JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION
Vol. 9, No. 1, March, pp. 2–14, 1998
ARTICLE NO. VC970378

Review of Postprocessing Techniques for Compression Artifact Removal

Mei-Yin Shen and C.-C. Jay Kuo

Integrated Media Systems Center, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564

Received August 22, 1997; accepted March 2, 1998

Low bit rate image/video coding is essential for many visual communication applications. When bit rates become low, most compression algorithms yield visually annoying artifacts that highly degrade the perceptual quality of image and video data. To achieve high bit rate reduction while maintaining the best possible perceptual quality, postprocessing techniques provide one attractive solution. In this paper, we provide a review and analysis of recent developments in postprocessing techniques. Various types of compression artifacts are discussed first. Then, two types of postprocessing algorithms based on image enhancement and restoration principles are reviewed. Finally, current bottlenecks and future research directions in this field are addressed. © 1998 Academic Press

Key Words: **postprocessing; compression artifacts; image enhancement; image restoration; data compression; low bit rate coding.**

1. INTRODUCTION

A wide range of new applications in visual communication are made possible due to rapidly evolving telecommunication and computer technologies. Almost all these applications, including mobile or PSTN videotelephony, videoconferencing, and video over the Internet, require very efficient data compression methods to fit a large amount of visual information into the narrow bandwidth of communication channels while acceptable quality of reconstructed data is preserved. This is, however, challenging for most existing coding algorithms and standards, since highly visible artifacts appear when the coding bit rate is low. A large amount of research has been done to improve the reconstructed image/video quality at low bit rates. Among them, postprocessing techniques which are performed after decoding provide an interesting solution, since they can be easily incorporated in existing standards. Postprocessing aiming at the reduction of compression artifacts improves the overall perceptual quality for a given bit rate or, equivalently, increases the compression ratio with respect to a given quality requirement.

The visual quality of compressed images and video is generally affected by three factors: data source, coding bit rates, and compression algorithms. For a given compression method, more information (motion or spatial details) contained in the source signal requires more bits for representation. When compressed at the same bit rate, images with more details usually degrade more than those with fewer details. The coding bit rate is another important factor that determines quality. In lossy compression, there is a trade-off between the bit rate and the resulting distortion. The lower the bit rate, the more severe the coding artifacts due to loss of information. Furthermore, the type of artifacts depends on the compression algorithm used. For coding algorithms based on block DCT, the major artifacts are characterized by blockiness in flat areas and ringing along object edges. For wavelet-based coding techniques, ringing is the most visible artifact. Reduction of these artifacts can result in a significant improvement in the overall visual quality of the decoded images.

There are two strategies commonly adopted to reduce compression artifacts. One is to solve the problem at the encoder end, which is known as the preprocessing technique. The other uses a postprocessing technique at the decoder end. For a given decoding algorithm and a bit rate constraint, the quality of reconstructed images can be enhanced by prefiltering techniques [11, 32] that remove unnoticeable details in source images so that less information has to be coded or by perceptual-based coding techniques [23, 28] that optimize bit allocation based on a human visual model so that the degradation is less visible by human observers. Preprocessing techniques have been widely used in modern speech and audio coding. In contrast, there is much less amount of work in the area of image/video coding. In still image coding, postprocessing techniques at the decoder end have received a lot of attention and become an active research area recently. It is adopted to remove compression artifacts without increasing the bit rate or modifying the coding procedure.

Most postprocessing algorithms have been proposed to

1047-3203/98 $25.00

reduce coding artifacts resulting from block DCT, since the transform has been adopted in the JPEG and MPEG compression standards. Generally speaking, they are derived from two different viewpoints, i.e., image enhancement and image restoration. For algorithms based on image enhancement, the goal is to improve the perceived quality subjectively. The special structure of artifacts and human visual sensitivities are taken into account in the design of image enhancement methods. One typical example is the application of filtering along block boundaries to reduce blockiness. Image enhancement methods are heuristic in the sense that no objective criterion is optimized. With the image restoration approach, one formulates postprocessing as an image recovery problem. Reconstruction is performed based on the prior knowledge of the distortion model and the observed data at the decoder. Several classical image restoration techniques, including constrained least squares (CLS), projection onto convex sets (POCS), and maximum a posteriori (MAP) restoration have been used to alleviate compression artifacts. In addition to quality improvement, the computational complexity of these postprocessing techniques is another important issue for applications that require real-time processing. In this survey paper, our primary objective is to review recent developments in postprocessing techniques, especially those developed to remove artifacts due to block DCT coding. A new hybrid postprocessing method, which is very simple and effective, is presented in Section 5. With this hybrid method, each image block is classified into one of three block types: flat, edge, and texture blocks, and then different postprocessing techniques are applied to each type of blocks.

This paper is organized as follows. In Section 2, various kinds of compression artifacts are presented and discussed. Two types of postprocessing algorithms based on image enhancement and image restoration are examined, respectively, in Sections 3 and 4. A new hybrid postprocessing technique which utilizes POCS in the smooth regions and nonlinear smoothing in edge regions is proposed in Section 5. Experimental results of the proposed method are given in Section 6. Finally, current bottlenecks, future research directions, and concluding remarks are given in Section 7.

2. REVIEW OF COMPRESSION ARTIFACTS

To have a better understanding of quality degradation in compressed images and video, we classify artifacts into several types and discuss each of them separately in this section.

2.1. Image Compression Artifacts

A. Blocking Artifact

The blocking effect is the most noticeable artifact associated with both JPEG and MPEG compression standards. Other coding techniques that involve block partitioning, such as vector quantization, block truncation coding, and fractal-based compression also suffer from this artifact. In block-based coding schemes, blockiness arises since each block is encoded without considering the correlation between adjacent blocks.

B. Ringing Artifact

The ringing effect is caused by a coarse quantization of high frequency components. When severely quantized, high frequency coefficients are filtered out. Through the inverse transform (or the synthesis filter bank), the quantization error appears as ringing noise along sharp edges. This noise shows up in the area influenced by the impulse response length of the filter in use. The longer the impulse response length is, the wider the ringing noise spreads. For example, in the JPEG and MPEG compression standards, the impulse response length of the DCT transform is eight. Because of the short impulse response length, ringing is less perceptible in JPEG and MPEG coded images, especially at medium or high bit rates. It becomes much more visible at low bit rates. Generally speaking, the ringing effect occurs in all coding schemes that involve quantization in the frequency domain. The ringing effect is the most noticeable artifact in the subband/wavelet coding schemes at low bit rates.

C. Blurring

Blurring also results from loss of high frequency components and occurs in all lossy coding techniques at low bit rates. However, a moderate degree of blurring is generally not annoying to human observers. As the viewing distance is increased, the effect of blurring is decreased. It can be completely removed when the viewing distance is increased to such a point that the missing frequencies fall outside the passband of the human visual system (HVS).

Examples of the above three compression artifacts are shown in Fig. 1. The original woman image given in Fig. 1a is one of the JPEG2000 test images. Two regions of the original image are boxed and enlarged in Figs. 1b and e for reference. When the image is encoded with JPEG at the rate of 0.26 bpp, significant blocking and ringing effects appear in these two regions as shown in Figs. 1c and f. The same image is also encoded with a wavelet image coder at the rate of 0.125 bpp, and the two enlarged regions are shown in Figs. 1d and g, where the ringing effect along sharp edges is quite noticeable. When compressed at such low bit rates, ''out of focus'' blurring is also observed in Figs. 1c, d, f, and g.

D. Texture Deviation

Another type of distortion is known as texture deviation, which is caused by the loss of fidelity in mid-frequency

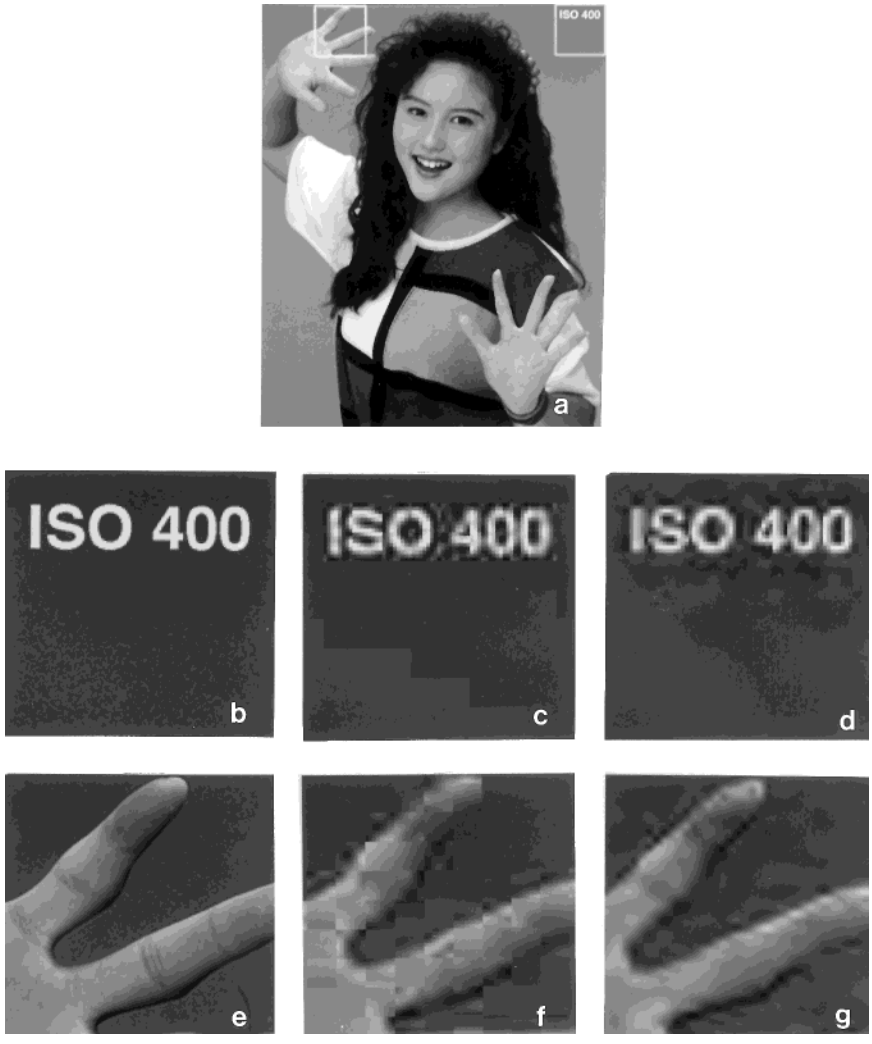

FIG. 1. Illustration of compression artifacts: (a) is the original woman image; (b) and (e) are the enlarged (original) text and hand regions; (c) and (f) show the text and hand regions coded by JPEG at 0.26 bpp; and (d) and (g) show the text and hand regions coded by a wavelet-based coder at 0.125 bpp.

components and appears as granular noise. In transform coding, it is less visible to the human visual system. However, in segmentation- or model-based coding [16] texture deviation often manifests itself as an oversmoothing of texture patterns that can be visually annoying.

2.2. Video Compression Artifacts

All image artifacts discussed above also appear in image sequences in a similar fashion. In addition, there are some artifacts that appear only in compressed video.

A. Flickering

One major temporal artifact in compressed video is flickering, which appears as background noise. It can be either the random noise due to digitization of original video data or the quantization noise caused by compression such as the ringing and blocking effects. Besides background noise, unequal quantization levels between adjacent frames may also cause flickering. For still image compression, the quality factor across the whole image is often approximately the same. However, different bit rates can be assigned to different frames or even to different regions in the same frame for the rate control purposes in video compression. The variation of this quality factor can also lead to temporal flickering.

B. Motion Jerkiness

An image sequence is a series of discrete frames displayed at a suitable rate. If the frame rate is high enough, the observer does not see the individual element of the

display but the continuous motion of the objects. However, at low bit rates, the frame rate may drop below the threshold which enables the perception of smooth motion so that objects in video have discontinuous jumps in their motion. It was shown in [2] that the maximum object displacement between adjacent frames without being noticed is a function of the spatial frequency. For example, with the same velocity, it is easier to see jerky motion in spatially varying regions such as edges or textures.

3. POSTPROCESSING VIA IMAGE ENHANCEMENT

While reduction of compression artifacts in still images has been studied extensively, little work has been done in the quality improvement of compressed video. Therefore, we will focus primarily on postprocessing techniques for still image artifact removal in this survey paper. We classify postprocessing algorithms according to their solution approach. Roughly speaking, a solution can be derived from two different principles: image enhancement and image restoration. Methods based on these two principles are examined in Sections 3 and 4.

In most applications of visual communication, the major concern about quality is how well images are perceived by a human observer, not how close they are to original images. Therefore, postprocessing methods can be designed to match the perception of the human visual system. For example, all artifacts described in Section 2 may appear in all regions of a DCT-coded image. However, due to perceptual masking [27], the high frequency noise in the flat region is more visible to human eyes and thus requires special attention for perceived quality improvement. A block classification method [15] which distinguishes different types of blocks can be very useful as an initial step. The image enhancement approach aims at smoothing visible artifacts instead of restoring the pixel back to its original value.

3.1. Postfiltering

Since block and ringing effects resulting from block DCT coding are high frequency artifacts, a straightforward solution is to apply lowpass filtering to the region where artifacts occur. A space-invariant filtering method to reduce blocking artifacts in image coding was first proposed by Reeves and Lim [34]. To maintain the sharpness of the decoded image, filter coefficients must be selected carefully. However, filtering without considering the local image statistics often causes the loss of high frequency details such as edges. Therefore, a number of adaptive spatial filtering techniques [12, 13, 19, 33, 37] have been proposed to overcome this problem. Generally speaking, as adaptive filtering technique uses classification and edge detection to categorize pixels into different classes for adaptation. Then, different spatial filters (either linear or nonlinear) are used to remove coding artifacts according to the label information. Classification is essential to adaptive filtering techniques which attempt to exploit local statistics of image regions and the sensivity of human eyes. Classification can be performed either in the spatial domain by using local variances [33] or in the transform domain by examining the distribution of DCT coefficients [15, 19]. For blocks of size 8×8, an example of the transform domain classification scheme can be expressed as

$$\sum_{i=k}^{63} |C_i| \leq T, \tag{1}$$

where C_i are the quantized DCT coefficients, k can be an integer between 0 and 63, and T is a threshold value. If the left-hand side of (1) is smaller than threshold T, the block is classified as a smooth block. Otherwise, it is classified as an edge or a texture block. Previous work [37, 13, 19] also attempted to estimate the edge information in compressed images to preserve the quality of edges. However, it is worthwhile to emphasize that edge detection is usually sensitive to quantization, and it is a very difficult task to extract edges accurately from highly compressed images since edges are somewhat distorted.

As each block of an image has been properly classified, one or more appropriate spatial filters can be used to smooth coding artifacts. To avoid blurring, lowpass filtering is only applied to areas that are not masked by the presence of detail information such as smooth regions. Two-dimensional separable lowpass FIR filters provide one popular choice to artifact reduction. To achieve better spatial adaptivity, the length of the lowpass filter may vary. Long-tap filters are used in a larger smooth region. For smooth areas close to an edge, short-tap filters such as a filter of size 3×3, are preferred. For block coded images, the jagged edge appearance can be efficiently smoothed by applying a 1D directional filter which is perpendicular to the direction of the edge. In some implementations, nonlinear filters are used to achieve the goal. In the work of Kundu [12], the Hodges–Lehman D filter [5] was used in large smooth areas while a 5×5 median filter or multistage median filter was applied to edges. It was reported that combination of these nonlinear filters can effectively reduce artifacts in block DCT coded images.

Adaptive spatial filtering is also an efficient method to reduce blocking and ringing effects in image sequences. However, considering additional temporal artifacts such as flickering and motion unsmoothness, it is usually integrated with temporal filtering [20, 40]. A temporal filter that smoothes noisy image sequences along motion trajectories is commonly used as the prefilter within the motion compensation loop at the encoder [4, 9, 38] to reduce the frame rate and increase the temporal correlation. In block

motion-compensated DCT coding, a temporal filter can also be used as the postfilter to smooth motion and reduce flickering. To reduce the memory requirement, temporal filters are often implemented as HR filters [40], since the HR filter can achieve noise attenuation with a lower order than the FIR filter. A typical temporal filter can be written as

$$y(m, n, k) = \alpha x(m, n, k) + (1 - \alpha) y(m - \Delta m, n - \Delta n, k - 1),$$

where $y(m, n, k)$ is the pixel at position (m, n) of the kth frame, $y(m - \Delta m, n - \Delta n, k - 1)$ is its best matching pixel in the previous reference frame, and α is the filter coefficient which can be adjusted according to statistics along the temporal domain. In general, a small value of α (typically smaller than 0.5) is adopted in the stationary area, and $\alpha = 1$ is used in scene changes or time intervals containing fast motion. However, there are several shortcomings with this approach. For example, a blocking artifact that remains at the same position for a few frames is difficult to reduce by using only temporal filtering. Yang *et al.* [40] incorporated adaptive spatial and temporal filtering processes and applied them iteratively to reduce video coding artifacts.

To conclude, adaptive postfiltering (spatial or temporal) has the advantage of a low computational complexity and the adaptation to local statistics relies on good classification schemes. However, difficulty may arise for image/video coded at low bit rates since adaptation may not be robust. Poor estimation and adaptation in postprocessing can sometimes cause image quality degradation.

3.2. AC Prediction

Niss [29] proposed a technique to reduce the blocking effect in JPEG by predicting the low-frequency AC coefficients from DC coefficients in a given block and its eight neighboring blocks. A quadratic surface, given by

$$P(x, y) = a_1x^2y^2 + a_2x^2y + a_3xy^2 + a_4x^2 + a_5xy + a_6y^2 + a_7x + a_8y + a_9 \quad (2)$$

is used to model the shape of a local surface. Coefficients a_i, $1 \le i \le 9$, can be determined by finding the best fit to the surface determined by nine DC values in an array of 3×3 blocks. Once coefficients $a_1, \ldots, a_9$ are determined, low-frequency AC coefficients required to reproduce the quadratic surface can be calculated by a set of equations represented in terms of these nine DC values. To avoid the introduction of a large distortion value, predicted AC coefficients are confined to their original quantization intervals. This method considerably reduces the blocking effect in the smooth area, but the prediction is not very reliable in the neighborhood of edges.

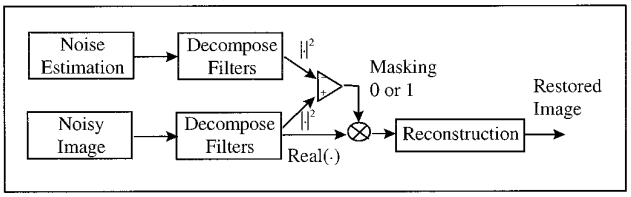

FIG. 2. The block diagram of perceptual-based image enhancement.

3.3. Enhancement Based on Perceptual Model

A postprocessing scheme based on properties of HVS was proposed by Macq *et al.* [24]. According to the perceptual model, a stimulus to a given perceptual channel is not perceived if the contrast value is below the visibility threshold corresponding to this channel. The visibility threshold for each channel depends on the content of the background in that channel, which is known as the masking effect [22]. Macq *et al.* assumed a simple masking model. That is, noise is visible if and only if its perceptual component power is equal to or greater than that of the background. The global scheme of the postprocessing procedure is shown in Fig. 2. To remove compression artifacts, the decompressed JPEG image is first decomposed into several perceptual channels. Then, unmasked artifacts are removed by simply setting the corresponding perceptual components to zero in each perceptual channel. After thresholding, all perceptual channels are integrated to obtain an artifact-free image. The visibility thresholds of artifacts are determined by the perceptual components of the noisy image contrast to the original image in each perceptual channel. However, since the original image is not available at the decoder, the contrast is actually obtained by computing the difference between the decoded image and estimated noise. Noise estimation is achieved by filtering the noisy image, applying the image codec to the filter image, and then subtracting the obtained image from the filtered image to regenerate the noise.

4. POSTPROCESSING VIA IMAGE RESTORATION

The image restoration approach treats artifact removal as an ill-posed image recovery problem. A number of classical image restoration algorithms have been tailored to the deblocking of DCT coded images. The diagram of the image restoration approach is shown in Fig. 3. In this section, we classify different postprocessing methods into the following three categories:

1. criterion-based methods such as minimum least

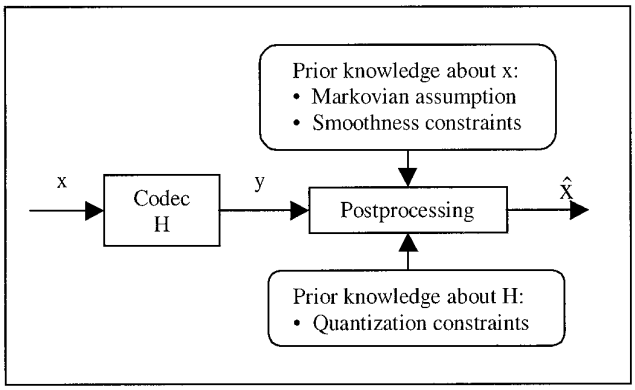

FIG. 3. An image restoration approach to postprocessing.

squares (MLS), minimum mean square error (MMSE), or maximum a posteriori (MAP) probability;

2. constrained optimization methods such as constrained least squares (CLS);

3. constraint-based methods such as projection onto convex sets (POCS) where constraints are established based on some prior knowledge of the original image and the compression process.

4.1. Criterion-Based Methods

The basic concept behind criterion-based methods is to find the solution (i.e., an artifact-free image), which satisfies some predefined optimality criterion. The choice of the criterion is critical to the design of postprocessing algorithms. The best choice for the criterion is an error measurement that provides a high correlation to perceptual image quality. However, since there is no widely agreed-upon mathematical expression of a human perception quality metric at present, the mean square error is still the commonly used criterion.

A. Minimum Mean Square Error (MMSE) Formulation

Let $\mathbf{y}$ be the decompressed (or reconstructed) image of the original image $\mathbf{x}$. The degradation due to compression can be written as

$$\mathbf{y} = \mathbf{x} + \mathbf{e}, \tag{3}$$

where $\mathbf{e}$ is the reconstruction error. We would like to find an estimate $\hat{\mathbf{x}}$ of the original image so that the cost

$$\mathbf{E}\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} \tag{4}$$

is minimized. The best estimate $\hat{\mathbf{x}}$ can be obtained by using a linear or a nonlinear estimator acting on the observed image $\mathbf{y}$.

An example of using the linear minimum mean square error (LMMSE) criterion to reduce coding artifacts in MPEG coded video is given by Nakajima *et al.* [26]. With the assumption that $\mathbf{e}$ is zero-mean white noise uncorrelated with $\mathbf{x}$, each pixel $y(i, j)$ of the decompressed image can be restored by the LMMSE estimate of the original signal $\hat{x}(i, j)$. The solution is given by

$$\hat{x}(i,j) = \mu_x(i,j) + \alpha[y(i,j) - \mu_y(i,j)], \tag{5}$$

where

$$\alpha = \frac{\sigma_x^2(i,j)}{\sigma_x^2(i,j) + \sigma_e^2(i,j)}, \tag{6}$$

and $\mu_x(i, j)$ and $\mu_y(i, j)$ are local means calculated by a local averaging operator. The calculation of the variance $\sigma_x^2(i, j)$ requires the estimate of the noise variance $\sigma_e^2(i, j)$. However, the actual value of $\sigma_e^2(i, j)$ is not available at the decoder end and has to be estimated from the observed image or other prior information. In the work of Nakajima *et al.* [26], $\sigma_e^2(i, j)$ was obtained empirically by training several test image sequences. The LMMSE estimator that achieves a compromise between artifact reduction and detail preservation should be selected. A similar approach that uses linear estimators to implement deblocking for JPEG coded images was proposed by Hong, Chan, and Siu [8]. In their work, the restoration technique was only applied to pixels at block boundaries to avoid blurring.

Wu and Gersho [39] proposed a nonlinear estimator that employed nonlinear interpolation and vector quantization to improve the reconstructed image quality coded with the JPEG baseline process. Instead of adopting the conventional JPEG decoding process, each quantized DCT coefficient in the 8×8 image block was used as an index to fetch a codeword from the corresponding codebook. The sum of all 64 codewords formed the final reconstructed image block, with the codebook generated from a sufficiently large training set based on the MMSE criterion. Simply speaking, their method learns the nonlinear statistical relationship from the training set and exploits this relationship to restore proper values. Since restored values are limited to a linear combination of codebook entries, this method is sensitive to the choice of the training set. Another nonlinear estimator based on wavelet decomposition and soft thresholding in the wavelet domain was proposed by Gopinath *et al.* [6], where a threshold was selected to minimize the mean square error in (4).

B. Maximum A Posteriori (MAP) Probability Formulation

In MAP-based restoration, the criterion is to maximize the probability density of the actual image conditioned on

the observation. Let $\mathbf{y}$ be the observed image that contains compression artifacts. The original image $\mathbf{x}$ can be estimated by a MAP estimate $\hat{\mathbf{x}}$ via

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}).$$

By applying Bayes' rule and the log-likelihood function, we obtain

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \{\log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x}) - \log p(\mathbf{y})\}. \quad (7)$$

Since $p(\mathbf{y})$ is constant with respect to the optimization parameter $\mathbf{x}$, the term $\log p(\mathbf{y})$ can be dropped. By assuming that the given image $\mathbf{x}$ is compressed to the same observation $\mathbf{y}$ every time (i.e., fixed compression method and compression ratio), (7) can be rewritten as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{X}} \{-\log p(\mathbf{x})\}, \quad (8)$$

where $\mathbf{X}$ is the set of images which are compressed to $\mathbf{y}$ and $p(\mathbf{x})$ is the a priori probability modeled by a general form of the Gibbs distribution associated with a certain Markov random field (MRF),

$$p(\mathbf{x}) = \frac{1}{Z}\exp\left\{\sum_{c\in C}\phi_c(\mathbf{x})\right\}. \quad (9)$$

In this expression, $\phi_c(\mathbf{x})$ is the energy function of a first-order MRF defined over transitions between adjacent neighbors [3] such that

$$\phi_c = \sum_{i,j}\varphi(x_i - x_j), \quad (10)$$

where $\varphi(\cdot)$ denotes a potential function. The form of the potential function can play a crucial role in the estimation process. It is chosen to reflect prior expectations and requirements of a given reconstruction problem. For example, the Huber minimax function

$$\varphi_\alpha(t) = \begin{cases} t^2, & |t| \le \alpha, \\ \alpha^2 + 2\alpha(|t| - \alpha), & |t| > \alpha, \end{cases} \quad (11)$$

can be used as a potential function to reduce coding artifacts [18, 20, 17]. The Huber function that allows some inconsistency is able to smooth artifacts while preserving the sharpness of edges. The convex property of $\varphi_\alpha(t)$ makes the optimization procedure tractable.

The estimate in (8) can be determined by several iterative techniques, including the gradient projection method [30], or the iterative conditional mode (ICM) [18]. Another MAP-based method that adopts a nonstationary Gauss–Markov model was proposed by Özcelik *et al.* [31]. The maximization of the posterior function of this approach was carried out by using mean field annealing. Li and Kuo [17] proposed a multiscale MAP method, in which the decoded image was enhanced from coarse to fine scales. Postprocessing at coarse scales improves the global appearance of the image and reduces long-range artifacts such as ringing, while postprocessing at finer scales preserves the sharpness of edges. It achieved a better performance in suppressing ringing artifacts and computational speed than the single-scale MAP method.

4.2. Constrained Optimization Methods

This class of methods optimizes an optimality criterion subject to constraints on the solution, which are obtained from a priori knowledge of the original image. A well-known method of this type is the method of constrained least squares (CLS). With the image degradation model in (3), the estimate is defined to be a member within the intersection of the following two sets:

$$C_\sigma = \{\mathbf{x}: \|\mathbf{x} - \mathbf{y}\|^2 \le \sigma^2\}, \quad C_\varepsilon = \{\mathbf{x}: \|\mathbf{A}\mathbf{x}\|^2 \le \varepsilon^2\}.$$

Computing the CLS estimate involves minimizing the objective function

$$J(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\|^2 + \left(\frac{\sigma}{\varepsilon}\right)^2 \|\mathbf{A}\mathbf{x}\|^2.$$

Operator $\mathbf{A}$ in the above expression is chosen such that C_ε summarizes the prior knowledge of the original image. In general, $\mathbf{A}$ is a highpass filter which imposes the smoothness constraint. Furthermore, σ and ε are selected so that the set intersection is nonempty [25]. The ratio of σ^2 and ε^2 is used as the regularization parameter that controls the degree of smoothness within the restored image. An iterative solution [10] of the estimate is given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \beta\left(\mathbf{y} - \left(\mathbf{I} + \left(\frac{\sigma}{\varepsilon}\right)^2 \mathbf{A}^t\mathbf{A}\right)\mathbf{x}_k\right), \quad (12)$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{A}^t$ is the transpose of $\mathbf{A}$, and the relaxation parameter β is chosen within the range

$$0 < \beta < \frac{2}{\left\| I + \left(\frac{\sigma}{\varepsilon}\right)^2 \mathbf{A}^t\mathbf{A}\right\|}$$

to ensure convergence. Yang *et al.* [41] solved (12) in the DCT domain with the quantization constraint. Hong, Chan, and Siu [7] incorporated CLS regularization in sub-

band decomposition/reconstruction to remove the blocking effect.

4.3. Constraint-Based Methods

The basic idea behind constraint-based methods is to impose a number of constraints on the coded image and to restore it to its artifact-free form accordingly. Constraints can be obtained from the prior knowledge of the compression algorithm, noise, or properties of the input image. One famous example in this class of solutions is called the projection onto convex sets (POCS). Generally speaking, the use of POCS for image restoration has a long history and can be found in the classical image restoration literature [43]. In POCS, a constraint set is defined as a closed convex set whose members are consistent with a priori knowledge of the original image. For m known properties, we may be able to construct m closed convex sets, C_i, $i = 1, 2, \ldots, m$. A feasible solution, which is in the intersection of all sets, can be found via an iterative process. That is, we start from a certain initial guess, and project the result onto each constraint set in sequence until it converges to the final solution.

In the context of image postprocessing, two constraint sets are commonly used in POCS. One is the quantization constraint set (C_q) determined from the known quantization intervals of each DCT coefficient in the DCT domain. The other is the smoothness constraint set that captures the smoothness properties of the original image (C_s) in the spatial domain. Let P_q and P_s be the two projection operators which project a solution onto C_q and C_s, respectively. Then, the artifact-free image can be obtained by performing the iteration

$$\mathbf{f}_{k+1} = P_q P_s \mathbf{f}_k, \tag{13}$$

where $\mathbf{f}_k$ is the postprocessed image at the kth iteration. The projection operator P_q simply restricts each DCT coefficient to its original quantization interval. The definition of P_s is important for noise smoothing. Since there is no well-defined quantitative measure of image smoothness, the choice of P_s varies. In the work of Rosenholtz and Zakhor [36], C_s was chosen to be the set of band-limited images. A projection onto this set is equivalent to a lowpass filtering process. However, the filter kernel used in their experiments is not an ideal lowpass filter, and it was pointed out by Reeves and Eddins [35] that the convergence of their algorithm should be justified from the viewpoint of constrained minimization rather than that of POCS. In the work of Yang *et al.* [41], the smoothness constraint set was selected to capture discontinuities along block partition boundaries to reduce the blocking effect. This set is of the form

$$C_s = \{\mathbf{f}: \|\mathbf{A}\mathbf{f}\| \le s\}, \tag{14}$$

where $\mathbf{A}$ is a highpass filter. The projection operator that maps an arbitrary image onto the smoothness constraint set in (14) can be derived and takes the form

$$P_s(\mathbf{f}_i) = \begin{cases} \mathbf{f}_{i+1} + \alpha \cdot (\mathbf{A}\mathbf{f}_i), & \text{if } i = 8n, \\ \mathbf{f}_i - \alpha \cdot (\mathbf{A}\mathbf{f}_i), & \text{if } i = 8n+1, \\ \mathbf{f}_i, & \text{otherwise}, \end{cases} \tag{15}$$

where $\mathbf{f}_i$ is the ith row or column, $\mathbf{A}\mathbf{f}_i$ is the pixel-difference between the i block boundaries, and $\alpha = 0.5(1 + s/\|\mathbf{A}\mathbf{f}\|)$. A weighted version of (14) is given by Yang *et al.* [42] to achieve spatial adaptivity.

Since POCS is an iterative process and requires one DCT and IDCT pair in each iteration, its computational complexity is extremely high. Many techniques have been proposed to reduce the complexity per iteration and/or to speed up the convergence rate. For example, Kwak and Haddad [14] improved the complexity of the algorithm of Yang *et al.* [41] by canceling out the IDCT and DCT pairs in POCS iterations. Lai *et al.* [15] proposed a spatially adaptive scheme to match the human perceptual characteristics and to improve the convergence rate by introducing frequency domain filtering.

All POCS methods described above aim at reducing the blocking effect. If there are other types of artifacts present in the image, additional constraint sets are required and the corresponding computational complexity increases as well. Also, it is challenging to determine the proper constraint for nonblocking artifacts. For example, the constraint set for the ringing artifact is very difficult to define. However, it is known that the ringing effect can be suppressed effectively by the median filter.

5. HYBRID POSTPROCESSING ALGORITHM

In this section, we will present a new hybrid method that demonstrates an excellent image enhancement capability with a very low computational cost. The proposed algorithm is detailed below.

Step 1: *Classification*. The coded bitstream is first dequantized. Each 8×8 block in an image is classified into two categories: flat block, and high activity block. There exist many classification rules. A simple DCT domain classification rule is adopted here and stated as follows. Let V, H, and D be three directional indices that correspond to different edge directions of the block in spatial domain. The values of V, H, and D for each block can be calculated as

$$\begin{aligned} V &= \sum_{i=m,n} C_{m,n}, && 2 \le m \le 7, 0 \le n \le 1, \\ H &= \sum_{i=m,n} C_{m,n}, && 0 \le m \le 1, 2 \le n \le 7, \\ D &= \sum_{i=k} C_{k,k} + \sum_{i=k} C_{k,k-1} + \sum_{i=k} C_{k-1,k}, && 2 \le k \le 5, \end{aligned} \tag{16}$$

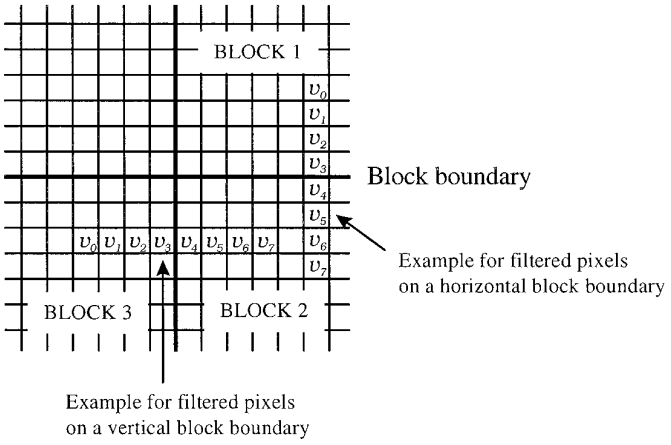

FIG. 4. Examples of positions of filtered pixels.

where $C_{m,n}$ is dequantized DCT coefficient at position (m, n). Let M be the maximum among three indices V, H, and D. If M is less than a threshold value, the block is classified as a flat block; otherwise, it is classified as a high activity block, which can be either a texture or an edge block. The threshold value T determines the sensitivity of classifier and it is selected from training images such that the probability of misclassification error is minimized. In our implementation, we chose T between 20 to 30.

Step 2: *Deblocking*. Our deblocking filter is a spatial adaptive low pass filter. In a very smooth region, filtering only on block boundaries is not good enough to reduce the blocking artifact due to dc offset between adjacent blocks. Therefore, we apply a stronger lowpass filter to flat blocks. As an example in Fig. 4, Blocks 1–3 are all classified as flat blocks. A 1D 7-tap filter (1, 1, 2, 4, 2, 1, 1) is applied along the vertical boundary to obtain eight new pixel values as indicated by v_0, v_1, v_2, v_3, v_4, v_5, v_6, v_7. The same filtering operation is also applied along the horizontal boundary. For high activity blocks such as texture and edge blocks, filtering over the entire block shows no visual improvement, but results in blurring. In the case of high activity blocks, it is desirable that the deblocking filter is applied once and restricted to the boundary pixels only. In this case, we apply a 1D 3-tap filter (1, 4, 1) to obtain new boundary pixel values v_3, v_4.

Step 3: *Deringing*. Although all high activity blocks may contain ringing artifacts, not all of them require deringing filtering. As mentioned in the previous section, the ringing artifact near the flat regions is more visible to human eyes due to perceptual masking. Thus, we divide the high activity blocks into two groups: edge blocks and texture blocks. For the block with all high activity blocks in its four connected neighbors, we treat it as a texture block; otherwise, it is treated as an edge block. Extraction of edge blocks can be done by a simple morphological process on the classification map. Since the ringing artifact is confined to blocks, the deringing filter is applied only to the block classified as an edge block and only within the block. An example of classification is shown in Fig. 5. The deringing filter consists of two steps: edge pixel extraction and adaptive smoothing. First, a simple gradient operator is applied to locate the position of edge pixels within the block. Second, to preserve the quality of the edge, no process is applied to edge pixels. A 5×5 multilevel median filter [1] is applied to pixels marked as nonedge pixels. The median filter is stated as follows. Let

$$z_k(n_1, n_2) = median[a(.,.) \in \mathbf{W}_k[a; (n_1, n_2)]], \quad 1 \le k \le 4,$$

where

$$\begin{aligned}
\mathbf{W}_1[a; (n_1, n_2)] &= \{a(n_1 + l_1, n_2) &&: -2 \le l_1 \le 2\},\\
\mathbf{W}_2[a; (n_1, n_2)] &= \{a(n_1 + l_1, n_2 + l_1) &&: -2 \le l_1 \le 2\},\\
\mathbf{W}_3[a; (n_1, n_2)] &= \{a(n_1, n_2 + l_1) &&: -2 \le l_1 \le 2\},\\
\mathbf{W}_4[a; (n_1, n_2)] &= \{a(n_1 + l_1, n_2 - l_1) &&: -2 \le l_1 \le 2\}.
\end{aligned}$$

The output of the median filter is given by

$$y_{med}(n_1, n_2) = med[y_{max}(n_1, n_2), y_{min}(n_1, n_2), a(n_1, n_2)], \tag{17}$$

FIG. 5. An example of classification. Light gray, dark gray, and black represent edge, texture, and flat block, respectively.

TABLE 1
PSNR Improvements of Images Processed by Various Postprocessing Algorithms

Image	Size	bpp	PSNR (dB)	PSNR improvement			Time (s)		
				POCS [41]	MMAP [17]	Proposed	POCS [41]	MMAP [17]	Proposed
Goldhill	720 × 576	0.198	29.493	0.502	0.679	0.617	6.4	87.3	6.1
Hotel	720 × 576	0.268	29.095	0.469	0.802	0.725	7.8	83.3	6.0
Woman	512 × 640	0.211	29.355	0.341	0.676	0.521	6.0	68.8	4.8
Cafe	512 × 640	0.536	22.961	0.156	0.493	0.342	9.7	67.2	4.3
Lena	512 × 512	0.246	30.704	0.685	1.011	0.897	4.6	54.1	3.8
Cameraman	256 × 256	0.376	25.176	0.180	0.599	0.133	2.4	10.7	0.9

where

$$y_{max}(n_1, n_2) = max_{1\le k\le 4}[z_k(n_1, n_2)],$$
$$y_{min}(n_1, n_2) = min_{1\le k\le 4}[z_k(n_1, n_2)], \tag{18}$$

Step 4: *Quantization constraint*. In the final step, we check if the DCT coefficients are still within their original quantization bins. If they are outside the original bins, we adjust their values to satisfy this constraint.

6. EXPERIMENTAL RESULTS

Experiments were carried out to evaluate the performance of various postprocessing algorithms. We used a number of images of different sizes selected from JPEG2000 test images and other *de facto* standard test images. These images were encoded with the JPEG standard, with the same quantization table used in [41]. For comparative study, three postprocessing algorithms: POCS [41], multiscale MAP(MMAP) [17], and the method proposed in Section 5 were applied to the decoded images.

Table 1 summarizes the PSNR improvement and computing time of different approaches with respect to the original JPEG-coded images. All computing was performed on a Pentium Pro 200 MHz Machine with 64 MB of memory. We see from the results that MMAP has the best PSNR performance, but the processing time is much longer than the other two methods due to the high complexity. The proposed method provides a slightly worse objective improvement, but with much less computing time. The POCS method only restores the boundary pixels so that the PSNR improvement is the worst among the three. Its complexity is also high since it requires several iterations to converge, where each iteration has to perform one DCT/IDCT pair. Furthermore, both MMAP and POCS require buffering the whole frame in the iterative process. Our new scheme only utilizes information in neighboring blocks, so that it saves memory bandwidth. It is clear that our new scheme is more suitable for on-line real-time process than the other two methods.

It is worthwhile to point out that PSNR is not the best measure to evaluate the postprocessing algorithm, since it does not completely correlate with human perception. We show the subjective quality of different methods in Fig. 6. It was found that both MMAP and the proposed method achieve better subjective quality. The POCS method does not remove blockiness completely.

7. CONCLUSION AND FUTURE WORK

Compression of digital image and video data plays an important role in reducing the transmission and storage cost for visual communication. However, when bit rates become very low, most compression algorithms yield visually annoying artifacts that highly degrade the perceptual quality of image and video data. To achieve high bit rate reduction while maintaining the best possible perceptual quality, postprocessing techniques provide one attractive solution. There has been a significant number of research activities in the postprocessing of compressed image/video over the last decade. In this paper, we provided a review and analysis of recent developments in postprocessing techniques. A discussion on compression artifacts was given first. Then, algorithms based on image enhancement and image restoration were examined in detail, and bottlenecks in each method were also addressed.

Finally, we would like to comment on future directions in this field. Generally speaking, the image enhancement approach has the advantages of being relatively simple and fast, while the image restoration approach produces the best quality at a higher computational cost. For both approaches, adaptively applying postprocessing to match human perceptual sensitivities should improve the performance. Thus, a better understanding of the psychovisual model may lead to an evolutionary improvement of the postprocessing scheme. Even though a number of good postprocessing algorithms are available for still images,

FIG. 6. Subjective quality of different approaches: (a) JPEG coded at 0.26 bpp; (b) POCS [41]; (c) MMAP [17]; (d) proposed method.

work in video postprocessing is still far from complete. Existing algorithms that perform efficiently on still images may not be extended to video in a straightforward manner. Coding artifacts in video are generally more complex and difficult to model. Nevertheless, a real-time video postprocessing scheme with a low computational cost is very desirable in modern digital visual communication.

ACKNOWLEDGMENTS

This research has been funded by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, with additional support from the Annenberg Center for Communication at the University of Southern California and the California Trade and Commerce Agency. The authors thank Dr. Jin Li of the Sharp Laboratory of America and Mr. Yung-Kai Lai of the University of Southern California for their valuable discussions and support.

REFERENCES

1. G. A. Arce and R. E. Foster, Multilevel median filters: Properties and efficiency, in *Proc. ICASSP, 1988*, pp. 824–827.
2. D. C. Burr, L. Ross, and M. C. Morrone, Smooth and sampled motion, *Vision Res.* **26,** 1986, 643–652.
3. J. Besag, Digital image processing: Towards Bayesian image analysis, *J. Appl. Statist.* **16,** No. 3, 1989, 395–407.
4. J. M. Boyce, Noise reduction of image sequences using adaptive motion-compensated frame averaging, in *Proc. IEEE ICASSP, 1992*, pp. III.461–III.464.
5. P. J. Bickel and J. L. Hodges, The asymptotic theory of Galton's test and a related simple estimate of location, in *Ann. Math. Statist.* **38,** 1967, 73–89.
6. R. A. Gopinath, M. Lang, H. Guo, and J. E. Odegard, Wavelet-based post-processing of low bit rate transform coded images, in *Proc. IEEE ICASSP, 1994*, 913–917.
7. S.-W. Hong, Y.-H. Chan, and W.-C. Siu, Subband adaptive regularization method for removing blocking effect, in *Proc. IEEE International Conference on Image Processing, Oct. 1995*, pp. II.523–II.526.
8. S.-W. Hong, Y.-H. Chan, and W.-C. Siu, A practical real-time post-processing technique for blocking effect elimination, in *Proc. IEEE International Conference on Image Processing, Sept. 1996*, pp. II.21–II.24.
9. D. S. Kalivas and A. A. Sawchuk, Motion-compensated enhancement of a noisy image, in *Proc. IEEE ICASSP, 1990*, pp. 2121–2124.

10. A. K. Katsaggelos, Iterative image restoration algorithms, *Opt. Eng.* **28,** No. 7, 1989, 735–748.
11. R. Kliehorst, R. Lagendijk, and J. Biemond, Noisy reduction of image sequences using motion compensation and signal decomposition, *IEEE Trans. Image Process.* **4,** 1995, 274–284.
12. A. Kundu, Enhancement of JPEG coded images by adaptive spatial filtering, in *Proc. IEEE International Conference on Image Processing, Oct. 1995*, pp. 187–190.
13. C. J. Kuo and R. J. Hsieh, Adaptive postprocessing for block encoded images, *IEEE Trans. Circuits Systems Video Technol. Aug. 1995*, pp. 298–304.
14. K. Y. Kwak and R. Haddad, Projection-based eigenvector decomposition for reduction of blocking artifacts of DCT coded image, in *Proc. IEEE International Conference on Image Processing, Oct. 1995*, pp. II.527–II.530.
15. Y.-K. Lai, J. Li, and C.-C. J. Kuo, Image enhancement for low bit rate JPEC and MPEG coding via postprocessing, in *SPIE: Visual Communication and Image Processing '96, Mar. 1996*.
16. H. Li, A. Lundmark, and R. Forchheimer, Image sequence coding at very low bit rate: A review, *IEEE Trans. Image Process.* **5,** 1994, 589–609.
17. J. Li and C.-C. J. Kuo, Coding artifact removal with multiscale postprocessing, in *IEEE International Conference on Image Processing, Santa Barbara, CA, Oct. 1997*.
18. J. Luo, C. W. Chen, K. J. Parker, and T. Huang, Artifacts reduction in low bit rate DCT-based image compression, *IEEE Trans. Circuits Systems Video Technol.* **5,** No. 9, 1996, 1363–1368.
19. W. E. Lynch, A. R. Reibman, and B. Liu, Postprocessing transform coded images using edges, in *Proc. IEEE ICASSP, May 1995*, pp. 2323–2326.
20. T. S. Liu and N. Jayant, Adaptive post-processing algorithms for low bit rate video signals, in *Proc. IEEE ICASSP, 1994*, pp. V.401–V.404.
21. H. S. Malvar and D. H. Staelin, The LOT: Transform coding without blocking effects, in *IEEE Trans. Acoust. Speech Signal Process.* **37,** 1989, 553–559.
22. S. Comes and B. Macq, Human visual quality criterion, in *SPIE: Visual Commun. Image Process. '90, Oct. 1990*.
23. B. Macq, Weighted optimum bit allocation to orthogonal transforms for picture coding, *IEEE J. Select. Areas Comm.* **10,** 1992, 875–883.
24. B. Macq, M. Mattavelli, O. V. Claster, E. van der Plancke, S. Comes, and W. Li, Image visual quality restoration by cancellation of the unmasked noise, in *Proc. IEEE ICASSP, 1994*, pp. V.53–V.56.
25. K. Miller, Least squares method for ill-posed problem with a prescribed bound, *SIAM J. Math. Anal.* **1,** 1970, 52–74.
26. Y. Nakajima, H. Hori, and T. Kanoh, A pel adaptive reduction of coding artifacts for MPEG video signals, in *Proc. IEEE ICASSP, 1994*, pp. 928–932.
27. A. N. Netravali and B. Haskell, *Digital Pictures: Representation and Compression*, Plenum, New York, 1988.
28. N. Nill, A visual model weighted cosine transform for image compression and quality assessment, *IEEE Trans. Commun.* **33,** 1985, 551–557.
29. B. Niss, *Prediction of AC coefficients from the DC values*, ISO/IEC JTC1/SC2/WG8 N745, May 1988.
30. T. P. O'Rourke and R. Stevenson, Improved image decompression for reduced transform coding artifacts, *IEEE Trans. Circuits Systems Video Technol.* **5,** No. 6, 1995, 490–499.
31. T. Özcelik, J. C. Brailean, and A. K. Katsaggelos, Image and video compression algorithms based on recovery techniques using mean field annealing, *Proc. IEEE,* **83,** No. 2, 1995, 304–315.
32. M. Ozjan, M. Sezan, and A. Tekalp, Adaptive motion-compensated filtering of noisy image sequences, *IEEE Trans. Circuits Systems Video Technol.* **37,** 1989, 83–98.
33. B. Ramamurthi and A. Gersho, Nonlinear space variant postprocessing of block coded images, *IEEE Trans. Acoust. Speech Signal Process.* **34,** 1986, 1258–1267.
34. H. C. Reeves and J. S. Lim, Reduction of blocking effects in image coding, *Opt. Eng.* **23,** 1984, 34–37.
35. S. J. Reeves and S. I. Eddins, Comments on iterative procedures for reduction of blocking effects in transform coding, *IEEE Trans. Circuits Systems Video Technol.* **3,** 1993, 439–440.
36. R. Rosenholtz and A. Zakhor, Iterative procedures for reduction of blocking effects in transform coding, *IEEE Trans. Circuits Systems Video Technol.* **1,** 1992, 81–94.
37. K. Sauer, Enhancement of low bit-rate coded images using edge detection and estimation, *Comput. Vision Graphics Image Process.* **53,** 1991, 52–62.
38. M. I. Sezan, M. K. Ozkan, and S. V. Fogel, Temporally adaptive filtering of noisy image sequences using a robust motion estimation algorithm, in *Proc. IEEE ICASSP, 1991*, pp. 2429–2432.
39. S.-W. Wu and A. Gersho, Improved decoder for transform coding with application to the JPEG baseline system, *IEEE Trans. Commun.* **40,** 1992, 251–254.
40. L. Yan, A nonlinear algorithm for enhancing low bit-rate coded motion video sequence, in *Proc. IEEE ICASSP, 1994*, pp. 923–927.
41. Y. Yang, N. Galatsanos, and A. K. Katsaggelos, Regularized reconstruction to reduce blocking discrete cosine transform compressed images, *IEEE Trans. Circuits Systems Video Technol.* **3,** No. 6, 1993, 421–432.
42. Y. Yang, N. Galatsanos, and A. K. Katsaggelos, Projection-based spatially adaptive reconstruction of block-transform compressed images, *IEEE Trans. Image Process.* **4,** No. 7, 1995, 896–908.
43. D. C. Youla, Generalized image restoration by the method of alternating orthogonal projections, *IEEE Trans. Circuits Systems* **25,** 1978, 694–702.

MEI-YIN SHEN received the B.S. and M.S. degree in 1990 and 1993 from National Cheng-Kung University, Tainan, Taiwan, and University of Southern California, respectively. She is currently enrolled in the Ph.D. degree program in the Department of Electrical Engineering—Systems at the University of Southern California. Her research interests are image video compression and postprocessing techniques for varies compression techniques.

C.-C. JAY KUO received the B.S. degree from the National Taiwan University, Taipei, in 1980 and the M.S. and Ph.D. degrees from the

Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively, all in electrical engineering. From October 1987 to December 1988, he was Computational and Applied Mathematics (CAM) Research Assistant Professor in the Department of Mathematics at the University of California, Los Angeles, USA. Since January 1989, he has been with the Department of Electrical Engineering—Systems and the Signal and Image Processing Institute at the University of Southern California, where he currently has a joint appointment as Associate Professor of Electrical Engineering and Mathematics. His research interests are in the areas of digital signal and image processing, audio and video coding, wavelet theory and applications, multimedia technologies and large-scale scientific computing. He has authored more than 280 technical publications in international conferences and journals. Dr. Kuo is a member of SIAM, ACM, a Senior Member of IEEE, and a Fellow of SPIE. He is the Editor-in-Chief for the *Journal of Visual Communication and Image Representation*, and served as Associate Editor for *IEEE Transaction on Image Processing* from 1995–1998 and *IEEE Transaction on circuits and Systems for Video Technology* from 1995–1997. Dr. Kuo received the National Science Foundation Young Investigator Award (NYI) and Presidential Faculty Fellow (PFF) Award in 1992 and 1993, respectively.