Pyramids for image compression with neural networks interpolators *

B. Simon, B. Macq and M. Verleysen †

April 1, 1993

Abstract

We present a new pyramidal decomposition of images for compression purposes. This new transformation is based on the Laplacian transform, where the classical half-band interpolators have been replaced by multi-layer perceptrons. The obtained scheme allows still a multiresolution access to the image, while its decorrelative properies have considerably been improved. Experimental results on real pictures show the validity of this approach.

1 Introduction

Pyramidal decomposition introduced by Burt and Adelson [1], has emerged as a powerful tool for signal analysis and compression. It has been shown in [2] that such a representation is consistent with the human visual perception. The pyramidal structure, as illustrated in Figure 1, has two operators.

- Operator D is a decimator, which takes a sequence x and produces a low-pass time-compressed version y. We will concentrate here on two-fold decimators. Generally the decimators are implemented by an half-band filter followed by a downsampler. The aim of the filter is to avoid aliasing artifacts on the decimated signal [3].
- Operator I is an interpolator, which produces an upsampled version z of its input y. We will also concentrate on two-fold interpolators. Generally, interpolators are implemented by upsamplers (which merely inserts zero-valued samples between adjacent samples of the input sequence y) followed by an half-band filter.

A difference signal d is computed between x and z. d can be viewed as the ***unpredictable*** information from the lower resolution. An important feature of the pyramidal decomposition is that the output signals, y and d permit a perfect reconstruction of the original signal x. A multiresolution representation is achieved by recursively applying the structure to y.

An important application of the Laplacian pyramid is image compression. For such an application, the encoded signals are d and y rather than x for the following reasons.

*The main part of this text has also been presented in ESANN'93

†REGARDS, Télécommunications et Télédétection, Université Catholique de Louvain, Place du levant 3, 1348 Louvain-la-Neuve, BELGIUM

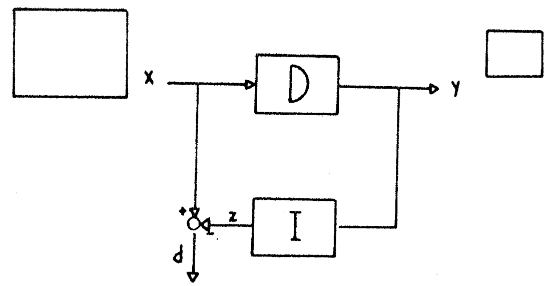

Figure 1: Laplacian pyramid

- d is largely decorrelated and so may be represented pixel per pixel with many fewer bits than x.
- y is a low-pass version of x with a reduced sample rate on which the same decorrelative process can be applied.
- When d is more roughly encoded, in order to achieve higher compression rate, the coding artifacts are less visible because the d signal is tuned with the multiresolution structure of the human visual system.

The Laplacian pyramid was the first description of a multiresolution image compression scheme : the code is constitued by an image at a low resolution and by the details at each further resolution. Such a coding is, as an example, very important for compatible high-definition TV coding [3] (the TV signal is directly reachable without decoding the whole bit stream of the HDTV signal). Further refinements of multiresolution image coding schemes were proposed in the litterature, among which the most popular is the wavelet transform [4].

Recently, it has been proposed to adapt the multiresolution transform to the signal contents : the main idea was to minimize the energy of the d signal by taking into account the correlation factors of the original signal, x. This work was performed either for wavelet transforms [5] and for Laplacian pyramid [6]. In both those two cases the decimators and the interpolators are implemented by linear filters. We suggest here to follow the same path (minimize the variance of d) by using

- linear decimators, in order to keep the low-pass version of x without anoying aliasing artifacts,
- interpolators implemented with multi-layer perceptrons, in order to improve the minimization of the variance of d.

As it will be shown, the aim is reached, with furthermore, an interpolated version of x, namely z, considerably more significant.

While other transforms implemented by artificial neural networks have already been studied for image compression purposes, as in [7] and [8], our approach is original in the sense that it keeps the multiresolution decompsition free of aliasing artifacts by using linear decimators.

The interpolators presented in this paper have been studied in the frame of the Laplacian pyramid for image compression. However they lead also to useful tools for image zooming.

2 Linear half-band interpolators

In classical Laplacian pyramids the decimators and interpolators are implemented by linear filters. Those operators are generally described in the Fourier domain. Let us denote by w the impulse response of the interpolation and decimation filters (we suppose as in [1] that they are the same). Denoting by $X(\Omega), D(\Omega), Z(\Omega)$ and $W(\Omega)$ the (2-D) Fourier transforms of x, d, z and w we have the following relations

$$Z(\Omega) = W(\Omega)\left(W(\Omega)X(\Omega) + W(\Omega - \pi)X(\Omega - \pi)\right) \tag{1}$$

and

$$D(\Omega) = X(\Omega)\left(1 - W^2(\Omega)\right) - W(\Omega)W(\Omega - \pi)X(\Omega - \pi) \tag{2}$$

We see that the first term of D is a half-band highpas filtered vesrion of X, while the second term is due to aliasing. Therefore, the D signal represents more or less the information of the input signal X in the pulsation interval $[\pi/2, \pi]$. According to the Fourier description, and in accordance with the Cramer-Loève theorem, the successive D signal would be perfectly decorrelated if the w filters are perfect half-band.

3 Limits of the half-band approach

The half-band approach results from the Fourier analysis. In the Fourier domain, it is supposed that the signals are stationary and unlimited in the spatial domain. Such a model is however very poor when it is required to describe features like contours, blobs or fractal textures. In such cases the image features generate a wide range of non-structured frequencies. This is the reason why we have searched for a wider range of operators than linear filters. Using multilayer perceptrons, we hope to handle the signal with an internal representation in the operators taking into account non-linearities like contours and particular textures.

At this stage of our study we have used a three-layer perceptron as an interpolator in a given direction. We have applied it separately in the vertical and horizontal directions. The operators are adapted to the signal (minimization of the variance of the d signal) by the backpropagation algorithm [9]. Further studies are planned in order to bind internal states of the interpolator to the features of the image.

4 Multi-layer perceptrons for interpolation

We have implemented separable interpolators using one hidden layer perceptrons. Each of the vertical and horizontal direction interpolators have 5 inputs 20 internal nodes and 2 ouputs. The weights of these NN are adapted by the well-known backpropagation algorithm. The output of the i-th neuron in the k-th layer is given by

$$v_k(j) = f\left[\sum_j w_{k-1}(i,j)v_{k-1}(i)\right] \tag{3}$$

where $w_{k-1}(i,j)$ is the weight associated to the link between the nodes $v_k(j)$ and $v_{k-1}(i)$ and $f[]$ is the hyperbolic tangent function.

The inputs nd ouputs are in the range $[-1, 1]$.

Our aim is to minimize the mean square of d, i.e.

$$\sigma_d^2 = \sum_{i,j}(x(i,j) - z(i,j))^2 \tag{4}$$

Such an error function allows to use the backpropagation algorithm and the weights can then be adjusted by

$$\Delta w_1(i,j) = \eta\, \delta_2(j)\, v_1(i) \tag{5}$$

where

$$\delta_2(j) = f'\left[\sum_k w_1(k,j)v_1(k)\right] \sum_k w(j,k)\delta_3(k) \tag{6}$$

and

$$\Delta w_2(i,j) = \eta\, \delta_3(j)\, v_2(i) \tag{7}$$

where

$$\delta_3(j) = f'\left[\sum_k w_2(k,j)v_2(k)\right] (v_3(j) - z(.,.)) \tag{8}$$

5 Experiments

The multlayer perceptrons descibed in the previous section have been used as interpolators in the Laplacian pyramids. We have worked with the CLAIRE picture, which is a CCIR 601 image (576×720 pixels of 8 bits). Figure 2 shows the corresponding decimated and re-interpolated signal, obtained with a linear (gaussian) filter. Figure 3 shows the image interpolated by multilayer perceptrons. The result is considerably better than the image obtained with linear filters.

6 Conclusions

The main conclusion is that the multi-layer perceptrons perform a better interpolation than a linear filter. The entropy gain is noticable (about 20% in the case of lossless compression). The interpolated image is considerably more pleasant to see.

As a second conclusion, it has been noticed that the network learns to interpolate every image and is not specially suited to a particular picture.

Further study should analyse more in deep the theoretical behaviour of the network for this particular task (Fourier analysis, handling of particular features, internal states, ...).

Biliography

1. P.J. Burt and E.H. Adelson, *The Laplacian Pyramid as a Compact Image Code*, IEEE Trans. on Communications, vol COM-31, no 4, pp.532-540, April 1983.
2. J.B.O. Martens and G.M.M. Majoor, *The Perceptual Relevance of Scale-Space Image Coding*, Siganl Processing, vol. 17, pp. 353-364, 1989
3. R.E. Crochiere and L.R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall Signal Processing Series, Englewood Cliffs, New Jersey, 1983.
4. S.G. Mallat *A Theory for Multiresolution Signal Decomposition : The Wavelet Representation*, IEEE Trans. Pattern Anal. Machine Intell., vol. 11, pp.674-693, July 1989.
5. P. Delsarte, B. Macq and D.T.M. Slock, *Signal-Adapted Multiresolution Transform for Image Coding*, IEEE Trans. on Information Theory, vol. 38, no. 2, pp. 897-904, March 1992.
6. G.G. Gurski, M.T. Orchard and A.W. Hull, *Optimal Linear Filters for Pyramidal Decomposition*, Int. Conf. on Accoustic, Speech and Signal Proc., ICASSP 92, vol II, pp. 381-384, 1992.
7. G.L. Sicuranza, *Artificial Neural Network for Image Compression*, Electronic Letters, vol. 26, no 7, March 1990
8. F. Arduini, S. Fioravanti and D.D. Giusto, *Adaptive Image Coding Using Multilayer Neural Networks*, Int. Conf. on Accoustic, Speech and Signal Proc., ICASSP 92, vol II, pp. 381-384, 1992.
9. D.E. Rumelhart, G.E. Hinton and R.J. Williams, *Learning Representation by Back-Propagating Errors*, Nature, vol. 323, October 1986.

Figure 2: Lineraly interpolated CLAIRE

Figure 3: N.N. interpolated CLAIRE