Image Downsampling and Upsampling Methods[1]

Abdou Youssef
Department of EECS
The George Washington University
Washington, DC 20052
Tel: (202 994-6569, Fax: (202) 994-0227, Email: ayoussef@gwu.edu

Abstract
Downsampling and upsampling are widely used in image display, compression, and progressive transmission. In this paper we examine new down/upsampling methods using frequency response analysis and experimental evaluation. We consider six classes of filters for down/upsampling: decimation/duplication, bilinear interpolation, least-squares filters, orthogonal wavelets, biorthogonal wavelets, and a new class that we term *binomial* filters. Our findings show that binomial filters and some biorthogonal wavelet filters are among the best filters for down/upsampling, and significantly outperform the standard methods.

Keywords: Down/upsampling, Filtering, Frequency Response, Progressive Transmission.

1 Introduction

Downsampling and upsampling are two fundamental and widely used image operations, with applications in image display, compression, and progressive transmission. Downsampling is the reduction in spatial resolution while keeping the same two-dimensional (2D) representation. It is typically used to reduce the storage and/or transmission requirements of images. Upsampling is the increasing of the spatial resolution while keeping the 2D representation of an image. It is typically used for zooming in on a small region of an image, and for eliminating the pixelation effect that arises when a low-resolution image is displayed on a relatively large frame. More recently, downsampling and upsampling have been used in combination: in lossy compression [5], multiresolution lossless compression [1], and progressive transmission [2, 5].

The standard methods for down/upsampling are decimation/duplication and bilinear interpolation [5], which yield low visual performance. The increasing use of down/upsampling, especially in combination, warrant the development of better methods for them.

In this paper, we examine the existing methods and propose new down/upsampling-combination methods, and formulate a frequency-response approach for evaluating them. The approach is validated experimentally. Our findings show that the best down/upsampling filters are what we term binomial filters and some well-chosen biorthogonal wavelets. Bilinear interpolation was found significantly inferior, and decimation duplication came last.

2 Formulation and Evaluation of Down/Upsampling

Consider a signal $x = (x_n)$. Downsampling x by two can be generally viewed as pre-filtering x with a linear filter $g = (g_k)$, yielding a signal $u = (u_n)$, and then decimating u by two, getting a signal $v = (v_n)$ where $v_n = u_{2n}$ for all n. Upsampling v by two, on the other hand, can be viewed as zero-upsampling followed by post-filtering. That is, v is zero-upsampled to a signal $w = (w_n)$ where $w_{2n} = v_n$ and $w_{2n+1} = 0$ for all n; then w is passed through a filter p yielding an approximation $\hat{x}$ of the original signal x. This general down/upsampling scheme is shown in Figure 1. Note that in the case of 2D images, the downsampling is applied row-wise then column-wise, and so is upsampling. Note also that this general scheme captures

[1]This research was performed in part at the National Institute of Standards and Technology.

the standard bilinear interpolation and decimation/duplication methods; their filters as well as other classes of filters will be presented in the next section.

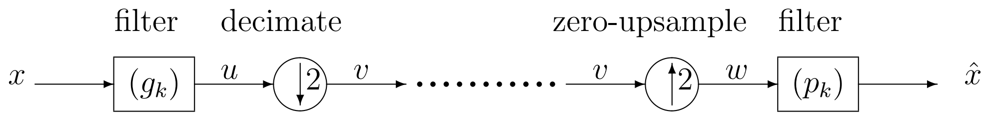

Figure 1: **The General Downsampling/Upsampling Scheme**

To understand the behavior of the scheme, we use the z-transform (i.e., frequency-domain analysis). For any sequence $a = (a_n)$, the z-transform of a is a complex function $A(z) = \sum_n a_n z^n$. The z-transforms of x, u, v, w, $\hat{x}$, g and p, are denoted $X(z)$, $U(z)$, $V(z)$, $W(z)$, $\hat{X}(z)$, $G(z)$ and $P(z)$, respectively. Without any loss of generality, we will limit z to the unit circle, i.e., $z = e^{i\omega}$, where ω is the frequency variable; $A(e^{i\omega})$ is the Fourier transform of a, which is periodic with period 2π, and its magnitude is symmetric around the origin.

We now derive the z-transform relation of the scheme of Figure 1. By the fundamental convolution theorem, $U(z) = G(z)X(z)$ and $\hat{X}(z) = P(z)W(z)$. Furthermore,

$$W(z) = \sum_n w_{2n} z^{2n} + \sum_n w_{2n+1} z^{2n+1} = \sum_n v_n z^{2n} = \sum_n u_{2n} z^{2n} = \tfrac{1}{2}[U(z) + U(-z)],$$

using the facts that $w_{2n+1} = 0$, $w_{2n} = v_n = u_{2n}$. Consequently,

$$\hat{X}(z) = P(z)W(z) = \tfrac{1}{2}P(z)[U(z) + U(-z)] = \tfrac{1}{2}P(z)[G(z)X(z) + G(-z)X(-z)],$$

yielding

$$\hat{X}(z) = [\frac{1}{2}P(z)G(z)]X(z) + [\frac{1}{2}P(z)G(-z)]X(-z). \quad (1)$$

The first term on the right hand side of equation (1) is what contributes to the true reconstructed signal, while the second term is aliasing distortion. A good down/upsampling scheme yields a reconstructed signal $\hat{x}$ that is very close to the original signal x, and in the ideal (though impossible) extreme, $\hat{x} = x$, or correspondingly $\hat{X}(z) = X(z)$, for any input signal x. Equation (1) reveals that to have an ideal scheme, the filters g and p must satisfy the two identities $P(z)G(z) = 2$ and $P(z)G(-z) = 0$. Although this is impossible, it serves as a yardstick for comparative evaluation of candidate schemes. More importantly, equation (1), along with the understanding of the relevant characteristics of both the human visual system and typical natural images, helps identify the desirable and more or less achievable properties of the two filters g and p of the scheme.

Therefore, we must have $|P(z)G(-z)|$ as close to zero as possible, and $P(z)G(z)$ as close to 2 as possible. It is well-known that the human visual system is most sensitive to low-frequency (i.e., low-variation) signals, and that this sensitivity decreases quickly at higher frequencies. Thus, it is desirable to have $P(z)G(z)$ very close to 2 in the largest low-frequency range possible, while still keeping $|P(z)G(-z)|$ near zero in the entire range of frequencies. That is, $P(z)G(z)$ should form a low-pass (LP) filter. In particular, g and p should be LP filters in their own right, especially when the downsampling and upsampling are to be used as two separate standalone operations and not just in combination. Fortunately, natural images contain largely low-frequency data, and thus the data loss incurred by such a scheme tends to be mathematically small and visually acceptable. It remains to determine how large we can make the low-pass band where $P(e^{i\omega})G(e^{i\omega})$ is near 2. This is addressed next.

Let g and p be two ideal low-pass filters, where $|G(e^{i\omega})|$ is 1 over $[0\ \alpha]$ and zero over $(\alpha\ \pi]$ for some α, and $|P(e^{i\omega})|$ is 2 over $[0\ \beta]$ and zero over $(\beta\ \pi]$ for some β. Observe that $|G(-z)| = |G(e^{i(\omega-\pi)})| = |G(e^{i(\pi-\omega)})|$ (because of the magnitude symmetry); therefore, $|G(-z)|$ is 1 over $[\pi - \alpha\ \pi]$ and 0 over $[0\ \pi - \alpha)$. Consequently, $|P(z)G(-z)|$ is 2 in the frequency range $[\pi - \alpha\ \beta]$ and zero in the remainder of $[0\ \pi]$. Since the ideal value of $|P(z)G(-z)|$ is zero over all $[0\ \pi]$, it follows that we should ideally have $\pi - \alpha \geq \beta$, that is,

$\alpha+\beta \leq \pi$. On the other hand, $P(z)G(z)$ is 2 in the frequency range $[0\ min(\alpha,\beta)]$, and so $min(\alpha,\beta)$ is desired to be as large as possible. Clearly, when $\alpha+\beta\leq\pi$, the maximum value of $min(\alpha,\beta)$ is $\frac{\pi}{2}$, achieved at $\alpha=\beta=\frac{\pi}{2}$. We have thus proved the following theorem.

Theorem 1 *The ideal down/upsampling scheme is one where both filters g and p are low-pass filters with pass-band equal to* $[0\ \frac{\pi}{2}]$. *In that scheme, the aliasing distortion is zero, and the reconstructed signal preserves the largest possible range of low frequencies.*

Ideal low-pass finite-length filters are not achievable, but many families of finite filters approximate ideal filters reasonably well. With realistic low-pass filters g and p, $|P(z)G(z)|$ is approximately 2 in the frequency range $[0\ \frac{\pi}{2}]$ and approximately zero in $(\frac{\pi}{2}\ \pi]$, and $|P(z)G(-z)|$ is nearly zero throughout the frequency range $[0\ \pi]$ except for a "hump" around the middle of that range. Clearly, the smaller and narrower that hump is, the better. Those two yardsticks will be the basis of our evaluation of various filters for down/upsampling.

3 The Classes of Filters to Be Considered

We consider six classes of filters for down/upsampling. The trigonometric polynomial

$$R_N(e^{i\omega}) = cos^{2N}(\frac{\omega}{2})\sum_{k=0}^{N-1}\begin{pmatrix} N-k+1 \\ k \end{pmatrix} sin^{2k}(\frac{\omega}{2}), for\ any\ positive\ integer\ N,$$

will be needed. The six classes of filters are:

- Biorthogonal wavelets [3]: Their filters g and p are characterized by $P(e^{i\omega})G(e^{i\omega}) = R_N(e^{i\omega})$, for $N\geq 1$. The combined length of g and p is $4N$.
- Binomial filters: They are characterized by $P(e^{i\omega}) = 2G(e^{i\omega}) = 2R_N(e^{i\omega})$, for $N\geq 1$. g and p are symmetric of effective length $2N+1$ each, and their coefficients are rational numbers with powers-of-2 denominators, making the arithmetic very efficient.
- Orthogonal wavelets: Daubechies well-known wavelets D_N [3]. The filters g and p are the length-N low-pass analysis and synthesis filters of D_N.
- Least-Squares filters [4]: $p=2g$ where g is a symmetric LP filter with cutoff at $\frac{\pi}{2}$.
- Bilinear interpolation: $g_{-1}=g_0=1/2$, $p_0=1$, $p_1=p_{-1}=1/2$. All others are 0.
- Decimation/duplication: $g_0=1$, and $p_0=p_1=1$. All undefined g_k and p_k are 0.

4 Frequency-Response Evaluation of the Six Classes

In this section the dual frequency responses $P(z)G(z)$ and $|P(z)G(-z)|$ of the six classes are evaluated and plotted in Figure 2, where $z=e^{i\omega}$ in the frequency range $[0\ \pi]$.

The biorthogonal wavelets form a large class since for any N there are many ways of factoring R_N into G and P, yielding filters of combined length $4N$. We generated all possible factorizations for all $N=2,3,...,14$, resulting in 4297 filter pairs (g,p). We measured the closeness of their frequency response to that of the ideal LP filter with passband $[0\ \pi/2]$, and accordingly selected the best 18 wavelets [7]. In this research, we evaluated the dual frequency responses of those 18 filter pairs; they all have good dual frequency responses, which improve as N increases. We selected three filter pairs for our presentation (2nd row of Fig. 2): a 9/7 filter pair (i.e., length(g)=9, length(p)=7), a 16/16 pair, and a 33/19 pair.

Binomial filters are very good and exhibit a consistent improvement in their dual frequency responses as N increases. We tested many members of this class, but we present here (1st row in Fig. 2) three binomial filter pairs (BN 7/BN 7, BN 15/BN 15 and BN 27/BN 27) whose lengths are comparable to the three biorthogonal filters chosen above, to simplify the comparison. As is evident in the Figure, the binomial filters have the best dual responses among all the classes considered, and the responses approach the ideal quickly as N increases.

The Daubechies orthogonal filter pairs also exhibit a consistent improvement in its dual frequency responses as N increases. We chose three members (D_8/D_8, D_{16}/D_{16} and D_{28}/D_{28}) of lengths comparable to those chosen from the other classes. Note that although the dual response of D_N approaches the ideal as N increases, the orthogonal filters are not linear phase (unlike all the other classes considered). Therefore, they incur phase distortion which becomes worse for larger N. This will be evident in the experimental evaluation.

The dual responses of least-squares (LS) filters also approach the ideal as the filter length increases. However, for short filters, the frequency response is quite far from ideal, and exhibits familiar ripples which increase in number as the filter length increases. We chose LS 8/LS 8, LS 16/LS 16, and LS 26/LS 26 for further evaluation (4th row of Fig. 2).

Finally, the dual frequency responses for bilinear interpolation and for decimation/duplication, shown the in the last row of Figure 2, are quite far from ideal, and very inferior to the first 5 classes. The decimation/duplication is the worst, especially with respect to its aliasing frequency response. The visual effect of that is the highly visible pixelation.

5 Experimental Evaluation and Validation

To validate our dual frequency response evaluation approach, and to get a quantified measure of the performance of the six classes, we tested them on two different images: the well-know image of Lena, and a fingerprint image. The test was carried out with three down/upsampling factors, 2, 4 and 8, where for 4 and 8 the downsampling (and upsampling) by two was performed twice and three times, respectively. For each filter pair and each sampling factor, the SNR between the original image and the down-then-upsampled image was computed.

The experimental results, both in objective SNR plots (Figure 3) and in visual form, are in agreement with the dual frequency response evaluation. The SNR plots show that the binomial filters and the selected biorthogonal wavelets are the best. The orthogonal filters perform well for short filters, but with long filters the phase distortion deteriorates the performance. Least-squares filters are bad at short lengths. The standard techniques, bilinear interpolation and decimation/duplication perform poorly — they are inferior to the binomial filters by about 4 and 6 decibels, respectively, which is significant.

We conclude by giving the coefficients of the winning filters in Tables 1 and 2.

BN† 5: $g_0 = \frac{16}{2^5},\ g_{\pm1} = \frac{9}{2^5},\ g_{\pm3} = \frac{-1}{2^5}$
BN 7: $g_0 = \frac{256}{2^9},\ g_{\pm1} = \frac{150}{2^9},\ g_{\pm3} = \frac{-25}{2^9},\ g_{\pm5} = \frac{3}{2^9}$
BN 15: $g_0 = \frac{8388608}{2^{24}},\ g_{\pm1} = \frac{5153148}{2^{24}},\ g_{\pm3} = \frac{-1288287}{2^{24}},\ g_{\pm5} = \frac{429429}{2^{24}},\ g_{\pm7} = \frac{-122694}{2^{24}}$ $g_{\pm9} = \frac{26026}{2^{24}},\ g_{\pm11} = \frac{-3549}{2^{24}},\ g_{\pm13} = \frac{231}{2^{24}}$

† p is implicit: $p_k = 2g_k\ \forall k$. Also, $g_{2k} = 0$ for all nonzero integers k.

Table 1: Three Binomial Filters

9/7 wavelet: $g_{-k} = g_k$ and $p_{-k} = p_k$									
$g_{0..4}$:	0.602949	0.266864	-0.078223	-0.016864	0.026749				
$p_{0..3}$:	1.115087	0.591272	-0.057544	-0.091272					
16/16 wavelet: $g_{-k} = g_{k-1}$ and $p_{-k} = p_{k+1}$									
$g_{0..7}$:	0.509000	0.036221	-0.094339	0.040131	0.027620	-0.017637	-0.003759	0.002763	
$p_{1..8}$:	0.939274	0.265629	-0.162821	-0.089211	0.031598	0.020993	-0.003148	-0.002314	
33/9 wavelet: $g_{-k} = g_k$ and $p_{-k} = p_k$									
$g_{0..16}$:	0.535998	0.300982	-0.040128	-0.071851	0.033196	0.026008	-0.017695	-0.007066	0.008725
	0.002367	-0.002558	-0.000504	0.000523	0.000068	-0.000065	-0.000003	0.000004	
$p_{0..4}$:	0.043307	0.033844	-0.154696	-0.072745	0.622617				

Table 2: Filter Coefficients of Three Good Biorthogonal Wavelets

6 Conclusions

In this paper we examined several classes of filters for down/upsampling, and found that binomial filters and certain biorthogonal filters are best. Considering also the short length and integer processing afforded by high-performance binomial filters, our single most important conclusion is that short-length binomial filters (5 or 7 coefficients) are the best choice. They can be used effectively in progressive transmission of images. They could also be used in variable-bandwidth video transmission as well as in prediction-based lossless compression. The latter two applications need further study, and are the subject of future research.

References

[1] S. R. Burgett and M. Das, "Multiresolution Multiplicative Autoregression Coding of Images," *Proc. IRRR Conference on Systems Engineering,* pp. 276–279, 1991.

[2] P. J. Burt and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications,* COM-31, pp. 532–540, Apr. 1983.

[3] I. Daubechies, *Ten Lectures on Wavelets,* Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[4] L. B. Jackson, *Digital Filters and Signal Processing,* Kluwer Academic Publishers, 1996.

[5] "Generic Coding of moving pictures and associated audio (MPEG-2)", ISO-13818-2.

[6] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding,* Prentice Hall, 1995.

[7] A. Youssef, "Selection of Good Biorthogonal Wavelets for Data Compression," *CISST '97,* Las Vegas, Nevada, pp. 323–330, June 1997.

Figure 2: The Frequency Response of Down/Upsampling Scheme for the Six Classes of Filters

Figure 3: Experimental Evaluation of The Down/Upsampling Schemes