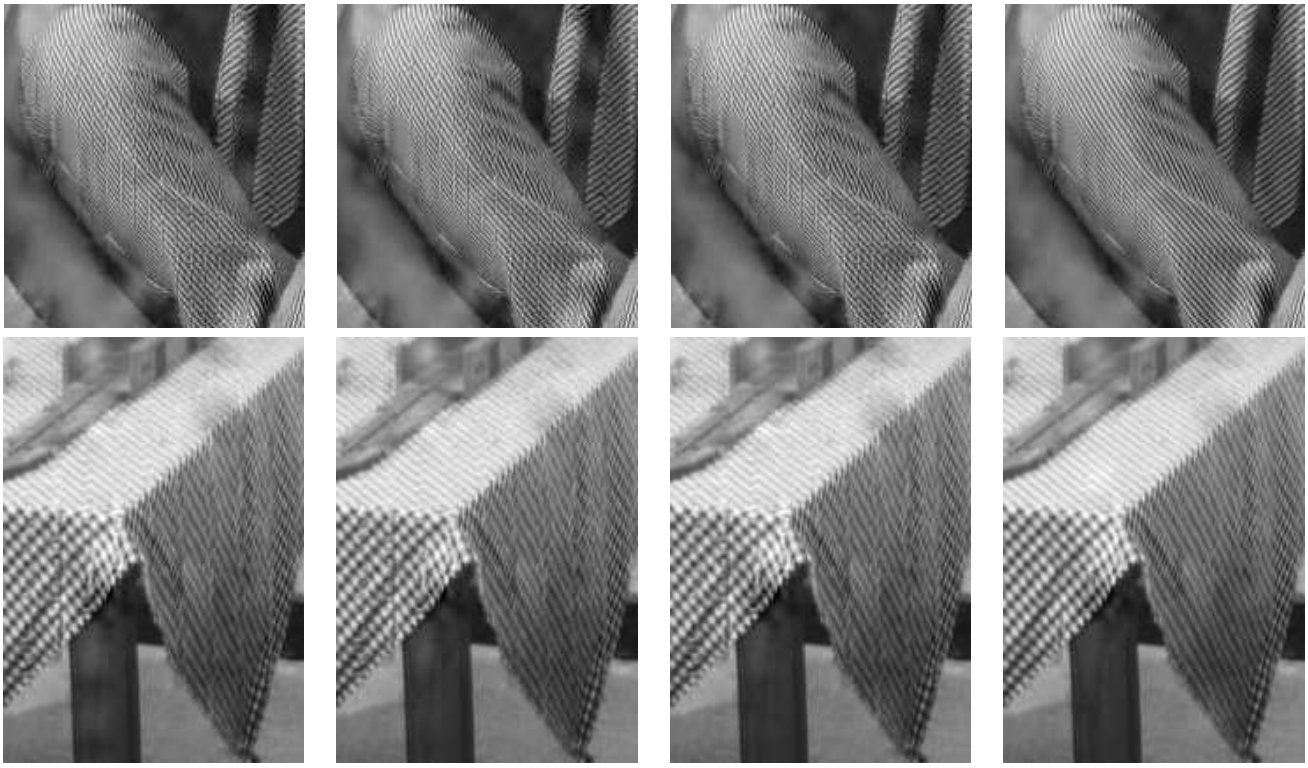


Fig. 7. Enlarged image parts from the results in Fig. 6. From left to right: Detail of the reconstruction with the bilinear interpolation, the methods of [6], [7], and the proposed method.

TABLE III

AVERAGE PSNR [DB] OF *all* THE 416 I-FRAMES FOR THE PROPOSED GMRF METHOD AND ITS FAST VERSION.

p	Proposed	Proposed (fast)
0	40.01	40.01
1	34.68	34.64
2	32.26	32.22
3	30.53	30.50
4	29.10	29.07

frequency information has gone missing in the compression stage, and cannot be recovered in the reconstruction stage anyway, so using the more complex method is less beneficial. However, if the image is encoded at a high bitrate (e.g., 2.0 bpp), almost all the high frequency information is still present after compression, and can be successfully recovered with the proposed method. Hence, in this case it is clearly beneficial to apply the *full version* proposed method, with GMRF-based reconstruction in all high frequency subbands. We illustrate this with visual examples in Fig. 8. For the *Lena*-image, there is no obvious difference between the results of the full version and the fast version of the proposed method. For the *Barbara*-image, the difference is visually obvious. In terms of PSNR, there is more than 1 dB loss for the fast version.

In practice, the importance of processing the coefficients of the highest-frequency subbands can be estimated by measuring the energy in these subbands from the correctly received coefficients. The threshold for reconstructing these coefficients must then be determined from the desired image quality, the available processing power and the desired latency (e.g., depending on the frame rate in real-time video communication).

D. Comparison with block-based error concealment approaches

In this section, we compare our method for the reconstruction of lost wavelet coefficients with the block-based error concealment approaches. In general, comparing the performance of the proposed scheme with schemes that are very different in nature such as the block-based error concealment approaches is not trivial. The comparison criteria should be carefully defined. One scheme may perform better in case of no packet loss, i.e., in terms of pure compression quality, while another scheme can perform better in case of quality loss. We selected two publications on block-based approaches [8], [36] that provide enough information for a fair comparison with the proposed method.

In [8], a reconstruction method for block-based wavelet domain image coders (such as, e.g., JPEG2000) is proposed. This method requires somewhat smaller tile sizes and smaller code-block sizes than in the standard JPEG2000 compression which results in a lower compression efficiency compared to

TABLE IV

AVERAGE PSNR [DB] FOR THE PROPOSED GMRF METHOD AND ITS FAST VERSION.

Lena (0.8 bpp)		
p	Proposed	Proposed (fast)
0	38.29	38.29
1	31.85	31.80
2	29.05	29.00
3	27.13	27.09
4	25.61	25.59

Couple (0.8 bpp)		
p	Proposed	Proposed (fast)
0	33.33	33.33
1	30.55	30.57
2	28.71	28.73
3	27.30	27.31
4	26.09	26.11

Peppers (0.8 bpp)		
p	Proposed	Proposed (fast)
0	36.69	36.69
1	30.45	30.45
2	27.70	27.70
3	25.82	25.83
4	24.33	24.34

Barbara (0.2 bpp)		
p	Proposed	Proposed (fast)
0	25.32	25.32
1	24.58	24.51
2	23.83	23.72
3	23.07	22.96
4	22.30	22.20

Barbara (0.8 bpp)		
p	Proposed	Proposed (fast)
0	32.81	32.81
1	29.66	28.98
2	27.47	26.72
3	25.72	25.05
4	24.22	23.67

Barbara (2.0 bpp)		
p	Proposed	Proposed (fast)
0	40.54	40.54
1	31.90	30.68
2	28.65	27.58
3	26.44	25.56
4	24.70	24.01

TABLE V

AVERAGE PSNR [DB] OF THE RECONSTRUCTED *Lena*-IMAGE (512×512) AND *Goldhill*-IMAGE (512×512). NO COMPRESSION IS USED.

Lena		
loss rate	block-based [8]	Proposed
6.25%	33.8	36.0
12.50%	31.2	32.6

Goldhill		
loss rate	block-based [8]	Proposed
6.25%	36.3	36.7
12.50%	32.3	33.4

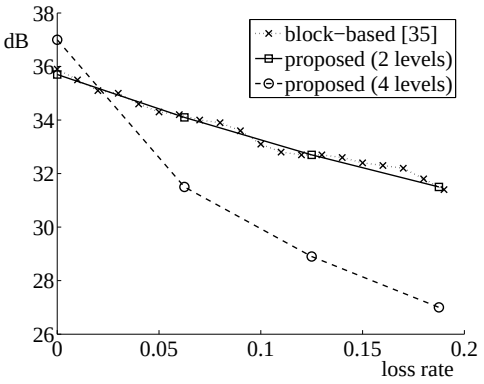


Fig. 9. Comparison of the average PSNR [dB] for the reconstruction of the *Lena*-image (512×512) compressed at 0.60 bits per pixel. The results for the block-based error concealment approach of [36] are comparable to the results for the proposed method with 2 wavelet decomposition levels. The results for the proposed method with 4 wavelet decomposition levels are better in case of no packet loss (i.e., a better compression quality) but worse in case of packet loss (i.e., a worse reconstruction quality).

the standard JPEG2000 coding. In [8], only the results for uncompressed images were reported and hence we compare our results with those from [8] on uncompressed *Lena* and *Goldhill* images in Table V. For this comparison, we used three levels of wavelet decomposition as in [8]. The results show that the proposed method outperforms the block-based one from [8].

The approach in [36] conceals the errors in block-based image coding systems (such as the standard JPEG system) by using neural network techniques in the spatial domain. In Fig. 9, we compare the results of our method to those from [36, Table 2] for the grayscale *Lena*-image (512×512) compressed at 0.60 bits per pixel. The results show that our method with 2 decomposition levels yields a similar performance as [36]. The results also show that using more decomposition levels in our approach provides better compression quality in case of no loss, but at the same time degrades its performance in case of packet loss.

In summary, a fair comparison of the proposed method with the different concept of block-based approaches implies weighting compression ability versus ability to reconstruct packet losses. The way in which the images are compressed and the compression factor influences the reconstruction ability. The compression schemes are sometimes altered to allow reconstruction in case of data loss, which yields a slightly worse compression efficiency. For example, in block-based DCT image coding, the dc values of adjacent blocks are usually compressed with differential coding. If a dispersive packetization is used, the efficiency of this differential coding will be lower because it cannot operate on adjacent blocks (which are spread into different packets) but on blocks that are further away. However, without a dispersive packetization, the reconstruction of the lost blocks will be difficult and it may even be impossible. The results presented in this paper show that the proposed error concealment method in combination with a dispersive packetization compares well with the more traditional block-based approaches.

V. CONCLUSION

This paper presents a novel locally adaptive error concealment method for subband coded I-frames based on a Gaussian Markov Random Field model. Our method estimates the optimal interpolation weights by imposing an inhomogeneous GMRF model and estimating its parameters from the local context of the lost coefficient. Experiments have demonstrated a significant improvement over the state-of-the-art methods including more complex MRF-based approaches. The proposed method is especially effective for the reconstruction of high

Fig. 8. Performance comparison between the proposed method and its simplified fast version. Left: the full version of the proposed method (The PSNR of *Lena* and *Barbara* is 32.14 and 28.27 dB respectively). Right: its fast version (The PSNR of *Lena* and *Barbara* is 32.08 and 27.25 dB respectively).

frequency information, where the existing methods usually fail.

ACKNOWLEDGEMENT

The authors wish to thank Prof. Dr. Ivan Bajić for generously sharing his code and simulation environment of [6], which we used in our experiments for comparative purposes.

APPENDIX

In this section, we explain how many additions (A) and multiplications (M) are needed in each step of the reconstruction of a lost coefficient with the proposed method. As introduced in Section III-D, N_{Ω_s} is the number of neighboring coefficients in Ω_s. For a lost LL-coefficient, $N_{\Omega_s} = 9$; for a high frequency coefficient, $N_{\Omega_s} = 21$.

A. Initial estimate

An initial estimate is made for both low and high frequency coefficients:

- a bilinear interpolation for a lost LL^J-coefficient requires 3A and 1M;
- a linear interpolation for a lost LH^j or HL^j-coefficient requires 1A and 1M;
- a replacement by zero for a lost HH^j-coefficient requires no operations.

B. Calculation of the interpolation weights $\widehat{\theta}^s_{(1,0)}$ and $\widehat{\theta}^s_{(0,1)}$

1) Calculation of $\mathbf{q}_t$: For all N_{Ω_s} neighbors, the upper and lower, and left and right neighbors have to be summed together. $\mathbf{q}_t = \begin{bmatrix} x_{(t1+1,t2)} + x_{(t1-1,t2)} & x_{(t1,t2+1)} + x_{(t1,t2-1)} \end{bmatrix}$: $2N_{\Omega_s}$A.
2) Calculation of $\mathbf{q}^T\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \begin{bmatrix} q_1 & q_2 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$: $3N_{\Omega_s}$M
3) Calculation of $\sum_{N_{\Omega_s}}(\mathbf{q}^T\mathbf{q}) = \begin{bmatrix} a_1 + \ldots + a_{N_{\Omega_s}} & b_1 + \ldots + b_{N_{\Omega_s}} \\ b_1 + \ldots + b_{N_{\Omega_s}} & c_1 + \ldots + c_{N_{\Omega_s}} \end{bmatrix}$: $3(N_{\Omega_s} - 1)$A.
4) Calculation of the inverse of a 2×2 matrix: $\sum_{N_{\Omega_s}}(\mathbf{q}^T\mathbf{q})^{-1} = \begin{bmatrix} d & e \\ e & f \end{bmatrix}^{-1} = \frac{1}{df-e^2} \begin{bmatrix} f & -e \\ -e & d \end{bmatrix}$. The determinant: 1A and 2M. The multiplication of the determinant and the matrix: 3M. Overall: 1A and 5M.
5) Calculation of $\sum(\mathbf{q}^T x) = \begin{bmatrix} g_1 \\ h_1 \end{bmatrix} X_1 + \ldots + \begin{bmatrix} g_{N_{\Omega_s}} \\ h_{N_{\Omega_s}} \end{bmatrix} X_{N_{\Omega_s}}$: $2(N_{\Omega_s} - 1)$A + $2N_{\Omega_s}$M.
6) Calculation of $\widehat{\boldsymbol{\theta}^s} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} D \\ E \end{bmatrix} = \begin{bmatrix} AD + BE \\ BD + CE \end{bmatrix}$: 2(1A+2M) = 2A+4M.

C. Calculation of the lost coefficient $\widehat{x}_s$

Calculation of the lost coefficient $\widehat{x}_s = \widehat{\theta}^s_{(1,0)}(x_{(s_1+1,s_2)} + x_{(s_1-1,s_2)}) + \widehat{\theta}^s_{(0,1)}(x_{(s_1,s_2+1)} + x_{(s_1,s_2-1)})$: 3A+2M.

D. Overview

In Table VI we summarize the number of additions and multiplications necessary for the reconstruction of low frequency (LL^J) and high frequency (LH^j, HL^j and HH^j) coefficients. For a low frequency coefficient, we need 67 additions and 57 multiplications. For an LH^j or HL^j coefficient, we need 149 additions and 117 multiplications. For an HH^j coefficient, we need 148 additions and 116 multiplications.

REFERENCES

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

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

[3] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *Proceedings of the IEEE*, vol. 78, no. 4, pp. 678–689, Apr. 1990.

[4] M. Barni, "A fast algorithm for 1-norm vector median filtering," *IEEE Transactions on Image Processing*, vol. 6, no. 10, pp. 1452–1455, Oct. 1997.

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

[6] I. 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.

[7] J. Rombaut, A. Pižurica, and W. Philips, "Locally adaptive passive error concealment for wavelet coded images," *IEEE Signal Processing Letters*, vol. 15, pp. 178–181, 2008.

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

[9] O. 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.

[10] S. Li, *Markov Random Field Modelling in Computer Vision*, T. Kunii, Ed. Tokyo: Springer-Verlag, 1995.

[11] K. Abend, T. Harley, and L. Kanal, "Classification of binary random patterns," *IEEE Transactions on Information Theory*, vol. 11, pp. 538–544, 1965.

[12] R. Chellappa, *Progress in pattern recognition 2*. NorthHolland, 1985, ch. Two-Dimensional Discrete Gaussian Markov Random Field Models for Image Processing, pp. 79–112.

[13] G. Cross and A. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 25–39, 1983.

[14] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, Nov. 1984.

[15] B. Gidas, "A renormalization group approach to image processing problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 2, pp. 164–180, Feb. 1989.

[16] C. Bouman and B. Liu, "Multiple resolution segmentation of textured images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 99–113, 1991.

[17] M. Comer and E. Delp, "Segmentation of textured images using a multiresolution gaussian autoregressive model," *IEEE Transactions on Image Processing*, vol. 8, pp. 408–420, Mar. 1999.

[18] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 39–55, 1987.

[19] S. Krishnamachari and R. Chellappa, "Multiresolution gauss-markov random field models for texture segmentation," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 251–267, February 1997.

[20] K. Pyun, J. Lim, C. Won, and R. Gray, "Image segmentation using hidden markov gauss mixture models," *Transactions on Image Processing*, vol. 16, no. 7, pp. 1902–1911, July 2007.

[21] C. Ballester, M. Bertalmio, V. Caselles, M. Garrido, A. Marques, and F. Ranchin, "An inpainting-based deinterlacing method," *Transactions on Image Processing*, vol. 16, no. 10, pp. 2476–2491, Oct. 2007.

[22] Y. Zhang and K.-K. Ma, "Error concealment for video transmission with dual multiscale markov random field modeling," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 236–242, February 2003.

[23] E. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *Proc. SPIE, 44th Annual Meeting*, vol. 3813, Denver, CO, July 1999, pp. 188–195.

[24] M. Malfait and D. Roose, "Wavelet-based image denoising using a Markov random field a priori model," *IEEE Trans. on Image Processing*, vol. 6, pp. 549–565, Apr. 1997.

[25] A. Pižurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for bayesian wavelet based image denoising,"

TABLE VI

NUMBER OF ADDITIONS (A) AND MULTIPLICATIONS (M) PER COEFFICIENT FOR LOW FREQUENCY (LL^J) AND HIGH FREQUENCY (LH^j, HL^j AND HH^j) COEFFICIENTS.

	LL^J ($N_{\Omega_s} = 9$)	LH^j and HL^j ($N_{\Omega_s} = 21$)	HH^j ($N_{\Omega_s} = 21$)
A	3A+1M	1A+1M	0A+0M
B1	$2N_{\Omega_s}$A	$2N_{\Omega_s}$A	$2N_{\Omega_s}$A
B2	$3N_{\Omega_s}$M	$3N_{\Omega_s}$M	$3N_{\Omega_s}$M
B3	$3(N_{\Omega_s} - 1)$A	$3(N_{\Omega_s} - 1)$A	$3(N_{\Omega_s} - 1)$A
B4	1A+5M	1A+5M	1A+5M
B5	$2(N_{\Omega_s} - 1)$A+$2N_{\Omega_s}$M	$2(N_{\Omega_s} - 1)$A+$2N_{\Omega_s}$M	$2(N_{\Omega_s} - 1)$A+$2N_{\Omega_s}$M
B6	2A+4M	2A+4M	2A+4M
C	3A+2M	3A+2M	3A+2M
Total	$(7N_{\Omega_s} + 4)$A+$(5N_{\Omega_s} + 12)$M 67A+57M	$(7N_{\Omega_s} + 2)$A+$(5N_{\Omega_s} + 12)$M 149A+117M	$(7N_{\Omega_s} + 1)$A+$(5N_{\Omega_s} + 11)$M 148A+116M

IEEE Transactions on Image Processing, vol. 11, no. 5, pp. 545–557, May 2002.

[26] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society B*, vol. 36, no. 2, pp. 192–236, 1974.

[27] R. Chellappa, Y. Hu, and S. Kung, "On two-dimensional markov spectral estimation," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 31, no. 4, pp. 836–841, Aug. 1983.

[28] R. Kashiap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Trans. on Information Theory*, vol. 29, pp. 60–72, 1983.

[29] P. Huber, "Robust estimation of a location parameter," *Annals of Mathematical Statistics*, vol. 35, pp. 73–101, 1964.

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

[31] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, and J. Cornelis, "Wavelet coding of volumetric medical datasets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 441–458, March 2003.

[32] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.

[33] I. Daubechies, *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[34] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1998.

[35] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Prentice Hall, 2000.

[36] Y. Huang, "Concealment of damaged block coded images using adaptive multilayer perceptrons," in *Proceedings of the 2000 ICS: Workshop on Image Processing and Pattern Recognition*, Chiayi, Taiwan, December 2000, pp. 70–77.

BIOGRAPHIES OF AUTHORS

Joost Rombaut was born in Dendermonde, Belgium, on July 4, 1981. He received the diploma degree in computer science engineering from Ghent University, Belgium, in 2004. He is currently pursuing the Ph.D. degree in computer science engineering in the Department of Telecommunications and Information Processing, Ghent University.

In September 2004, he joined the Department of Telecommunications and Information Processing, Ghent University. His research interests include image and video processing, error resilient coding, and transmission.

Aleksandra Pižurica (M'98) received the diploma degree in electrical engineering from the University of Novi Sad, Serbia, in 1994, the M.Sc. degree in telecommunications from the University of Belgrade, Serbia, in 1997, and the Ph.D. degree from Ghent University, Belgium, in 2002.

From 1994 until 1997, she was with the Department of Telecommunications, University of Novi Sad, and in 1997, she joined the Department of Telecommunications and Information Processing, Ghent University. Since 2005, she has been a postdoctoral research fellow of the Fund for the Scientific Research in Flanders (FWO). Her research interests include image restoration, multiresolution representations, Markov random field models, signal detection and estimation, multimedia applications, and remote sensing.

Wilfried Philips (M'92) was born in Aalst, Belgium, on October 19, 1966. He received the diploma degree in electrical engineering and the Ph.D. degree in applied sciences from Ghent University, Belgium, in 1989 and 1993, respectively.

From October 1989 until October 1997, he was with the Department of Electronics and Information Systems, Ghent University, for the Flemish Fund for Scientific Research (FWO-Vlaanderen), first as a research assistant and later as a postdoctoral research fellow. Since November 1997, he has been with the Department of Telecommunications and Information Processing, Ghent University, where he is currently a full-time Professor and heads the research group "Image Processing and Interpretation," which has recently become part of the virtual Flemish ICT research institute IBBT. Some of the recent research activities in the group include image and video restoration and analysis and the modeling of image reproduction systems. Important application areas targeted by the group include remote sensing, surveillance, and industrial inspection.