

Digital restoration of defocused images in the wavelet domain

Edmund Y. Lam

Many cameras nowadays are equipped with an autofocus mechanism that attempts to take pictures at the best possible focus. However, there are situations in which the pictures are still out of focus, such as when the photographer has mistakenly focused at a wrong position or when the focusing region consists of objects of different depths and therefore confuses the autofocus system. With a digital camera, we can attempt to use digital image-restoration techniques to bring the pictures back into focus. We design an algorithm that restores the image by digitally enhancing the corresponding frequency bands. We employ the restoration in the wavelet domain so that this restoration scheme can be compliant with JPEG 2000, which is positioned to succeed JPEG as the next image-compression standard and has the potential to be widely adopted by the digital photography industry owing to its many advanced features. © 2002 Optical Society of America

OCIS codes: 100.3020, 100.7410, 110.5200, 220.1000, 220.2560.

1. Introduction

Although the JPEG international standard was officially adopted in 1994, the JPEG committee had already been crafting it for many years beforehand.¹ The purpose of the standard was for digital compression and coding of continuous-tone still images. A block-based discrete cosine transform (DCT) was incorporated in its core, which was, at the time, the state-of-the-art compression technology, with many nice features such as frequency discrimination and energy-compaction efficiency. Much research has been carried out since then to reduce the computational complexity of calculating the DCT. Many dedicated codecs and application-specific integration circuits were also developed to optimize the speed for performing the compression and decompression. In many respects, the standard was quite successful, as evident from its proliferation in many areas of image representation, storage, and transmission.

However, two major forces together have prompted the development of a new international standard. First, there is the demands from new applications,

which welcome new or better features. The Internet has become widely popular in the past few years. It is desirable that images be progressive in quality and resolution, with fast decoding as limited by the bandwidth of the transmissions. This is especially the case for mobile applications. At the same time, the emergence of electronic commerce pushes for better image security. Meanwhile, medical imaging desires a decoding scheme from which the user can interactively select a region of interest, as well as a range of lossy to lossless transmission of the image. As for digital photography, the goal is to have a more efficient compression scheme with a relatively low complexity. Progressive coding of images is also desirable because of the limited memory available in the camera, which allows for the discard, if necessary, of the less significant information of existing pictures to make room for a new photograph.

Second, continuous advancement in research since the adoption of the JPEG standard has made many of these goals realizable. Most noticeably, wavelet-based compression schemes have been shown to produce better image quality than DCT-based compression at the same bit rate. Alternatively, at the same quality (usually measured by mean-square error), wavelet compression often has better compression efficiency than JPEG. Furthermore, it is also ideally suitable for multiresolution transmission because of its subband decomposition nature. At the same time, a breakthrough in the use of bit-plane coding gives rise to progressive decoding, allowing

The author was with Information Systems Laboratory, Stanford, California 94305. He is currently with the University of Hong Kong, Hong Kong. His e-mail address is elam@ieee.org.

Received 4 June 2001; revised manuscript received 29 April 2002.

0003-6935/02/234806-06\$15.00/0

© 2002 Optical Society of America

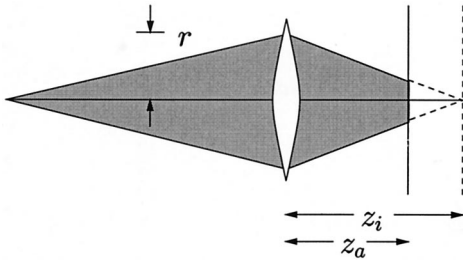


Fig. 1. Image formation with the out-of-focus aberration.

the possibility of providing a range of lossy to lossless transmission of an image. A more elaborate bit-stream syntax could also permit better region-of-interest coding. It is therefore not surprising that a new image-compression standard called JPEG 2000 has been developed that uses the wavelet as its core.²

Currently, there is not a commercially available digital camera that uses JPEG 2000, mostly because it is still a new standard. However, there are already chip sets dedicated for its functionalities, and efforts are underway to produce fast computations of its core, such as with the use of parallel computers.³ It is therefore not far-fetched to explore new image-processing algorithms for digital photography that can be incorporated with the new standard.

One particular application we investigate here is the possibility of restoring out-of-focus images. Even with the capability of autofocus in most cameras, we still encounter situations in which the pictures are not sharply focused, because of either photographer mistakes or ambiguity in the focusing region that confuses the autofocus system. With digital photography, we can attempt to use computation to compensate for the out-of-focus aberration. In Section 2 we begin with a deeper understanding of the nature of the out-of-focus blur using Fourier optics and the particular conditions for our problem. Then in section 3 we give a brief outline of the JPEG

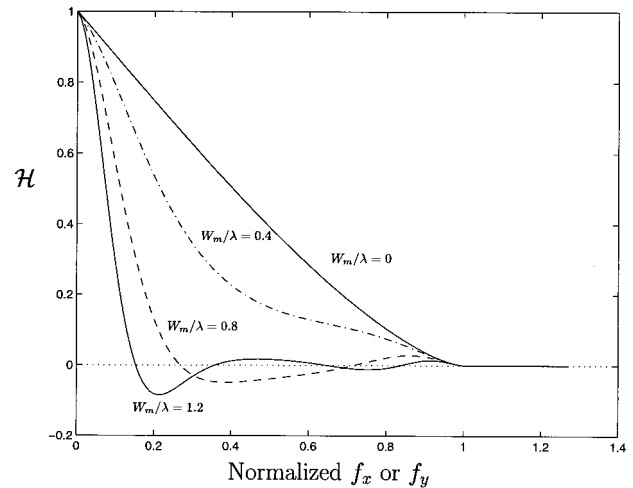


Fig. 2. Cross-sectional view of the OTFs for a circular pupil with different amounts of defocus.

2. Defocusing

In incoherent imaging of a diffraction-limited system, the object and the image are related by⁴

$$\mathcal{G}_i(f_x, f_y) = \mathcal{H}(f_x, f_y)\mathcal{G}_g(f_x, f_y), \quad (1)$$

where \mathcal{G}_i is the normalized frequency spectrum of the image intensity I_i , \mathcal{G}_g is the normalized frequency spectrum of the object intensity I_g , and \mathcal{H} is the optical transfer function (OTF). Let the ideal focusing plane be at a distance z_i from the lens, while we focus at z_a as shown in Fig. 1. Assuming the radius in the pupil is r , we define

$$W_m = \frac{1}{2} \left(\frac{1}{z_a} - \frac{1}{z_i} \right) r^2, \quad (2)$$

which is an indication of the severity of the focusing error when normalized by wavelength λ .⁵ The path-length error is therefore

$$W(x, y) = W_m \frac{x^2 + y^2}{r^2}, \quad (3)$$

and the OTF is

$$\mathcal{H}(f_x, f_y) = \frac{\iint_{\mathcal{A}(f_x, f_y)} \exp\{jk[W(x+x_0, y+y_0) - W(x-x_0, y-y_0)]\} dx dy}{\iint_{\mathcal{A}(0,0)} dx dy}, \quad (4)$$

2000 standard, emphasizing the components relevant to our algorithm that would be described in the following sections. Section 4 provides the main details of the algorithm, with Section 5 showing some experimental results. We conclude with a discussion of the strength and weakness of this algorithm in Section 6.

where $\mathcal{A}(f_x, f_y)$ is the area of overlap of the pupil displaced by $\pm(x_0, y_0)$ and

$$x_0 = \frac{\lambda z_i f_x}{2}, \quad y_0 = \frac{\lambda z_i f_y}{2}. \quad (5)$$

Figure 2 shows a cross section of the OTF for a cir-

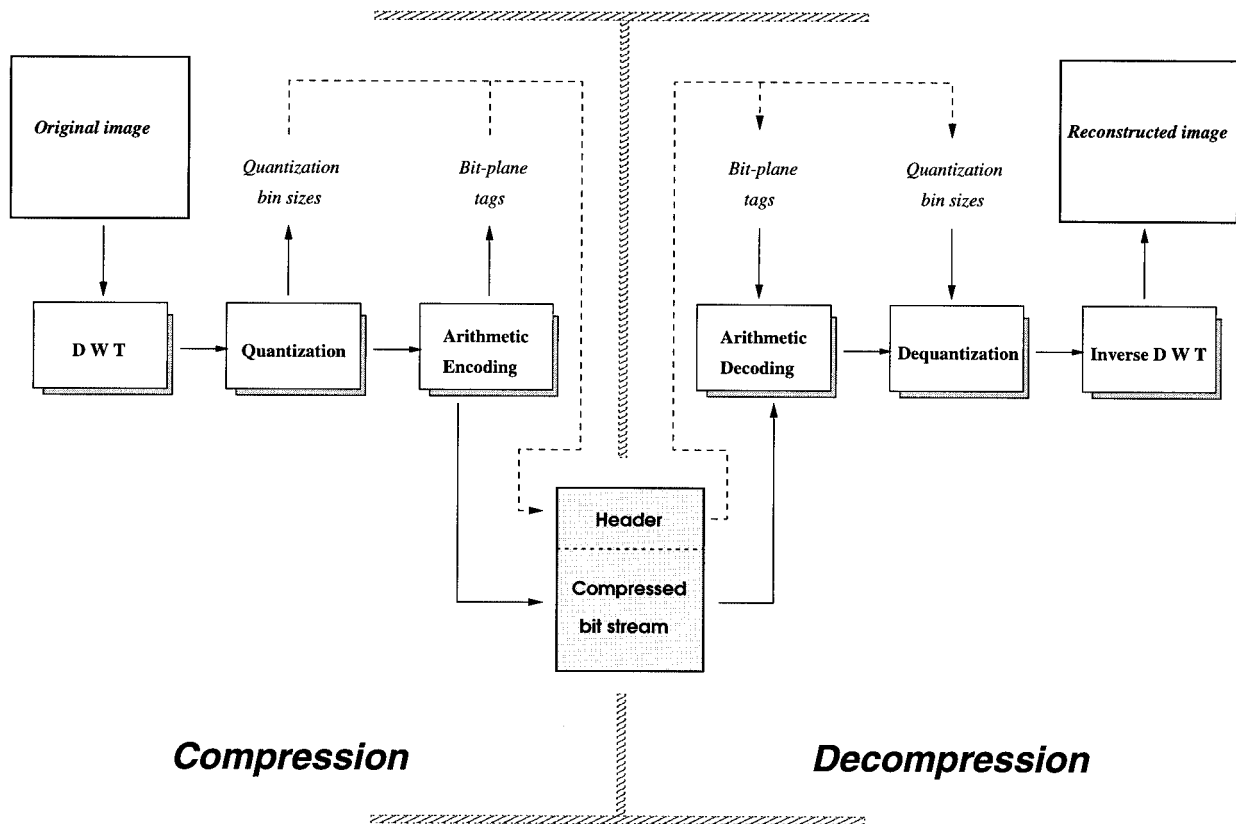


Fig. 3. Block diagram of the JPEG 2000 compression and decompression system.

cular pupil with various amounts of defocus. We see that even for in-focus imagery, we have a cutoff because of the finite size of the pupil, and we also see attenuation at frequencies before cutoff. For restoration of a defocused image, our goal is to modify it to resemble the in-focus image rather than to reproduce the object itself.

Note also that implicit in Eq. (1) is the assumption that the imaging system is linear space invariant. This means all points in the object lie at the same distance from the lens in order to have the same OTF; moreover, the field-dependent in-focus image quality should not show large variations that are due to intrinsic lens aberrations. Such an assumption is generally acceptable when the depth of field is shallow.

3. JPEG 2000

The basic structure of the JPEG 2000 encoding and decoding scheme is shown in Fig. 3. The encoder consists of a mathematical transform called the discrete wavelet transform (DWT), followed by quantization and entropy coding. In the decoder the three processes are reversed. The DWT is a series expansion for discrete signals. The major advantage of a wavelet transform is that its basis functions enjoy good space and frequency localizations.⁶ This property is especially attractive for image processing because typical images have localized edges that contain a lot of high-frequency components. Therefore they would typically take more coefficients to

represent in a DCT than in a DWT. In part I of the JPEG 2000 standard, which defines the essential components of the compression scheme and is royalty free, there are only two particular choices of the wavelet bases, one reversible with integer taps and one irreversible with floating point taps.² Both belong to the family of biorthogonal filters and therefore share many similar properties.⁷ Our algorithm does not alter the wavelet basis to remain compliant to the standard.

For lossy compression, we improve coding efficiency by discarding some of the small wavelet coefficients, especially in the high frequency. This is achieved in the quantization process by use of a uniform quantizer with a dead zone in the middle. For example, if I is the wavelet coefficient and Q_e is the quantization bin size, then the quantized coefficient q is given by

$$q = \begin{cases} \left\lfloor \frac{I}{Q_e} \right\rfloor & \text{if } I \geq 0 \\ \left\lceil \frac{I}{Q_e} \right\rceil & \text{if } I < 0 \end{cases}, \quad (6)$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x and $\lceil x \rceil$ denotes the smallest integer larger than or equal to x . This is illustrated in Fig. 4. In dequantization, we recover \hat{I} as an estimate of I by

$$\hat{I} = [q + \alpha \operatorname{sgn}(q)]Q_d, \quad (7)$$

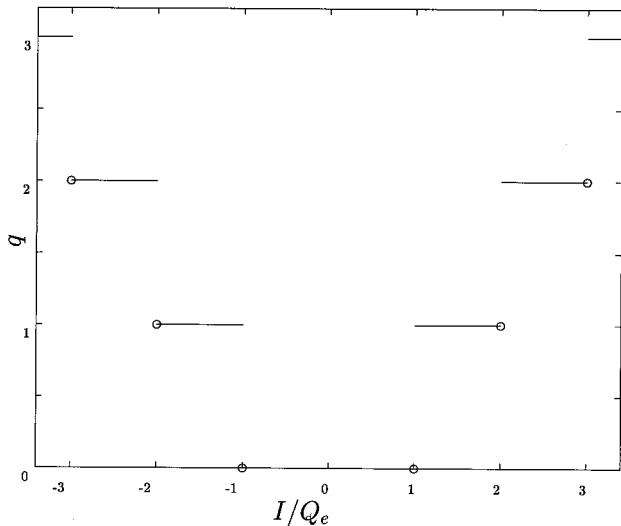


Fig. 4. Quantization of wavelet coefficients with a dead zone.

where $\text{sgn}(q)$ denotes the sign of q and is equal to zero if $q = 0$. Q_d is usually taken to be the same as Q_e . However, as we will see, the flexibility of choosing $Q_d \neq Q_e$ allows us to implement restoration together with compression. Furthermore, α is a design parameter. We usually pick $\alpha = 0.5$ in the absence of knowledge about the probability density function of q . However, because q is usually modeled as Laplacian,^{8,9} a smaller value of α may be preferable to reduce the overall distortion. We will find the optimal value of α in Section 4 for two cases of the distortion metrics.

The last part of the compression pipeline is the arithmetic coding. Here we are not too concerned with it because it is an entropy coder and therefore does not introduce any loss in the signal content. As with Huffman coding for the JPEG algorithm, entropy coding does not affect our design of image-restoration algorithms.

4. Wavelet-Domain Restoration

In our wavelet-domain-restoration algorithm, we assume that the focus defect is known or estimated by another method beforehand, possibly by detection of the distance between the lens and the object plane. Our algorithm also requires training, first, by the collection of a set \mathcal{C}_{in} of in-focus images and another set \mathcal{C}_{out} of the corresponding out-of-focus images as our training set.¹⁰ Let $J_{\text{in}}(b)$ be the vector containing the coefficients from the wavelet subband labeled b of the in-focus images and $J_{\text{out}}(b)$ be the corresponding vector for the out-of-focus images. We seek a multiplicative factor for each subband such that

$$J_{\text{out}}(b)a(b) \approx J_{\text{in}}(b). \quad (8)$$

The best $a(b)$, in the mean-square sense, is given by

$$a(b) = \frac{\langle J_{\text{in}}(b), J_{\text{out}}(b) \rangle}{\langle J_{\text{out}}(b), J_{\text{out}}(b) \rangle}, \quad (9)$$

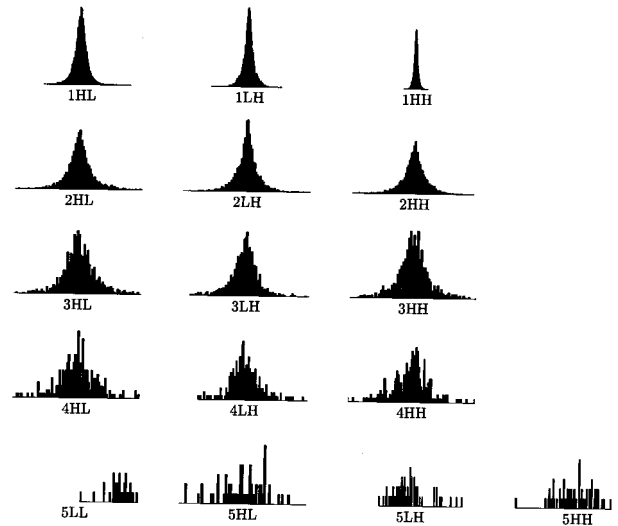


Fig. 5. Wavelet coefficient distribution of the bridge image by use of the reversible wavelet basis. The number below each distribution refers to the subband level, and the letters L and H denote whether they are for the low frequency subband or the high frequency subband, respectively. The first letter is for the horizontal decomposition, while the second letter is for the vertical decomposition.

with $\langle \cdot, \cdot \rangle$ denoting the inner product. Unfortunately, we have no guarantee that $a > 0$, and indeed it may not be if its calculation is dominated by noise. Therefore we clip a to be ≥ 0.001 . Note that we do not have to multiply the DWT coefficients by $a(b)$ before quantization but could simply adjust the encoding quantization step size, $Q_e(b)$, by

$$Q_e(b) = \frac{Q_d(b)}{a(b)}, \quad (10)$$

where $Q_d(b)$ is the decoding quantization step size.

In the dequantization step, the JPEG 2000 standard allows for a parameter α in Eq. (7), which we can optimize if we know the probability density function of the wavelet coefficient distributions. Figure 5 shows their distributions for the bridge image decomposed for five levels. Empirically, we see that the coefficients at the first few levels can be reasonably modeled by the Laplacian distribution, given by

$$p(I) = \frac{\mu}{2} \exp(-\mu|I|), \quad (11)$$

where μ controls the variance of the distribution. The higher levels have too few pixels, so the probability density function is hard to observe.

Consider the case in which we choose α for each subband b . Our first task is to find μ for the particular subband. Note that for a Laplacian distribution, the variance v^2 equals $2/\mu^2$. For any image, we could readily obtain v^2 for each subband; we estimate the corresponding μ by

$$\mu(b) = \left[\frac{2}{v^2(b)} \right]^{1/2}. \quad (12)$$



Fig. 6. In-focus and out-of-focus images for simulation.

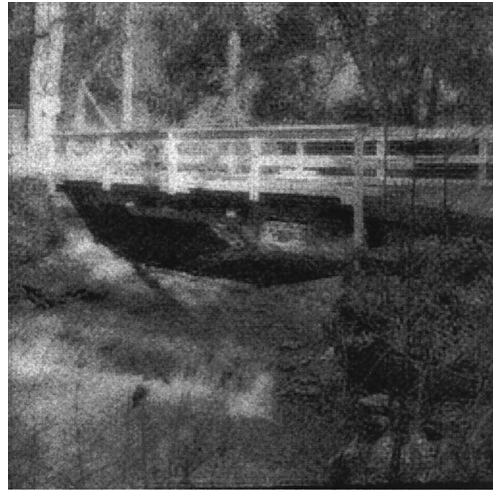


Fig. 7. Simulation results for the bridge image by use of the reversible wavelet basis.

We can then design a better codebook for dequantization to minimize the distortion other than using the midpoint of the partition. If an l_2 norm is used in the distortion measure, the centroid of the partition is optimal, i.e.,

$$\hat{I} = \frac{\int_{qQ_d}^{(q+1)Q_d} Ip(I)dI}{\int_{qQ_d}^{(q+1)Q_d} p(I)dI}, \quad (13)$$

where, as before, Q_d is the size of the dequantization partition, q is the compressed value, I is the original wavelet coefficient, and \hat{I} is the optimal decoded value for that partition.¹¹ Solving Eq. (13), we get¹²

$$\hat{I} = (q + 0.5)Q_d - \coth\left(\frac{\mu Q_d}{2}\right) \frac{Q_d}{2} + \frac{1}{\mu}. \quad (14)$$

In comparison with Eq. (7), we find that the best value of α is

$$\alpha = 0.5 - \left[\frac{1}{2} \coth\left(\frac{\mu Q_d}{2}\right) - \frac{1}{\mu Q_d} \right]. \quad (15)$$

Note that $0 < \alpha \leq 0.5$ because $1 > \coth(\theta) - (1/\theta) \geq 0$, where $\theta = \mu Q_d/2$ in this case.

Other distortion measures can also be used. For instance, if an l_1 norm is used, we require that

$$\int_{qQ_d}^{\hat{I}} p(I)dI = \int_{\hat{I}}^{(q+1)Q_d} p(I)dI; \quad (16)$$

i.e., we seek the median of the partition.¹¹ Again solving Eq. (16), we get¹²

$$\hat{I} = (q + 0.5)Q_d - \frac{1}{\mu} \ln \cosh\left(\frac{\mu Q_d}{2}\right). \quad (17)$$

Table 1. Improvement in SNR by Use of New Quantization Levels in the Wavelet Domain

Conditions	Signal-to-Noise Ratio	
	Without Noise (dB)	With Noise (dB)
Before compression	25.1	24.3
With the standard quantization	23.4	22.9
With new quantization	24.5	24.5

In comparison with Eq. (7), we have

$$\alpha = 0.5 - \left[\frac{1}{\mu Q_d} \ln \cosh \left(\frac{\mu Q_d}{2} \right) \right]. \quad (18)$$

As before, $0 < \alpha \leq 0.5$.

5. Simulation Results

We test our algorithm on a 256×256 pixel bridge image blurred by an out-of-focus aberration with $W_m/\lambda = 0.4$. In the second experiment, we also add some additive white Gaussian noise with the signal-to-noise ratio (SNR) = 30 dB. The in-focus and out-of-focus images are shown in Figs. 6(a) and 6(b).

In this simulation we decompose the out-of-focus image using the reversible wavelet as specified in part I of JPEG 2000, to five levels. Figure 7(a) shows the case in which the noisy out-of-focus image is compressed and decompressed with the reference quantization level. The image in Fig. 7(b) is the same as in Fig. 7(a), except that we have used Eqs. (10) and (15) to improve the quantization. We can see that the image in Fig. 7(b) is sharper than that in Fig. 7(a). We tabulate the SNR calculations for the noiseless and noisy cases in Table 1. The numerical results agree with our observation on the improvement in quality with the new quantization method.

6. Conclusions

In this paper we have presented an image-restoration algorithm that is compliant with the JPEG 2000 scheme by taking advantage of the design of the quantization bin sizes. Experimental results show that the restored images have improved in sharpness, and a numerical comparison based on the SNR also confirms enhancement of image quality. The

goal is to embed the restoration as part of the on-board compression in the application-specific integrated circuit of a digital camera because this method also has little demand for additional computational power.

One drawback of this algorithm, however, is that it requires training with images that have similar spectral characteristics as our target images. Further research is needed for a more direct approach that can perform the restoration without training beforehand.

References

1. W. Pennebaker and J. Mitchell, *JPEG Still Image Data Compression Standard* (Van Nostrand Reinhold, New York, 1992).
2. D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice* (Kluwer Academic, Boston, 2001).
3. E. Y. Lam, P. Y. Tam, and X. Ouyang, "An efficient parallel implementation of the 2-D pyramidal discrete wavelet transform," in *2001 International Conference on Imaging Science, Systems, and Technology* (CSREA Press, Las Vegas, Nev., 2001), Vol. 1, pp. 178–183.
4. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).
5. E. Y. Lam and J. W. Goodman, "Discrete cosine transform domain restoration of defocused images," *Appl. Opt.* **37**, 6213–6218 (1998).
6. M. Vetterli and J. Kovačević, *Wavelets and Subband Coding* (Prentice Hall, Englewood Cliffs, N. J., 1995).
7. G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge, Wellesley, Mass., 1996).
8. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 674–693 (1989).
9. E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Process.* **9**, 1661–1666 (2000).
10. K. Panchapakesan, A. Bilgin, M. W. Marcellin, and B. R. Hunt, "Joint compression and restoration of images using wavelets and non-linear interpolative vector quantization," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, New York, 1998), Vol. 5, pp. 2649–2652.
11. A. Gersho and R. Gray, *Vector Quantization and Signal Compression* (Kluwer Academic, Boston, 1992).
12. E. Y. Lam, "Image restoration algorithms for monochrome digital photography," Ph.D. dissertation, (Stanford University, Stanford, Calif., 2000).