Complexity Scalable Video Decoding Scheme for H.264/AVC

Hoyoung Lee[1], Jaehwan Kim[2], Luong Pham Van[3], Bongsoo Jung[4], Kwangpyo Choi[5], Younghun Joo[6], and Byeungwoo Jeon[7]
[1,2,3,7] Departments of Electronic and Electrical Engineering, Sungkyunkwan University, Suwon, KOREA
[1]hoiing@skku.edu, [2]kjh4759@skku.edu, [3]pvluong@skku.edu,[7]bjeon@skku.edu
[4,5,6] Samsung Electronics Co., LTD., Suwon, KOREA
[4]bongsoo.jung@samung.com, [5]kp5.choi@samsung.com, [6]mrzoo@samsung.com

***Abstract*— Recent proliferation of portable devices requires video contents playable on virtually any portable devices. However, their limited computing resources pose significant restriction to accomplish real-time decoding of high resolution or high quality video contents. To solve the problem, we propose a complexity scalable video decoding scheme for real-time playback on portable devices. In this paper, we analyze the complexity of H.264/AVC video decoding elements and develop a complexity scalable decoding scheme by simplifying motion compensation and deblocking filtering. Experimental results with the H.264/AVC main profile coded bitstream show that its decoding complexity can be reduced up to 26% without significant loss in subjective quality as compared to the conventional H.264/AVC decoder.**

Keywords-H.264/AVC; video decoder; complexity scalable decoding

I. INTRODUCTION

The development of portable devices and prevalence of wireless communication infrastructures make the multimedia services very popular. The video services such as IPTV, video conferencing, or video telephony are the most popular multimedia services. However, the limited battery capacity and the computational performance of the portable devices are major restriction to implement real-time playback of video contents on portable devices. Moreover, increasing demand of playback of higher resolution or quality video contents on portable platforms makes it quite difficult to attain real-time video.

To attain the real-time video decoding, we have to decrease its workload by reducing computational complexity of decoding process. The significant problem in such reduction is the possible huge degradation of objective quality due to processing mismatch between encoder and decoder. Since distorted pictures caused by reduced complexity process are used again for reference picture for the following pictures to be decoded, propagation of distortion error will become larger so that it causes even more degradation of objective quality as time goes on. However, due to some characteristics of human visual system, certain degradation in video quality might be tolerable to a certain extent.

Therefore, many complexity scalable video decoding algorithms have been developed to make a good compromise between complexity reduction and subjective quality drop [1]-[4]. Peng [1] proposed a discrete-cosine transform (DCT)-based complexity scalable video decoder via pruning the DCT data. Chen *et al.* [2] expended the IDCT pruning approach [1] by using a simpler interpolation filtering method according to frame types in motion compensation. Lei *et al.* [3] proposed a complexity scalable algorithms in AVS video codec (Audio and Video coding Standard in China) using a loop filter and luminance interpolation in motion compensation scaling method. Its encoder sends some information about the loop filter and the luminance interpolation to a decoder for complexity control of a decoder. H. Nam *el at* [4] proposed complexity scalable H.264 decoder with downsized decoding. W. Ji *el at* [5] proposed energy-scalable video decoding algorithms.

The H.264/AVC standard achieves high coding efficiency with many advanced coding tools such as variable block size motion compensation, multiple reference frames, quarter-pel motion vector accuracy, context adaptive entropy coding, etc. [6]. To implement a complexity scalable H.264/AVC decoder, we analyze its decoding complexity by decoding tool by tool and develop some complexity control parameters which our proposed complexity scalable decoding scheme utilizes in its decoding complexity control.

The rest of the paper is organized as follows. In Section II, we analyze H.264/AVC video decoding functions from the viewpoint of complexity control. In Section III, we describe the proposed method for complexity control of decoder. Experimental results are given in Section IV. Finally, we make some conclusions in Section V.

II. VIDEO DECODING ELEMENTS FOR COMPLEXITY CONTROL

H.264/AVC decoder performs variable length decoding (VLD) of incoming bitstream and then reconstructs various syntax elements such as motion vector, reference index, quantization parameter, and residual data of slices. The residual data are obtained through inverse quantization (IQ) and inverse transform (IT). Following, they are combined with a predictor which is generated either by motion compensation or by intra prediction. Subsequently, reconstructed picture is generated through deblocking filtering process. To evaluate the video decoding elements from the view point of complexity control, we analyze the complexity of them. Table I depicts the complexity profiling

TABLE I. COMPLEXITY PROFILE OF DECODING ELEMENTS

Decoding elements	Complexity rate(%)
Motion compensation	27.51
Variable length decoding(VLD)	25.19
Deblocking filter	16.65
IQ/IT	10.65
Reconstruction	3.08
Intra prediction	0.57
Others	16.34

result of H.264/AVC decoder in terms of decoding time. Its analysis is based on bitstreams conforming to the H.264/AVC main profile with IBBPBBP structure where every 60th frame is coded as I picture.

As shown in Table I, motion compensation and VLD are the major complex elements in a video decoder. This means, other than the VLD, motion compensation is the most complex process in decoder. Second major complex element is the deblocking filter, which is applied to reconstructed picture after finished the decoding process. It is an important element to improve a subjective quality of reconstructed video, especially, in low bit rates bitstreams. By the way, IQ/IT takes only about 10% of complexity in which IT occupies more computational complexity than IQ. In the previous investigation [1], it is found out that complexity control of IT process brings a significant quality loss. Therefore, in this paper, we decide two decoding functions of motion compensation and deblocking filter for complexity control.

A. Motion compensation

In H.264/AVC motion compensation for luma component, pixel value at fractional quarter-pel positions generated according to motion vectors. Fig. 1 depicts their positions. Predicted values at half-pel positions are generated by an one-dimensional 6-tap FIR interpolation filtering horizontally or vertically. A sample value at quarter-pel position is generated by averaging values at two nearest half-pel and integer positions. Computational complexity to generate the fractional samples is different depending on sample positions as depicted in Table II. Samples at quarter-pel positions labeled as *f, i, k, q* are the most complex positions. On the other hands, half-pel samples labeled *b*, *h* are the least complex ones. To reduce the complexity of interpolation filtering, we simplify the interpolation filtering of each sample position using adjacent integer-pel samples. For example, a simplified sample value at quarter sample position labeled as *a* is derived,

$$\begin{aligned} a &= (G+b+1)/2 \\ &= \{G+(E-5\cdot F+20\cdot G+20\cdot H-5\cdot I+J+16)/32+1\}/2 \\ &= (E-5\cdot F+52\cdot G+20\cdot H-5\cdot I+J+48)/64 \\ &\simeq (48\cdot G+16\cdot H+32)/64 \end{aligned} \quad (1)$$

where G is at an integer position as depicted in Fig. 1, and b is at half-pel position derived by 6-tap FIR interpolation filtering. To use a shift operation, we adjust a rounding value appropriately. In Table II, we propose a simplified luma interpolation filtering of each fractional sample. Motion

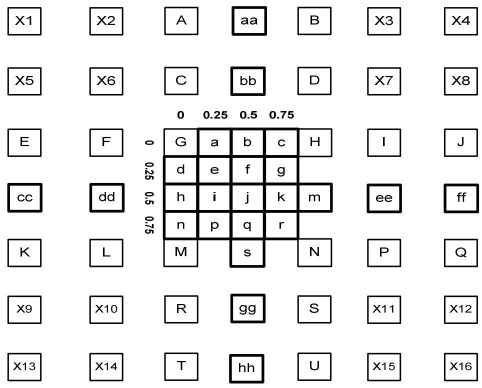

Figure 1. Fractional sample positions for quarter sample luma interpolation

compensation for chroma components are generated by a bi-

TABLE II. COMPLEXITY COMPARISON OF INTERPOLATION FILTERING FOR H.264/AVC AND THE PROPOSED METHOD

	Sample position	Interpolation by H.264/AVC	Simplified interpolation by the proposed
I	b(0.5,0)	6-tap	*(G+H+1)/2*
	h(0.5,0)	6-tap	*(G+M+1)/2*
II	a(0.25,0)	6-tap+2-tap	*(48G + 16H+32)/64*
	c(0.75,0)	6-tap+2-tap	*(16G + 48H+32)/64*
	d(0,0.25)	6-tap+2-tap	*(48G + 16M+32)/64*
	n(0,0.75)	6-tap+2-tap	*(16G + 48M+32)/64*
III	e(0.25,0.25)	(6-tap)×2+2-tap	*(2G + H+M+2)/4*
	g(0.25,0.75)	(6-tap)×2+2-tap	*(G + 2M+N+2)/4*
	p(0.75,0.25)	(6-tap)×2+2-tap	*(2H + N+G+2)/4*
	r(0.75,0.75)	(6-tap)×2+2-tap	*(H + 2N+M+2)/4*
IV	j(0.5,0.5)	(6-tap)×6+6-tap	*(G + M + H + N+2)/4*
V	f(0.5,0.25)	(6-tap)×6+6-tap+2-tap	*(3G + 3H + M + N+4)/8*
	i(0.5,0.75)	(6-tap)×6+6-tap+2-tap	*(G + H + 3M + 3N+4)/8*
	k(0.25,0.5)	(6-tap)×6+6-tap+2-tap	*(3G + H + 3M + N+4)/8*
	q(0.75,0.5)	(6-tap)×6+6-tap+2-tap	*(G + 3H + M + 3N+4)/8*

linear interpolation of four neighboring integer samples as,

$$\begin{aligned} a=&((8-xFrac_c)\times(8-yFrac_c)\times A+xFrac_c\times(8-yFrac_c)\times B+ \\ &(8-xFrac_c)\times yFrac_c\times C+xFrac_c\times yFrac_c\times D+32)/64 \end{aligned} \quad (3)$$

where a is a predicted chroma sample value and A,B C,D are the integer-pel position samples. $xFrac_c$ and $yFrac_c$ are the fractional offsets of fractional samples. To reduce the complexity for chroma interpolation filtering, a predicted chroma sample is copied from nearest neighboring integer-pel sample.

B. Deblocking filter

In H.264/AVC video coding, its deblocking filter consists of three processing phases: boundary strength decision, filtering decision, and actual pixel filtering.

In the boundary strength decision of H.264/AVC, boundary strength parameter (BS) is determined by the rules in [6] for each boundary of 4×4 block. BS can be 0 ~ 4 according to the rules. In our experiment, we found out that most often selected BS value is 0 or 2. Furthermore blocking artifact is more noticeable in flat and simple regions than in complex textured regions [7]. In H.264/AVC, flat and simple regions are often predicted in large partitions such as 16×16 or 16×8, 8×16. On the other hand, complex regions are mostly coded under 8×8 sub-block partitions. Using these observations, we re-design a boundary strength decision process as shown in Fig. 2.

Filtering process for sample sets (p_0, q_0) only takes place when following condition is satisfied [6]:

$$\begin{aligned} &BS > 0 \ and \\ &|p_0 - q_0| < \alpha \,\&\&\, |p_1 - p_0| < \beta \,\&\&\, |q_1 - q_0| \leq \beta \end{aligned} \tag{4}$$

where α, β are thresholds dependent on quantization parameter QP. Since filtering decision process has lots of comparison operation and it is performed for each edge, complexity would be much increased. In our simplified deblocking filter, filtering decision is performed just one time per each 4×4 block using an average value of samples such as:

$$\begin{aligned} &BS' > 0 \ and \\ &|p_{0Avg} - q_{0Avg}| < \alpha \,\&\&\, |p_{1Avg} - p_{0Avg}| < \beta \,\&\&\, |q_{1Avg} - q_{0Avg}| \leq \beta \end{aligned} \tag{5}$$

where p_{Avg} and q_{Avg} is an average value of samples in P and Q 4 4 blocks.

After the boundary strength decision and filtering decision, actual filtering process is applied to each block boundary. In H.264/ AVC, filtering strength is different depending on BS value. If BS < 4, a 4-tap FIR filter is applied with input samples p_0, p_1, q_0, q_1, and producing

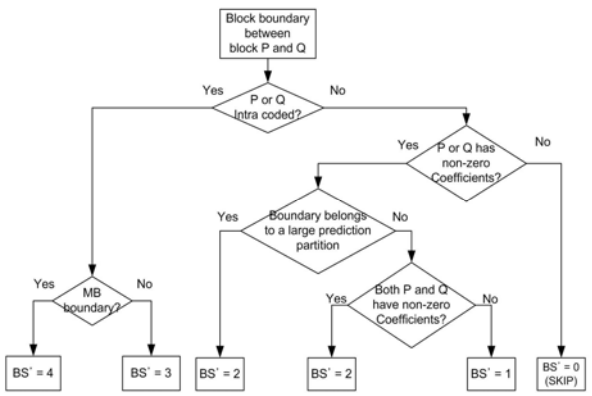

Figure 2. Simplified boundary strength decision

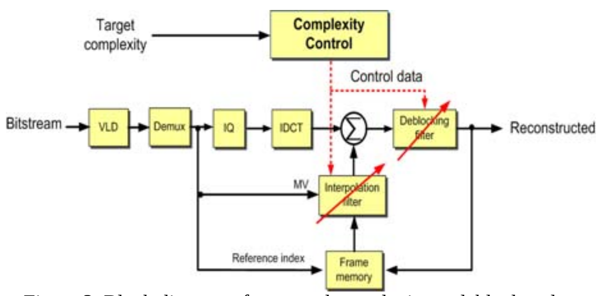

Figure 3. Block diagram of proposed complexity scalable decoder

TABLE III. PROPOSED MOTION COMPENSATION COMPLEXITY REDUCTION LEVEL (luma)

MCR_{Level}	Complexity reduction sample position
0	No reduction
1	*f,i,q,k*
2	*j*+ (MCR_{Level}=1)
3	*e, p, r, g*+ (MCR_{Level}=2)
4	*a, c, d, n*+ (MCR_{Level}=3)
5	*b, h*+ (MCR_{Level}=4)

TABLE IV. PROPOSED COMPLEXITY REDUCTION LEVEL FOR DEBLOCKING FILTER

DFR_{Level}	I Slice	P Slice	B Slice
0	No reduction		
1	a	a	b
2	a	b	b
3	b	b	b
4	b	b	c
5	c	c	c

(a: conventional deblocking filter, b: simplified deblocking filter, c: forced deblocking filter off)

outputs p_0' and q_0'. In case of BS being 4, a 5-tap or 4-tap filter is applied according to conditions [6]. However, the proposed simplified deblocking filter applies a 2-tap FIR weak filter when BS' of a current block boundary is 1 according to the proposed boundary strength decision rules. When BS' value is larger than 1, we apply same filtering method like to that of H.264/AVC.

III. COMPLEXITY SCALABLE VIDEO DECODING SCHEME

Fig. 3 depicts a block diagram of the proposed complexity scalable decoding scheme. The complexity scalable video decoder has a minimum quality loss within the maximum complexity reduction by controlling the control parameters. Therefore, we have to find an optimal complexity control level of parameters which satisfy the minimum quality loss. To find an optimum control level of parameters, we evaluate complexity-distortion (C-D) performance according to various complexity control parameters.

A. Motion compensation complexity control

In this paper, we apply a simplified interpolation filtering method according to fractional sample positions as proposed in Table II. The complexity scalability for motion compensation can be attained by controlling the number of luma samples which are involved in the simplified interpolation filtering. Table III depicts a proposed motion compensation complexity reduction level (MCR_{Level}). As depicted in Table III, we reduce the complexity of interpolation filtering for luma samples from the most complex sample position labeled as *f*, *i*, *k*, *q* to the least complex position labeled as *b*, *h* according to MCR_{Level}. Furthermore in order to control the motion compensation for chroma samples, we apply the proposed simplified chroma interpolation method when the MCR_{Level} is larger than 0.

 ISBN: 978-1-61208-129-8

B. Deblocking filter complexity control

To control the complexity of deblocking filter, we apply a simplified deblocking filter and selectively make the deblocking filter off according to slice type as depicted in Table IV. We define a complexity control level for deblocking filter, namely deblocking filter reduction level (DFR_{Level}). When DFR_{Level} is 0, we use the conventional H.264/AVC deblocking filter without any complexity reduction. For increasing DFR_{Level}, we apply the proposed simplified deblocking filter from B slice to I slice. To attain the maximum complexity reduction of deblocking filter, we switch off the deblocking filtering process for all slice types.

C. Complexity control scheme

We define two complexity control parameters as above; those are DFR_{Level}, MCR_{Level} respectively. To find an optimum control level which satisfies a minimum distortion loss, we evaluate complexity-distortion(C-D) performance [8]. The complexity reduction and distortion is measured by following equations:

$$AST[\%] = \frac{DecodingTime(\text{reference}) - DecodingTime(\text{proposed})}{DecodingTime(\text{reference})} \times 100 \qquad (11)$$

$$\Delta PSNR = PSNR(\text{proposed}) - PSNR(\text{reference})$$

where AST is an average saving time and ΔPSNR is a difference in objective quality between reference and the proposed method and used for quality distortion.

Fig. 4 depicts complexity-distortion performance in terms of (DFR_{Level}, MCR_{Level}). In Fig. 4, each point represents MCR_{Level}s according to DFR_{Level}. And the optimal complexity control points are drawn with a line. Table V shows the complexity control parameter level *g* determined by (DFR_{Level}, MCR_{Level}) which minimize distortion and maximize complexity reduction as shown in Fig. 4. We have set up 6 different complexity levels and expected distortion and complexity reduction according to complexity level *g* as in Table V.

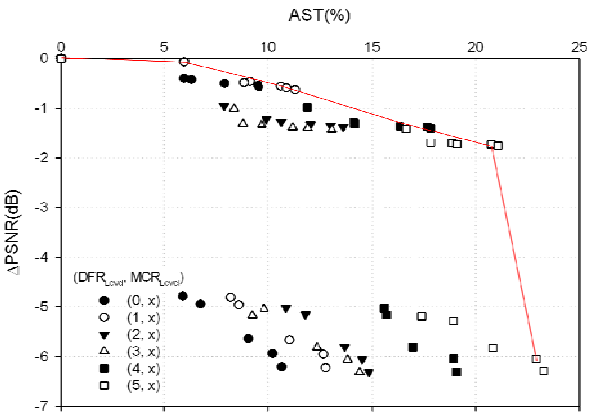

Figure 4. Joint complexity-distortion curve

TABLE V. JOINT COMPLEXITY REDUCTION LEVEL

Complexity level(*g*)	DFR_{Level}	MCR_{Level}	Relative Distortion	AST(%)
0	0	0	0	0
1	1	0	-0.072	5.9533
2	1	3	-0.5625	10.6242
3	4	3	-1.3635	16.325
4	5	4	-1.735	20.7446
5	5	5	-6.2835	23.2606

TABLE VI. EXPERIMENTAL RESULTS OF THE PROPOSED METHOD

Seq.	*g*	ΔPSNRY	ΔPSNRCb	ΔPSNRCr	AST(%)
city	0	0	0	0	0
	1	-0.06	-0.01	0.00	5.41
	2	-0.70	-3.97	-4.39	12.89
	3	-1.37	-4.03	-4.48	18.61
	4	-1.33	-4.76	-5.16	22.84
	5	-7.04	-4.76	-5.16	26.06
harbour	0	0	0	0	0
	1	-0.11	-0.03	-0.03	5.89
	2	-0.91	-3.99	-2.84	10.47
	3	-1.77	-4.43	-3.26	17.95
	4	-2.24	-5.64	-4.40	22.68
	5	-7.93	-5.64	-4.40	25.08

IV. EXPERIMENTAL RESULTS

To evaluate performance of the proposed scheme, we implemented it on JM17.0 H.264/AVC reference software. Bitstreams for experiments are coded as H.264/AVC main profile with GOP size = 60 under IBBPBBP structures. The number of reference frames is 5, and one picture is coded as one slice. QP is set to 22, 27, 32, and 37. The number of total encoded pictures is set to 300. We used two standard definition sequences for experiments: city and harbor.

The performance of the proposed scheme is measured by AST(%) and ΔPSNR(dB). Table VI shows experimental results of the proposed scheme. We can see that the proposed scheme attains similar complexity scalability and complexity reduction as estimated. However, we can find a huge distortion when complexity control level *g* is 5. Since P slice is used to as a reference slice to another slices, the error according to complexity control propagates to another slices. When *g* is 5, we reduce the complexity of motion compensation in P slices to attain the maximum complexity reduction. The errors in motion compensation in P slices are propagated to other slices. Fig. 5 shows the subjective qualities according to complexity control level *g*. As depicted in Fig. 5, we can identify that there is no significant subjective quality loss until *g*=4. However, the complexity control level *g* is 5, we can see some subjective quality loss, but it also maintains acceptable visual quality.

V. CONCLUSIONS

In this paper, a complexity scalable H.264/AVC decoding scheme is proposed using two control parameters. The proposed scheme can control complexity of a decoder with variable complexity control levels with the minimum quality loss. However, the proposed scheme has a restriction according to slice type. Our future work will find another complexity control variables which can control the complexity regardless of slice types and reduce more complexity as well. We will also develop complexity estimation and control method.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 20100008030).

REFERENCES

[1] S. Peng, “Complexity scalable video decoding via IDCT data pruning,” in Proc. of IEEE conference on consumer electronics 2001. Los Angeles, CA, June 2001.

[2] Y. Chen, Z. Zhong, T. Lan, S. Peng, and K. van Zon, "Regulated complexity scalable MPEG-2 video decoding for media processors," IEEE Transactions on Circuits and Systems for Video Technology, vol.12, no.8, pp. 678- 687, Aug 2002.

[3] C. Lei, Y. Chen, and W. Ji., “A complexity scalable decoder in an AVS video codec,” in Proc. of the 6th International Conference on Advances in Mobile Computing and Multimedia (MoMM '08), ACM, New York, USA.

[4] H. Nam, J. Jeong, K. Byun, J. Kim, and S. Ko , "A complexity scalable H.264 decoder with downsizing capability for mobile devices," Consumer Electronics, IEEE Transactions on , vol. 56, no. 2, pp. 1025-1033, May 2010.

[5] W. Ji, M. Chen, X. Ge, P. Li, and Y. Chen, “ESVD: An Integrated Energy Scalable Framework for Low-Power Video Decoding Systems,” EURASIP Journal on Wireless Communications and Networking, vol. 2010, Article ID 234131, 14 pages, Jun. 2010.

[6] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” IEEE Trans. Circuits Syst. VideoTechnol., vol. 13, no. 7, pp. 560–576, Jul. 2003.

[7] S. D. Kim, J. Yi, H. M. Kim, and J. B. Ra, “A deblocking filter withtwo separate modes in block-based video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp. 156–160, Feb. 1999.

[8] H. Lee, B. Jung, J. Jung, and B. Jeon, “Computational complexity scalable scheme for power-aware H.264/AVC encoding,” in Proc. of IEEE International Workshop on Multimedia Signal Process (MMSP '09), Rio de Janeiro, Brazil, pp. 1–6, Oct. 2009.

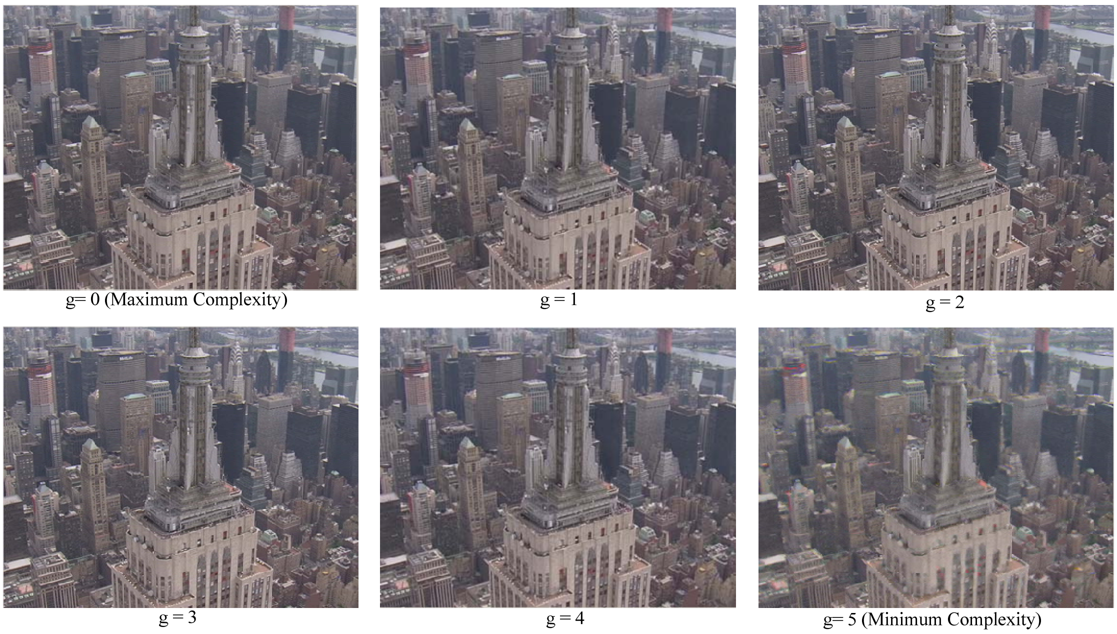

Figure 5. Subjective quality of proposed method (City sequence, QP 22 299th slice)