RESOLUTION ENHANCEMENT OF IMAGES USING WAVELET TRANSFORM EXTREMA EXTRAPOLATION

S. Grace Chang, Zoran Cvetković, and Martin Vetterli

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, CA 94720, USA
grchang,zoran,martin@eecs.berkeley.edu

ABSTRACT

One problem of image interpolation refers to magnifying a small image without loss in image clarity. We propose a wavelet based method which estimates the higher resolution information needed to sharpen the image. This method extrapolates the wavelet transform of the higher resolution based on the evolution of the wavelet transform extrema across the scales. By identifying three constraints that the higher resolution information needs to obey, we enhance the reconstructed image through alternating projections onto the sets defined by these constraints.

1. Introduction

Given a small image, the classic problem of image interpolation is to magnify the image many times without loss in the sharpness of the picture. Some existing methods such as bilinear and spline interpolations generate blurred images since they do not utilize any information relevant to preserving the image clarity. To deblur these images, one could use the standard approach of unsharp masking [1]. Other methods include modeling the edges or filtering with nonlinear filters to boost the high frequencies needed to make an image look sharper. In this paper, we propose a wavelet based method which estimates the higher resolution information needed to sharpen the image.

This information is obtained by extrapolating the wavelet transform of the higher resolution based on the evolution of wavelet transform extrema across the scales. The motivation for this algorithm comes from the fact that wavelet transform modulus maxima capture the sharp variations of a signal, and that their evolution across the scales characterizes the local regularity of the signal [2]. Discussion will be focused on magnifying the data by a factor of 2 for simplicity, although larger magnification could be achieved through iteratively performing the algorithm.

2. Enhancement Algorithm

The 1-D scenario is described first, and the 2-D problem will be treated as an extension of the 1-D case. Figure 1 illustrates the problem model. We model the available waveform f as being obtained from the high resolution signal f_0, which we want to recover, by lowpass filtering followed by downsampling by a factor of 2. Denote by $H_0(z)$ a lowpass filter, and by $H_1(z)$ a highpass filter such that the two filters, together with a synthesis pair $G_0(z)$ and $G_1(z)$, constitute a perfect reconstruction nonsubsampled filter bank. Note that the filter bank in the model is arbitrary, but we conjecture that as long as it is reasonable (i.e. a good lowpass/highpass pair of filters), the result of our algorithm will not depend strongly on the filter bank. The perfect reconstruction condition on nonsubsampled filter banks has the form

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1. \qquad (1)$$

In order to perfectly reconstruct the high resolution signal, we need to know both its highpass component g_s and its lowpass component f_s. However, only f, the downsampled version of f_s, is available. A standard approach would be to interpolate f using, for example, linear or spline interpolation, and possibly followed by some enhancement algorithm such as highpass filtering to deblur the result. The enhancement algorithm presented here is based on estimating the high frequency component g_s which is then combined with an estimate of f_s, through the synthesis filter bank, to give a reconstructed version of the high resolution signal.

An initial estimate $\hat{f}_s$ of the low frequency component f_s can be obtained by simply interpolating f, for instance, using linear interpolation. The approach to estimating g_s is to find its local extrema by analyzing the available data f. It is based on the fact that local extrema of the wavelet transform propagate across the scales, which can be used for extrapolation of higher frequency scales. Figure 2 shows an example of a waveform and its wavelet transform. The wavelet transform here denotes a linear operator implemented by an iterated nonsubsampled filter bank, as shown in Figure 3. The waveform in Figure 2 consists of a step edge, a single impulse, and their smoothed versions. It can be seen that the local extrema of the wavelet transform

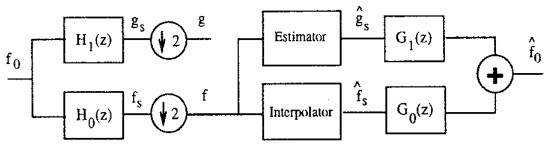

Figure 1: Problem Model

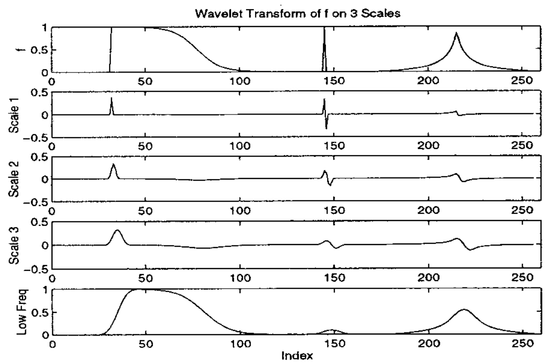

Figure 2: A synthetic waveform and its transforms, showing the propagation of the wavelet transform extrema.

are induced by singularities in the signal and that they propagate across the scales. The evolution of the extrema values corresponding to a singularity is described by the equation

$$W_j h[x_n^{(j)}] = K_n 2^{j\alpha_n}\ ,\ j = 1, 2, ..., J, \tag{2}$$

where $W_j h$ is the wavelet transform of the input signal h at scale j, $x_n^{(j)}$ is the location of the local extremum of scale j corresponding to the nth singularity, α_n is the Lipschitz regularity of f at the singular point, and K_n is a nonzero constant. Note that equation (2) actually holds for continuous time signals and the discretization introduces some deviation.

Consider now again the problem of finding an estimate $\hat{g}_s$ of g_s. It can be shown that the wavelet transform of f (Figure 1) is the decimated version, by a factor of 2, of the wavelet transform of f_0 starting from the second scale. The main idea of our algorithm is to extrapolate the higher scale signal g from the evolution of local extrema across the coarser scales. The signal $\hat{g}_s$ is then obtained by interpolating (linearly) between these local extrema.

Signal enhancement is performed by recognizing that the initial estimates of f_s and g_s can be further improved by identifying constraints that they should obey. The algorithm alternatingly projects the signal to satisfy three basic constraints:

1. The waveforms $(\hat{f}_s, \hat{g}_s)$ must be in the subspace V of $l^2(\mathbb{Z}^2)$, where V denotes the range of the wavelet transform.

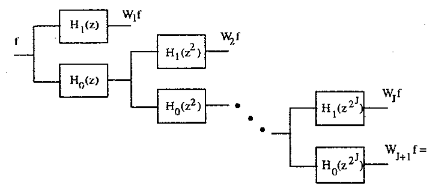

Figure 3: Wavelet Transform

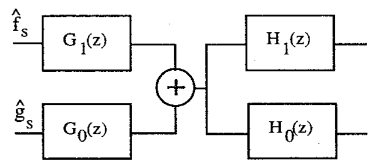

Figure 4: The projection operator onto the subspace V.

2. The downsampled version of $\hat{f}_s$ must be equal to f, which is the original signal that is available.
3. The local extrema of $\hat{g}_s$ should reflect sharp variations in f_0, i.e. their values and locations are determined by singularities in f_0.

Let V, S, and E be sets in $l^2(\mathbb{Z}^2)$ denoting, respectively, the sets of points which satisfy these three constraints. The reconstructed pair $(\hat{f}_s, \hat{g}_s)$ should belong to the sets V and S, and projecting it onto E improves the signal clarity.

Projecting $(\hat{f}_s, \hat{g}_s)$ onto the subspace V consists of filtering by the synthesis filter bank, followed by the analysis filter bank (see Figure 4). The synthesis filters used in the implementation of this projection operator should be selected as $G_0(z) = H_0(z^{-1})/P(z)$, $G_1(z) = H_1(z^{-1})/P(z)$, where $P(z) = H_0(z)H_0(z^{-1}) + H_1(z)H_1(z^{-1})$. In the case of power complementary analysis filters, $H_0(z)H_0(z^{-1}) + H_1(z)H_1(z^{-1}) = 1$, this reduces to $G_0(z) = H_0(z^{-1})$, $G_1(z) = H_1(z^{-1})$. For a detailed analysis of the projection operator onto V for both 1-D and 2-D signals, the reader is referred to [3].

Projection of $(\hat{f}_s, \hat{g}_s)$ onto S amounts to assigning values of f to even samples of $\hat{f}_s$.

The subspace V and the convex set S are well-defined, but the convex set E depends on our knowledge of the singularities of f_0. In Section 3, we discuss constraining the set E with varying degrees of leniency on the values and locations of the wavelet transform extrema, and finding a corresponding projection of $(\hat{f}_s, \hat{g}_s)$ onto E.

3. Implementation

For the extrapolation scheme, we need to first select important singularities and associate the corresponding extrema across the scales. Since the highest scale contains an abundance of extrema and is more sensitive to noise, selection of the extrema is done at the second scale. Due to discretization, the estimated Lipschitz regularity α disagrees from what it should be for a continuous time signal. Hence, scaling constants are multiplied to each scale of the wavelet transform, $W_j f$, and these constants are found empirically so as to make the discrete time step function have $\alpha = 0$. Furthermore, the parameters α_n and K_n of a singular point are estimated from the associated extrema from (2) using the least squared error (LSE) criterion.

Those extrema of the first scale that do not propagate to the coarser scales are simply assigned to the extrapolated scale using the same values and locations.

One of the constraints on the estimated waveform $\hat{g}_s$ is that its local extrema should reflect sharp variations in f_0. From f, we have some knowledge of what the extrema values of $\hat{g}_s$ should be. Hence, the set E can be thought of as the set of waveforms minimizing a specified cost function which penalizes when the extrema values do not conform to this knowledge. There are various cost functions that could be used. We can either (a) constrain $\hat{g}_s$ to retain the initial extrema estimates throughout the reconstruction, (b) allow the values to be within an interval (the rationale being that the estimates of the extrema values may not be very reliable), or (c) have no constraints at all on the values. Approaches (a) and (c) are extreme cases, assigning either infinite cost for wrong values or no cost at all. The allowed interval of approach (b) serves as a moderation, and one way of determining the interval is to make it proportional to the confidence interval of the LSE model fitting.

Since $\hat{g}_s$ is interpolated from the estimate of the subsampled waveform g, the sampling may be such that we miss the true extrema and obtain instead the points next to the extrema. Hence for each extremum of $\hat{g}_s$, the points next to it are also allowed to be extrema to account for this ambiguity. Once the extrema locations and values are determined, the points in between them must obey monotonicity. The algorithm used in constraining the points to obey monotonicity is described in [3].

In general, analyzing a 2-D problem treating the two coordinates independently is not an optimal approach. However, for computational feasibility, we propose here to treat the two coordinates separately. Hence, for the wavelet transform, the data is filtered by the separable 2-D filter bank using algorithms proposed by Mallat [2], and each row of the row component (and similarly for the column component) of the image is processed as in the 1-D case.

4. Experimental Results

To obtain a test image, the original 512×512 Lenna is lowpass filtered, subsampled by 2, and the process

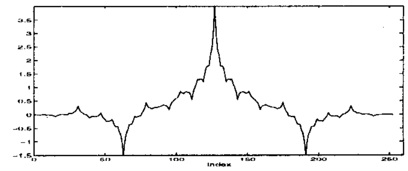

Figure 5: Iterating the process of upsampling by 2, followed by filtering by linear interpolator and unsharp masking using the discrete Laplacian gradient with $\lambda = 1$.

is repeated to obtain a low resolution image of 128×128, from which a 64×64 subimage is extracted as the available data for all of the interpolation methods. The lowpass filter used in obtaining the test image is a separable 2-D filter, $H(\omega_1, \omega_2) = H_1(\omega_1)H_1(\omega_2)$, where the impulse response of $H_1(\omega)$ is $[-1, 0, 9, 16, 9, 0, -1]$ normalized to $H_1(0) = 1$.

Figure 6 shows a 256×256 image of Lenna obtained from performing the enhancement algorithm iteratively twice on the 64×64 test image (i.e. magnify the 64×64 image, and then magnify the resulting 128×128 image again using the algorithm). Convergence occurs rather quickly and the reconstruction after obtaining the initial estimates or after 1-2 iterations is acceptable. After 5-10 iterations the image quality is quite good, and the images either do not change discernibly afterwards or they become slightly more blocky. Filters proposed in [2] are used in the filter bank. We observe that using regular filters yields better results (less blocky images) than non-regular filters (such as Haar). The values of the extrema are allowed to be within a confidence interval during reconstruction. Because the image data represents intensity values between 0 and 255, the pixel values are clipped to be within this interval during reconstruction. Comparing this image to the bilinearly interpolated image in Figure 7, the enhanced image looks sharper. Figure 8 shows a bilinearly interpolated image followed by unsharp masking with the commonly used discrete Laplacian gradient [1] with $\lambda = 1$, iterated twice. The resulting image has very blocky edges which are absent with our method, and it is also not as sharp.

For the algorithm variant that does not constrain the signal extrema, ringing effects tend to occur around the edges. On the other hand, when constraining the values to be the initial estimates, the resulting image is not very good because the initial estimates are not very reliable. A direct $4\times$ magnification was also experimented, but the result was not as good as performing two $2\times$ magnification.

A remark should be made on comparing our enhancement method with unsharp masking. The procedure of iteratively upsampling a signal, performing linear interpolation and unsharp masking (using the commonly used Laplacian gradient with $\lambda = 1$) does

Figure 6: 4× magnified Lenna using the 2× interpolation algorithm iteratively.

Figure 7: 4× magnification using bilinear interpolation.

not converge to a regular filter. Figure 5 shows the 1-D case of the iterated filter where the linear interpolation and unsharp masking impulse responses are, respectively, [.5, 1, .5] and $[-.5\lambda,\ 1+\lambda,\ -.5\lambda]$, with λ being a free parameter. Furthermore, because of the change in sampling rate, this operation not only accentuates the high frequency components but also other frequency components from replicates of the original spectrum. Hence, this method may not be adequate for image interpolation.

5. Conclusion

This paper proposes a wavelet based method for image interpolation. Using properties of wavelet transform extrema across the scales, we extrapolate the extrema needed at a higher scale for reconstruction of a higher resolution image. The result shows that the enhanced image is sharper and less blocky than simple schemes such as linear interpolation and unsharp masking.

The better performance comes at an expense of higher complexity and more computation than the linear methods, and the nonlinearity of our method makes it difficult to characterize the behavior of the algorithm analytically. Because the theoretical framework is geared towards isolated singularities, this method is not necessarily appropriate for, say, texture images.

For future research, we could explore the potential of processing the image with 2-D neighborhoods instead of with a separable 1-D approach. Since the method proposed here is for isolated singularities, a more comprehensive interpolation algorithm would be to segment the images into regions of isolated singularities and textures and process them differently.

References

[1] A. Jain, "Fundamentals of Digital Image Processing," Prentice Hall, 1989.

[2] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans. on PAMI*, Vol.14, No.7, July 1992, pp.2207-2232.

[3] Z. Cvetković and M. Vetterli, "Discrete-Time Wavelet Extrema Representation: Properties and Consistent Reconstruction" To appear in *IEEE Trans. on Signal Processing,* March 1995.

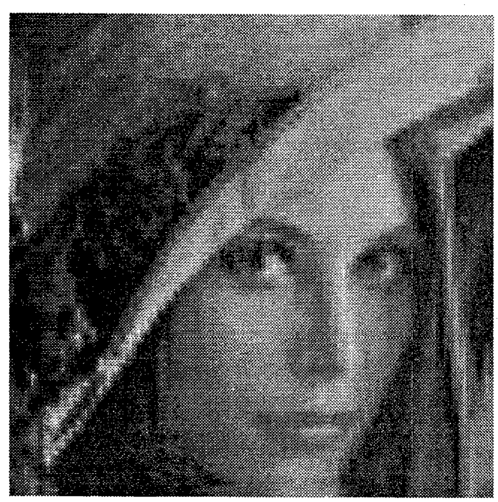

Figure 8: Lenna obtained from performing unsharp masking on a bilinearly interpolated image.