

Combined Scalable Video Coding Method for Wireless Transmission

Kalvein Rantelobo^{*1}, Wirawan², Gamantyo Hendrantoro³, Achmad Affandi⁴

^{1,2,3,4}Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember (ITS)
Kampus ITS Sukolilo, Surabaya 60111, Phone/Fax. +62 31 5947302/5931237

¹Department of Electrical Engineering, FST, Universitas Nusa Cendana
Jl. Adisucipto, Penfui, Kupang, NTT 80111 Phone/Fax. +62 380 881557

e-mail: rantelobo@ieee.org^{*1}, wirawan@its.ac.id², {gamantyo, affandi}@ee.its.ac.id^{3,4}

Abstrak

Video streaming bergerak adalah salah satu layanan multimedia yang mengalami perkembangan sangat pesat. Penggunaan bandwidth yang berfluktuatif pada jaringan kanal nirkawat merupakan kendala utama dalam komunikasi multimedia dewasa ini. Penelitian ini menawarkan kombinasi metode yang dapat diskala sebagai pemecahan yang paling atraktif untuk masalah tersebut. Metode yang dapat diskala untuk komunikasi nirkawat seharusnya disesuaikan dengan runtun video input yang akan diproses. Standard amandemen ITU (International Telecommunication Union) - Joint Scalable Video Model (JSVM) dalam penelitian ini, dipergunakan untuk menghasilkan metode penyandian video yang dapat diskala (CSVC) yang sesuai dengan kualitas layanan streaming video yang diperlukan pada transmisi nirkawat. Hasil penelitian menunjukkan penggunaan teknik yang dapat diskala terkombinasi menghasilkan kinerja lebih baik dari yang tidak dapat diskala secara keseluruhan pada penggunaan kapasitas laju bit pada lapis tertentu.

Kata kunci: JSVM, pengkodean video dapat diskala terkombinasi, transmisi nirkawat, transmisi video

Abstract

Mobile video streaming is one of multimedia services that has developed very rapidly. Recently, bandwidth utilization for wireless transmission is the main problem in the field of multimedia communications. In this research, we offer a combination of scalable methods as the most attractive solution to this problem. Scalable method for wireless communication should adapt to input video sequence. Standard ITU (International Telecommunication Union) - Joint Scalable Video Model (JSVM) is employed to produce combined scalable video coding (CSVC) method that match the required quality of video streaming services for wireless transmission. The investigation in this paper shows that combined scalable technique outperforms the non-scalable one, in using bit rate capacity at certain layer.

Keywords: combined scalable video coding, JSVM, video transmission, wireless transmission

1. Introduction

The utilization of wireless network in multimedia communication has increased significantly in recent years. The needs of mobility and accuracy of information trigger higher bit rate utilization. Many researchers in information and communication fields had focused on this topic.

Wireless communication has low reliability due to bandwidth fluctuations that lead to degradation of video quality significantly. Video streaming needs wide bandwidth to transmit the video data but the bandwidth is limited in the wireless environment in early stage. One attractive method to overcome this problem is by scale bit stream into a number of scales of priority (layers), in the form of base-layer and enhancement layer, which known as scalable video coding (SVC). For this reason, The Joint Video Team (JVT) formed by the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) - Moving Picture Experts Group (MPEG) and The International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) develops the extension of H.264/Advanced Video Coding (AVC)-SVC, which provides different video quality according to the bit stream is received [1], [2].

The rest of this paper is organized as follows. Section 2 presents JSVM (Joint Scalable Video Model) standards, i.e. SVC (Scalable Video Coding) and combined scalable method. Wireless channel model is presented in section 3, followed by results and conclusions in section 4 and 5 respectively.

2. JSVM (Joint Scalable Video Model)

2.1 Scalable Video Coding (SVC)

SVC is part of standard development H.264/MPEG-4 part 10 AVC (Advanced Video Coding), or H.264/AVC [3]-[6]. The development processes took a long 20 years, starting from H.262 and MPEG-2, followed by H.263+ and MPEG-4. Until now, SVC standard is still in amendment, and is a cooperative work of many parties to establish JSVM standard. Since January 2005, MPEG and VCEG join in JVT to carry into completion the amendment H.264/AVC as an official standard [7]-[11].

Scalability has proposed for the first time to reduce packet (cell) loss in ATM networks [2]. It created 2 (two) groups of bit stream or layer, base-layer and enhancement-layer. Layer containing vital information is base-layer, while enhancement-layer loaded by residual information to improve image or video quality, as shown in Figure 1.

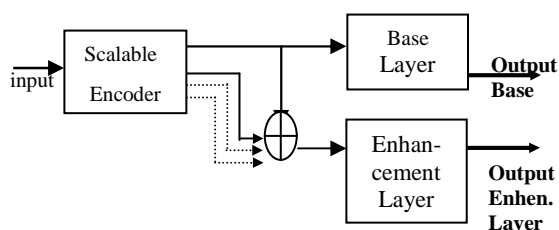


Figure 1. Scalability Coding System

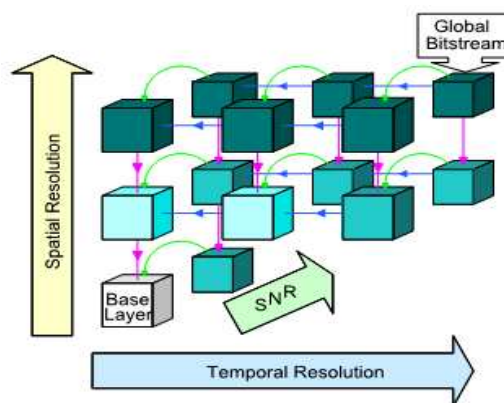


Figure 2. Concept of extracted link in combined scalable [14]

Within data transmission processes, in case of congestion in transmission channel, at least base-layer containing vital information still can be transmitted. There are three types of scalability method as in Figure 1: SNR (Signal to Noise Ratio) Scalability, Spatial Scalability, and Temporal Scalability. Beyond those, there is one type of scalability, which is combination the first three. This combined method is the main topic of this paper. This research will focus to the Combined Scalable Video Coding (CSVC) as the development of our previous work [12]. In this work, we compare a CSVC to non-SVC streaming framework based on JSVM version 9.8 [10].

2.2 Combined Scalability Video Coding (CSVC)

Implementation of the combined scalable coding is based on the structure and efficiency of SVC coding. In case that the inter layer resolution changed, spatial scalable will be dominant. SNR scalable is dominant in changes of SNR, and temporal scalability dominant in changes of rate [12]-[17]. Combination of those three is likely caused by varying in sequences characteristics, fluctuating network condition and multi terminal [11], [14], and [15].

This research utilizes combination of three scalable layers, which includes 1 base layer and 2 enhancement layers, as shown in Figure 2. Block diagram of encoder-decoder with 3 layers combined scalable is shown in Figure 3 [13], [14].

3. Wireless Transmission

In this research, we will use video transmission over wireless channel. Transmission channel is assumed to be slow and flat fading Rayleigh, with AWGN noise, detailed in Figure 4. Firstly, transmitted signal is multiplied by Rayleigh distribution amplitude factor, as fading effects, and then AWGN is added to the channel. Rayleigh fading source is generated by Jake method simulator [18].

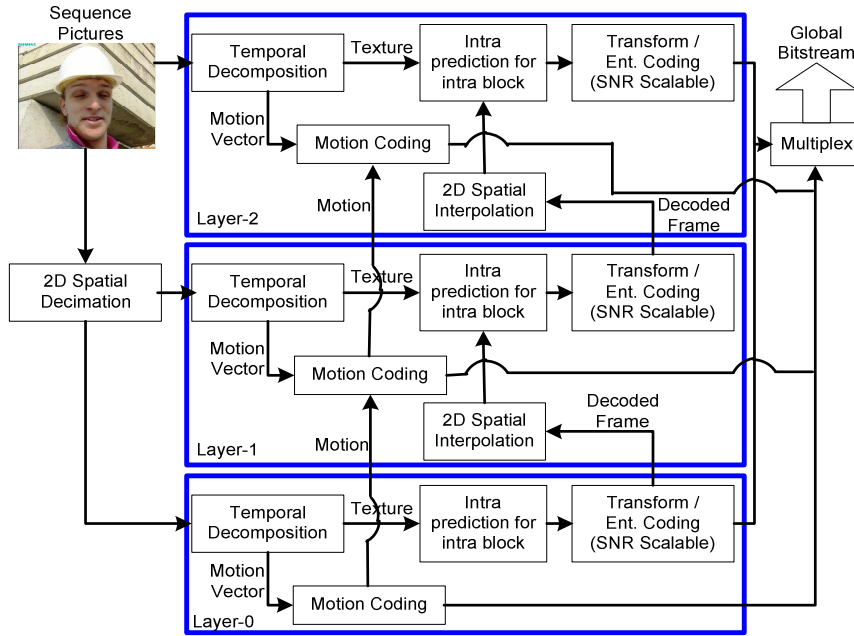


Figure 3. Structure of Encoder 3 layer SVC

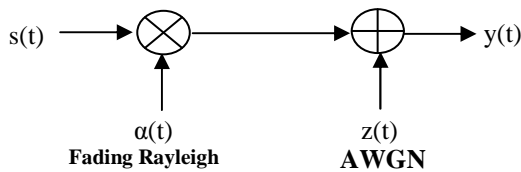


Figure 4. Wireless channel model

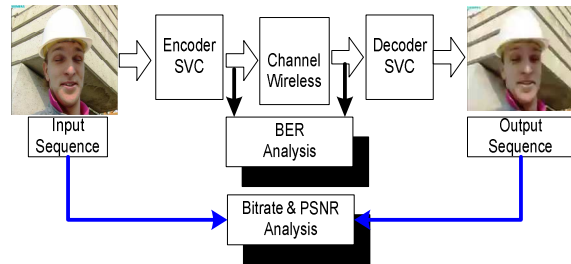


Figure 5. Research Configurations

The transmitted signal $S(t)$ is first multiplied by a factor of amplitude $\alpha(t)$ distribution of the effects of Rayleigh fading and white Gaussian noise $z(t)$ added to the channel. The wireless channel condition modeling is using the method of Jakes and the revision [19]. The receiver equivalent low pass signal in one signaling interval is

$$y(t) = \alpha e^{j\theta} s(t) + z(t) \tag{1}$$

The expression for the error rate of binary PSK (Phase-Sift Keying) as a function of the average probability density function (pdf) P_b of received SNR (γ_b) when α is random, defined as

$$P_b = \int_0^\infty P_b(\gamma_b) p(\gamma_b) d\gamma_b \tag{2}$$

When α is Rayleigh-distributed; then

$$p(\gamma_b) = \frac{1}{\gamma_b} \exp\left[-\frac{\gamma_b}{\gamma_b}\right]; \gamma_b \geq 0 \quad (3)$$

Where $\overline{\gamma_b}$ is the average SNR, defined as

$$\overline{\gamma_b} = \frac{\mathcal{E}_b}{N_0} E(\alpha^2) \quad (4)$$

The term $E(\alpha^2)$ is simply the average value of α^2 .

Distribution Rayleigh fading channel used in the analysis, i.e. to describe phenomena of small-scale fading. Rayleigh distribution is the resultant of random variables (α), with Gaussian distribution having zero mean, and the same standard deviation [20].

Wireless environment, especially in mobile multimedia communication services, has become main issue in many researches [21]-[24]. Almost all of them still use SVC in general. This research adopts some results from those previous works, with focus on impact on wireless channel from fading and AWGN in three layer combined scalable video transmission.

To the best of our knowledge, Scalable Video Coding scheme using the combined scalability video coding to enhance the quality of video streaming for wireless transmission systems has not been well studying yet.

4. Results and Analysis

4.1 Research Configurations

Experimental setup for simulation as in Figure 5, compute and analyzes BER, bit rate, and PSNR, which utilize separate software. In order to examine video sequences, both input sequences and output sequences, we use MFC YUVviewer software and VLC media player. Channel transmission analysis utilizes Matlab 6.5 source. Source code of JVSM is written in C++, run on Microsoft Visual Studio.Net 2005.

The configuration of combined scalable that we use are as follows: spatio-temporal scalable resolution and MGS scalability as scalable SNR ; 15 fps (frame per second) QCIF format picture on layer 1 (as base layer) ; 30 fps CIF on layer 2 (as enhancement layer 1) and 60 fps 4CIF on layer 3 (as enhancement layer 2). Table 1 lists general conditions and parameters of this research.

Table 1. Research's Conditions and Parameters

Parameters	Description
GOP size	16 frames
Spatial and Temporal Scalable	Spatio-temporal Scalable
SNR Scalability	MGS (Medium Grain Scalability)
Input Sequence	(1). Bus (150 frame); (2). Foreman (300 frame) (3). City (600 frame); (4). Soccer (600 frame)
QP	24, 32, 40
Motion Search Range	± 32
Number of Layer	3 layer (1 base, 2 enhancement)
CPU Test Platform	Intel® Core™2 Duo CPU@2.4GHz, 2 GB Main Memory

4.2 Pictures Quality Parameters

Peak Signal-to-Noise Ratio (PSNR) was been used to objectively measure the quality between an original sequence and reconstructed sequence. This metric depends on the Mean Squared Error (MSE) given by

$$MSE = \frac{1}{W.H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} |f(x,y) - g(x,y)|^2 \tag{5}$$

where W is number of pixel per row, H is number of row per frame, $f(x,y)$ is pixel's luminance intensity in the original frame, and $g(x,y)$ is pixel's luminance intensity in the reconstructed frame. PSNR is defined as

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \tag{6}$$

where n is number of bits per pixel. Bit rate estimator has computed by

$$B_r = \frac{N_b}{N_f} \times M_f \tag{7}$$

where B_r is bit rate, N_b is total bit, N_f is number of frame, and M_f is mean frame rate. All input videos sequence has analyzed by quantization parameters (QP) 24, 32 and 40.

4.3 Experimental Results and Analysis of the Bit rate and PSNR-Y

Observation carried out by comparing bit rate utilization in each layer from combined scalable with non-scalable layer. Simulation results show that:

- a. At input sequence BUS CIF 30 fps (Figure 6). Non-scalable layer compared to layer 2 results a close value of PSNR. Meanwhile, PSNR-Y compared to layer 3 has 0.5 to 2 dB of different at the same bit rate. Bit rate utilization in layer 3 clearly shows significant different (start from 96 kbps at PSNR-Y 27 dB to 1024 kbps at PSNR-Y 36 dB).
- b. Input sequence Foreman CIF 30 fps (Figure 7). Non-scalable layer compared to layer 2 and 3 at the same bit rate shows PSNR-Y value greater (0.5 dB to 2 dB). Meanwhile bit rate utilization in layer 1 shows significant different at 96 kbps at PSNR-Y 30 dB to 640 kbps at PSNR-Y 40 dB.

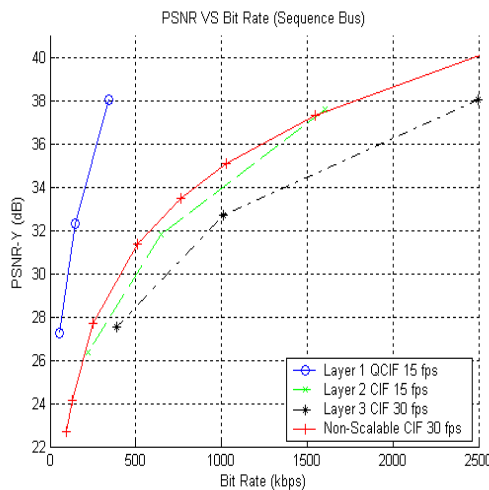


Figure 6. Graph of PSNR VS Bit Rate from Bus Sequence

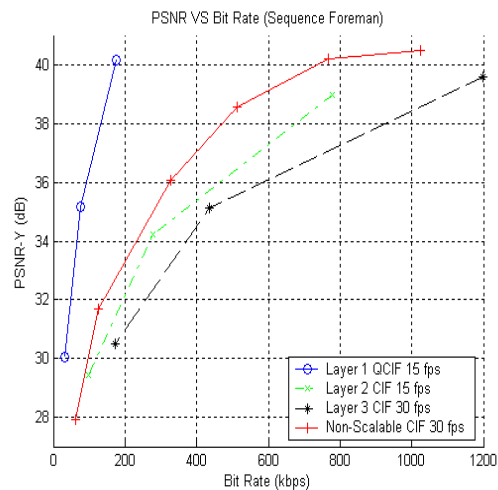


Figure 7. Graph of PSNR VS Bit Rate from Foreman Sequence

- c. At input sequences City 4CIF 60 fps (Figure 8), layer non-scalable has higher PSNR-Y value 1 to 2 dB than layer 3 at the same bit rate. Layer 1 and 2 clearly show that bit rate utilization

are significantly different, start from 384 kbps at PSNR-Y 33 dB to around 6000 kbps at PSNR-Y 38 dB.

- d. At input sequence Soccer 4CIF 60 fps (Figure 9) : comparison of non-scalable layer to layer 1 and 2 show significant different on bit rate utilization, that are around 256 kbps at PSNR-Y 32 dB to around 5000 kbps at PSNR-Y 39 dB. However, non-scalable layer has PSNR-Y value higher 1 dB to 2 dB than layer 3 at the same bit rate.

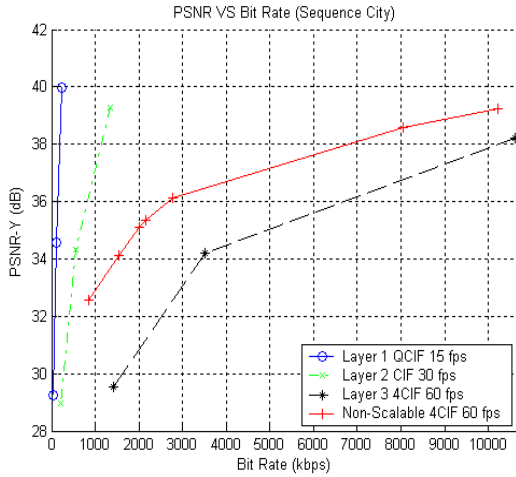


Figure 8. Graph of PSNR VS Bit Rate from City Sequence

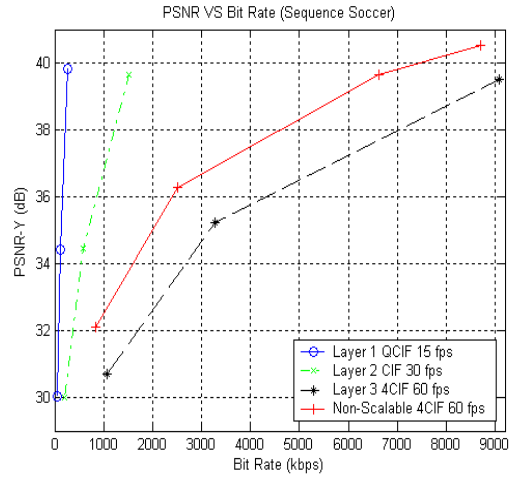


Figure 9. Graph of PSNR VS Bit Rate from Soccer Sequence

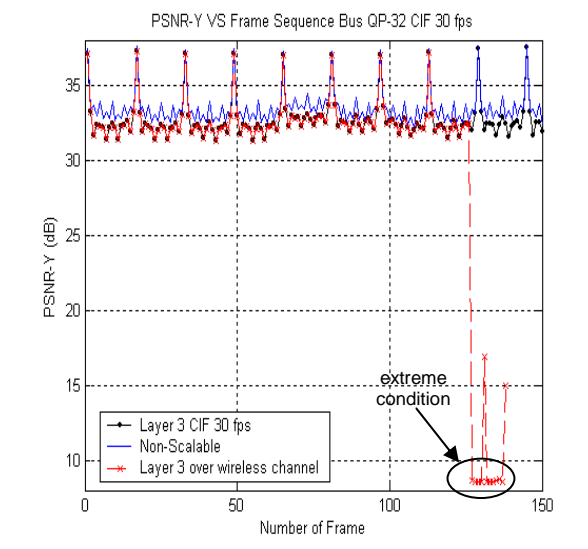


Figure 10. Graph of PSNR segmental from Bus sequence

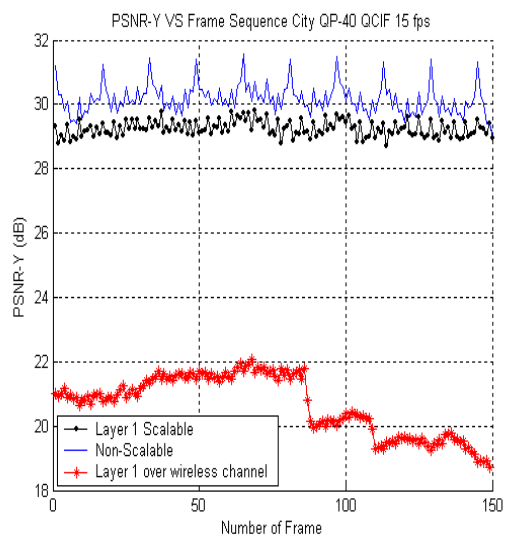


Figure 11. Graph of PSNR segmental from City sequence

The obtained results show that utilization of high input sequence (600 frames) and high input rate (60 fps) of City and Soccer sequence at layer 3 when compared to non-scalable layer has lower PSNR-Y (0.5 dB to 2 dB).

From the results above, we know that the use of CSVC on multicast system provides solutions with options to user to use layer 1 or 2 in case that layer 3 is not possible to be used because of error or bandwidth fluctuation within channel or networks. This will significantly

reduce connection failure even though signal degrades from flat fading and AWGN, which make it different to simulcast system (non-scalable). In other words, this will significantly lower connection failures.

4.4 Experimental Results and Analysis of the Impact on Wireless Channel

Impact of errors that come from wireless channel in bit stream of encoder output has compared to bit stream without error. It is shown in Figure 10 (Bus sequence) and Figure 11 (City sequence) that error effect degrades PSNR-Y, 9 dB (certain frame as an extreme condition show in after 125 frames in Figure 10) for Bus sequence and on average PSNR-Y is degrading 32.67 dB to 30.66 dB. Impact of wireless channel on sequence City (Figure 11) is significant, degrading PSNR-Y from 29.25 dB to 20.23 dB. Increasing the variance of AWGN value will increase the BER value, while the PSNR value will decrease. Table 2 shows that variance of AWGN from 0.01 to 0.001 plus Rayleigh fading channel degrades the average PSNR-Y (Figure 10 and Figure 11) from 2 dB to 10 dB. In some extreme case, reconstructed bit stream is fail, especially at AWGN variance 0.01.

Extreme condition as in Figure 12 is compared to normal condition (without error) as in Figure 13. Those two figures show the influence of error due to wireless channel as described in Figure 10 (Bus sequence) above.

Table 2. Results from impact of wireless channel

Input	AWGN Variance	Number of Bit Error	SNR (dB)	BER
BUS (QP 32)	0,01	89958	19.9978	0.0177
	0,001	27925	30.0001	0.0055
FOREMAN (QP 32)	0,01	78380	19.9985	0.0179
	0,001	24431	29.9980	0.0056
CITY (QP 40)	0,01	7455	20.0101	0.0180
	0,001	2418	29.9909	0.0058
SOCCER (QP 40)	0,01	7770	19.9909	0.0178
	0,001	2558	29.9972	0.0059



Figure 12. Extreme conditions of frame 135 in layer 3 Bus Sequence (error due to wireless channel)



Figure 13. Normal condition of frame 135 in layer 3 Bus Sequence (without error)

5. Conclusion

We have presented CSVC method for wireless transmission. We also investigated the impacts of the AWGN and Rayleigh fading channel on performance of CSVC. Our experiments illustrate the following in general combined scalable save bit rate in certain layer more than non-scalable. Layer 3 of combined scalable that has lower PSNR-Y value than non-scalable layer (mainly in City and Soccer sequence), range from 0.2 dB to 2 dB. Increasing in AWGN variance and Rayleigh fading channel degrade SNR and BER. Variance from 0.001 to 0.01 degrades PSNR-Y from 2 dB to 10 dB.

In the future, we are about to exploit wireless communication services using CSVC bit stream. The research will be simulating by Network Simulator II (NS-2). Transmission video sequence in data packet is using systems MIMO-OFDM over multipath fading.

Acknowledgment

The authors would like to thank Budi Setiawan (PT. INTI-Bandung) for his fruitful discussions about JSVM source-code and Dr. Adit Kurniawan (ITB-Bandung) for his discussions about fading channel. The authors gratefully acknowledge the team in Multimedia Communication Labs (B-304) of Electrical Engineering ITS Surabaya for the many insightful discussions and contributions. This work has been supported by a Postgraduate Scholarship from Directorate General of Higher Education (DGHE), Ministry of National Education, Republic of Indonesia.

References

- [1] Wu D, Hou YT, Zhang YQ. Scalable Video Coding and Transport over Broadband Wireless Networks. *Proceeding of the IEEE*. 2001; 89(1): 6-20.
- [2] Ghanbari M. Video Coding an Introduction to Standard Codec. London: IEE Telecommunications Series 42. 1999.
- [3] ITU-T Recommendation H.264. *Advanced video coding for generic audiovisual services*. Nopember, 2007.
- [4] Ostermann J. Video coding with H.264/AVC: Tools, Performance, and Complexity. *IEEE Circuits and Systems Magazine*. 2004: 7-28.
- [5] GJ Sulliva, T Wiegand. *Video Compression – from Concepts to the H.264/AVC Standard*. Proc. IEEE. 2005; 93(1): 18-31.
- [6] T Wiegand, GJ Sullivan. The H.264/AVC Video Coding Standard. *IEEE Signal Processing Mag*. 2007; 24(2): 148–153.
- [7] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. Joint Draft ITU-T Rec. H.264 ISO/IEC 14496-10 / Amd.3. *Scalable Video Coding*. Document: JVT-X201. 24th Meeting: Geneva, Switzerland. 2007.
- [8] Schwarz H, D Marve, T Wiegand. Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Transaction on Circuits and Syatems for Video Technology*. 2007; 17(9): 1103-1120.
- [9] Kouadio A, Clare M, Noblet L, Bottreau V. *SVC – a highly-scalable version of H.264/AVC*. EBU TECHNICAL REVIEW – Q2. 2008.
- [10] Wien M, H Schwarz, T Oelbaum. Performance Analysis of SVC. *IEEE Transaction on Circuits and Syatems for Video Technology*. 2007; 17(9): 1194-1203.
- [11] ITU-T and ISO/IEC (JVT). *Joint Scalable Video Model – JSVM (Version 9.8)*. 2006. (<http://ftp3.itu.int/av-arch/jvt-site>).
- [12] Rantelobo K, Hendrawan, S Tjondronegoro. *Analysis Performance of Video Coding H.263+ with Temporal Scalability*. Proceeding of Indonesia Conference on Telecommunications. Bandung. 2005.
- [13] Wien M, Schwarz H. *Testing Conditions for SVC Coding Efficiency and JSVM Performance Evaluation*. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Document: JVT-Q205, 16th Meeting: Poznan, PO. 2005.
- [14] Schwarz H, Schierl DMT, T Wiegand. *Combined Scalability Support for the Scalable Extension of H.264/AVC*. Proc. IEEE Int. Conference on Multimedia and Expo. 2005.
- [15] Schaar MVD, Radha H. A Hybrid Temporal-SNR Fine-Granular Scalability for Internet Video. *IEEE Transactions on Circuits and Syst. for Video Technol*. 2001; 11(3).
- [16] Błaszak Ł, Domański M, Maćkowiak S. *Spatio-Temporal Scalability in AVC codecs*. ISO/IEC JTC1/SC29/WG11, MPEG2003/M9469. Pattaya. 2003.
- [17] JR Ohm. *Advances in scalable video coding*. Proc. IEEE. 2005; 93(1): 42–56.
- [18] William CJ Jr. *Microwave Mobile Communications*. North Andover Massachusetts: Bell Telephone Laboratories, John Wiley & Sons. 1974.
- [19] P Dent, GE Bottomley, T Croft. Jakes fading model revisited. *IEEE Electronics Letters*. 1993; 29(13): 1162–1163.
- [20] Proakis JG, Salehi M. *Digital Communications*. Fifth Edition. NY: McGraw-Hill. 2008: 846-849.
- [21] T Schierl, T Wiegand. Mobile Video Transmission using SVC. *IEEE Trans. Circuits Syst. Video Technol*. 2007; 17(9): 1204-1217.
- [22] Wang Y, Chau LP, Yap KH. *Broadcast of scalable video over wireless networks*. IEEE International Symposium on Circuits and Systems (ISCAS). Taipei. 2009.
- [23] Bansal M, Kondi MJ. *Scalable H.264 Wireless Video Transmission over MIMO-OFDM Channels*. Proc. International Conference ACIVS 12th. Sydney, Australia. 2010.
- [24] Shiddiqi AM, Pratama P, Ciptaningtyas HT. Video Streaming Application Using Mobile Media Application Programming Interface. *TELKOMNIKA: Indonesian Journal of Electrical Engineering*. 2010; 8(3): 293-300.