Proceedings of APCC2008 copyright © 2008 IEICE 08 SB 0083

A Spatially Adaptive Gradient–Projection Algorithm to Remove Blocking Artifacts of H.264 Video Coding Standard

Woo-Seok Seo, Kwon-Yeol Choi, and Min-Cheol Hong, *Member, IEEE*
School of Electronic Engineering Soongsil University
Sangdo5-dong Dongjak-gu, Seoul, 156-743, Korea
E-mail : seo1301@vipl.ssu.ac.kr, tantis@vipl.ssu.ac.kr, mhong@ssu.ac.kr.

***Abstract-* In this paper, we present a spatially adaptive gradient-projection algorithm for H.264 video coding standard to remove blocking artifacts using local statistics. The proposed approach is based on the regularized iterative techniques. In order to complement the weak points of the Constraint Least Squares (CLS) method, we combine the projection method with the CLS method. In order to adaptively process each pixel according to the Human Visual System (HVS), the modified local variance is defined when the CLS method is applied. What's more, the location-based laplacian operators are proposed to efficiently remove the blocking artifacts between the blocks. And the difference of vertically or horizontally adjoining pixels is utilized to constrain the solution space when the projection method is applied. Quantization Index (QP) in H.264 is also used to control the degree of smoothness. The simulation results show that the proposed post-processing filter works better than the loop filter in H.264 with preserving the edges information at once.**

Keywords: blocking artifacts, CLS, Deblocking Filter, H.264/AVC Coding Standard, projection.

I. INTRODUCTION

The various visual communications, applications, and services become possible due to the image compression techniques such as JPEG, MPEG, and H.264. These techniques use the Block-based Discrete Cosine Transform (BDCT) because it exhibits very high energy compaction, and does not need the supplementary information comparing with Karhune-Loeve Transform and Fourier Transformation [1]. When the images are highly compressed BDCT based compression results in the annoying blocking artifacts which are the discontinuous effect between the neighborhood blocks. As the bit rate is lower, the blocking artifacts become more visible and the reconstructed images are more seriously degraded [1, 2].

There are many works to remove the artifacts . The related works can be classified into three cases, such as pre-processing, in-loop processing, and post-processing approaches.

Among them, post-processing has the advantages on that it is unnecessary to modify the existing coding algorithms [1,2].

H.264 video coding standard has been developed jointly by ITU-T and ISO to improve the coding performance. It adopts the new methods such as 4×4 block based integer transform, spatial intra prediction, variable block size motion estimation, and in-loop filter to increase compression performance. Since 4×4 block-based integer transformation is used in H.264 video coding standard, more visible artifacts are appeared in the reconstructed image. Therefore, loop filter is used to effectively remove the artifacts. However, it is not efficient since it distorts the important edge information [2].

The reduction of the blocking artifacts or the recovery of the quantization information is a typical ill-posed problem because the quantization process in image compression is a many-to-one mapping operation. Therefore, in order to overcome the defects of H.264 loop filter. We introduce a spatially adaptive gradient-projection algorithm. Among the many regularized iterative techniques, the CLS method has been used to solve these kinds of problems [4,5]. We introduce an algorithm to combine the projection with the CLS method to complement some problems of the CLS method. Particularly, the weighted norm CLS is used to consider HVS(Human Visual System) which differently reacts to the same noise in the flat region and the complex region. As a tool to measure the local activity of the each pixel, the modified local variance is used when the CLS method is applied. In addition, to control the bound of the projection set, the difference of vertically or horizontally adjoining pixels is utilized when the projection method is applied.

This paper is organized as follows. The regularized iterative techniques such as the weighted norm CLS and the projection method are reviewed in Section 2. In Section 3, we describe how to extract the local spatial activity and how to define the projection set. Experimental results are given in Section 4, and finally conclusions are presented in Section 5.

II. BACKGROUND

In general, the degradation model caused by image compression can be written as

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

where y, x, and n are of size $MN\times 1$, and represent the lexicographically ordered reconstructed blocky image, original image, and the additive noise respectively [1]. Since the additive noise due to the quantization is typically assumed the Gaussian noise that is uncorrelated with the original image, the least squares approach has been widely used.

The CLS approaches have been used to obtain solutions to Eq. (1), where a smoothing constraint is imposed into the solution space [4,5]. The regularized solutions are found by minimizing the following objective functional

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

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

where C is the 2-D Laplacian operator to impose the smoothness constraint. The first terms in the right side of Eqs. (2) and (3) represent fidelity to the data, and the second terms represent smoothness of x [6,7]. The regularization parameter α controls the trade-off between them. In particular, R and L which are diagonal weighted matrices with size $MN \times MN$ are added in Eq. (3) as compared with Eq. (2) for taking into account HVS.

By minimizing the objective functional (3) with respect to x, the regularized iterative solution is obtained as shown in Eq. (4) [5,6]. The iterative techniques are very useful because parameters determining the solution can be updated at each iteration step. When the relaxation parameter controlling the convergence speed is defined such that the convergence is guaranteed, the iteration solution can be written as

$$\begin{aligned} x_{k+1} &= x_k + \beta(R^t Ry - (R^t R + \alpha C^t L^t LC)x_k) \\ &= Gx_k \end{aligned} , \tag{4}$$

where A^t represents the transpose of matrix A, β represents the speed of convergence if $0 < \beta < \dfrac{2}{\|R^t R + \alpha C^t L^t LC\|}$.

The partially restored image in Eq. (4) can be projected on a certain set with some constraints, which can be obtained from the prior knowledge of the image compression or noise. The projection process is expressed such as Eq. (5)

$$\begin{aligned} \hat{x}_k &= Px_k \\ x_{k+1} &= G\hat{x}_k = GPx_k \end{aligned} , \tag{5}$$

where P denotes a projection operator of a signal onto a set of signals with desirable properties [6,7].

III. PROPOSED ALGORITHM

In the following section, we will describe how to define diagonal weighted matrices of the weighted norm CLS method, and then present the constraint set used in the projection method. Finally, the simple and effective method for chrominance signals is described.

A. The weighted norm CLS method

Hong et al. [5] use the maximum local variance to decide the elements of the weighted matrices. However, it is dangerous that the maximum local variance is used in all pixels filtering process, since the maximum local variance can be degraded under the influence of noise, the blocking artifact in the block boundary or local statistics. For example, if a large maximum local variance is selected due to the blocking artifact in the block boundary, the diagonal elements of L will be assigned a relatively larger value. It results in over-smooth image on the whole. The higher QP in H.264 is, the bigger the variance in the block boundary becomes. As a result of the bigger variance, the blocking artifact which is regarded as the edge is preserved. Moreover, Ref[5] do not consider the discontinuity degree between pixels within the block and pixels in the block boundary, and do not consider the effect of the quantization step size.

Therefore, we propose the modified variance, which exactly extracts local activity and has low computational cost using by shift operation instead of divider operation. The modified variance of the pixels within the block and in the block boundary is differently computed to minimize the influence of the blocking artifacts. A partition to classify pixels of the 4×4 block is shown in Figure 1, where region V, region H, region c, and region M include the vertical boundary pixels, the horizontal boundary pixels, the corner boundary pixels, and the pixels within the block, respectively.

C	V	V	C
H	M	M	H
H	M	M	H
C	V	V	C

M={(1,1),(1,2),(2,1),(2,2)}
V={(0,1),(0,2),(3,1),(3,2)}
H={(1,0),(2,0),(1,3),(2,3)}
C={(0,0),(0,3),(3,0),(3,3)}

Figure 1. Partition of pixels in the 4×4 block

The Modified Local Variances (MLV) of region M, region H, and region V are shown in Eq. (6), (7), and (8) respectively. Also MLV of region C is defined as an average between the vertical MLV and the horizontal MLV.

$$\begin{aligned} MLV(i,j) &= |x(i,j) - m(i,j)| + |x(i-1,j) - m(i,j)| \\ &\quad + |x(i+1,j) - m(i,j)| + |x(i,j-1) - m(i,j)| \\ &\quad + |x(i,j+1) - m(i,j)| \qquad for (i,j) \in M, \\ m(i,j) &= (4x(i,j) + 3x(i-1,j) + 3x(i+1,j) \\ &\quad + 3x(i,j-1) + 3x(i,j+1))/2^4 . \end{aligned} \tag{6}$$

$$MLV(i,j) = \sum_{p=-2}^{2} \left| x(i,j+p) - m(i,j) \right| \quad for(i,j) \in H,$$
$$m(i,j) = (4x(i,j) + 3x(i,j-2) + 3x(i,j-1) + 3x(i,j+1) + 3x(i,j+2))/2^4 . \quad (7)$$

$$MLV(i,j) = \sum_{p=-2}^{2} \left| x(i+p,j) - m(i,j) \right| \quad for(i,j) \in V ,$$
$$m(i,j) = (4x(i,j) + 3x(i-2,j) + 3x(i-1,j) + 3x(i+1,j) + 3x(i+2,j))/2^4 . \quad (8)$$

Finally, the elements of R and L are independently obtained using the weight function ω(MLV) as shown in Eq. (9). ω(MLV) is a kind of the visual function. We select a linear visual function to control a value between 0 and 1, because it is simple and flexible to design comparing with exponential functions or trigonometric functions. Particularly, we put pixel location information and QP into ω(MLV) to control the smoothness. The parameter M is used to reflect such a prior knowledge. As a result of M, ω(MLV) relatively rises in the block boundary and at high QP. It is also possible to implement integer operation and shift operation instead of dividers.

$$\omega(MLV(i,j)) = \begin{cases} (M+90-MLV(i,j))/2^8 & if\ MLV(i,j) \le 10 \\ (M+110-3\times MLV(i,j))/2^8 & if\ 10 < MLV(i,j) < 50 \\ (M+10-MLV(i,j))/2^8 & otherwise \end{cases} \quad (9)$$

where

$$M = \begin{cases} 4\times QP & for\ pixels\ within\ the\ block \\ 5\times QP & for\ pixels\ in\ the\ block\ boundary \end{cases}$$

$$L(m,n) = \begin{cases} 1 & if\ \omega(MLV(i,j)) > 1 \\ 0 & if\ \omega(MLV(i,j)) < 0 \\ \omega(MLV(i,j)) & otherwise \end{cases} \quad (10)$$

$R(m,n) = 1 - L(m,m)$,

where $0 \le i \le M-1, 0 \le j \le N-1, m = i \times N + j$

ω becomes close to 1 when the variances are low so that the high frequency components in the corresponding flat regions are strongly regularized. Inversely, ω becomes close to 0 when the variances are high, as a result, the high frequency components in the corresponding complex regions are preserved. This is in agreement with HVS.

Furthermore, the various laplacian operators are applied depending on the location of the pixels within the 4×4 block to efficiently remove the blocking artifacts in the block boundary. In this work, a directional lapacian operator is used as shown Figure 2.

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

Figure 2. Directional laplacian operators

B. The projection method

The CLS method has some problems. First, as shown in Figure 2, it just refers to the neighborhood pixels whose order is one. It is very short to manage the blocking artifacts. Second, it is difficult to measure the local activity of the pixels in the block boundary. In other words, it is hard to sharply distinguish the blocking artifacts from the edge in the block boundary. As a result of the above two reasons, the blocking artifacts are not clearly removed. Moreover it needs a great number of iterations to be regularized at high QP.

In this work, we use the projection method as shown in Eq. (5) for the fast convergence and the perfect removal of the blocking artifacts. The assumption to define the projection set is shown in Figure 3. In conclusion, the pixels in the block boundary are projected onto the set which is consisted of the pixels whose difference is similar to the difference of the pixels within the block.

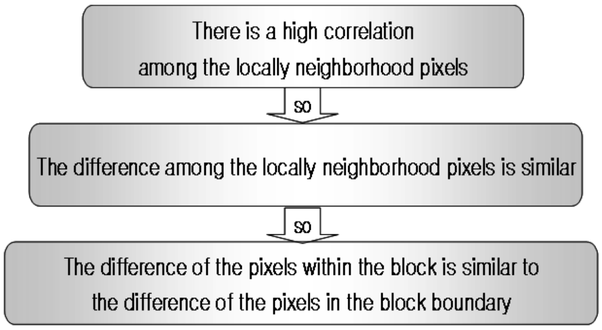


Figure 3. assumption about adjoining pixels

An important factor of the proposed projection method is the average of the differences of the horizontally or vertically adjoining pixels of the partially restored image at each iteration step. For example, we consider the horizontal case as shown in Figure 4.

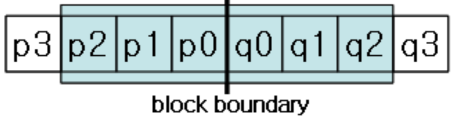


Figure 4. Horizontally adjoining pixels

MDB(Mean of Difference Between p0 and q0), MDA(Mean of Difference of Adjoining pixels), and BD(Block boundary Difference) are defined as

$$MDB = (|p2 - p1| + |p1 - p0| + |q0 - q1| + |q1 - q1|)/2^2 , \quad (11)$$

$$MDA=(3\times|p2-p1|+3\times|p1-p0|+4\times|p0-q0| \\ +3\times|q0-q1|+3\times|q1-q2|)/2^4 , \qquad (12)$$

$$BD=|p0-q0| . \qquad (13)$$

MDB is used to extract the local activity of the neighborhood pixels. MDA and BD represent the lower bound and upper bound of the projection set, respectively. For example, if MDB is large, that block boundary is the high activity region. As a result, a permissible range of the projection set is wide, namely looser bound. Inversely, if MDB is small, that block boundary is the smooth region. As a result, a permissible range of the projection set is limited, namely tighter bound. γ (MDB) is defined to adaptively control the bound of the projection set.

QP and pixel position as well as MDB are used to control a range of the projection set. The higher QP is, the more visible blocking artifacts become and the bigger the difference of pixels in the block boundary becomes. So the tighter bound is necessary to the pixels between block boundary and it is also necessary at higher QP is, the more visible blocking artifacts become and the bigger the difference of pixels in the block boundary becomes. The tuning parameter T is used to reflect them. According to these properties, the projection operator P of the difference between p0 and q0 is defined as

$$P(BD)=\begin{cases} BD\_BOUND & if\ BD>BD\_BOUND \\ BD & else \end{cases}$$

where $BD\_BOUND=(1\times\gamma)\times MDA+\gamma\times BD$

$$\gamma=\frac{T\times MDB^2}{T\times MDB^2+1} \qquad (14)$$

$$T=\begin{cases} \frac{8}{QP}\times 2 & for\ pixels\ within\ the\ block \\ \frac{8}{QP}\times 1 & for\ pixels\ in\ the\ block\ bouondary \end{cases}$$

Using P(BD) in Eq. (14), the final solution is obtain as if(p0<q0)

$$P_f(p0)=p0+P(BD),\ P_f(q0)=q0-P(BD)$$

else (15)

$$P_f(p0)=p0-P(BD),\ P_f(q0)=q0+P(BD)$$

Through the similar method, the pixels within the block are projected on the solution space. However the projection set bound of the pixels within block is relatively looser than that of the pixels in the block boundary.

C. *The simple method for chrominance signal*

The HVS is less sensitive to color than to luminance [8]. So it is unnecessary to use the regularized iterative techniques with the high computing cost. So chrominance signals (Cb, Cr) are only once projected using Eq. (14) and (15). It is sufficient to obtain the sub-optimal solution. Its adequateness is verified through the experiment results.

IV. EXPERIMENT RESULTS

We tested the proposed post-processing filter under the various conditions. The regularization parameter is set to $\alpha=\frac{J(x)^2}{Q(x)^2}$ like [4,5]. In particular, $J(x)^2$ is derived from QP_{step} . The termination criterion is defined as

$$(\frac{\|x_{k+1}-x_k\|_2^2}{\|x_k\|_2^2}<\mathrm{E})\ \text{or (ITERATION NUMBER > IN)}$$

$$\text{where } E=\begin{cases} 5\times10^{-6} & for\ proposed\ method \\ 5\times10^{-7} & for\ only\ CLS\ method \end{cases} , \qquad (16)$$

$$IN=\begin{cases} 5 & for\ poposed\ method \\ 15 & for\ only\ CLS\ method \end{cases} .$$

PSNR is used for the objective measure. For M×N dimensional 8 bits image, it is defined as

$$PSNR=10\log_{10}(\frac{MN\times255^2}{\|x-\hat{x}\|^2}) , \qquad (17)$$

where $\|\cdot\|$ represents the Euclidean norm, x and $\hat{x}$ denote the original frame and the recovered frame, respectively.

PSNR and the Average Iteration Number (AIN) comparisons as a function of QP are shown in Tables 1 and 2. Although the only proposed filter is applied at the decoder end unlike loop filter of H.264, PSNR of results with the proposed post-processing filter is higher than PSNR of results with loop filter of H.264. In particular, PSNR gain is bigger at high QP. Moreover, PSNR of results with the proposed post-processing filter is higher than PSNR of results with the only CLS method, at the same the proposed post-processing filter converges faster than the only CLS method. And for chrominance signals, PSNR of them is similar to PSNR of results with loop filter. So one-iteration step projection is sufficient to obtain sub-optimal solution.

In Figure 5 and 6 the image filtered with loop filter is degraded by over-blurring, and the loop filter misses the some blocking artifacts. But the missed blocking artifacts are clearly removed by using the proposed filter. Particularly, the ringing artifacts near the edge are also removed.

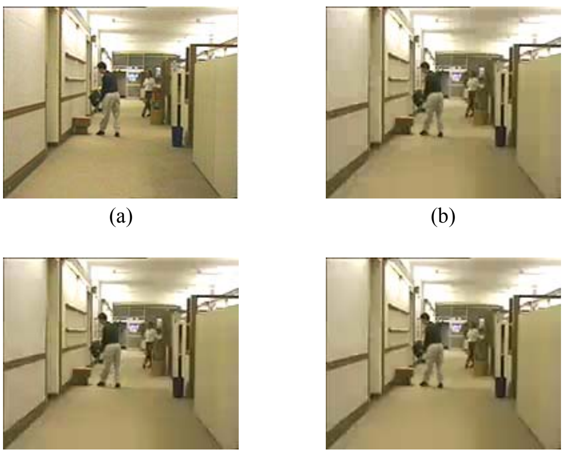
V. CONCLUSIONS

In this paper, we proposed the spatially adaptive post-processing deblocking filter to eliminate the blocking artifacts with preserving the high frequency components. In order to overcome defects of loop filter of H.264 and the CLS method, we combined the projection method with the weighted norm CLS method. The modified variance is differently computed according to the pixel location to minimize the effect of the blocking artifacts, and the mean of differences of neighbor pixels is also used to adaptively control the bound of the projection set.

As a result of combination, the post-processing filter not only removes the blocking artifacts without distortion, but also converges more quickly. Since the post-processing filter is located in the only decoder end, it can be adopted to the other BDCT based compression techniques.

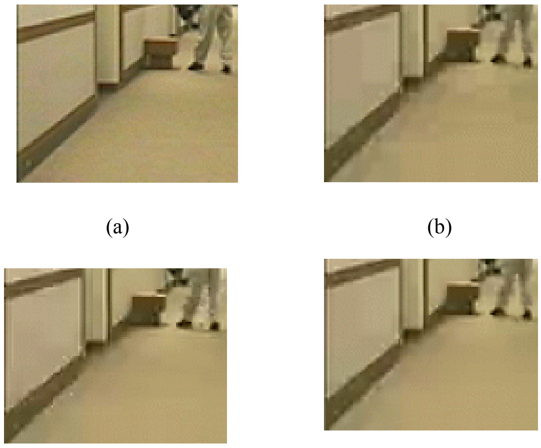
TABLE I
PSNR & AIN COMPARISON OF QCIF HALL_MONITOR

QP	Filter type	AIN	SNRY(dB)	SNRU(dB)	SNRV(dB)
QP 31	No filter	-	34.96	38.21	40.46
	Loop filter	-	35.34	38.66	40.73
	Only CLS	4	35.16	38.77	40.77
	Proposed filter	3	35.44	38.77	40.77
QP 36	No filter	-	31.21	36.58	39.20
	Loop filter	-	31.60	36.94	39.45
	Only CLS	6	31.56	37.06	39.47
	Proposed filter	3	31.77	37.06	39.47
QP 41	No filter	-	27.64	36.01	38.69
	Loop filter	-	28.06	36.42	38.88
	Only CLS	10	28.06	36.51	38.93
	Proposed filter	5	28.25	36.51	38.93
QP 46	No filter	-	24.22	35.21	37.42
	Loop filter	-	24.58	35.50	37.52
	Only CLS	15	24.71	35.59	37.56
	Proposed filter	5	24.84	35.59	37.56



(a) (b) (c) (d)

Figure 5. (a) 92th reconstructed frame of QCIF Hall_monitor, (b) corresponding CLS filtered frame, (c) corresponding loop filtered frame, (d)corresponding proposed post-filter frame



(a) (b) (c) (d)

Figure 6. (a) 92th reconstructed bottom left frame of QCIF Hall_monitor (zoom × 2), (b) corresponding CLS filtered bottom left frame (zoom × 2), (c) corresponding loop filtered bottom left frame (zoom × 2), (d)corresponding proposed post-filtered bottom left frame (zoom × 2)

ACKNOWLEDGMENT

This work was supported by the Seoul Future Contents Convergence (SFCC) Cluster established by Seoul R\BD Program and Soongsil University Research Fund.

REFERENCES

[1] M.-Y. Shen and C.-C. J. Kuo, "Review of postprocessing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation*, 9(1), pp.2-14, Mar. 1998.
[2] Z. Li, 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
[3] ITU-T SG16/Q6, JVT-G050r1 Draft ITU-T Recommendation on Final Draft International Standard of Joint Video Specification, May 2003.
[4] 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.
[5] S-W Hong, Y-H Chan, W-C Siu, "an adaptive constrained least square approach for removing blocking artifact," *IEEE ISCAS 95*, vol. 2, pp.773-776, May 1995.
[6] A. K. Katsaggelos, "Iterative Image Restoration Algorithm," *Optical Engineering*, vol. 28, no. 7, pp.735-748, Jul. 1989.
[7] M-C Hong, H-T Cha, H-S Hahn, "A Spatially Adaptive Gradient-Projection Image Restoration," *IEICE Trans. Information and Syst.*, vol. E85-D, no. 5, pp.910-913, May. 2002
[8] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. Circuits Syst.*, vol. 3, no. 6, pp.421-432, Dec. 1993.