PAPER

# A Spatially Adaptive Gradient-Projection Algorithm to Remove Coding Artifacts of H.264

**Kwon-Yul CHOI**[†], *Nonmember and* **Min-Cheol HONG**[†a)], *Member*

**SUMMARY** In this paper, we propose a spatially adaptive gradient-projection algorithm for the H.264 video coding standard to remove coding artifacts using local statistics. A hybrid method combining a new weighted constrained least squares (WCLS) approach and the projection onto convex sets (POCS) approach is introduced, where weighting components are determined on the basis of the human visual system (HVS) and projection set is defined by the difference between adjacent pixels and the quantization index (QI). A new visual function is defined to determine the weighting matrices controlling the degree of global smoothness, and a projection set is used to obtain a solution satisfying local smoothing constraints, so that the coding artifacts such as blocking and ringing artifacts can be simultaneously removed. The experimental results show the capability and efficiency of the proposed algorithm.

***key words:*** *H.264, coding artifacts, weighted constrained least squares, human visual system, projection onto convex sets*

## 1. Introduction

A variety of visual communications, applications, and services have become possible due to image compression techniques such as JPEG, MPEG, and H.263. These techniques use well-known block-based discrete cosine transform (BDCT) because it exhibits very high energy compaction, and does not need supplementary information as do Karhunen-Loeven transform and Fourier transform [1], [2]. When the images are highly compressed due to the limit of the channel or memory bandwidth, these techniques result in annoying artifacts. The loss of the frequency components caused by the independent block-based quantization process leads to the annoying coding artifacts in the reconstructed images. The lower the bit rate is, the more visible the artifacts become and the reconstructed images are more seriously degraded [1]–[3].

There are many potential approaches to solve the problem. Among them, post-processing and in-loop filters have been widely used. Post-processing techniques have the advantage of not needing to modify the existing coding procedure and it does not require an increase in the bit rate to obtain the satisfactory compressed images [2]–[7].

H.264 video coding standard was developed jointly by ITU-T and ISO to obtain better compression than the other video coding standards. It adopts the new methods such as 4×4 size block based integer transform, spatial intra prediction, variable block size motion estimation, and so on [1], [5]. H.264 uses an in-loop filter to remove blocking artifact in both the encoder and decoder. However, the in-loop filter distorts important information such as edges and moving objects, resulting in over-smoothness in high activity regions. Therefore, it is necessary to exploit a new approach to remove coding artifacts for higher visual quality applications.

In this paper, a spatially adaptive gradient-projection post-processing algorithm is proposed to simultaneously remove coding artifacts such as blocking and ringing artifacts while preserving edge information. We define a hybrid approach of the projection method coupled with the WCLS method to obtain a better solution space. In particular, a new regularized WCLS functional is defined to take into account the fact that HVS reacts different to the same noise in the flat region and the high activity region. In order to measure the local activity of each pixel, a local variation function is defined. In addition, a projection set is defined to control the bound of the solution space that is determined by the difference of adjoining pixels and quantization index. The proposed algorithm is designed to obtain a better solution that is an element of intersection set between the solution spaces determined by WCLS and POCS, so that the coding artifacts are effectively removed.

This paper is organized as follows. The background of the gradient-projection method is explained in Sect. 2. In Sect. 3, we describe a new gradient-projection algorithm to reduce the coding artifacts for the H.264 video coding standard. A way of determining weighting matrices based on a new linear visual function is introduced. In addition, the projection set using the difference between neighboring pixels and QI is described. Experimental results are given in Sect. 4, and finally conclusions are presented in Sect. 5.

## 2. Background

For a $M \times N$ size image, a common degradation model caused by quantization process can be written as [6], [7]:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where the vectors $\boldsymbol{y}$, $\boldsymbol{x}$, and $\boldsymbol{n}$ are of size $MN \times 1$, and represent the lexicographically ordered observed image, the original image, and the additive quantization noise respectively [6]. Under the assumption that the additive noise has typically Gaussian distribution and it is uncorrelated with the original image, the constrained least squares (CLS) approach has been widely used [6], [7]. The CLS approach has

Manuscript received October 13, 2009.
Manuscript revised December 14, 2010.
[†]The authors are with Soongsil University, Korea.
a) E-mail: mhong@ssu.ac.kr
DOI: 10.1587/transinf.E94.D.1073

been used to obtain a possible solution to Eq. (1) by minimizing the following functional [6], [7]

$$M(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \alpha\|\boldsymbol{C}\boldsymbol{x}\|^2, \tag{2}$$

where $\boldsymbol{C}$ is typically a high-pass operator used to impose the smoothness constraint. The first term in the right side of Eq. (2) represents the fidelity to the data, and the second term denotes the smoothness of $\boldsymbol{x}$ [6], [8]. Also, the regularization parameter $\alpha$ controls the trade-off between them. When the local statistics of the noise and the smoothness are incorporated to the functional, Eq. (2) can be rewritten as

$$\begin{aligned} M(\boldsymbol{x}) &= \|\boldsymbol{R}(\boldsymbol{y} - \boldsymbol{x})\|^2 + \alpha\|\boldsymbol{L}\boldsymbol{C}\boldsymbol{x}\|^2 \\ &= J(\boldsymbol{x})^2 + \alpha Q(\boldsymbol{x})^2 \end{aligned} \tag{3}$$

where $\boldsymbol{R}$ and $\boldsymbol{L}$ are diagonal weighting matrices with size $MN \times MN$. The diagonal weighting matrices can be determined in various ways. One is to determine them in advance using the observed image. Alternatively, they can be estimated iteratively using the partially restored image. For an observation image $\boldsymbol{y}$, the gradient iteration applied to Eq. (3) results in

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \beta[\boldsymbol{R}^t\boldsymbol{R}\boldsymbol{y} - (\boldsymbol{R}^t\boldsymbol{R} + \alpha\boldsymbol{C}^t\boldsymbol{L}^t\boldsymbol{L}\boldsymbol{C})\boldsymbol{x}_k] \\ &= [\boldsymbol{I} - \beta(\boldsymbol{R}^t\boldsymbol{R} + \alpha\boldsymbol{C}^t\boldsymbol{L}^t\boldsymbol{L}\boldsymbol{C})]\boldsymbol{x}_k + \beta\boldsymbol{R}^t\boldsymbol{R}\boldsymbol{y} \\ &= G\boldsymbol{x}_k + \beta\boldsymbol{R}^t\boldsymbol{R}\boldsymbol{y}, \end{aligned} \tag{4}$$

where $\boldsymbol{I}$ represents the identity matrix and $G$ denotes the gradient operator. In Eq. (4), $\boldsymbol{A}^t$ represents the transpose of matrix $\boldsymbol{A}$, and $\beta$ is the relaxation parameter to control the speed of convergence. According to the contraction mapping theorem [8], Eq. (4) converges if

$$0 < \beta < \frac{2}{\lambda_{\max}(\boldsymbol{R}^t\boldsymbol{R}) + \alpha\lambda_{\max}(\boldsymbol{C}^t\boldsymbol{L}^t\boldsymbol{L}\boldsymbol{C})}$$

where $\lambda_{\max}(\boldsymbol{A})$ represents the maximum eigen-value of matrix $\boldsymbol{A}$.

Constraints can be imposed on the partially restored image to obtain better feasible solution space. Then, the iteration takes the form

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1} &= G\boldsymbol{x}_k + \beta\boldsymbol{R}^t\boldsymbol{R}\boldsymbol{y} \\ \boldsymbol{x}_{k+1} &= P\hat{\boldsymbol{x}}_k = PG\hat{\boldsymbol{x}}_k + \beta\boldsymbol{P}\boldsymbol{R}^t\boldsymbol{R}\boldsymbol{y}, \end{aligned} \tag{5}$$

where $P$ denotes a projection operator (or concatenation of operator) of a signal onto a set of signals with desirable properties [9]. In Eq. (5), $\hat{\boldsymbol{x}}_k$ and $\boldsymbol{x}_k$ represent the $k$-th iterative solutions by gradient algorithm and by gradient-projection algorithm, respectively.

## 3. Proposed Algorithm

In this section, we describe a new hybrid approach combining WCLS and POCS to remove coding artifacts. A new visual function is defined to determine an element of the diagonal weighting matrices that are utilized to control the degree of global smoothness. Also, a convex set is defined to incorporate local smoothness into the feasible solution space that is obtained by WCLS, where the bounds of the convex set are determined by the difference between the pixels and QI.

### 3.1 Weighted CLS Method

There are various ways to determine the weighting matrices in Eq. (3). In general, they are determined to control the relative degree of the smoothness to the reconstructed image. The local variance was used to define the element of the diagonal matrix $\boldsymbol{R}$ in Ref. [6]. On the other hand, the element of the matrix $\boldsymbol{L}$ takes the inverse of the corresponding element of $\boldsymbol{R}$. In order to remove the coding artifacts, relatively large smoothness should be imposed to block boundary pixels rather than pixels within a block. However, as the degree of smoothness increases, the important features can be seriously blurred. Therefore, it is very important to determine an adequate way for $\boldsymbol{R}$ and $\boldsymbol{L}$.

The well-known regularization approach was used to remove the blocking artifacts in Ref. [6], where different degrees of smoothness were applied to the blocky image according to pixel position and quantization step size. However, the desirable local statistics were not considered in the restoration process and therefore the blocking and ringing artifacts were not effectively removed. Furthermore, it is not adequate for the H.264 video coding standard, since the algorithm is based on an $8 \times 8$ block- based image and different quantization approach. Reference [7] used the maximum local variance to decide the elements of the weighting matrices to reduce the blocking artifacts. However, this may result in over-smoothness of the reconstructed image, since the diagonal element of matrix $\boldsymbol{L}$ has a large value when the quantization noise is very significant.

In this work, we define a location-based local variation so that local activity is incorporated into the solution space, leading to preservation of edge information. It is well-known that a blocking artifact comes from the independent processing of each block of an image and the loss of low- frequency information. On the other hand, a ringing artifact is due to information loss in the high-frequency component. Therefore, the different in the smoothness should be applied to different pixel position. According to the above reason, the variation of within-block pixels and between-block pixels should be determined to minimize the influence of the coding artifacts. As shown in Fig. 1, since the degree of signal discontinuity between adjacent pixels is different from the pixel's position, pixels within a $4 \times 4$ block can be classified according to the location in the block.

Then, the local variation of an element pixel of each classification is determined as

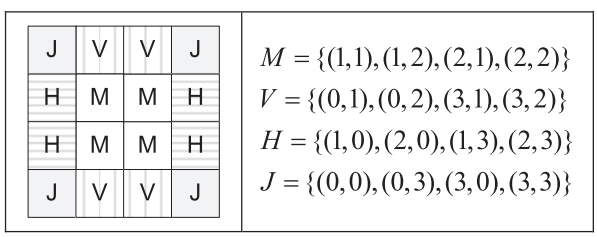


**Fig. 1** Pixel classification in a $4 \times 4$ block.

$$LV(i,j) = \begin{cases} \sum_{p=-1}^{1} |x(i+p,j) - m_M(i,j)| + \\ \sum_{q=-1(q\neq 0)}^{1} |x(i,j+q) - m_M(i,j)| \text{ for } (i,j) \in M \\ \sum_{q=-2}^{2} |x(i,j+q) - m_H(i,j)| \text{ for } (i,j) \in H \\ \sum_{p=-2}^{2} |x(i+p,j) - m_V(i,j)| \text{ for } (i,j) \in V \end{cases} \tag{6}$$

where $m_M(i,j)$, $m_H(i,j)$, and $m_V(i,j)$ represent the location-based mean values of each classification shown in Fig. 1. They are defined as

$$\begin{aligned} m_M(i,j) &= \frac{1}{5}\left(\sum_{p=-1}^{1} x(i+p,j) + \sum_{q=-1(q\neq 0)}^{1} x(i,j+q)\right), \\ m_H(i,j) &= \frac{1}{5}\sum_{q=-2}^{2} x(i,j+q), \\ m_V(i,j) &= \frac{1}{5}\sum_{p=-2}^{2} x(i+p,j). \end{aligned} \tag{7}$$

For a pixel belonging to $(i,j) \in J$, the local variation is determined as the average value of the vertical variation and the horizontal variation.

Then, the problem at hand is to obtain the element of $\boldsymbol{R}$ and $\boldsymbol{L}$ using the values of the location-based variation. In order to build the relation between the location-based variation and weighting matrix $\boldsymbol{L}$, we define a linear visual function that controls the visual degree, such that an element of $\boldsymbol{L}$ adaptively controls the degree of local smoothness of the coded image. The proposed linear visual function is defined to have a value between '0' and '1' and it has the following properties. As QI is higher (lower bit rate), the coding artifacts are more visible, and therefore larger smoothness is required to remove the coding artifacts. In addition, when the local variation is low, the high frequency components in the corresponding flat regions should be strongly regularized, leading to removal of coding artifacts. According to the above properties, a linear visual function can be written as

$$\omega(i,j) = \Phi_1(QI) - \mu\Phi_2(LV), \tag{8}$$

where $\omega(i,j)$ is a linear visual function of $(i,j)$ pixel, and $\Phi_1(.)$ and $\Phi_2(.)$ represent the monotonically increasing functions, respectively. In addition, $\mu$ denotes a parameter to control the relative contribution of the functions, and QI means the quantization index of the block to which the $(i,j)$ pixel belongs. A simple linear visual function satisfying Eq. (10) can be written as

$$\omega(i,j) = c_1 QI - c_2 LV(i,j) + c_3, \tag{9}$$

where $\omega(i,j)$ denotes the linear visual function of $(i,j)$ pixel, and $c_1$, $c_2$, and $c_3$ represent the constants, respectively. As pointed out, $c_1$ is higher, as the coding artifacts are more significant. Also, $c_2$ should be determined in such a way that the high frequency components in the low activity regions are strongly regularized. This is in agreement with HVS properties in Ref. [10]. In our experiments, $c_1$ and $c_2$ in Eq. (9) is differently determined according to the local variation so that the coding artifacts are effectively removed according to the pixel position and the degree of local activity. According to the above properties and the observation, the proposed visual function is written as

$$\omega(i,j) = \begin{cases} (Z - LV(i,j) + 2(T_2 - T_1) + c_3) \text{ for } LV(i,j) \le T_1 \\ (Z - 3LV(i,j) + 2T_2 + c_3) \text{ for } T_1 < LV(i,j) < T_2 \\ (Z - LV(i,j) + c_3) \text{ otherwise} \end{cases} \tag{10}$$

where $T_1$ and $T_2$ represent the thresholds for the linear function, respectively. In Eq. (10), constants of each range are chosen to satisfy the boundary conditions. The artifacts are more visible in block boundary regions, and therefore the parameter $Z$ in Eq. (10) is defined in a different way from pixel position. It is

$$Z = \begin{cases} 4 \times QI \text{ for } (i,j) \in M \\ 5 \times QI \text{ for } (i,j) \in \{H,J,V\} \end{cases}.$$

Then, the elements of the weighting matrices are determined as

$$\begin{aligned} \boldsymbol{L}(m,m) &= \begin{cases} 1 \text{ for } \omega(i,j) > 1 \\ 0 \text{ for } \omega(i,j) < 0 \\ \omega(i,j) \text{ otherwise} \end{cases}, \\ \boldsymbol{R}(m,m) &= 1 - \boldsymbol{L}(m,m), \end{aligned} \tag{11}$$

where $m = i \times N + j$ for $M \times N$ size image.

In Eq. (4), a high-pass operator $\boldsymbol{C}$ is used to impose the global smoothness constraint. Since the degree of coding artifacts is different according to the pixel's position, we use the directional Laplacian operators to efficiently remove the artifact in the block boundary pixels as well as within-block pixels. As shown in Fig. 2, each directional Laplacian operator is parallel to the orientation of $LV$. Since the output of the directional Laplacian operator increases in the block boundary, it is more effective than a general 2-D Laplacian operator. In this work, according to the pixel classification defined in Fig. 1, a different operator is used for $\boldsymbol{C}$ in Eq. (4). For example, when a pixel is an element of the set $H$ in Fig. 1, the third directional Laplacian operator of Fig. 2 is used for $\boldsymbol{C}$.

$$\underset{(i,j)\in(J\cup M)}{\begin{bmatrix} 0 & -0.25 & 0 \\ -0.25 & 1 & -0.25 \\ 0 & -0.25 & 0 \end{bmatrix}} \quad \underset{(i,j)\in V}{\begin{bmatrix} 0 & -0.5 & 0 \\ 0 & 1 & 0 \\ 0 & -0.5 & 0 \end{bmatrix}} \quad \underset{(i,j)\in H}{\begin{bmatrix} 0 & 0 & 0 \\ -0.5 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix}}$$

**Fig. 2** Directional Laplacian operator.

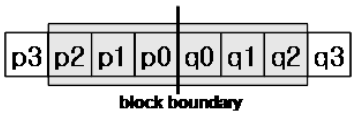


**Fig. 3** Pixel's location.

### 3.2 Weighted CLS Method

As noted, a more desirable solution can be obtained by imposing reasonable constraints into the solution space that is determined by WCLS. POCS has been widely used in various areas [6], [8], [9]. The proposed projection set is defined by the local statistics of the partially restored image at each iteration step. In order to define the local smoothing constraint, it is necessary to determine the parameters which describe the local properties of an image. In our work, we use the local difference between adjacent pixels and the quantization index.

As shown in Fig. 3, it is assumed that $p_0$ and $q_0$ are block boundary pixels of adjacent $4 \times 4$ blocks. Then, we define the following metrics

$$\begin{aligned} MPD &= \frac{1}{4}\left[\sum_{i=0}^{1}(|p_i - p_{i+1}| + |q_i - q_{i+1}|)\right], \\ MAD &= \frac{1}{5}(4 \times MPD + BD), \end{aligned} \tag{12}$$

where $BD = |p_0 - q_0|$ that means the difference between block boundary pixels. In Eq. (12), $MPD$ represents the mean of the difference of adjacent pixels excluding the difference between $p_0$ and $q_0$, and $MAD$ denotes the mean of the difference of adjacent pixels, including the difference between block boundary pixels. If we assume that the local activity of pixels within a block is similar to that of the neighboring block boundary pixels, a projection set having the following properties can be defined.

As $MPD$ is higher, the activity of the corresponding block boundary is higher. Therefore, a looser bound is necessary to preserve the important features. On the other hand, as $MPD$ is lower, the activity of the block boundary pixels is lower, so that a tighter bound is desirable to effectively remove coding artifacts. In addition, the visual degree of the coding artifacts is dependent of the quantization index QI. As QI is higher (lower bit rate), the coding artifacts are more visible, so that the tighter bounds are necessary. According to the above properties, the projection operator $P$ of the difference between $p_0$ and $q_0$ is defined as

$$P(BD) = \begin{cases} BD\_BOUND \text{ for } BD > BD\_BOUND \\ BD \text{ otherwise} \end{cases} \tag{13}$$

where the bound is defined as

$$\begin{aligned} BD\_BOUND &= (1-\gamma)MAD + \gamma BD, \\ \gamma &= \frac{T \times MPD^2}{T \times MPD^2 + 1}, \end{aligned} \tag{14}$$

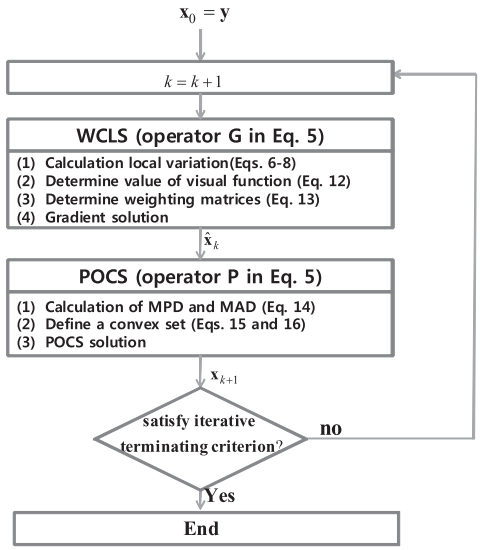


**Fig. 4** Flow chart of proposed algorithm.

$$T = \begin{cases} \dfrac{8}{QI} \times 2 \text{ for within-block pixels} \\ \dfrac{8}{QI} \text{ for between-block pixels} \end{cases},$$

where $\gamma$ represents a relative contribution between $MAD$ and $BD$ to determine the projection set. It is very clear how $\gamma$ and $T$ affect $BD\_BOUND$ which is a bound of the proposed projection set. As QI is lower (blocking artifact is less visible), $T$ becomes higher so that $\gamma$ is closer to 1. In such cases, the projection set has looser bounds. Also, as $MPD$ is higher, $\gamma$ is closer to 1, leading to looser bounds. This is in agreement in the requirement for removal of the blocking artifact [6]. In this work, the tuning parameter $T$ of Eq. (15) is experimentally determined. Using Eq. (14), the final solution is obtained as

$$\begin{aligned} P(p_0) &= \begin{cases} p_0 + P(BD) \text{ if } p_0 < q_0 \\ p_0 - P(BD) \text{ otherwise} \end{cases}, \\ P(q_0) &= \begin{cases} q_0 - P(BD) \text{ if } p_0 < q_0 \\ q_0 + P(BD) \text{ otherwise} \end{cases}. \end{aligned} \tag{15}$$

The projection set and operation are defined in a similar way for the pixels within a block. The whole procedure of the proposed algorithm is shown in Fig. 4.

## 4. Experimental Results

The proposed algorithm was tested with H.264 video coding standard (reference software JM12.1), and a number of experiments have been conducted with various video sequences, several resolutions and bit rates. Among them, QCIF and CIF 'Hall Monitor', 'News', and 'Foreman' sequences are described in the experimental results when the frame rate of the sequences is 10 frames per second. The important coding conditions of H.264 JM12.1 follow

- GOP structure: IPPP (baseline profile)

- Intra Picture Rate: 30 frames
- Rate Control: not used
- RD Optimization: not used
- Reference Frame Number for Motion Estimation: 1
- Motion Search Range: 16 for QCIF and 32 for CIF.

In addition, the following video coding expert group (VCEG) common condition is used to evaluate the coding performance [11]:

- QI of I-Slice: 22, 27, 32, 37
- QI of P-Slice: 23, 28, 33, 38.

The proposed hybrid algorithm was compared with the WCLS reported in Ref. [7] and the H.264 in-loop standard filter. In addition, the proposed WCLS was separately tested to discuss the effectiveness. In these experiments, the algorithms except for H.264 in-loop filter are used as post-processing filters which only require QI information as well as the decoded image from the H.264 decoder. Therefore, any extra overhead information is not necessary.

For evaluating the performance of the algorithms, peak signal to noise ratio (PSNR) was utilized. For $M \times N$ dimensional 8 bit image, it is defined as

$$PSNR = \frac{MN \times 255^2}{\|f - \hat{f}\|^2}, \tag{16}$$

where $\|\,.\,\|$ represents the Euclidean norm, and $f$ and $\hat{f}$ are an original frame and the corresponding reconstructed frame, respectively. BD-PSNR (ΔSNR) and BD-rate (Δbitrate) are used to compare the average PSNR difference and average bit rate difference between RD-curves of the no filter method and the other approaches [12]. Also, the criterion

$$\frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|^2}{\|\boldsymbol{x}_k\|^2} \leq 10^{-6} \tag{17}$$

In addition, we observed that $10 \leq T_1 \leq 15$ and $50 \leq T_1 \leq 55$ are good ranges to remove the coding artifacts. In the experiments, $T_1 = 10$, $T_2 = 50$, and $c_3 = 30$ were used for the thresholds and the constant of the linear visual function of Eq. (10).

Performance comparisons such as bit rate, PSNR of luminance channel, ΔSNR, and Δbitrate are shown in Tables 1–6, where the results without the H.264 in-loop filter are used as the reference outcomes to compare the performance in terms of ΔSNR and Δbitrate. The results show that the proposed algorithm consistently outperforms the other approaches. It is observed that PSNR gain is higher as QI is higher. Overall, it is observed that the average performance gain of the proposed hybrid algorithm over H.264 in-loop filter is 0.05 dB (ΔSNR) or 1.0% (Δbitrate). The results mean that a better solution (more effective removal of blocking and ringing artifacts) can be obtained by finding an

**Table 1** Performance comparisons of QCIF 'Hall monitor' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 25.09 | 30.34 | N/A | N/A |
| | 32,33 | 43.09 | 33.83 | | |
| | 27,28 | 79.30 | 37.40 | | |
| | 22,23 | 166.87 | 40.61 | | |
| in-loop filter | 37,38 | 24.86 | 30.71 | 0.399 | -7.08 |
| | 32,33 | 42.53 | 34.19 | | |
| | 27,28 | 77.72 | 37.70 | | |
| | 22,23 | 161.24 | 40.76 | | |
| Ref. [7] | 37,38 | 25.09 | 30.67 | 0.121 | -2.38 |
| | 32,33 | 43.09 | 34.09 | | |
| | 27,28 | 79.30 | 37.42 | | |
| | 22,23 | 166.87 | 40.61 | | |
| WCLS | 37,38 | 25.09 | 30.70 | 0.182 | -3.39 |
| | 32,33 | 43.09 | 34.11 | | |
| | 27,28 | 79.30 | 37.52 | | |
| | 22,23 | 166.87 | 40.65 | | |
| Hybrid method | 37,38 | 25.09 | 30.89 | 0.419 | -7.57 |
| | 32,33 | 43.09 | 34.35 | | |
| | 27,28 | 79.30 | 37.78 | | |
| | 22,23 | 166.87 | 40.85 | | |

**Table 2** Performance comparisons of QCIF 'News' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 32.21 | 29.81 | N/A | N/A |
| | 32,33 | 55.66 | 33.28 | | |
| | 27,28 | 98.86 | 37.01 | | |
| | 22,23 | 173.41 | 40.76 | | |
| in-loop filter | 37,38 | 31.49 | 30.08 | 0.270 | -4.11 |
| | 32,33 | 55.03 | 33.51 | | |
| | 27,28 | 98.06 | 37.20 | | |
| | 22,23 | 172.44 | 40.86 | | |
| Ref. [7] | 37,38 | 32.21 | 30.11 | 0.158 | -2.39 |
| | 32,33 | 55.66 | 33.50 | | |
| | 27,28 | 98.86 | 37.11 | | |
| | 22,23 | 173.41 | 40.77 | | |
| WCLS | 37,38 | 32.21 | 30.15 | 0.182 | -2.75 |
| | 32,33 | 55.66 | 33.52 | | |
| | 27,28 | 98.86 | 37.14 | | |
| | 22,23 | 173.41 | 40.77 | | |
| Hybrid method | 37,38 | 32.21 | 30.28 | 0.336 | -5.08 |
| | 32,33 | 55.66 | 33.64 | | |
| | 27,28 | 98.86 | 37.34 | | |
| | 22,23 | 173.41 | 40.91 | | |

**Table 3** Performance comparisons of QCIF 'Foreman' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 48.90 | 29.49 | N/A | N/A |
| | 32,33 | 84.01 | 32.64 | | |
| | 27,28 | 152.44 | 36.01 | | |
| | 22,23 | 294.10 | 39.57 | | |
| in-loop filter | 37,38 | 47.38 | 29.63 | 0.200 | -3.56 |
| | 32,33 | 82.58 | 32.81 | | |
| | 27,28 | 150.43 | 36.10 | | |
| | 22,23 | 293.34 | 39.61 | | |
| Ref. [7] | 37,38 | 48.90 | 29.62 | 0.063 | -1.13 |
| | 32,33 | 84.01 | 32.77 | | |
| | 27,28 | 152.44 | 36.02 | | |
| | 22,23 | 294.10 | 39.58 | | |
| WCLS | 37,38 | 48.90 | 29.67 | 0.080 | -1.42 |
| | 32,33 | 84.01 | 32.77 | | |
| | 27,28 | 152.44 | 36.04 | | |
| | 22,23 | 294.10 | 39.60 | | |
| Hybrid method | 37,38 | 48.90 | 29.81 | 0.232 | -4.08 |
| | 32,33 | 84.01 | 32.92 | | |
| | 27,28 | 152.44 | 36.21 | | |
| | 22,23 | 294.10 | 39.71 | | |

**Table 4** Performance comparisons of CIF 'Hall monitor' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 73.45 | 31.90 | N/A | N/A |
| | 32,33 | 141.02 | 34.84 | | |
| | 27,28 | 344.63 | 37.63 | | |
| | 22,23 | 961.33 | 40.24 | | |
| in-loop filter | 37,38 | 71.36 | 32.26 | 0.399 | -11.65 |
| | 32,33 | 136.56 | 35.17 | | |
| | 27,28 | 328.81 | 37.89 | | |
| | 22,23 | 911.20 | 40.36 | | |
| Ref. [7] | 37,38 | 73.45 | 32.24 | 0.117 | -3.96 |
| | 32,33 | 141.02 | 35.10 | | |
| | 27,28 | 344.63 | 37.65 | | |
| | 22,23 | 961.33 | 40.26 | | |
| WCLS | 37,38 | 73.45 | 32.24 | 0.164 | -5.40 |
| | 32,33 | 141.02 | 35.14 | | |
| | 27,28 | 344.63 | 37.71 | | |
| | 22,23 | 961.33 | 40.31 | | |
| Hybrid method | 37,38 | 73.45 | 32.45 | 0.414 | -12.57 |
| | 32,33 | 141.02 | 35.34 | | |
| | 27,28 | 344.63 | 38.02 | | |
| | 22,23 | 961.33 | 40.58 | | |

**Table 5** Performance comparisons of CIF 'News' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 87.52 | 31.66 | N/A | N/A |
| | 32,33 | 153.13 | 34.94 | | |
| | 27,28 | 275.86 | 38.32 | | |
| | 22,23 | 501.41 | 41.59 | | |
| in-loop filter | 37,38 | 85.88 | 31.92 | 0.273 | -4.74 |
| | 32,33 | 150.30 | 35.16 | | |
| | 27,28 | 271.85 | 38.48 | | |
| | 22,23 | 497.71 | 41.68 | | |
| Ref. [7] | 37,38 | 87.52 | 31.97 | 0.172 | -3.00 |
| | 32,33 | 153.13 | 35.18 | | |
| | 27,28 | 275.86 | 38.44 | | |
| | 22,23 | 501.41 | 41.60 | | |
| WCLS | 37,38 | 87.52 | 31.99 | 0.186 | -3.23 |
| | 32,33 | 153.13 | 35.20 | | |
| | 27,28 | 275.86 | 38.44 | | |
| | 22,23 | 501.41 | 41.63 | | |
| Hybrid method | 37,38 | 87.52 | 32.14 | 0.323 | -5.58 |
| | 32,33 | 153.13 | 35.35 | | |
| | 27,28 | 275.86 | 38.57 | | |
| | 22,23 | 501.41 | 41.74 | | |

element of the intersection set between a solution space by WCLS and a solution space by POCS. Furthermore, we observed that the proposed POCS algorithm increased convergence speed by effectively suppressing the coding artifacts.

Figure 5 (a)–(f) shows the enlarged 87th original frame of the 'Hall monitor' sequence, the reconstructed frame without the H.264 in-loop filter, the reconstructed frame with the H.264 in-loop filter, the reconstructed frame with Ref. [7], the reconstructed frame with the proposed WCLS, and the reconstructed frame with the proposed hybrid method of WCLS and POCS, respectively. The results with the H.264 in-loop filter effectively remove the blocking artifact, but there is over-smoothness along the edge information. Furthermore, the ringing artifact is visible since pixels within block are not effectively suppressed. On the other hand, the proposed hybrid algorithm effectively reduces both artifacts. Similar results are consistently obtained with the 'News' and 'Foreman' sequences as shown in Figs. 6 and 7. Almost blocking artifacts as well as the ringing artifacts have been well removed by WCLS. However, the detailed blocking and the ringing artifacts still

**Table 6** Performance comparisons of CIF 'Foreman' sequence.

| Method | QI (I,P) | Bit rate (kbps) | PSNR_Y (dB) | ΔSNR (dB) | Δbitrate (%) |
|---|---|---|---|---|---|
| no-filter | 37,38 | 144.73 | 30.35 | N/A | N/A |
| | 32,33 | 256.36 | 33.20 | | |
| | 27,28 | 523.76 | 36.28 | | |
| | 22,23 | 1142.50 | 39.61 | | |
| in-loop filter | 37,38 | 138.23 | 30.57 | 0.264 | -5.84 |
| | 32,33 | 248.70 | 33.40 | | |
| | 27,28 | 515.21 | 36.43 | | |
| | 22,23 | 1134.40 | 39.68 | | |
| Ref. [7] | 37,38 | 144.73 | 30.55 | 0.076 | -1.68 |
| | 32,33 | 256.36 | 33.34 | | |
| | 27,28 | 523.76 | 36.30 | | |
| | 22,23 | 1142.50 | 39.63 | | |
| WCLS | 37,38 | 144.73 | 30.56 | 0.095 | -2.11 |
| | 32,33 | 256.36 | 33.35 | | |
| | 27,28 | 523.76 | 36.33 | | |
| | 22,23 | 1142.50 | 39.64 | | |
| Hybrid method | 37,38 | 144.73 | 30.71 | 0.322 | -7.00 |
| | 32,33 | 256.36 | 33.52 | | |
| | 27,28 | 523.76 | 36.63 | | |
| | 22,23 | 1142.50 | 39.81 | | |

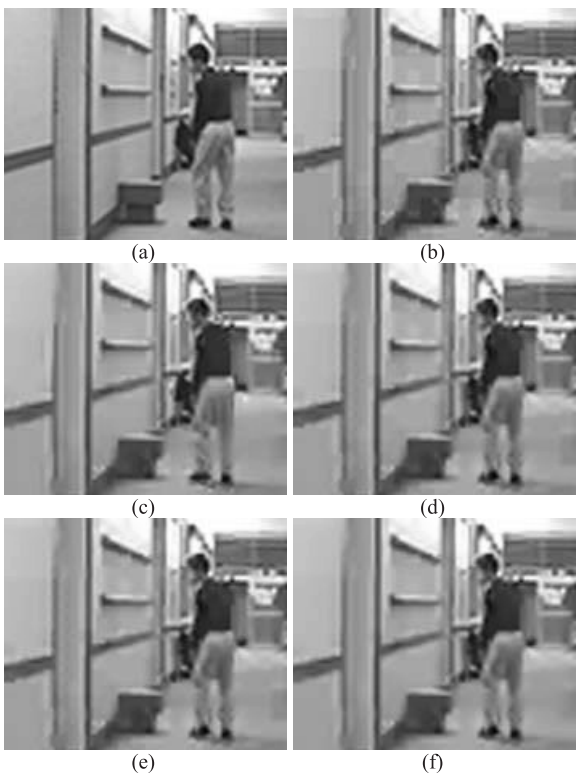


**Fig. 5** (a) Enhanced 87th original frame of QCIF 'Hall monitor' sequence, (b) reconstructed frame without filter (QI: I-37, P-38), (c) reconstructed frame with H.264 in-loop filter, (d) reconstructed frame with Ref. [7], (e) reconstructed frame with proposed WCLS, (f) reconstructed frame with proposed hybrid method.

**Fig. 6** (a) Enhanced 136th original frame of QCIF 'News' sequence, (b) reconstructed frame without filter (QI: I-37, P-38), (c) reconstructed frame with H.264 in-loop filter, (d) reconstructed frame with Ref. [7], (e) reconstructed frame with proposed WCLS, (f) reconstructed frame with proposed hybrid method.

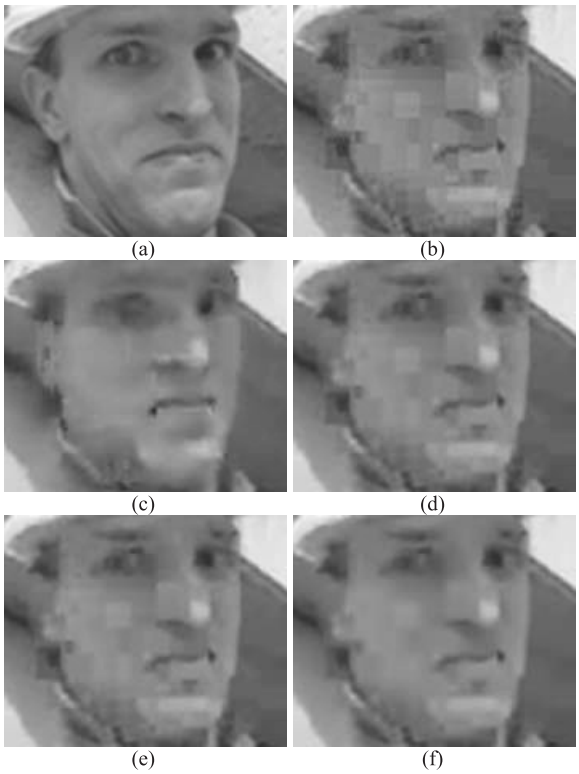

**Fig. 7** (a) Enhanced 138th original frame of QCIF 'Foreman' sequence, (b) reconstructed frame without filter (QI: I-37, P-38), (c) reconstructed frame with H.264 in-loop filter, (d) reconstructed frame with Ref. [7], (e) reconstructed frame with proposed WCLS, (f) reconstructed frame with proposed hybrid method.

appear, since WCLS has its restriction to reflect the local smoothness. On the other hand, the detailed coding artifacts are well removed with the proposed POCS algorithm, which incorporates the local smoothing constraint into the solution space determined by WCLS.

In order to evaluate the performance of visual quality, the double stimulus continuous quality scale (DSCQS) as described in ITU-R Recommendation BT.500-11 was used [13]. Thirty assessors were involved. Test scores range between 1.0 (poor) and 5.0 (excellent). DSCQS comparisons are shown in Tables 7 and 8. In all cases, the test score of the proposed algorithm overcomes the other approaches. However, the visual quality improvement is lower as QI is higher, since the compressed video is more seriously degraded and therefore, there is the limitation to recover the original information.

In addition, normalized average CPU running time was used to compare the computational complexity, as shown in Table 9. As expected, the H.264 in-loop filter is the most competitive among the methods used in our experiments with respect to computational cost. The computational cost of the proposed algorithm increases as QI is higher, since relatively many pixels need the filtering process. However, the computational cost of the proposed hybrid algorithm is less than that of WCLS as reported in Ref. [7] and our WCLS, since the hybrid approach needs fewer iterations for convergence. Even though the proposed hybrid algorithm is computationally more expensive than the H.264 in-loop filter, it can be applied to various higher visual quality services such as digital video monitoring system.

**Table 7** DSCQS comparisons with QCIF video sequences.

| QI (I,P) | Method | Hall Monitor | News | Foreman |
|---|---|---|---|---|
| 22,23 | no-filter | 3.7 | 3.7 | 3.8 |
| | in-loop filter | 4.3 | 4.4 | 4.2 |
| | Ref. [7] | 4.0 | 3.8 | 3.9 |
| | WCLS | 4.0 | 3.9 | 3.9 |
| | Hybrid method | 4.5 | 4.4 | 4.3 |
| 27,28 | no-filter | 3.4 | 3.3 | 3.4 |
| | in-loop filter | 4.0 | 3.9 | 3.9 |
| | Ref. [7] | 3.7 | 3.4 | 3.5 |
| | WCLS | 3.8 | 3.5 | 3.5 |
| | Hybrid method | 4.3 | 4.0 | 4.0 |
| 32,33 | no-filter | 2.9 | 2.6 | 2.7 |
| | in-loop filter | 3.5 | 3.2 | 3.3 |
| | Ref. [7] | 3.2 | 2.8 | 2.9 |
| | WCLS | 3.2 | 2.9 | 3.0 |
| | Hybrid method | 3.8 | 3.5 | 3.5 |
| 37,38 | no-filter | 2.1 | 2.0 | 1.8 |
| | in-loop filter | 2.6 | 2.4 | 2.3 |
| | Ref. [7] | 2.4 | 2.1 | 2.0 |
| | WCLS | 2.4 | 2.1 | 2.0 |
| | Hybrid method | 2.8 | 2.6 | 2.5 |

The coding artifacts were reduced with respect to ob-

**Table 8** DSCQS comparisons with CIF video sequences.

| QI (I,P) | Method | Hall Monitor | News | Foreman |
|---|---|---|---|---|
| 22,23 | no-filter | 4.0 | 3.6 | 3.8 |
| | in-loop filter | 4.4 | 4.4 | 4.3 |
| | Ref. [7] | 4.1 | 4.0 | 3.9 |
| | WCLS | 4.1 | 4.1 | 3.9 |
| | Hybrid method | 4.6 | 4.5 | 4.3 |
| 27,28 | no-filter | 3.6 | 3.4 | 3.4 |
| | in-loop filter | 4.1 | 4.0 | 3.9 |
| | Ref. [7] | 3.8 | 3.7 | 3.5 |
| | WCLS | 3.8 | 3.8 | 3.5 |
| | Hybrid method | 4.2 | 4.2 | 4.0 |
| 32,33 | no-filter | 3.0 | 3.0 | 2.7 |
| | in-loop filter | 3.6 | 3.5 | 3.3 |
| | Ref. [7] | 3.3 | 3.2 | 2.9 |
| | WCLS | 3.3 | 3.2 | 3.0 |
| | Hybrid method | 3.9 | 3.8 | 3.5 |
| 37,38 | no-filter | 2.2 | 2.4 | 1.8 |
| | in-loop filter | 2.8 | 2.6 | 2.5 |
| | Ref. [7] | 2.5 | 2.4 | 2.1 |
| | WCLS | 2.5 | 2.4 | 2.2 |
| | Hybrid method | 3.1 | 2.9 | 2.8 |

**Table 9** Normalized average running time comparisons.

| QI (I,P) | Method | QCIF | CIF |
|---|---|---|---|
| 22,23 | in-loop filter | 1.0 | 1.0 |
| | Ref. [7] | 1.6 | 1.7 |
| | WCLS | 1.4 | 1.3 |
| | Hybrid method | 1.3 | 1.5 |
| 27,28 | in-loop filter | 1.0 | 1.0 |
| | Ref. [7] | 2.6 | 2.9 |
| | WCLS | 2.1 | 2.0 |
| | Hybrid method | 1.9 | 2.2 |
| 32,33 | in-loop filter | 1.0 | 1.0 |
| | Ref. [7] | 4.1 | 4.5 |
| | WCLS | 3.3 | 3.2 |
| | Hybrid method | 3.0 | 3.2 |
| 37,38 | in-loop filter | 1.0 | 1.0 |
| | Ref. [7] | 6.2 | 6.8 |
| | WCLS | 5.0 | 4.8 |
| | Hybrid method | 4.2 | 4.4 |

jective criterion and subjective criterion by effectively controlling the degree of the global smoothness and the local smoothness. The proposed algorithm is to obtain a solution that is an element of the intersection set between a solution space using the weighted gradient approach incorporating the global smoothing constraint and a projection onto sets incorporating the local smoothing constraint.

## 5. Conclusions

In this paper, we propose a spatially adaptive post-processing algorithm to eliminate coding artifacts by preserving important information such as edge regions. In order to solve the over-smoothness problem of in-loop filter of H.264 and typical CLS approaches, a hybrid method combining POCS and WCLS is presented. The new local variation is defined according to the pixel location to minimize the coding artifact, and the local mean of difference of adjacent pixels is used to adaptively control the bound of the projection set. The proposed post-processing algorithm not only efficiently removes the coding artifacts without the loss of the edge information, but also more quickly converges.

In order to more rigorously preserve edge information, the incorporation of a robust function as a smoothing constraint term is under investigation. With this incorporation, it is expected that a more sophisticated formulation can be derived and better performance can be obtained.

## Acknowledgments

This work was supported by the Korean Science and Engineering Foundation (KOSEF) grant fund by the Korea Government (MEST) (No. 2011-0000148) and the Ministry of Knowledge Economy, Korea under the Information Technology Research Center support program supervised by the National IT Industry Promotion Agency (NIPA-2010-C1090-1021-0010).

### References

[1] I. Richardson, H.264 and MPEG-4 Video Compression, Wiley, 2003.
[2] M.-Y. Shen and C.-C.J. Kuo, "Review of postprocessing techniques for compression artifact removal," J. Vis. Commun. Image Represent, vol.9, no.1, pp.2–14, March 1998.
[3] M.-C. Hong, "Low computing post processing to suppress annoying artifacts of compressed video sequences," IEICE Trans. Inf. & Syst., vol.E89-D, no.3, pp.1214–1220, March 2006.
[4] Z. Li and E.J. Delp, "Blocking artifact reduction using a transform-domain Markov Random Field model," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.12, pp.1583–1593, Dec. 2005.
[5] ITU-T SG16/Q6, Draft ITU-T Recommendation on Final Draft International Standard of Joint Video Specification, May 2003.
[6] Y. Yang, N. Galatsanos, and A. Katsaggelos, "Regularized reconstruction to reduce blocking artifact of block discrete cosine transform compressed images," IEEE Trans. Circuits Syst. Video Technol., vol.3, no.6, pp.421–432, Dec. 1993.
[7] S.-W. Hong, Y.-H. Chan, and W.-C. Siu, "An adaptive constrained least square approach for removing blocking artifact," Proc. Int. Symposium on Circuit and System, vol.2, pp.773–776, May 1995.
[8] A.K. Katsaggelos, "Iterative image restoration algorithm," Optical Engineering, vol.28, no.7, pp.735–748, July 1989.
[9] M.-C. Hong, H.-T. Cha, and H.-S. Hahn, "A spatially adaptive Gradient-Projection image restoration," IEICE Trans. Inf. & Syst., vol.E85-D, no.5, pp.910–913, May 2002.
[10] G.L. Anderson and A.N. Netravali, "Image restoration based on a subject criterion," IEEE Trans. Syst. Man Cybern., vol.SMC-6, no.12, pp.845–853, Dec. 1976.
[11] T. Tan, G. Sullivan, and T. Wedi, "Recommended simulation common conditions for coding efficiency experiments," ITU-T SG16 Q.6 Document, VCEG-AE010r1, 31st VCEG Meeting, Marrakech, Morocco, Jan. 2007.
[12] G. Bjointegaard, "Calculation of average PSNR difference between RD-curves," ITU-T SG16 Q.6 Document, VCEG-M33, 13th VCEG Meeting, Austin, Texas, USA, April 2001.
[13] Recommendation ITU-T, BT.500-11, Methodology for the subjective assessment of the quality of television pictures, ITU-T, 2002.

**Kwon-Yul Choi** received the B.S. degree from Soongsil University, Korea in 2006. He is currently pursuing M.S. degree in information and telecommunication engineering at Soongsil University, Korea. His research interests include image restoration, image filtering, and video coding.

**Min-Cheol Hong** received B.S. and M.D. degrees from Yonsei University, Korea in 1988 and 1990, and Ph.D. degree from Northwestern University in 1997, respectively. He worked for LG Information and Communication Research Center and LG Electronics Corporate Research Center in 1990–1991 and 1998–2000, respectively. Also, he worked as a postdoctoral research Fellow at Northwestern University in 1997–1998. He joined Soongsil University in 2000, where he is currently an associate professor. His research areas include image/video restoration, nonlinear signal processing, motion modeling and analysis, blind image deconvolution, and video coding.