On the POCS-Based Postprocessing Technique to Reduce the Blocking Artifacts in Transform Coded Images

Hoon Pack, Rin-Chul Kim, Associate Member, IEEE, and Sang-Uk Lee, Member, IEEE

Abstract—In this paper, we propose a novel postprocessing technique, based on the theory of projections onto convex sets (POCS), to reduce the blocking artifacts in transform-coded images. It is assumed, in our approach, that the original image is highly correlated. Thus, the global frequency characteristics in two adjacent blocks are similar to the local ones in each block. We consider the high-frequency components in the global characteristics of a decoded image, which are not found in the local ones, as the results from the blocking artifact. We employ $N$-point discrete cosine transform (DCT) to obtain the local characteristics, and $2N$-point DCT to obtain the global ones, and then derive the relation between $N$-point and $2N$-point DCT coefficients. A careful comparison of $N$-point with $2N$-point DCT coefficients makes it possible to detect the undesired high-frequency components, mainly caused by the blocking artifact. Then, we propose novel convex sets and their projection operators in the DCT domain. The performances of the proposed and conventional techniques are compared on the still images, decoded by JPEG. The results show that, regardless of the content of the input images, the proposed technique yields significantly better performance than the conventional techniques in terms of objective quality, subjective quality, and convergence behavior.

Index Terms—Blocking artifact, DCT, POCS, postprocessing.

I. INTRODUCTION

The discrete cosine transform (DCT) is the most popular transform technique in the data compression area because its performance on highly correlated signals is close to that of the Karhunen–Loève transform (KLT), which is known to be optimal in the mean-squared error sense [3]. Thus, image coding techniques based on the block discrete cosine transform (BDCT) have been widely used and have found many applications. In particular, the BDCT is the recommended transform technique for both still and moving image coding standards, such as JPEG [4], H.261 [5], and MPEG [6]. But, a major problem relating to the BDCT techniques is that the decoded images, especially at very low bit rates, exhibit the visually annoying blocking artifact. That is, coarse quantizations of DCT coefficients, after dividing an image into nonoverlapped blocks, cause visual discontinuities between two adjacent blocks.

Until now, there have been many approaches to alleviate the blocking artifact in the BDCT image coding techniques, which can be classified into two categories. One is variations of transform structure, such as the interleaved block transform [7], [8], the lapped transform [9], [10], and the combined transform [11]. The other is postprocessing techniques, such as filtering approaches [12], [13], the maximum a posteriori (MAP) probability approach [14], [15], and iterative approaches based on the theory of projections onto convex sets (POCS) [1], [2], [16].

The interleaved block transform techniques were suggested by Farrelle [7] and Pearson [8]. The lapped transform approach, instead of overlapping the data blocks, is to overlap the transform basis functions, such as the short-space fast Fourier transform [9] and the lapped orthogonal transform (LOT) [10]. In the combined transform coding approaches [11], the input image is divided into two sets: highly correlated and less correlated sets. The lossless compression technique is then applied to the highly correlated set, while the BDCT is applied to the less correlated set. Since the interblock correlation in the less correlated set is significantly reduced, the blocking artifact is also reduced.

Each technique described so far requires coding schemes of its own, such as transform, quantization, bit allocation, etc. Thus, unfortunately, these techniques cannot be applied to commercial coding system products, such as JPEG [4], H.261 [5], and MPEG [6]. On the other hand, the postprocessing technique is very attractive because this technique is independent of the coding schemes. The simplest approach among the postprocessing techniques is low-pass filtering. Since the blocking artifact is the result of an artificial boundary discontinuity between two adjacent blocks, it can be characterized by high-frequency components. The low-pass filtering technique smoothes out the block boundary, at the expense of possible loss in high-frequency components, resulting in unnecessary blurring of the decoded image. A number of variations of low-pass filtering, such as block boundary filtering [12], nonlinear filtering, and statistical filtering techniques [13], have been proposed. In [14] and [15], a probabilistic model is assumed, and the compressed image is reconstructed using the MAP approach.

On the other hand, the iterative approach based on the theory of POCS was first introduced for image restoration.
[17]. The images degraded by "noises" are restored iteratively by using a priori information about the original image, and the convergence behavior can be explained by the theory of POCS. Thus, if we view the blocking artifact as the noises, then we can apply the image restoration techniques to alleviate the blocking artifact. In this context, iterative postprocessing techniques, based on the theory of POCS, have been proposed [1], [2], [16].

In this paper, based on the theory of POCS, we will propose a novel postprocessing technique to reduce the blocking artifacts in decoded images. Assuming that the original image is highly correlated, the global frequency characteristics in two adjacent blocks should be similar to the local ones in each block. Thus, we can consider the high-frequency components in the global characteristics of a decoded image, which are not found in the local ones, as the results from the blocking artifact. In our approach, we will employ $N$-point DCT to obtain the local characteristics, and $2N$-point DCT to obtain the global ones, and derive the relation between $N$-point and $2N$-point DCT coefficients. Then, by comparing $N$-point with $2N$-point DCT coefficients, we can detect the undesired high-frequency components, mainly caused by the blocking artifact.

In this paper, based on the notion described so far, we propose novel closed convex sets, which include the images free from the blocking artifact, and their projection operators in the DCT domain.

This paper is organized as follows. In Section II, based on the theory of POCS, we explain the basic notion behind the postprocessing technique to reduce the blocking artifact. In Section III, we propose novel closed convex sets and their projection operators, and describe the proposed postprocessing technique in detail. In Section IV, we present the simulation results on still images, and compare the proposed technique with the conventional ones in terms of objective quality, subjective quality, and convergence behavior. Finally, in Section V, we provide the conclusion.

II. POSTPROCESSING TECHNIQUES
BASED ON THE THEORY OF POCS

A. Iterative Approach Based on the Theory of POCS

Given $n$ closed convex sets $C_i, i = 1, 2, \cdots, n$, the iteration

$$f_k = P_{f_{k-1}} \cdots P_2 P_1 f_{k-1}, \quad k = 1, 2, \cdots \quad (1)$$

will converge to a limiting point $f^*$ of the intersection $C^* = \cap_{i=1}^{n} C_i$, as $k \to \infty$, for an arbitrary initial element $f_0$ in Hilbert space $H$. In (1), $P_i$ is the projection operator onto $C_i$, i.e.,

$$\|f - P_i f\| = \min_{g \in C_i} \|f - g\| \quad (2)$$

where $\| \cdot \|$ denotes the norm in $H$.

The theory of POCS was first introduced in the field of image restoration by Youla and Webb [17]. In [17], it is assumed that every known property of the original image $f$ can be formulated into a corresponding closed convex subset in $H$. Therefore, such $n$ properties will generate $n$ well-defined closed convex sets $C_i, i = 1, \cdots, n$, and necessarily, the original image $f$ should be included in all the $C_i$’s and also the intersection of all the $C_i$’s $C^*$, i.e.,

$$f \in C^* = \cap_{i=1}^{n} C_i. \quad (3)$$

It is clear that the intersection $C^*$, which includes $f^*$, is a closed and convex set. Consequently, irrespective of whether $C^*$ includes the elements other than $f^*$, the problem of restoring $f_0$ from its $n$ properties is equivalent to finding at least one point belonging to $C^*$.

If the projector $P_i$ projecting onto $C_i$ is known, the problem can be solved easily since $P_i f_0 \in C_i$ for any $f_0$. Unfortunately, $C^*$ may be nonlinear and complex in structure so that a direct realization of $P_i$ is almost infeasible. However, if each $P_i$ projecting onto its respective $C_i$ is effectively realizable, then the problem is recursively solvable, due to the theory of POCS. It is noted that every point of $C^*$ is clearly a fixed point of every $P_i$, and the composition operator $T = P_n P_{n-1} \cdots P_1$ as well, i.e., if $f^* \in C^*$, then $P_i f^* = f^*$ for all $i$ and $T f^* = f^*$. Therefore, to be convinced of whether $f^*$ is a limiting point, it should be satisfied that $P_i f^* = f^*$ for all $i$.

Fig. 1 shows the block diagram of a coding system with a postprocessor to reduce the blocking artifact. On the decoder side, note that it is impossible to know what the original image was. However, if we have a priori information about the original image, we can define our own "desired image," and minimize the distance between the desired and the postprocessed image by using the iterative approach based on the theory of POCS. Thus, the decoded images are reconstructed by using not only the received data, but also a priori information about the original image, which is certainly free from the blocking artifact.

B. The Conventional Postprocessing Techniques

The POCS-based iterative postprocessing techniques have been intensively investigated in [1], [2], and [16]. From these studies, it is known that the postprocessing techniques based on the theory of POCS satisfy two requirements. First, the closed convex sets, which capture the properties of the image free from the blocking artifact and close to the original image, should be defined. Second, it is necessary to derive the projection operators onto them. Then, the decoded image converges to the postprocessed image, which is free from the blocking artifact and close to the original image.

Fig. 1 shows the block diagram for the postprocessing technique, based on the theory of POCS, is given in Fig. 2. As shown in Fig. 2, two different closed convex sets are defined. One is the quantization
constraint set (QCS) $C_q$, which denotes the set including the images close to the original image. The other is the smoothness constraint set (SCS) $C_s$, which denotes the set including the images free from blocking artifacts. In Fig. 2, notice that the $P_q(P_s)$ is the projection operator onto $C_q(C_s)$.

The QCS was first introduced by Rosenholtz and Zakhor [1]. In [1], the QCS is formulated from the quantization interval for the DCT coefficient. The conjecture that the QCS includes the images close to the original image can be justified by considering the encoding process. More specifically, the quantization interval includes the original DCT coefficient, and it was determined to minimize the mean-square error between the original and the quantized coefficients. Since the quantization parameters are transmitted from the encoder, the quantization intervals for each DCT coefficient are already known to the decoder. Thus, in the DCT domain, we can define a closed convex set $C_q$ as the quantization interval and its projection operator $P_q$, which moves the DCT coefficient into the quantization interval. Fig. 3(a) describes the notion behind $C_q$ and $P_q$. For a detailed mathematical notation and its explanation for $C_q$ and $P_q$, refer to [2].

It is a very important problem to define the $C_s$ theoretically. However, it would be very difficult to analyze the blocking artifact quantitatively in the spatial or frequency domain since it is strongly related to the human visual system. Thus, instead of defining the $C_s$ precisely, various efforts [1], [2] have been made to find the set closely related to the blocking artifact.

Instead of defining the $C_q$, Rosenholtz and Zakhor defined $C_f$, which is a set including band-limited images with a given cutoff frequency. Assuming that the high-frequency components are mainly due to the blocking artifact, the $C_f$ would include the images free from the blocking artifact. The projection operator onto $C_f$ is equivalent to an ideal low-pass filter ($P_f$). Fig. 3(b) describes the notion behind $C_f$ and $P_f$. However, since an ideal low-pass filter cannot be implemented in the real world, a $3 \times 3$ low-pass filter ($L$) was used in [1] instead of the $P_f$. Thus, the iterative postprocessing technique in [1] is described by

$$ f_k = P_f L f_{k-1}, \quad k = 1, 2, \cdots $$

where $f_k$ denotes the postprocessed image at the $k$th iteration. But, since the $L$ is not the projection operator onto $C_f$, the convergence cannot be guaranteed. However, Reeves et al. [18] showed that, by considering the iterative postprocessing technique in [1] as the constrained minimization problem [18], the iteration in (4) converges for any low-pass filter $L$ whose magnitude response is less than or equal to unity.

In [1], it was argued that the frequency components, which are higher than a given cutoff frequency, are mainly due to blocking artifacts. But, the original image may contain the high-frequency components above the given cutoff frequency. Thus, although iterative application of the low-pass filtering indeed reduces the high-frequency components caused by the blocking artifact, it could also degrade the original high-frequency components. As a result, the technique in [1] may yield excessively smoothed images, as will be shown in the simulation results. Moreover, Rosenholtz’s technique requires a heavy computational burden since a considerable number of iterations is necessary to achieve the convergence.

On the other hand, by defining new closed convex sets ($C_y$) instead of $C_q$, Yang et al. [2] derived the projection operator ($P_y$) onto $C_y$ theoretically. More specifically, Yang et al. defined $C_y$, capturing the intensity variations between adjacent two pixels’ values at the block boundaries of an image. In [16], Yang et al. also defined the modified $C_y$, incorporating the spatial adaptivity based on the previous work [2]. The $P_y$ onto $C_y$ reduces the difference of two pixels’ values only at the block boundary to alleviate the blocking artifact. Fig. 3(c) explains the $P_y$ and $C_y$ graphically. However, it should be noted that the blocking artifact is mainly caused by...
the difference between intensities of two adjacent blocks, not by two adjacent pixels only at the block boundary. Thus, the adjustment of two pixels’ values only at the block boundary is not sufficient to reduce the blocking artifact effectively, and may generate undesirable high-frequency components inside the block. Moreover, it was reported that Yang’s technique also requires a number of iterations to achieve the convergence [2].

Therefore, both Rosenholtz’s and Yang’s techniques seem to be inappropriate to reduce the blocking artifact effectively. Also, since both techniques converge with a number of iterations, they cannot be easily applied to real-time postprocessing of moving image sequences.

III. A PROPOSED POSTPROCESSING TECHNIQUE

A. Smoothness Constraint Sets and Their Projection Operators

In this section, we will propose the smoothness constraint sets capturing the images free from the blocking artifact, and their projection operators in the DCT domain.

The blocking artifact is strongly related to the human visual system. It would be very difficult to quantitatively analyze the blocking artifact either in the spatial or frequency domain. It is believed, however, that the blocking artifact augments the high-frequency components rather than the low ones. In this context, in [1], the high-frequency components above a certain cutoff frequency are considered as the results from the blocking artifacts in a decoded image. But it is noted that the blocking artifact is observed over two adjacent blocks, and the frequency distribution in two adjacent blocks varies according to its local statistics. Thus, we cannot be sure whether or not the high-frequency components above a certain cutoff frequency are indeed due to the blocking artifact. Thus, in this paper, attempts are made to detect the high-frequency components, mainly caused by the blocking artifact, based on the frequency distribution in two adjacent blocks.

In our approach, it is assumed that the original image is highly correlated. Thus, the global frequency characteristics of two adjacent blocks are very similar to the local ones of each block. If the blocking artifact appears between two adjacent blocks, then the global frequency characteristics of two adjacent blocks may be quite different from the local ones. Thus, a careful comparison of the local frequency characteristics with the global one makes it possible to detect the blocking artifact in the frequency domain.

It is well known that N-point DCT C(k) is closely related to the 2N-point discrete Fourier transform (DFT) F(k) for symmetrically extended data [19]. In other words

\[ C(k) = e^{j\pi k/2N} \cdot F(k), \quad 0 \leq k \leq N - 1. \]  

In (5), we can see that the magnitude response in the DFT domain is identical to that in the DCT domain. In our approach, we employ N-point DCT to find the local frequency characteristics of each block, and 2N-point DCT to find the global ones of two adjacent blocks. Now, we shall describe the proposed technique based on the DCT in more detail.

First, let us denote an \( xN \times yN \) input image and its pixel by \( f \) and \( f(m,n) \), respectively, where \((m,n)\) is the pixel coordinate. Also, let \( f_{Rij} \) be a \( 2N \times 1 \) vector, representing 2N elements in the row direction, given by

\[
\begin{align*}
f_{Rij} &= \{u_{Rij}(0),\ldots,u_{Rij}(N-1),v_{Rij}(0),
\ldots,v_{Rij}(N-1)\}
\end{align*}
\]

where \( u_{Rij}(n) = f((i-1)N+n,j) \) and \( v_{Rij}(n) = f(iN+n,j), i = 1,\ldots,x-1 \) and \( j = 0,1,\ldots,yN-1 \), respectively. Also, let \( w_{Rij}(n) \) be a concatenated signal of \( u_{Rij}(n) \) and \( v_{Rij}(n) \), defined as

\[
w_{Rij}(n) = \begin{cases} u_{Rij}(n), & 0 \leq n \leq N - 1 \\ v_{Rij}(n-N), & N \leq n \leq 2N - 1. \end{cases}
\]

Similar to \( f_{Rij} \), we can define a \( 1 \times 2N \) vector \( f_{C_{ij'}} \), representing 2N elements in the column direction, \( u_{C_{ij'}}(n) \), \( u_{C_{ij'}}(n) \), and \( w_{C_{ij'}}(n) \) as follows:

\[
f_{C_{ij'}} = \{u_{C_{ij'}}(0),\ldots,u_{C_{ij'}}(N-1),v_{C_{ij'}}(0),
\ldots,v_{C_{ij'}}(N-1)\}
\]

\[
w_{C_{ij'}}(n) = \begin{cases} u_{C_{ij'}}(n), & 0 \leq n \leq N - 1 \\ v_{C_{ij'}}(n-N), & N \leq n \leq 2N - 1. \end{cases}
\]

For the sake of simplicity in notation, we will denote \( u_{Rij} \) and \( u_{C_{ij'}} \) by \( u \) and \( u \), \( v \) and \( v \), and \( u \) and \( v \) by \( u \) and \( v \), respectively. Thus, \( \{u(n)\} \) and \( \{v(n)\}\), \( n = 0,\ldots,N - 1 \), are two adjacent blocks of size \( N \), and \( \{w(n)\}\), \( n = 0,\ldots,2N - 1 \), is the block concatenated from \( u(n) \) and \( v(n) \) of size \( 2N \), respectively. And let us also denote the DCT of \( u(n), v(n) \), and \( w(n) \) by \( U(k), V(k) \), and \( W(k) \), respectively. Then, we have

\[
U(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos \left[ \frac{\pi(2n+1)k}{2N} \right]
\]

\[
V(k) = \alpha(k) \sum_{n=0}^{N-1} v(n) \cos \left[ \frac{\pi(2n+1)k}{2N} \right]
\]

\[
W(k) = \frac{1}{\sqrt{2}} \alpha(k) \sum_{n=0}^{2N-1} w(n) \cos \left[ \frac{\pi(2n+1)k}{2N} \right]
\]
where

\[ \alpha(k) = \begin{cases} \sqrt{\frac{1}{N}}, & k = 0 \\ \frac{\sqrt{2}}{N}, & \text{otherwise}. \end{cases} \]

From (7) or (9), \( W(k) \) can be rewritten as

\[
W(k) = \frac{1}{\sqrt{2}} \alpha(k) \left\{ \sum_{n=0}^{N-1} u(n) \cos \left[ \frac{\pi(2n+1)k}{2(2N)} \right] + \sum_{n=N}^{2N-1} v(n-N) \cos \left[ \frac{\pi(2n+1)k}{2(2N)} \right] \right\}
= \frac{1}{\sqrt{2}} \alpha(k) \sum_{n=0}^{N-1} \left\{ u(n) \cos \left[ \frac{\pi(2n+1)k}{2N} \right] + v(n) \cos \left[ \frac{\pi(2n+1)k}{2N} + \frac{k}{2}\pi \right] \right\},
\]

(13)

By substituting (10) and (11) into (13), \( W(k) \) can be expressed as

\[
W(k) = \frac{1}{\sqrt{2}} \alpha(k) \left\{ \sum_{n=0}^{N-1} u(n) \cos \left[ \frac{\pi(2n+1)k}{4N} \right] + (-1)^{k/2} V \left( \frac{k}{2} \right) \right\},
\]

(14)

From (14), it is observed that the 2\( k \)th DCT coefficient of \( W \) depends only on the 4\( k \)th DCT coefficients of \( U \) and \( V \), while the odd-numbered DCT coefficients of \( W \) are expressed as a weighted sum of \( u(n) \) and \( v(n) \). Hence, only the odd-numbered DCT coefficients of \( W \) are affected by the discontinuities in the blocks. Note that the DCT is closely related to the DFT as shown in (5), and the DFT is a sampled value of the Fourier transform. If the original image is correlated highly enough so that the block discontinuities are invisible, then the odd-numbered DCT coefficients of \( W \) of the original image can be approximated by interpolating adjacent even-numbered coefficients. In this case, by considering (14), the meaningful (nonzero) DCT coefficients of \( W \) are kept within the range two times larger than those of \( U \) and \( V \). In the decoded image, if nonzero coefficients of \( W \) occur at higher locations than the predetermined value \( K_{R_{ij}} \), we can obtain the projected element \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \). It is noted that \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \) is a unique element in \( C_{R_{ij}}(W_{C_{ij}}) \) that minimizes the norm of difference since the coefficients of \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \) at lower locations than the predetermined value remain unchanged. Here, the projection operator \( \tilde{P}_{R_{ij}}(P_{C_{ij}}) \) on \( C_{R_{ij}}(C_{ij}) \) can be defined as

\[
\mathcal{F}P_{R_{ij}}f_{R_{ij}} = \left\{ \begin{array}{ll}
0, & k > K_{R_{ij}} \\
\frac{1}{\sqrt{2}} \left\{ U_{R_{ij}}(\frac{k}{2}) + (-1)^{k/2} V_{R_{ij}}(\frac{k}{2}) \right\}, & k = 0, 2, 4, \ldots, \text{and } k \leq K_{R_{ij}}
\end{array} \right.
\]

(17)

\[
\mathcal{F}P_{C_{ij}}f_{C_{ij}} = \left\{ \begin{array}{ll}
0, & k > K_{C_{ij}} \\
\frac{1}{\sqrt{2}} \left\{ U_{C_{ij}}(\frac{k}{2}) + (-1)^{k/2} V_{C_{ij}}(\frac{k}{2}) \right\}, & k = 0, 2, 4, \ldots, \text{and } k \leq K_{C_{ij}}
\end{array} \right.
\]

(18)

where \( K_{R_{ij}} = 2 \cdot \max[\text{NZ}(U_{R_{ij}}), \text{NZ}(V_{R_{ij}})] + \rho_k \) and \( \text{NZ}(\cdot) \) represents the location of the last nonzero DCT coefficient. Similar to \( C_{R_{ij}} \), we can define \( (y-1) \times xN \) closed convex sets \( C_{C_{ij}} \)’s as

\[
C_{C_{ij}} = \left\{ f_{C_{ij}} \subset f, \text{NZ}(W_{C_{ij}}) \leq K_{C_{ij}} \right\}
\]

(16)

where \( K_{C_{ij}} = 2 \cdot \max[\text{NZ}(U_{C_{ij}}), \text{NZ}(V_{C_{ij}})] + \rho_k \).

It is noted that \( p \) in (15) and (16) is a factor representing how many DCT coefficients of \( W \) whose locations are greater than \( 2 \cdot \max[\text{NZ}(U), \text{NZ}(V)] \) contribute to the blocking artifact, i.e., to what degree we tolerate the frequency components, which are higher than \( 2 \cdot \max[\text{NZ}(U), \text{NZ}(V)] \). In our approach, we choose 2 for \( \rho_k \).

From a decoded image, we can determine the values for \( K_{R_{ij}} \) and \( K_{C_{ij}} \) for every \( C_{R_{ij}} \) and \( C_{C_{ij}} \), respectively. Thus, we can define \( (x-1) \times yN \) closed convex sets. It is noted that the sets \( C_{R_{ij}} \) and \( C_{C_{ij}} \) are closed intervals in the DCT domain that are convex.

By observing the smoothness constraint sets defined above, the projection operators onto those sets can be easily found as follows. As shown in (2), the projection operator \( P_{g} \) of an element \( f \) on the set \( C_{g} \) should be chosen such that \( P_{g}f \) yields the minimal norm \( \|f - g\| \) among \( g \)’s in \( C_{g} \). In our approach, by discarding the nonzero coefficients of \( W_{R_{ij}}(W_{C_{ij}}) \) at higher locations than the predetermined value \( K_{R_{ij}}(K_{C_{ij}}) \), we can obtain the projected element \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \). It is noted that \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \) is a unique element in \( C_{R_{ij}}(C_{ij}) \) that minimizes the norm of difference since the coefficients of \( \tilde{W}_{R_{ij}}(W_{C_{ij}}) \) at lower locations than the predetermined value remain unchanged. Here, the projection operator \( \tilde{P}_{R_{ij}}(P_{C_{ij}}) \) on \( C_{R_{ij}}(C_{ij}) \) can be defined as

\[
\mathcal{F}P_{R_{ij}}f_{R_{ij}} = \left\{ \begin{array}{ll}
0, & k > K_{R_{ij}} \\
\frac{1}{\sqrt{2}} \left\{ U_{R_{ij}}(\frac{k}{2}) + (-1)^{k/2} V_{R_{ij}}(\frac{k}{2}) \right\}, & k = 0, 2, 4, \ldots, \text{and } k \leq K_{R_{ij}}
\end{array} \right.
\]

(17)

\[
\mathcal{F}P_{C_{ij}}f_{C_{ij}} = \left\{ \begin{array}{ll}
0, & k > K_{C_{ij}} \\
\frac{1}{\sqrt{2}} \left\{ U_{C_{ij}}(\frac{k}{2}) + (-1)^{k/2} V_{C_{ij}}(\frac{k}{2}) \right\}, & k = 0, 2, 4, \ldots, \text{and } k \leq K_{C_{ij}}
\end{array} \right.
\]

(18)
In other words, the projection operators in (17) and (18) also can be defined as (for notational convenience, notice that the subscripts are all omitted)

$$P = \mathcal{F}^{-1} \mathcal{P} \mathcal{F}$$

(19)

where $\mathcal{F}^{-1}$ is the inverse DCT operator, and $\mathcal{P}$ is a $2N \times 2N$ matrix, given by

$$\mathcal{P} = \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{pmatrix} K.$$

(20)

It is noted that $\mathcal{P}$ is a linear transform, and $P \cdot P = P$, i.e., $P$ is the projection operator.

**B. Proposed Technique for Reducing the Blocking Artifacts**

Given a decoded image $\mathbf{f}_0$, we determine $xN \times yN$ closed convex sets $C_q$'s by considering the quantization levels for each DCT coefficient of each block. Then, according to (15) and (16), we define $(x-1) \times yN$ closed convex sets $C_R$'s and $xN \times (y-1)$ sets $C_C$'s by computing $N$-point DCT to each block in the horizontal and vertical directions, respectively, as shown in Fig. 4. Note that these sets can be defined from a priori information. Then, the projection operators on smoothness constraint sets are applied to each two adjacent blocks in the horizontal and vertical direction, according to (17) and (18) or (19), respectively. Note that this computation is carried out on all of the smoothness constraint sets $C_{Rij}$ and $C_{C_{ijf}}$. Next, the projection operators on quantization constraint sets are applied [2], yielding the postprocessed image $\mathbf{f}_k$. As shown in Fig. 2, the procedure described so far is carried out iteratively to obtain the postprocessed image $\mathbf{f}_k$. 

Fig. 5. Four test images. (a) LENA image. (b) ZELDA image. (c) BABOON image. (d) BRIDGE image.
The computational complexity of the proposed technique can be examined as follows. The projection onto QCS requires \( N \times N \) point DCT/IDCT operations, as is the same in the conventional techniques [1], [2]. And, the projection onto SCS requires the transform operation on \( 2N \)-point data, i.e., about \( 8N \) multiplications per pixel since the SCS’s are overlappedly defined in both the horizontal and the vertical directions as shown in Fig. 4. Thus, the proposed technique seems to require a heavy computational burden, compared with Rosenholtz’s technique, which requires nine multiplications per pixel, and Yang’s technique. However, as will be shown in the next section, the proposed technique converges with fewer iterations than the conventional ones [1], [2]. Thus, the computational complexity of the proposed technique is comparable to those of the conventional techniques [1], [2].

### IV. SIMULATION AND RESULTS

In this section, the performances of three postprocessing techniques, namely, Rosenholtz’s, Yang’s, and the proposed techniques, are evaluated on still images. Fig. 5 shows four 512 \( \times \) 512 original test images, namely, “LENA,” “ZELDA,” “BABOON,” and “BRIDGE.” As is observed, the four test images could be divided into two categories. More specifically, the first two images, “LENA” and “ZELDA,” contain more low-frequency components and smooth areas than the second two images. On the other hand, the second two images, “BABOON” and “BRIDGE,” contain more high-frequency components and edge components than the first two images. Thus, we will be able to examine how much the high-frequency components in the original image affect the performance, and the robustness of the content of the input images. Four decoded images, with visible blocking artifacts, are obtained by JPEG [4] with the same quantization table as shown in Table I.

In Fig. 6, the PSNR (peak signal-to-noise ratio) performance of the proposed technique is compared with those of Rosenholtz’s and Yang’s techniques, with iteration increasing. One iteration of postprocessing is carried out by applying two projection operators, that is, first \( P_Q \) on \( C_s \) and next \( P_Q \) on \( C_Q \), as described in the previous section. Thus, it is noted that in Fig. 6, the PSNR performance is presented by the unit of a 0.5 iteration. As shown in Fig. 6, it is observed that the proposed technique converges with a single iteration, regardless of the content of the input images, while Yang’s technique converges with a few iterations. Note that to be convinced of whether or not the postprocessed image \( f^* \) has converged, \( P_Qf^* = f^* \) and \( P_Qf^* = f^* \) should be satisfied. Thus, from Fig. 6, we...
Fig. 7. Comparison of subjective quality on LENA. (a) Decoded image by JPEG: 30.093 dB. (b) Postprocessed image by Rosenholtz’s technique: 29.62 dB. (c) Postprocessed image by Yang’s technique: 30.78 dB. (d) Postprocessed image by proposed technique: 31.02 dB.

can see that the Rosenholtz’s technique fails to converge, in terms of the theory of POCS. Also, the performance of the proposed technique is better than that of Yang’s technique on the “LENA” and “ZELDA,” and almost comparable on the “BABOON” and “BRIDGE,” in terms of the PSNR.

The nonconvergence nature of Rosenholtz’s technique, in terms of the theory of POCS, is due to the use of the global smoothing operator $L$ [18]. Moreover, if the input image contains relatively large high-frequency components, Rosenholtz’s technique fails to converge, but oscillates instead, as shown in Fig 6.

For a comparison of the subjective quality, we present in Fig. 7 four photographs of LENA, which are converged after 30 iterations. Note that the highest PSNR performance already can be achieved at the first iteration as shown in Fig. 6. However, it is fair to compare the converged images since the techniques, based on the POCS theory, feature the iteration of the projection operations until the convergence is attained. To make a comparison of the subjective quality more clear, we also present enlarged photographs of the details of LENA, such as areas of mouth, hair, hat, and shoulder in Fig. 8. As is observed in Figs. 7 and 8, respectively, although Rosenholtz’s technique alleviates the blocking artifact, it yields an excessively smoothed image due to the use of a global low-pass filter. For Yang’s technique, although the edges are well preserved, the blocking artifacts are still observed in smooth areas, such as the shoulder and hat in Fig. 8. In Yang’s technique, notice that only the pixels at the block boundary are adjusted to reduce the blocking artifact. However, as mentioned previously, the blocking artifact is mainly caused by the
difference between two adjacent blocks, not by two adjacent pixels only at the block boundary. Thus, it is believed that after adjusting the pixels' values only at the block boundary, if the $P_q$ could not make any changes in the pixel values inside block, the blocking artifact still remains, although convergence is achieved.

On the other hand, Figs. 7 and 8 reveal that the blocking artifacts are alleviated more effectively by the proposed technique than the other two techniques. It is also observed that the high-frequency components, such as edges, are preserved very faithfully by the proposed technique. Similar results are also observed on the other three test images. Hence, it is believed that the smoothness constraint sets and their projection operators, proposed in this paper, are properly defined, providing good performance in terms of both the PSNR and subjective quality. Moreover, as is observed in Fig. 6, the proposed technique achieves convergence at only a single iteration, while the others require a few iterations. Thus, the proposed technique can be easily implemented by applying the projections only once.

V. CONCLUSION

In this paper, based on the theory of POCS, we have proposed a novel postprocessing technique to alleviate the blocking artifact in decoded images. In our approach, it was assumed that the input image is highly correlated so that similar frequency characteristics are maintained between adjacent blocks. If we are able to detect the high-frequency components that are not present within blocks, we can consider them as the result of the blocking artifact. In our approach, we
employed N-point DCT to obtain the frequency components inside blocks, and 2N-point DCT to obtain those in two adjacent blocks. Thus, by comparing the N-point with 2N-point DCT coefficients, we could detect the undesired high-frequency components, mainly caused by the blocking artifact. Based on the above notion, we have proposed novel closed convex smoothness constraint sets, which capture images free from the blocking artifact, and projection operators onto them in the DCT domain.

The simulation results showed that the blocking artifacts could be significantly alleviated, and the original high-frequency components, such as edges, are faithfully preserved by the proposed technique. It was shown that the performance of the proposed technique is better than that of the conventional ones, in terms of objective quality, subjective quality, and convergence behavior. Moreover, the proposed technique was shown to be robust in the content of the input images, in terms of convergence and performance improvement. Finally, due to the noniterative nature of the proposed technique, it can be easily employed for the postprocessing of moving image sequences.

REFERENCES


Hoon Paek was born in Seoul, Korea, on July 27, 1966. He received the B.S., M.S., and Ph.D. degrees, all in control and instrumentation engineering, from Seoul National University, Seoul, Korea, in 1990, 1992, and 1997, respectively.

He is currently working in LSI Division 1, System LSI Business, Samsung Electronics. His research interests include image processing and communication, specifically algorithms for image recovery/enhancement, image sequence coding techniques over noisy channels, and their VLSI architectures.

Rin-Chul Kim (M’95–A’96) was born in Seoul, Korea, on February 27, 1963. He received the B.S., M.S., and Ph.D. degrees, all in control and instrumentation engineering, from Seoul National University, in 1985, 1987, and 1992, respectively.

In 1992–1994, he was with the Daewoo Electronics Company, Ltd. In 1994, he joined the School of Information and Computer Engineering, Hunsung University, where he is an Assistant Professor. His current research interests include image processing, image data compression, and VLSI signal processing.

Sang-Uk Lee (S’75–M’79) was born in Seoul, Korea, on August 11, 1949. He received the B.S. degree from Seoul National University, Seoul, Korea, in 1973, the M.S. degree from Iowa State University, Ames, in 1976, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1980, all in electrical engineering.

In 1980–1981, he was with the General Electric Company, Lynchburg, VA, and worked on the development of digital mobile radio. In 1981–1983, he was a Member of Technical Staff, MA-COM Research Center, Rockville, MD. In 1983, he joined the Department of Control and Instrumentation Engineering, Seoul National University, as an Assistant Professor, where he is now a Professor in the School of Electrical Engineering. Currently, he is also affiliated with the Automatic System Research Institute and Institute of New Media and Communications, Seoul National University. His current research interests are in the areas of image and video signal processing, digital communication, and computer vision.

Dr. Lee is a member of Phi Kappa Phi.