SPATIAL ERROR CONCEALMENT OF CORRUPTED IMAGE DATA USING FREQUENCY SELECTIVE EXTRAPOLATION

Katrin Meisinger and André Kaup

Chair of Multimedia Communications and Signal Processing,
University of Erlangen-Nuremberg
Cauerstr. 7, 91058 Erlangen, Germany
{meisinger,kaup}@LNT.de

ABSTRACT

This paper introduces a method for spatial error concealment of lost image data in erroneous image transmission. The image content of the correctly received surrounding blocks is successively approximated by a weighted linear combination of basis functions and the missing block is obtained by extrapolation. An implementation in the frequency domain allows an efficient realization. Investigations show that particularly 2D DFT basis functions are suited for the signal extrapolation in order to be able to reconstruct both monotone areas and edges.

1. INTRODUCTION

Transmission of images or videos coded by block based techniques like JPEG or MPEG in error prone environments may lead to block loss. Therefore error concealment at the decoder side has to be applied. Commonly, two approaches are used in error concealment. On the one hand spatial error concealment uses the surrounding correctly received image information. Temporal error concealment on the other hand exploits motion information. In image transmission and intraframe coded video transmission spatial error concealment is applied due to the lack of motion information.

A standard approach of Wang et. al [1] assumes that the image content is changing smoothly. Hence the algorithm tries to restore the transition across the block boundary as smooth as possible.

The pixel-based method of [2] predicts each pixel from the available next neighbors. The missing block is pixel-based reconstructed from eight directions and given by a weighted linear combination of these reconstructed blocks. Hence, the method is computationally very complex.

Additionally methods implemented for comparison are described in [3], [4] and [5].

2. SPATIAL ERROR CONCEALMENT

2.1. Discrete Linear Approximation

In [6] a method of object based coding was developed. The texture of an area is successively approximated and then cut to the shape of the object. This principle is used in this paper for error concealment by estimating the missing image content using the surrounding area. The image content of the known blocks is successively approximated and the missing block obtained by extrapolation (see Fig. 1).

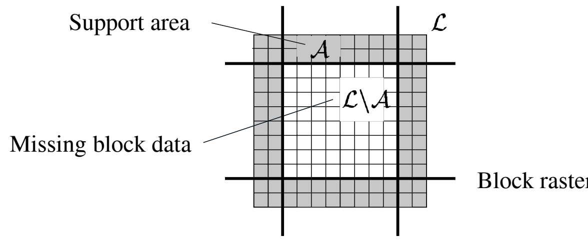

Fig. 1. The known area $\mathcal{A}$ (gray) is approximated by a parametric model and the missing area $\mathcal{L}\backslash\mathcal{A}$ (white) obtained by extrapolation.

The gray values of the pixels are denoted by $f[m,n]$ whereas m indicates the row and n the column. In order to approximate the known blocks, a parametric model $g[m,n]$ is defined being a weighted linear combination of basis functions. Assuming two-dimensional DFT basis functions

$$\varphi_{k,l}[m,n] = \mathrm{e}^{j\frac{2\pi}{M}mk}\mathrm{e}^{j\frac{2\pi}{N}nl} \tag{1}$$

the following parametric model is obtained

$$g[m,n] = \frac{1}{2MN}\sum_{(k,l)\in\mathcal{K}}(c_{k,l}\varphi_{k,l}[m,n] + c_{M-k,N-l}\ \varphi_{M-k,N-l}[m,n]). \tag{2}$$

$c_{k,l}$ denote the complex expansion coefficients and $\mathcal{K}$ the set of used basis functions. The image signal $f[m,n]$ and thus also its parametric model $g[m,n]$ are real signals. Therefore the following holds

$$c_{M-k,N-l} = c^*_{k,l} \quad \text{as well as} \tag{3}$$

$$\varphi^*_{M-k,N-l}[m,n] = \varphi_{k,l}[m,n] \tag{4}$$

whereas the considered block has size $M \times N$.

An error criterion $E_{\mathcal{A}}$ between the original signal and the parametric model is established

$$E_{\mathcal{A}} = \sum_{m,n\in\mathcal{A}}(f[m,n] - g[m,n])^2\,, \tag{5}$$

however, only evaluated at known pixels $(m,n)\in\mathcal{A}$. The set $\mathcal{L}$ is the set of all pixels whereas the gray area $\mathcal{A}$ denotes the known pixels in Fig. 1. In order to minimize the error criterion, (5) is partially derived with respect to the expansion coefficients and set to zero

$$\frac{\partial E_{\mathcal{A}}}{\partial c_{k,l}} = 0 \quad \text{and} \quad \frac{\partial E_{\mathcal{A}}}{\partial c_{M-k,N-l}} = 0 \tag{6}$$

This leads to a set of linear equations which is unfortunately not uniquely resolvable because more basis functions, i.e.($M \times N$), are available in the approximation than pixels in $\mathcal{A}$. In order to obtain a solution the principle of successive approximation is applied. This iterative algorithm computes one expansion coefficient per iteration. Hence, the image content is described by a few dominant features, i.e. by a weighted linear combination of a few selected basis functions. Therefore a suitable basis function has to be selected which is described in Sec. 2.3 and the update of the respective expansion coefficient in Sec. 2.2.

2.2. Principle of Successive Approximation

In the following the computation of the expansion coefficient $c_{u,v}^{\nu+1}$ in iteration $\nu+1$ is described assuming a suitable basis function with index u, v has already been selected. Hence, the following approximation is available in iteration ν

$$g^{(\nu)}[m,n] = \frac{1}{2MN} \sum_{(k,l)\in\mathcal{K}_\nu} (c_{k,l}^{(\nu)}\,\varphi_{k,l}[m,n] + c_{M-k,N-l}^{(\nu)}\,\varphi_{M-k,N-l}[m,n]) \tag{7}$$

with $\mathcal{K}_\nu$ denoting the set of used basis functions. Defining the window function

$$w[m,n] = \begin{cases} 1 & , \quad m,n \in \mathcal{A} \\ 0 & , \quad m,n \in \mathcal{L}\backslash A \end{cases} \tag{8}$$

we can express the residual error signal in the known area $(m,n) \in \mathcal{A}$ as

$$r^{(\nu)}[m,n] = (f[m,n] - g^{(\nu)}[m,n]) \cdot w[m,n]. \tag{9}$$

In the following the residual error signal in the next iteration is derived. The residual error signal in the known area is further approximated by a weighted suitable basis function

$$r^{(\nu+1)}[m,n] = r^{(\nu)}[m,n] - \frac{1}{2MN}\left(\Delta c\varphi_{u,v}[m,n] + \Delta c^*\varphi_{M-u,N-v}[m,n]\right)\cdot w[m,n]. \tag{10}$$

The coefficient Δc is obtained by minimizing the error criterion for the residual error according to (5) which is after some computations equivalent to solving

$$\Delta c \sum_{(m,n)\in\mathcal{A}} \varphi_{u,v}[m,n]\varphi_{M-u,N-v}[m,n] + \Delta c^* \sum_{(m,n)\in\mathcal{A}} \varphi_{M-u,N-v}[m,n]\varphi_{M-u,N-v}[m,n] = 2MN \sum_{(m,n)\in\mathcal{A}} r^{(\nu)}[m,n]\varphi_{M-u,N-v}[m,n]. \tag{11}$$

Analogously, a conjugate complex equation is obtained to (11).

In order to update $c_{u,v}^{(\nu+1)}$, the coefficient $c_{u,v}^{(\nu)}$ is modified by Δc

$$c_{u,v}^{(\nu+1)} = c_{u,v}^{(\nu)} + \Delta c \tag{12}$$

$$c_{M-u,N-v}^{(\nu+1)} = c_{M-u,N-v}^{(\nu)} + \Delta c^* \tag{13}$$

The conjugate complex coefficient is updated due to symmetry requirements.

The index u, v is included in the set of selected basis functions

$$\mathcal{K}_{\nu+1} = \mathcal{K}_\nu \cup \{u,v\} \quad \text{if} \quad u,v \notin \mathcal{K}_\nu. \tag{14}$$

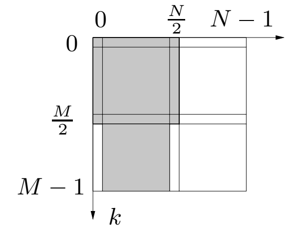

Fig. 2. Search area for DFT basis functions

2.3. Selection of Suitable Basis Function

In the last section the expansion coefficient corresponding to a selected basis function was computed. In the following, the selection of a suitable basis function $\varphi_{u,v}[m,n]$ is derived.

First the energy of the residual error signal is computed in the known area

$$\sum_{(m,n)\in\mathcal{A}} (r^{(\nu+1)}[m,n])^2 = \sum_{(m,n)\in\mathcal{A}} (r^{(\nu)}[m,n])^2 - \frac{1}{(2MN)^2} \sum_{(m,n)\in\mathcal{A}} (\Delta c\varphi_{u,v}[m,n] + \Delta c^*\varphi_{M-u,N-v}[m,n])^2 \tag{15}$$

taking into account that the residual error in the next iteration is orthogonal to the selected basis function. Furthermore, the residual error energy is decreased in each iteration as (15) shows.

In the next iteration that basis function is selected which is minimizing the residual error energy. The residual error energy becomes minimum if the decrease of the residual error energy becomes maximum. Therefore the index u, v is selected maximizing

$$\Delta E_{\mathcal{A}}^{(\nu)} = \frac{1}{2(MN)^2}\left(|\Delta c|^2 \sum_{m,n\in\mathcal{A}} \varphi_{u,v}[m,n]\varphi_{u,v}^*[m,n] + \operatorname{Re}\{\Delta c^2 \sum_{m,n\in\mathcal{A}} (\varphi_{u,v}[m,n])^2\}\right) \tag{16}$$

However, the search area for DFT basis function is limited to the gray area in Fig. 2 due to symmetry requirements (3), (4).

The algorithm is initialized by

$$g^{(0)}[m,n] = 0. \tag{17}$$

The algorithm terminates when the decrease of the residual error energy drops below a prespecified threshold ΔE_{min}.

3. IMPLEMENTATION IN FREQUENCY DOMAIN

All derivations are done so far by an approximation of a series expansion in the spatial domain. However, the selection of a basis function by the evaluation of the sums (16) is computationally complex. Hence, all computations are expressed in the frequency domain in order to allow an efficient implementation.

The transform with respect to the special shape $\mathcal{A}$ can be expressed by a shifted window in the frequency domain

$$\sum_{(m,n)\in\mathcal{A}} \varphi_{u,v}[m,n]\varphi_{k,l}^*[m,n] = \sum_{(m,n)\in\mathcal{L}} w[m,n]\varphi_{u,v}[m,n]\varphi_{k,l}^*[m,n] = W[k-u,l-v] \tag{18}$$

Hence, we can express the computation of Δc (11) in the frequency domain with help of (3), (4)

$$\Delta c W[0,0] + \Delta c^* W[2k, 2l] \quad = \quad 2MN \cdot R[k,l]. \quad (19)$$

Analogously we obtain a second conjugate complex equation. Solving the equations with respect to Δc yields

$$\Delta c = \begin{cases} MN \frac{R[u,v]}{W[0,0]} & , u,v \in \mathcal{M} \\ 2MN \frac{R[u,v]W[0,0] - R^*[u,v]W[2u,2v]}{W[0,0]^2 - |W[2u,2v]|^2}, & \text{else} \end{cases} \quad (20)$$

with $\mathcal{M} = \{(0,0), (0, \frac{N}{2}), (\frac{M}{2}, 0), (\frac{M}{2}, \frac{N}{2})\}$. The case differentiation is necessary due to the symmetry requirements (3), (4) and the definition of $g[m,n]$ according to (2).

The basis function with index u, v is selected maximizing

$$\Delta E_{\mathcal{A}}^{(\nu)} = \begin{cases} 2\frac{R[k,l]^2}{W[0,0]} & , \quad k,l \in \mathcal{M} \\ 2\frac{|R[k,l]|^2 W[0,0] - \mathrm{Re}\{R[k,l]^2 W^*[2k,2l]\}}{W[0,0]^2 - |W[2k,2l]|^2}, & \text{else} \end{cases} \quad (21)$$

In iteration $\nu + 1$ we obtain a residual error signal of

$$R^{(\nu+1)}[k,l] = R^{(\nu)}[k,l] - \frac{1}{2MN}(\Delta c W[k-u, l-v] + \Delta c^* W[k-(M-u), l-(N-v)]) \quad (22)$$

The iteration terminates when the decrease of the residual error energy drops below a prespecified threshold. The parametric model is then obtained by inverse DFT transform

$$g[m,n] = \mathrm{IDFT}_{M,N}\{G[k,l]\}. \quad (23)$$

Finally, the missing block is cut out of $g[m,n]$.

Since all equations can be expressed in the DFT domain, only one transform into the DFT domain is required in the beginning and one at the end.

The method described in this section is also applied in the field of missing X-ray data interpolation in [7] and modeling the human sense of hearing by signal processing in [8].

4. SIMULATION RESULTS

In general, periodic functions are suited for signal approximation and extrapolation. The comparison of different periodic basis functions (DCT, DFT) shows that DFT basis functions are better suited for the extrapolation. The DCT contains only vertical and horizontal basis images (see Fig. 3, left side) in contrary to the DFT having also diagonal ones (real part is shown in Fig. 3, right side).

Simulated block losses of 8×8 and 16×16 pixels are investigated in order to test the error concealment algorithm. Table 1 shows the PSNR results evaluated at concealed blocks comparing different error concealment techniques. The images Lena, Baboon and Peppers with a size of 512×512 pixels are investigated.

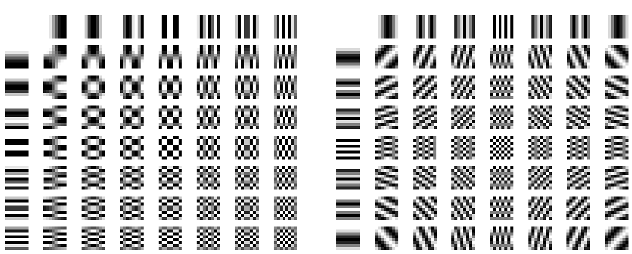

Fig. 3. 8×8 basis images. Left: DCT. Right: Real part of DFT.

First of all we take a look at images concealed with the proposed method. Due to the limited space only parts of images are shown (full images can be seen on [9]). Fig. 4 shows on the left side the Lena image with a 8×8 block loss and on the right the concealed image. The missing block and a surrounding of known pixels form a block of 12×12 pixels like in Fig. 1. The FFT transforms this block in the DFT domain. A size of 64×64 is chosen in order to apply the FFT algorithm. Additionally, with a larger FFT size a better spectral resolution is obtained. The algorithm terminates when either the termination threshold ΔE_{min} per pixel in the known area or a maximum number of iterations Max_it is exceeded. $\Delta E_{min} = 24$ and Max_it equals 4 iterations for the Lena image resulting in 3.1 iterations on average per block. However, a FFT size of 128×128 is applied in the case of a 16×16 block loss. The terminations thresholds are chosen to $\Delta E_{min} = 6$ and $Max\_it = 7$. Fig. 5 shows on the right side the concealed Peppers image from the left side. The frame of known pixels is 6 and the average iteration length is 5.7. In the case of the damaged Baboon image (Fig. 6, left) a frame of 12 pixels is chosen and on average 6.9 iterations are run. Fig. 6 shows the result on the right.

Obviously, the parameters frame size and number of iterations depend on the image content. In principle it holds that the best solution for the first iteration is achieved by a frame of one pixel. The DC component is chosen in the first iteration and can be best extrapolated from the direct surrounding corresponding to the color of the missing block. However, in order to restore details, larger frames are necessary as well as more iterations.

Generally, the image content of the missing block becomes more uncorrelated to the image content of the surrounding as the distance from the missing block increases. A small frame is chosen for the Lena image because the signals are partly highly uncorrelated. When the frame contains details which do not belong to the content, the performance decreases. However, the frame can be chosen larger for the Peppers image due to the clear structures and monotone areas. A larger frame is anyway necessary for a detailed reconstruction. For the image Baboon an even larger frame is chosen since the fur shows a noise-like frequency behavior. In the case of a larger frame and more iterations even the reconstruction of the structure of the fur is possible (see Fig. 6, right side).

In the following the behavior of the algorithm is described in comparison to the reference techniques. The introduced algorithm is superior to the other techniques for the image Lena concealing 8×8 block losses with 26.0 dB except for [2] (see Table 1). However, [2] is impractical due to the excessive computational complexity. For comparison Fig. 7 left shows the concealed image according to [1] with a PSNR of 24.7 dB. Edges can not be reconstructed and the concealed parts appear blurred. In contrary, the proposed method can restore edges, especially also diagonal ones, which is obvious from Fig.4, right side (e.g. hat).

Table 1 shows the convincing results of the method [2] except for the image Baboon. The algorithm exploits the correlation between pixels in order to predict the missing pixel. If this assumption is not fulfilled like for the fur in Fig. 7, right side, the algorithm fails. This area cannot be predicted appearing in white spots. Additionally, the performance applying method [2] decreases for larger block losses from 20.6 dB to 18.7 dB in contrary to the introduced method. For comparison Fig. 6 shows on the right side the concealed image with the proposed method. The subjective result is even more convincing than the PSNR performance.

	8×8 Block loss			16×16 Block loss		
	Lena	Peppers	Baboon	Lena	Peppers	Baboon
Maximally smooth recovery [1]	24.7 dB	26.0 dB	20.0 dB	23.7 dB	24.2 dB	19.5 dB
POCS [3]	24.7 dB	25.7 dB	19.5 dB	22.3 dB	22.1 dB	18.9 dB
Spatial domain interpolation [4]	24.0 dB	26.1 dB	17.8 dB	21.2 dB	23.3 dB	16.4 dB
Reconstruction [5]	24.5 dB	26.3 dB	18.8 dB	-	-	-
Sequential error-concealment [2]	28.1 dB	29.5 dB	20.6 dB	23.9 dB	26.9 dB	18.7 dB
Introduced algorithm	26.0 dB	26.4 dB	19.4 dB	22.8 dB	24.5 dB	19.1 dB

Table 1. Error concealment techniques in comparison.

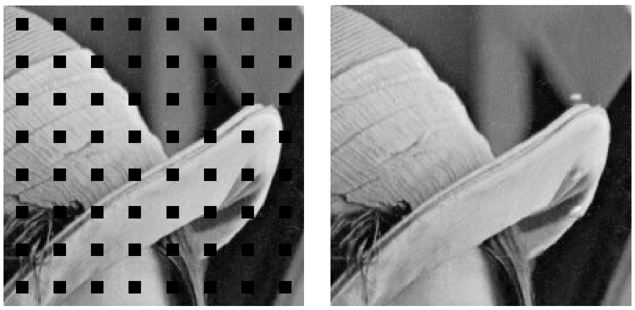

Fig. 4. Left: Simulated 8×8 block loss. Right: Concealed image, frame: 2, iterations: 3.1 per block.

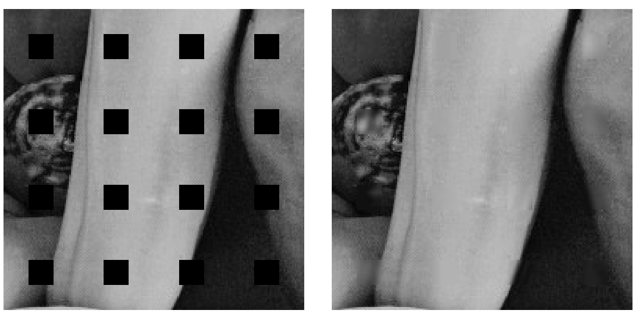

Fig. 5. Left: Simulated 16×16 block loss. Right: Concealed image, frame: 6, iterations: 5.7 per block.

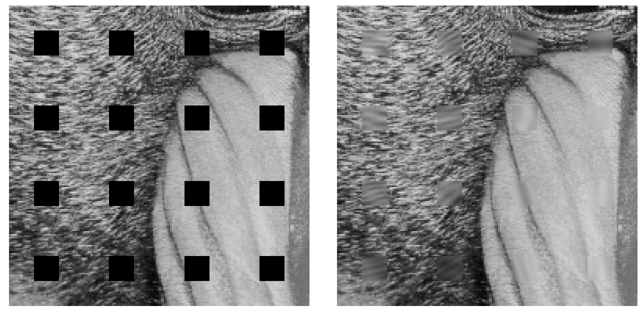

Fig. 6. Left: Simulated 16×16 block loss. Right: Concealed image, frame: 12, iterations: 6.9 per block.

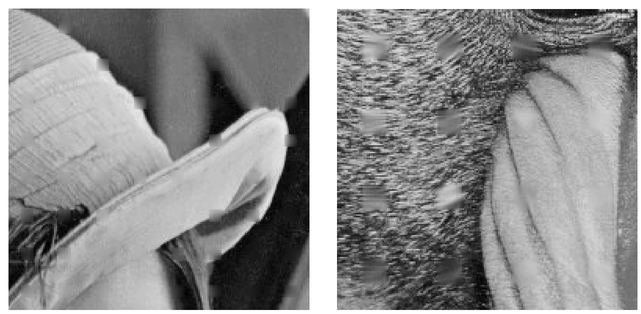

Fig. 7. Left: Concealed with [1]. Right: Concealed with [2].

5. CONCLUSION

We presented a method for spatial error concealment. The computational complexity is in contrary to the reference methods controlled by the image content, i.e. monotone areas require a few, edges and noise-like areas more iterations. The frame of known pixels depends also on the image content and controls the performance. So far the frame size has to be preselected and therefore it is desirable to have an adaptive frame constituting future work.

6. REFERENCES

[1] Y. Wang, Q.-F. Zhu, and L. Shaw, “Maximally smooth image recovery in transform coding,” *IEEE Trans. on Commun.*, vol. 41, no. 10, pp. 1544–1551, Oct. 1993.

[2] L. Xin and M. T. Orchard, “Novel sequential error-concealment techniques using orientation adaptive interpolation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 857–864, Oct. 2002.

[3] H. Sun and W. Kwok, “Concealment of damaged block transform coded images using projections onto convex sets,” *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 470–477, April 1995.

[4] Z. Alkachouh and M. Bellanger, “Fast DCT-based spatial domain interpolation of blocks in images,” *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 729–732, April 2000.

[5] S. Shirani, F. Kossentini, and R. Ward, “Reconstruction of baseline JPEG coded images in error prone environments,” *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1292–1299, July 2000.

[6] A. Kaup and T. Aach, “Coding of segmented images using shape-independent basis functions,” *IEEE Trans. Image Process.*, vol. 41, no. 10, pp. 1544–1551, Oct. 1998.

[7] T. Aach, “Missing data interpolation by transform-based successive approximation,” in *Proc. Workshop on Spectral Methods and Multirate Signal Processing (SMMSP-2001)*, Pula, Croatia, June 2001, pp. 3–10.

[8] R. Sottek, *Modelle zur Signalverarbeitung im menschlichen Gehör*, Ph.D. thesis, University RWTH Aachen, Germany, 1993.

[9] “http://www.lnt.de/lms/research/projects/concealment/,” .