2-D empirical mode decompositions - in the spirit of image compression

Anna Linderhed

Image Coding Group, Department of Electrical Engineering, Linköpings Universitet
S-58183 Linköping, SWEDEN
anna@isy.liu.se

Swedish Defence Research Agency, P.O. Box 1165, SE-581 11 Linköping, Sweden
anna.linderhed@foi.se

ABSTRACT

The Empirical mode decomposition (EMD) is an adaptive decomposition of the data, as is the *Wavelet packet best basis* decomposition. This work present the first attempt to examining the use of EMD for image compression purposes. The Intrinsic Mode Function (IMF) and their Hilbert spectra are compared to the wavelet basis and the wavelet packet decompositions expanded in each of its best bases on the same data. By decomposing the signal into basis functions, the waveforms in the signal is represented by the basis and a set of decorrelated discrete values in a vector. A coding scheme is presented where the idea is to decompose the signal into its IMF:s where only the max and min values for each IMF is transmitted. The reconstruction of the IMF in the decoder is done with spline interpolation. We have in the two-dimensional EMD an adaptive image decomposition without the limitations from filter kernels or cost functions. The IMF:s are, in the two-dimensional case, to be seen as spatial frequency subbands, with various center frequency and bandwidth along the image.

1. INTRODUCTION

The Empirical mode decomposition (EMD) is an adaptive decomposition with which any complicated signal can be decomposed into its Intrinsic Mode Functions (IMF:s) for which we can have a well defined instantaneous frequency. The original purpose for the EMD was to find a decomposition which made it possible to use the instantaneous frequency, defined as the derivative of the phase of an analytic signal, in the time frequency analysis of the signal [Huang]. In this paper the focus is on compression, but we start with a description of the original work and its purpose as we understand it. In chapter 2 is the EMD and the sifting process for the IMF described. In chapter 3 we present how a one dimensional coder can be implemented and last, in chapter 4, the extension to two dimensions is done.

1.1 Time-frequency analysis

Whenever we use any form of the fourier transform to represent frequencies we are limited by the uncertainty principle. For an infinite time signal we can have exact information of the frequencies in the signal but the moment we restrict ourselves to analyse a signal of finite time length we have put a bound on the precision in the frequencies we can discover. The uncertainty principle formulated as

$$\Delta f \cdot \Delta t \geq \frac{1}{4\pi}$$

defines the relation between the bandwidth Δf and the size of the time window Δt in which the signal is represented. In the fourier transform the time window Δt is infinitely long hence the frequency information is exact, the bandwidth Δf is infinitely narrow. In the short time fourier transform the time window Δt is finite and that means that the bandwidth Δf have to be wider in order to fulfil the uncertainty principle. In subband representation of a signal the subsampled *approximation* of the original image is produced by the correlation of the signal and a low pass filter and a following subsampling by a factor two, the *detail signal* is produced by a high pass filtering of the original signal with a high pass filter and a following subsampling by a factor two. Applied recursively to the approximation part of the resulting signal we have the hierarchical subband representation. This gives good frequency information for the lower frequencies and good time

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

information for the higher frequencies. The reconstruction or synthesis of the signal from the subband coefficients is done by upsampling and filtering with the reconstruction filters. The wavelet transform is now the name used for the subband decomposition, the difference compared to subband coding is that the wavelet theory has equipped us with a tool for proper filter construction [Daubechies], and the wavelet transform uses scaled and translated versions of a prototype wavelet as basis functions to represent a signal. The best basis techniques was introduced by Coifman as a natural extension to the wavelet techniques [Coifman]. The wavelet packet tree is constructed when each subband is recursively filtered and sub-sampled into two new signals. The wavelet packet tree representation generates a complete hierarchy of segmentations in frequency and it is a redundant expansion of the image. The best basis is chosen by a cost function applied to each of the nodes, deciding if it is to be kept or split.

The instantaneous frequency represents the frequency of the signal at one time, without any knowledge needed about the signal at other times. There exists different definitions for instantaneous frequency, here we use the derivative of the phase of the signal as described in [Huang]. This is not possible for real signals so there is a need for creating a complex valued signal to represent our real valued signal. Another problem with using the instantaneous frequency is that it only provides one value at each time. A signal usually consists of many intrinsic frequencies and this is where the EMD is used, to decompose the signal into its IMF where only one frequency component is present at each time so we can have a well defined instantaneous frequency. By applying the Hilbert transform to each IMF we get a set of analytical signals representing our input signal. We can plot the instantaneous frequencies as defined for each IMF by the Hilbert spectrum.

1.2 Example

We create a composite chirp by adding together a linear chirp for 2 seconds at 1kHz sample rate with start at DC and cross 150Hz at t=1sec, and a "convex" quadratic chirp for 1 second at 1kHz sample rate with start at 25Hz increasing to 100Hz, and a "concave" quadratic chirp for1 second at 1kHz sample rate with start at 100Hz decreasing to 25Hz. In Figure 1.1 is the time frequency plot of the composite chirp signal by the windowed fourier transform to the left and the corresponding plot of the wavelet packet best basis decomposition of the same signal to the right.

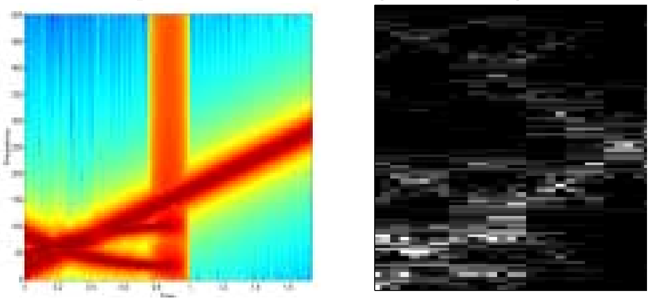

Figure 1.1. Time-frequency plot of the composite chirp signal by the windowed fourier transform to the left and the corresponding plot of the wavelet packet best basis decomposition of the same signal to the right.

Both the *short time fourier transform* and the *wavelet packet best basis* representation suffers from the uncertainty principle. In the left plot the signal response is smeared in the frequency axis as a result from the narrow time window and hence a broad frequency window used to analyse the signal. In the right figure the wavelet packet best basis representation is plotted. Here all the small boxes are of the same area and the grayscale indicates the correlation of the signal at the time and the analysis filter associated to the box. The form and position of the box is dictated by the *best basis* search algorithm originally designed by [Coifman].

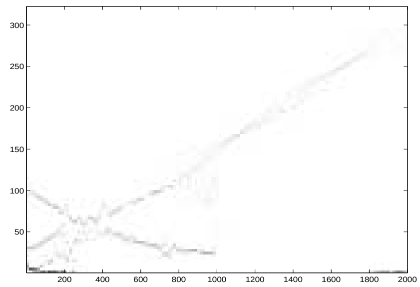

Figure 1.2. The hilbert spectrum of the composite chirp signal.

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

The Hilbert spectrum in Figure 1.2 is composed by Hilbert spectra of each IMF of the composite chirp signal shown in Figure 2.3. The instantaneous frequency is plotted against the time axis and it shows not only the inter frequencies generated by the construction of the signal, but also the intra frequencies which comes from the change of curve form by the frequency change and the addition of signals.

2. EMPIRICAL MODE DECOMPOSITION

It is the EMD itself that is of interest from the compression perspective. It is a totally adaptive decomposition of the signal into its intrinsic modes, independent of any filters, cost functions or uncertainty principles. The IMF:s catches the signal components in a very compact way in well behaved signals leading to a clean representation by a few components.

2.1 Sifting process for IMF

A function is an IMF if it fulfills the following demads: It has the same number of zero crossings and extrema. Or the number of zero crossings and extrema differs only by one. Futher, the envelopes defined by the local maxima and minima respectively are symmetric.

The IMF is found by the sifting process:

1. Find all the local max points and all the local min points of the signal.
2. Create upper envelope by spline interpolation of the local maxima and the lover envelope by spline interpolation of the local minima of the input signal.
3. For each time, take the mean of the upper envelope and the lover envelope.
4. Subtract the mean signal from input signal.
5. Check if mean signal is close enough to zero. If not, repeat the process from 1. with the result signal from 4. as input signal. If it is the result is IMF, define the residue as result from IMF subtracted from input signal
6. Find next IMF by starting over from 1. with the residue as input signal.

The process for finding the first IMF is shown in Figure 2.1

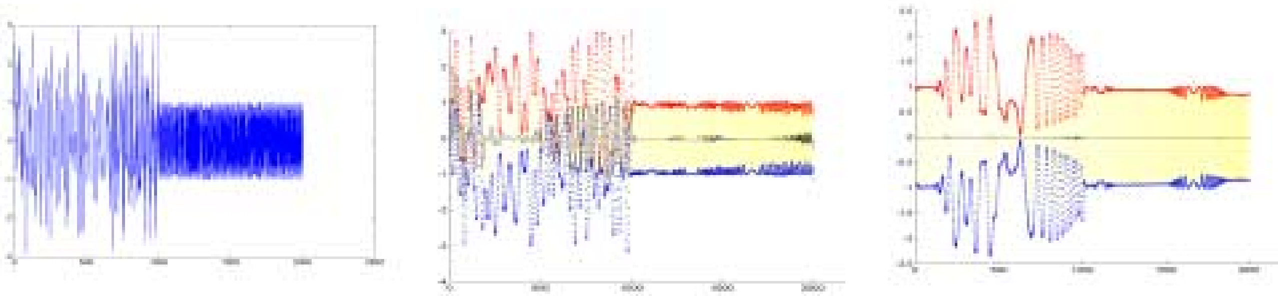

Figure 2.1. The composite chirp to the left and the result from one iteration in the IMF sifting process with the upper and lover envelope and their mean plotted in the same figure in the middle and the resulting first IMF with the upper and lover envelope and their mean plotted in the same figure to the right.

In the original work by [Huang] the stop criteria for the sifting process is when the difference between two consecutive siftings is smaller than some selected threshold. For compression purposes the stop criteria will be based on when the envelope mean signal is close enough to zero. There is two reasons for this, first it is easier to calculate and, second, forcing the envelope mean to zero will guarantee the symmetric envelopes and the demand on number of zero crossings and number of extremas that defines the IMF.

When one IMF is found the residual is created by subtracting this IMF from the input signal of the sifting process. The next IMF is found by the sifting process with the residual as the new input signal. The last residual does not have any extremas, it is a constant or a trend.

In Figure 2.2 the first 6 IMF:s are plotted together with the residual and the signal on top for the composite chirp. The

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

compact representation of the signal is clear when we compare the IMF:s to the corresponding wavelet decomposition. Only the three first IMF:s contribute substantially to the signal representation while all six of the wavelet detail bands is needed, the residual of the EMD is a constant while the approximation signal is essential for the signal representation in the wavelet case.

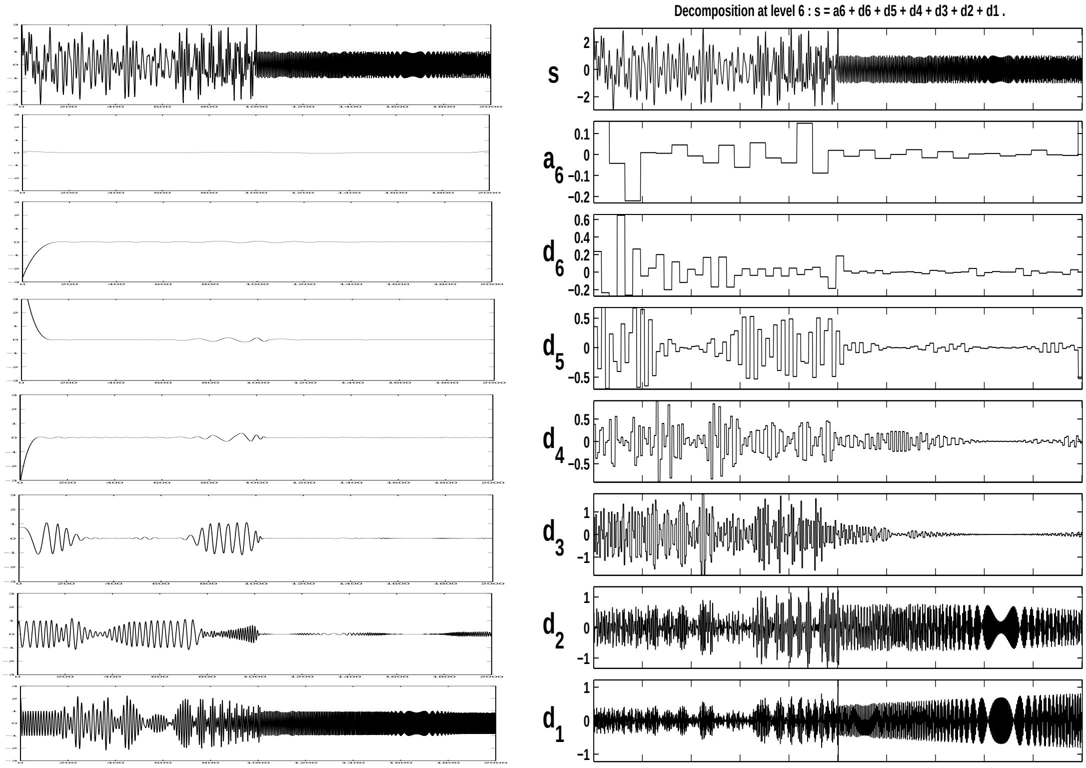

Figure 2.2. The first 6 IMF:s, the residual and the signal on top for the composite chirp to the left and the wavelet decomposition of the same signal to the right.

The same is shown in the Hilbert spectra of the IMF:s, only the first three shows frequency content.

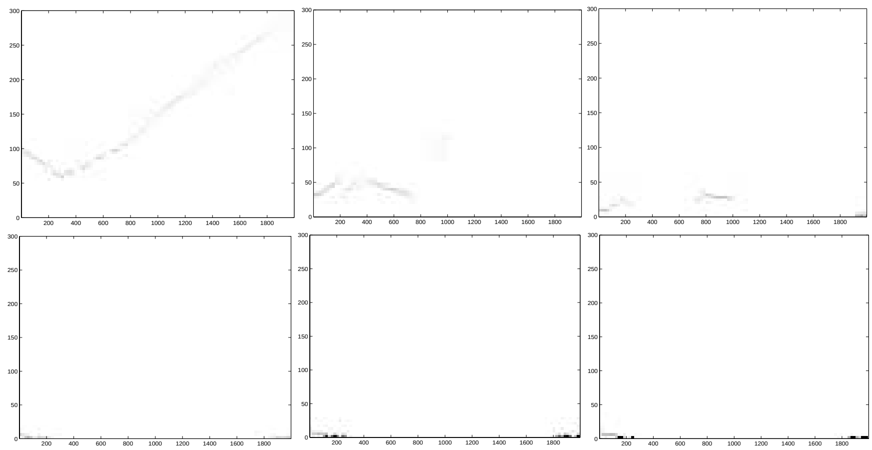

Figure 2.3. The Hilbert spectrum of the IMF:s respectively, the first on the top left to the sixth on the bottom right.

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

3. DECOMPOSING FOR COMPRESSION

3.1 The IMF as a subband

The first IMF catches all the highest frequencies in the signal. The range of frequencies contained in the IMF varies over time but the first IMF always contain the highest frequency in the signal at each time. The EMD decomposes the signal into its intrinsic modes instead of into its correlations to a filter as is done in transform decompositions like Discrete Fourier Transform (DFT) and wavelet transform for example. This gives a very compact representation in that there are no more decomposition signals than there are components in the signal. The IMF:s would be nice to use encoding audio. The basis function (IMF) follows the frequency variations of the signal instead of trying to approximate the function linearly with fixed basis functions.

3.2 Coding

A transform can be seen as a change of basis, a change of point of view. Only the more uncorrelated representation in the transform domain render possible compression of the signal. Even more compression is achieved if the transform coefficients are quantized and some intelligent source coding algorithm is applied [Forchheimer]. By decomposing the signal into basis functions, the waveforms in the signal is taken out into the basis and the signal is represented by a set of decorrelated discrete values in a vector.

The idea is to decompose the signal into its IMF:s and transmit only max and mean values for each IMF. Then reconstruct the IMF in the decoder with spline interpolation. The scheme is shown in Figure 3.1.

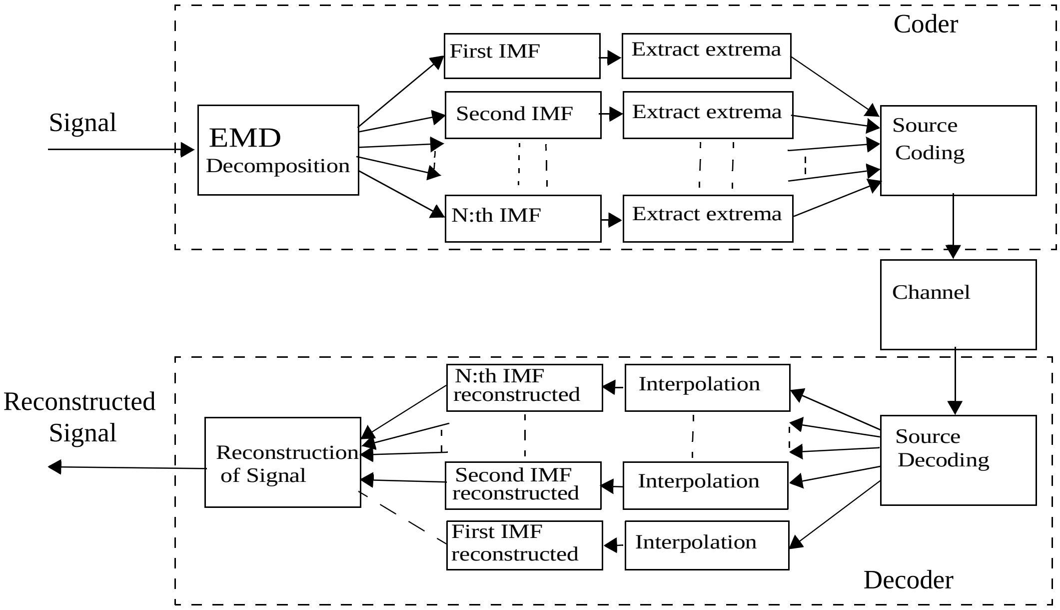

Figure 3.1. The EMD coding scheme.

To get a feeling for the behaviour of the IMF:s in a coding application, the entropy where measured for the IMF:s themselves, their extrema vector and for the quantized extrema vector and listed in Figure 3.2. The entropy of a signal serves as a indicator of performance of a lossless compression algorithm. The sum of entropy for the IMF:s is much higher than the entropy for the signal itself. Yet, the IMF:s are well behaved signals and the lossless compression algorithm ZIP compresses the much longer sequence of IMF:s into the same number of bytes as when compressing the original signal. Choosing to use only the extrema points gives the opportunity to code very sparse signals. The IMF extrema vectors are plotted in Figure 3.3. As can be seen, both in the plots and in the listing in Figure 3.2, the number of extremas is low compared to the length of the original signal to code. The fewer extrema points to code the better. The cost of coding these sparse signals is estimated by the entropy of the extrema in Figure 3.2 and even more compression is achieved with quantization.

The control of the envelope mean in the sifting process is important. It controls the symmetricity of the envelopes and a properly executed sifting process generates fewer IMF:s because the difference between a real IMF and the IMF produced by the sifting process is tranferred to the sifting for the next IMF in the residual.

Signal	**# sample**	**entropy**	**# extrema**	**entropy for extrema**	**entropy for quantized extrema**
Composite chirp	2001	7.08			
IMF 1	2001	7.19	637	2.72	2.03
IMF 2	2001	6.73	345	1.56	0.99
IMF3	2001	5.13	189	0.68	0.23
IMF4	2001	3.37	101	0.46	0.11
IMF5	2001	2.69	54	0.28	0.04
IMF6	2001	3.41	33	0.20	0.02

Figure 3.2. Listing of entropies of the IMF:s and their extrema vectors.

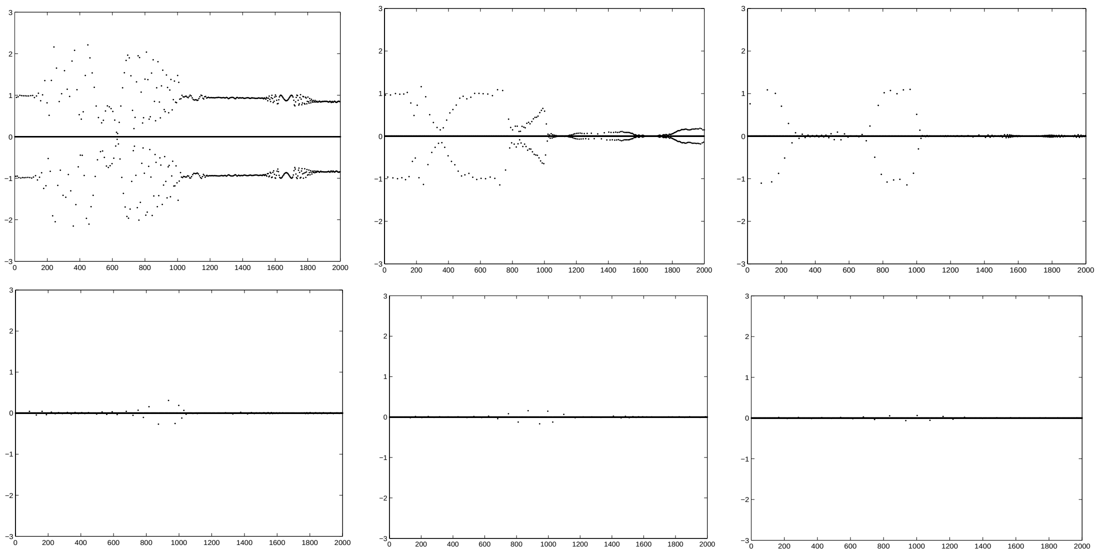

Figure 3.3. The extremas of the IMF:s respectively, the first on the top left to the sixth on the bottom right.

3.3 Reconstruction

The reconstruction is done by a spline interpolation of the coefficients received and then add the reconstructed IMF:s together. The interpolation introduces an error plotted in Figure 3.4. This is because the IMF itself is not a spline function but approximated by a spline interpolation of the extrema points when reconstructed. When no quantization is done on the

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

extrema vectors and no source coding is done, the spline interpolation causes the least reconstruction error of several tested interpolation methods. In combination with quantization of the extrema vector and some source coding algorithm another interpolation may be better.

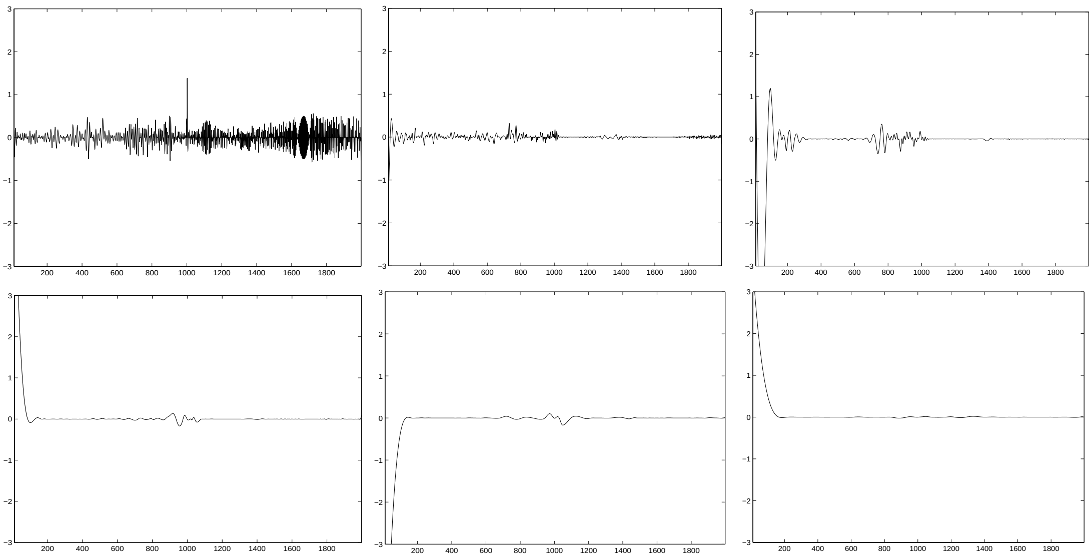

Figure 3.4. The reconstruction error for the IMF:s respectively, the first on the top left to the sixth on the bottom right.

4. IMAGE EMD

The results and ideas in chapter 3 applies to two-dimesional signals as well. The extension of the EMD to two dimensions rely on proper spline interpolation in two dimensions. The edge constraints are important as the errors introduced there traverses into the image. We have in the two-dimensional EMD an adaptive image decomposition without the limitations from filter kernels or cost functions. The IMFs are to be seen as spatial frequency subbands, with various center frecuency and bandwidth along the signal. A test image and its first four IMF:s are shown in Figure 4.1 and Figure 4.2. It shows the same rapid decomposition into the intrinsic images and not many IMF:s are needed to describe the image. Besides the coding scheme presented in chapter 3, the image IMF:s are well behaved functions and would be easier to compress with standard image compression algorithms modified to work in specified frequency bands defined by the IMF.

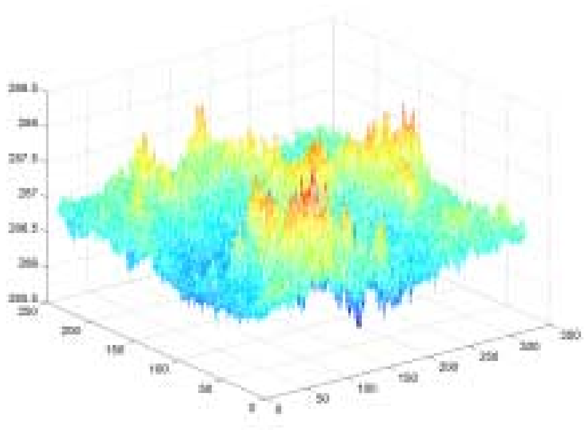

Figure 4.1. Infrared photo of a minefield.

Presented at SPIE International symposium on Aerosense, Orlando, Florida, USA, 1-5 April, 2002. Proceeding of SPIE,Wavelet and Independent Component Analysis Applications IXI, Volume 4738, ISBN 0-8194-4488-X

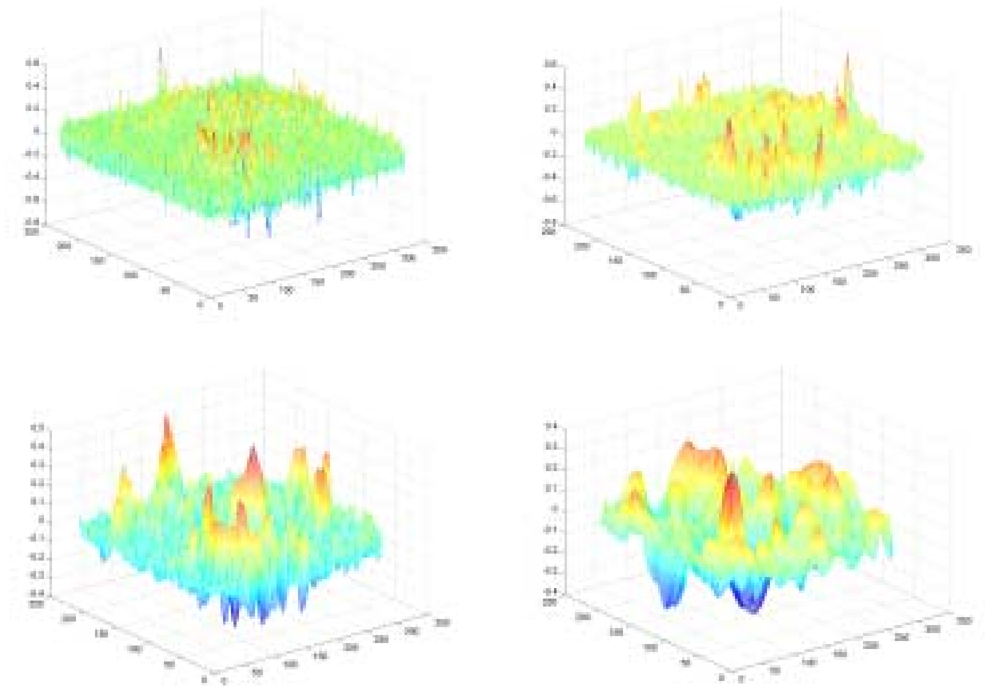

Figure 4.2. The first four IMF:s of the image in Figure 4.1.

5. CONCLUSION

It is the EMD itself that is interesting in the compression perspective. It is a totally adaptive decomposition of the signal into its intrinsic modes, independent of any filters, cost functions or uncertainty principles. The IMF:s catches the signal components in a very compact way, in well behaved signals, leading to a clean representation by a few components. By decomposing the signal into basis functions, the waveforms in the signal is represented by the basis and a set of decorrelated discrete values in a vector. A coding scheme is presented where idea is to decompose the signal into its IMF:s and transmit only max and mean values for each IMF. Then the reconstruction of the IMF in the decoder is done with spline interpolation.

We have in the two-dimensional EMD an adaptive image decomposition without the limitations from filter kernels or cost functions. The IMFs are to be seen as spatial frequency subbands, with various center frequency and bandwidth along the signal.

REFERENSES

[Coifman] Coifman, R.R.; Wickerhauser, M.V. Entropy-based algorithms for best basis selection., IEEE Transactions on Information Theory, Volume: 38 Issue: 2 Part: 2,March 1992 Page(s): 713 -718

[Daubechies] Ingrid Daubechies. Orthonormal Bases of Compactly Supported Wavelets. Comm.of Pure and Applied Mathematics, Vol XLI 909-996 (1988)

[Forchheimer] Robert Forchheimer. Image Coding and Data Compression. Dept. of Electrical Engineering. Linköping University. Linköping 1993.

[Huang] Huang N. E. et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A (1998) 454, 903-995.