COMPRESSION BY IMAGE EMPIRICAL MODE DECOMPOSITION

Anna Linderhed

Swedish Defence Research Agency, Department of IR Systems, Division of Sensor Technology, Linköping, Sweden
Image Coding Group, Department of Electrical Engineering, Linköpings Universitet, Sweden
anna.linderhed@foi.se

ABSTRACT

Empirical Mode Decomposition (EMD) in two dimensions provides a tool for image processing by its special ability to locally separate spatial frequencies. The tendency is that the Intrinsic Mode Functions (IMFs) other than the first are low frequency images. Variable sampling of the EMD is used for image compression. This is done blockwise using the non-uniformly located extrema points of the IMF to steer the uniform sampling rate of the block. The variable sampling process results in a number of samples for each block. These can be squeezed into blocks of smaller size and DCT coded. The DCT components are then quantized and thresholded, leaving us with even fewer components to represent the block. The coding method presented in this paper is a new approach to image compression that opens doors for future applications.

1. INTRODUCTION

EMD-based time-frequency analysis, called Hilbert-Huang Transform (HHT) [2], is only one of many applications made possible by EMD. The results and ideas in time domain applications using EMD apply to two-dimensional signals, such as images, as well. EMD decomposes the spatial frequency components into a set of IMFs where the highest spatial frequency component of *each spatial position* is in the first IMF and the second highest spatial frequency component of each spatial position is in the second IMF, etc. An IMF is defined as a function in which the number of extrema points and the number of zero crossings are the same or differ by one [2]. In the two-dimensional case this demand is relaxed. The upper and lower envelope of the IMF are symmetric with respect to the local mean, which is used to define the IMF instead of the number of extrema points and zero crossings. In two dimensions there are many possibilities to define extrema, each one yielding a different decomposition. In this work we simply extract the extrema points by comparing the candidate data point with its nearest 8-connected neighbours. Clearly, more sophisticated methods exist, but the extrema points defined by an 8-connected neighbourhood serves the purpose for EMD at this stage, with further improvement being possible. The extension of the EMD to two dimensions relies on proper two-dimensional spline interpolation of the scattered extrema points. We use the thin-plate smoothing spline interpolation [1] for this implementation of two-dimensional EMD. This method gives a surface with continuous second derivative everywhere and turns out to successfully decompose an image into its IMFs and a smooth residue with no or only a few extrema points [3]. This paper starts with a review of the sifting process used in the EMD in Section 2, and a review of the variable sampling in Section 3. In Section 4 these methods are used in an image coder, and results are presented in Section 5.

2. SIFTING FOR THE TWO-DIMENSIONAL IMF

To find the first IMF, start with the image itself as input signal $h_{10}(m, n) = x(m, n)$. The first index is the IMF number, l=1,..,L, and the second index is the iteration number, k=1,..,K, in the sifting process. m and n represent the two spatial dimensions. To find the next IMF, use the residue corresponding to the previously found IMF as input signal $h_{20}(m, n) = r_1(m, n)$.

The sifting process to find the IMFs of a signal $x(m,n)$, comprises the following steps:

(1) Find the positions and amplitudes of all local maxima, and find the positions and amplitudes of all local minima in the input signal.

(2) Create the upper envelope by spline interpolation of the local maxima and the lower envelope by spline interpolation of the local minima. Denote the envelopes $e^+(m, n)$ and $e^-(m, n)$ respectively.

(3) For each position (m,n), calculate the mean of the upper envelope and the lower envelope.

$$\bar{e}_{lk}(m, n) = \frac{e^+(m, n) + e^-(m, n)}{2} \tag{1}$$

The signal $\bar{e}_{lk}(m, n)$ is referred to as the envelope mean.

(4) Subtract the envelope mean signal from the input signal

0-7803-9134-9/05/$20.00 ©2005 IEEE

Figure 1. The Lenna image at 128x128 pixel size.

$$h_{lk}(m, n) = h_{l(k-1)}(m, n) - \bar{e}_{lk}(m, n) \qquad (2)$$

This is one iteration of the sifting process. The next step is to check if the signal $h_{lk}(m, n)$ from step 4 is an IMF or not. The process stops when the envelope mean signal is close enough to zero as proposed in [5].

$$|\bar{e}_{lk}(m, n)| < \varepsilon \qquad \forall(m, n) \qquad (3)$$

The value of ε in the stop criterion affects the EMD in such a way that if it is not small enough there will not be a sufficient number of IMFs to separate all intrinsic modes in the signal. On the other side, if the number ε is too small the iterations will take long time.

Forcing the envelope mean to zero will give us the wanted symmetry of the envelope and the correct relation between the number of zero crossings and the number of extremes that define the IMF. This way we will find the IMF without actually having to check for symmetric envelopes.

(5) Check if the mean signal is close enough to zero, based upon the stop criterion. If not, repeat the process from step 1 with the resulting signal from step 4 as the input signal a sufficient number of times.

When the stop criterion is met the IMF $c_l(m, n)$ is defined as the last result of (4).

$$c_l(m, n) = h_{lK}(m, n) \qquad (4)$$

After the IMF is found, define the residue $r_l(m, n)$ as

$$r_l(m, n) = h_{l0}(m, n) - c_l(m, n) \qquad (5)$$

(6) The next IMF is found by starting over from step 1, now with the residue as the input signal.

$$h_{(l+1)0}(m, n) = r_l(m, n) \qquad (6)$$

Steps (1) to (6) can be repeated for all the subsequent r_j. The EMD is completed when the residue, ideally, does not contain any extrema points. The signal can be expressed as the sum of IMFs and the last residue

$$x(m, n) = r_L(m, n) + \sum_{j=1}^{L} c_j(m, n) \qquad (7)$$

The Lenna image, shown in Figure 1, is decomposed with the EMD method described above. The image's four IMFs and their corresponding residues are shown in Figure 2. The last residue has only a very few extrema points.

3. VARIABLE SAMPLING OF OVERLAPPING BLOCKS

EMD is a truly empirical method, not based on the Fourier frequency approach but related to the locations of extrema points and zero crossings. Based on this we use the concept of *empiquency* [4], short for empirical mode frequency, instead of a traditional Fourier-based frequency measure to describe the signal oscillations. The special property of the IMF that the empiquency varies is used to control the variable sampling of the EMD. Areas with many extrema points have high empiquency, while areas with a few or no extrema points have low empiquency. The IMFs are smoother than the image itself; only the first IMF holds the nonsmooth parts of the image. This means that it should be possible to subsample the IMFs. Due to the different empiquencies in the different parts of the IMF, the subsampling can be different in different parts of the IMF. The significant (with respect to a defined zero zone) extrema points define the maximum empiquency in the IMF [3]. Maximum empiquency is found by examining the space between the significant extrema points. As suggested in [4] we treat the IMF blockwise. This way the sampling rate for each block can be defined according to its empiquency content. The high-empiquency blocks that cannot be subsampled are not modified. The remaining ones are subsampled.

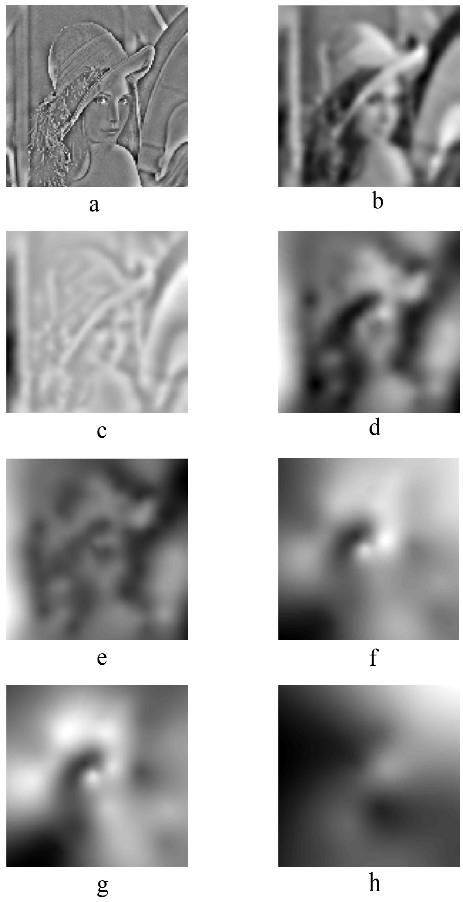

Figure 2. a) First IMF, b) first residue, c) second IMF, d) second residue, e) third IMF, f) third residue, g) fourth IMF, h) fourth residue.

In the implementation of the blocking process we choose to use overlapping blocks of size 7x7 pixels, details can be found in [3]. The purpose of the overlap is to minimize the artifacts from the blocking and to further reduce

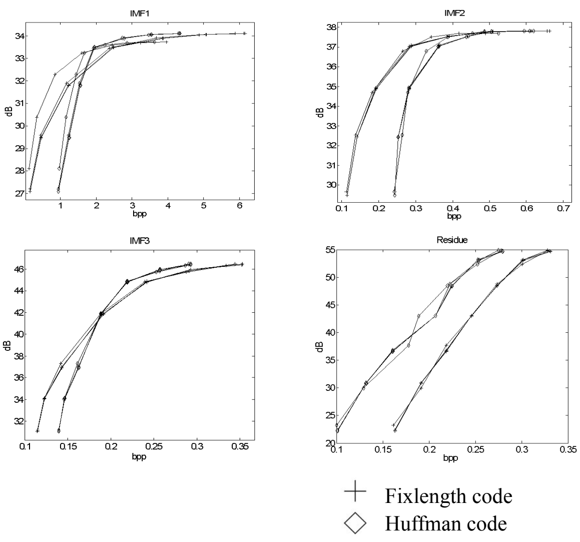

Figure 3. VSDCTEMD Coding result of Detail two. Zero zone=1, 10, 20 quantization varies from 256 levels down to 8 levels

the number of the samples used to represent the IMF. The size of the block is chosen so that the corners of the block are always represented, regardless of the chosen sampling rate. The sampling pattern within a block consists of every pixel, every second pixel, every third pixel, and every sixth pixel in both directions, to represent 1/1, 1/4, 1/9, and 1/36 of the pixels, respectively. The overlapping pixels in two neighbouring blocks will be the same, but used several times. This will ensure that the concatenated blocks have the same values at the edge pixels. For the reconstruction the uniformly sampled points of the block are connected by a surface created by the use of an interpolating cubic spline extended to two dimensions.

4. CODING OF THE EMD USING DCT OF THE VARIABLE SAMPLED BLOCKS (VSDCTEMD)

This image coder use the basic concept of the EMD for the decomposition of the image and apply the variable sampling and DCT coder on the IMFs and residue. A similar approach is used in [6] but with a wavelet coder instead of a DCT coder. Each IMF is divided into overlapping blocks. The maximum empiquency in the block determines the sampling rate for the subsampling of the block. The samples are represented by one sample alone or 6x6, 3x3, or 2x2 blocks of samples. These are DCT coded and the components are quantized and thresholded before a two-bit block header is added. The resulting component stream is Huffman coded or fixlength coded.

Reconstruction is done by reading the two-bit header which indicates the sampling rate used. With this information the components achieved from the inverse DCT transform of the stream can be placed in their proper place in the 7x7 pixel block. The stream only contains samples for the 6x6 block, the missing samples are found in the reconstructed neighbouring blocks. Since the missing samples are located in the rightmost column and the lowest row of the block, the reconstruction starts with the block in the lower right corner working row-wise through the blocks. For the blocks with no neighbours holding missing samples dummy samples are used. The 7x7 blocks of samples are interpolated and the non-overlapping 6x6 part of the reconstructed block is used to generate the output IMF. The image is then reconstructed by the adding of reconstructed IMFs and the reconstructed last residue.

We have found [3][5][6] that the first IMF is almost as hard to compress as the image itself while the rest of the IMFs and all the residues are smooth and can be effectively represented by only a small part of a full size DCT. In a slightly modified version of the coder presented here we only decompose the image into one IMF and one residue. We use the VSDCTEMD on the first IMF and threshold coding of the full size DCT on the first residue.

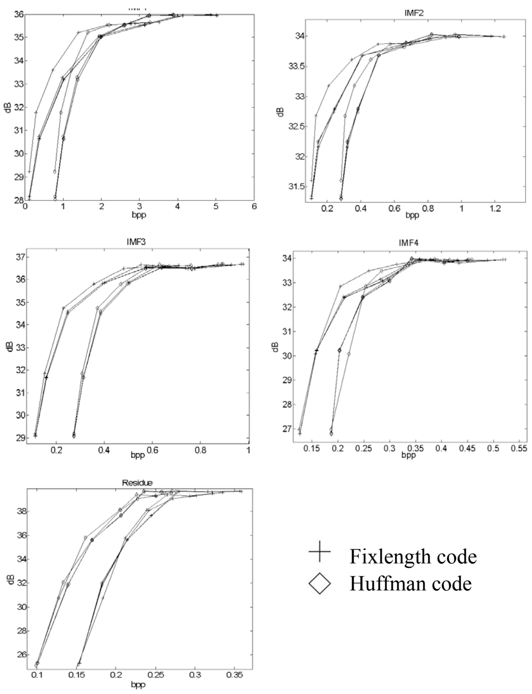

Figure 4. VSDCTEMD Coding results of Detail one. Zero zone=1, 10, 20 quantization varies from 256 levels down to 8 levels.

5. RESULTS

The VSDCTEMD coder is tested on two different images, Detail one and Detail two, using zero zones 1, 10, and 20. The quantization varies from 256 levels down to 8 levels. The DCT components are thresholded leaving only those having a value larger than or equal to the threshold. The threshold varies from 0.1% to 100% of the maximum value. Figure 3 and Figure 4, show the coding results for the EMD of the test images, respectively. The coding method gives good results for all IMFs and the residue. Even for the difficult first IMF the result is over 30 dB for a bitrate of 1 bpp. In Figure 5 example reconstructions of the images are presented along with the original image.

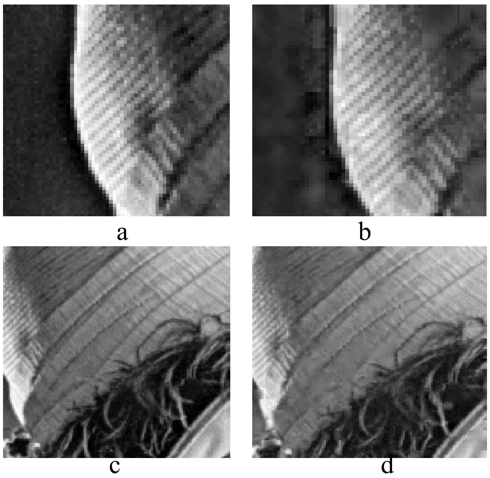

Figure 5. VSDCTEMD. a) Detail one original, b) Detail one image reconstructed to 28.27 dB using 4.79 bpp. c) Detail two original, d). Detail two image reconstructed to 30.76 dB using 3.56 bpp.

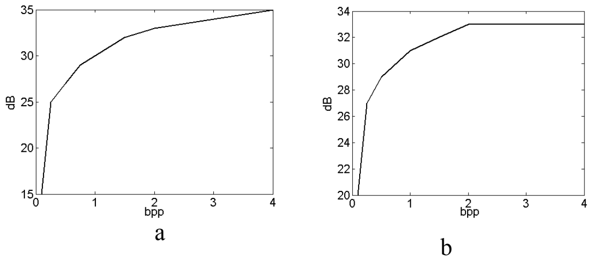

Figure 6. Result of combining the coding by VSDCTEMD in first IMF and DCT threshold coding on first residue, a) Detail one b) Detail two.

Although the coding result of each IMF and the residue are satisfactory, the image reconstruction by the addition of the reconstructed IMFs and reconstructed residue sum the bitrates to large numbers. Detail one image reconstructs to 28.27 dB using 4.79 bpp and Detail two image reconstructs to 30.76 dB using 3.56 bpp.

The result of the coding of the first IMF using VSD-CTEMD and DCT threshold coding on the first residue is shown in Figure 6. Two different reconstruction examples for each image are presented in Figure 7. This way the sum bitrate can be held low. The Detail one image reconstructs to 30.09 dB using 1.00 bpp, or 29.23 dB using 0.79 bpp, and the Detail two image reconstructs to 30.02 dB using 1.28 bpp, or 28.88 dB using 0.69 bpp.

6. SUMMARY

The EMD in two dimensions provides a tool for image processing. We show how an image can be decomposed into a set of IMFs and a residue with a minimum of extrema points. The tendency is that the IMFs other than the first are low frequency images. This can be used for coding by the DCT of the whole image. The variable sampling of the EMD has been used for image compression, blockwise, using the non-uniformly located extrema points of the IMF

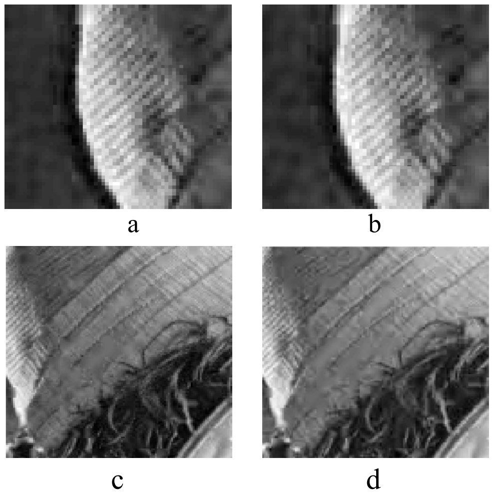

Figure 7. Result of coding by VSDCTEMD in first IMF and DCT threshold coding on first residue.a) Detail one 1.00 bpp 30.09 dB, b)Detail one 0.79 bpp 29.23 dB, c) Detail two 1.28 bpp 30.02 dB, d) Detail two 0.69 bpp 28.88 dB

to steer the uniform sampling rate of the block. The DCT components from the variable sampling of each block are quantized and thresholded, leaving us with even fewer components to represent the block. The result from the full VSDCTEMD is not impressive due to the addition of many IMFs, while for the use of VSDCTEMD on only the first IMF and a threshold coded full sized DCT on the first residue we get at better result. These results are of the same quality as the results presented in [6].

7. REFERENCES

[1] F. L. Booksstein, “Principal warps: Thin-plate splines and the Decomposition of deformations”, *IEEE Trans. on Pattern analysis and machine intelligence*, vol 11, no 6, june 1989.

[2] N. E. Huang et al. ”The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”, *Proc. R. Soc. Lond. A* , 1998.

[3] A. Linderhed, “Adaptive Image Compression with Wavelet Packets and Empirical Mode Decomposition”, PhD thesis, Linköping University, Sweden, 2004.

[4] A. Linderhed, “Variable Sampling of the Empirical Mode Decomposition of Two-Dimensional Signals”, *International Journal of Wavelets, Multiresolution and Information Processing (IJWMIP) special issue on Sampling Problems Related to Wavelet Theory and Time-Frequency Analysis*, forthcoming.

[5] A. Linderhed, “2D empirical mode decompositions in the spirit of image compression”, *Wavelet and Independent Component Analysis Applications IX* , SPIE Proceedings Vol. 4738, p. 1-8, April 2002.

[6] A. Linderhed, “Image compression based on empirical mode decomposition”. *Proceedings of SSBA symposium on image analysis,* SSBA’04, Uppsala, p. 110-113, 11-12 March 2004.