論文 96-21-5-5

Adaptive Rate Control Scheme for Very Low Bit Rate Video Coding

Hwang-Seok Oh*, Heung-Kyu Lee*, Joon Hyeon Jeon** *Regular Members*

초저속 전송 매체용 비디오 코딩을 위한 적응적 비트율 제어에 관한 연구

正會員 오 황 석*, 이 흥 규*, 전 준 현**

ABSTRACT

In video coding systems, an effective rate control method is one of the most important issues for the good video quality. This paper presents an adaptive rate control scheme based on buffer fullness, quantization, and buffer utilization for very low bit rate communication lines, such as 16kbit/s, 24kbit/s, and so on. The strategy is implemented on H.263, which is a video coding algorithm for narrow band telecommunication channels up to 64kbit/s recommended by ITU-T SG15, to show the effectiveness. The simulation result shows that the suggested rate control scheme has better SNR performance and buffer utilization of source coder than those of linear and non-linear[9] buffer control strategies.

Keyword: video coding, very low bit rate, rate control

요 약

비디오 부호화 시스템에서 효율적인 비트율 제어 기법은 좋은 화질을 얻기 위한 매우 중요한 요소중의 하나이다. 본 논문에서는 16 kbit/s, 24 kbit/s 등과 같은 전송율을 갖는 초저속 전송 매체를 위해 버퍼의 상태, 양자화 그리고 버퍼 이용율을 기반으로 하는 적응적 비트율 제어 방식을 제안한다. 제안한 기법의 효율성을 평가하기 위해서 ITU-T SG15에 의하여 제안된 초저속 전송 매체용 비디오 부호화 알고리즘인 H.263에 제안한 기법을 구현하였다. 실험결과는 제안된 비트율 제어 기법이 SNR 성능 측면과 버퍼 이용율 측면에서 선형 기법과 [9]가 제안한 비선형 기법보다 좋음을 보여준다.

I. Introduction

One of the important technologies to guarantee constant video quality in audio-visual communication applications such as video-telephony, video-conferencing through the public switched telephone networks or mobile communication lines is the rate control method. Because decoded video quality depends on

* 한국과학기술원
** 한국통신

論文番號: 95382-1106
接受日字: 1995年 11月 6日

the coding parameters which are controlled by the rate control method, an effective bit rate control method is one of the most important issues in video coding systems for the transmission through the very low bit rate communication lines with limited transmission rate such as 16kbit/s, 24kbit/s, and so on. In video coding system, each of video frames generates different amount of data according to its activity, so it is necessary for the rate control schemes to handle the bit rate generated by source coder to preserve constant image quality[8].

In general, rate control scheme can be classified into two types[6], which are forward rate control and backward rate control method. In video coding system with forward rate control, the quantization parameters and all of the rate control variables are determined by examining input image's activity such as variance, number of object, etc. On the other hand, in backward rate control scheme, bandwidth of transmission line, buffer size, and amount of generated data for previous frames are critical factors for the decision of the coding parameters.

In this paper, we present an adaptive rate control scheme which determines the coding parameters such as quantization parameter, threshold value for block classifications based on buffer status. The buffer status is modeled as a non-linear system to get high adaptivity to input sequence in accordance with previous history, amount of bit generated in current frame, scene changes and other circumstances. The presented scheme is implemented on H.263 video coding algorithm suggested by ITU-T SG15 to show the effectiveness.

The paper is organized as follows. The overview of narrow band video coding algorithm H.263 is described in Section 2. In Section 3, we explain a new adaptive buffer control scheme using buffer_fullness, quantization parameter, and block classification information in some detail. Then experimental results are given in Section 4 to show the effectiveness of the presented strategy. Finally, we draw conclusions with a few future works in Section 5.

Ⅱ. Very Low Bit Rate Video Codec H.263

The draft recommendation H.263 is designed for compressing the moving picture component of audio-visual services at very low bit rate up to 64kbit/s[10]. The generalized block diagram of H. 263 source coder is shown in Figure 1. A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The main elements are prediction, block transform, quantization, and variable length coding.

The video multiplex is arranged in a hierarchical

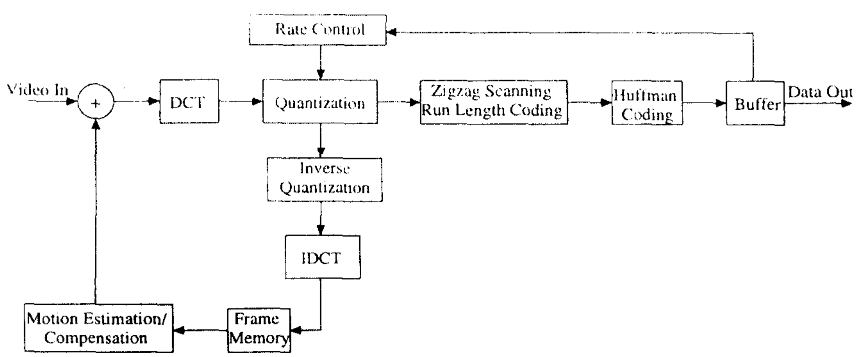

Figure. 1 Block diagram of H.263 source coder

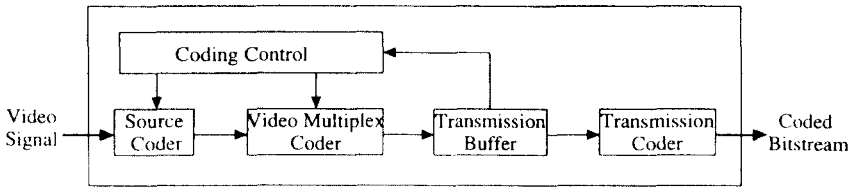

Figure. 2 Simple buffer control scheme in H.263

structure with four layers. From top to bottom the layers are picture, group of blocks, macroblock, block [10]. Each picture is divided into macroblocks. For each macroblock, the coder finds best matched macroblock by minimizing the displaced block difference (DBD) to eliminate temporal redundancy. According to the result of prediction the coder decides which coding modes, intra-or inter-mode, should be used. If small DBD has occurred the macroblock is coded with inter-mode which transmits coded stream of the displaced vectors and DBD. Otherwise macroblock is coded by intra-mode which transmits block information. DBD and block information are transformed by discrete cosine transform(DCT), quantized, and then coded by variable length coder(VLC), in inter- and intra-mode, respectively.

For realistic simulations of video codec with limited buffer and coding delay, a buffer regulation is needed. Several parameters may be varied to control the rate of generation of coded video data. These include processing prior to the source coder, the quantizer, block significance criterion, and temporal subsampling. The conventional buffer control strategy is characterized in Figure 2[9].

To regulate the output bit rate, H.263 uses the following picture quantization parameters as initial, but for now do not specify any specific buffer regulation method as a recommendation[10].

$$QP_i = \overline{QP}_{i-1}\left(1 + \frac{\Delta_1 B}{2\bar{B}} + \frac{12\Delta_2 B}{R}\right) \tag{1}$$

$$\Delta_1 B = B_{i-1} - \bar{B} \tag{2}$$

$$\Delta_2 B = B_{i,mb} - \frac{mb}{MB}\bar{B} \tag{3}$$

where QP_i is the quantization parameter for i^{th} frame, $\overline{QP}_{i-1}$ is the mean value of the quantization parameter for previous picture, B_{i-1} is the number of bits spent for the previous picture, $\bar{B}$ is the target number of bits per picture, mb is present macroblock number, MB is the number of macroblocks in a picture, $B_{i,mb}$ is the number of bits spent until now for the picture, and R means bit rate. The first two terms of this formula are constant for all macroblocks within a picture. The third term adjusts the quantization parameter during coding of the picture. So Equation (2) is characterized as a linear rate control scheme to determine the quantization parameter QP for each macroblock, that is $QP = QP_{i-1} \times f$(*rate of bit generated from previous picture to current macroblock*). For this buffer regulation, it is assumed that the process of encoding is temporarily stopped when the physical transmission buffer is nearly full. This means that buffer overflow and forced-to-fixed blocks will not occur. However, this also means that no minimum frame rate and delay can be guaranteed[10].

In this paper, we consider linear control and nonlinear rate control schemes characterized as Equation (4) and (5) to compare the effectiveness of proposed scheme at the next section. In linear rate control strategy, quantization parameter QP is determined as follows

$$QP(n) = b(n-1) \tag{4}$$

where $QP(n)$ is the normalized quantization parameter of the n^{th} macroblock and $b(n-1)$ is the normalized rate buffer occupancy at the $(n-1)^{th}$ macroblock. On the other hand, [9] suggested the non-linear rate control scheme using a nonlinear relationship between the buffer occupancy and quantization parameter. In this scheme, quantization parameter QP is determined as follows

$$QP(n) = \begin{cases} \alpha(\alpha^{-1} b(n-1))^k & \text{if } 0 \le b(n-1) < \alpha \\ 1-(1-\alpha)((1-\alpha)^{-1}(1-b(n-1)))^k & \text{if } \alpha \le b(n-1) \le 1 \end{cases} \tag{5}$$

where $QP(n)$ and $b(n-1)$ are the same as described in Equation (4). If $k=1$, it is equivalent to the linear rate control scheme. In these two methods, the quantization parameter QP is decided by only buffer occupancy without any consideration of input video characteristics when the constant α and k are fixed.

Ⅲ. Adaptive Rate Control Scheme

We present more adaptive rate control scheme based on buffer_fullness and quantization parameter for very low bit rate video coding applications. In our method two parameters are controlled. They are quantization parameter QP and block classification information COD that represents whether a macroblock is coded or not. In the area of very low bit rate video communication with limited physical buffer size, the significant problem for rate control is a buffer overflow which causes data loss or great coding delay. So it is very important to predict the amount of data exactly when a macroblock is coded in given quantization parameter. Because a good prediction of data amount to be generated makes source coder know the best matched quantization parameter QP with considering the relationship between QP and amount of bits at that QP, it is possible to use physical buffer effectively by taking QP and COD values without buffer overflow and underflow. Thus, in presented scheme, we use the history of relationship between quantization parameter and amount of bit generated at that quantization value to get the most proper QP. The quantization parameter QP is calculated as Equation (6).

$$\begin{aligned} &QP(n) = \min\{QP \mid C(n-1) + f(QP) - \overline{mb} < BS \times \mu\} \\ &C(n) = C(n-1) + B(n) - \overline{mb} \end{aligned} \tag{6}$$

where $QP(n)$ is the quantization parameter for n^{th} macroblock, $C(n-1)$ is the amount of buffer content until $(n-1)^{th}$ macroblock, $\overline{mb}$ is the mean of out going bit rate to channel for a macroblock, BS means physical buffer size, μ means buffer utilization factors, and $f(QP)$ denotes the amount of bit to be generated, which is described in QP-Bitrate table, and $B(n)$ means number of bits generated by source coder for n^{th} macroblock. In Equation (6), QP-Bitrate Table is composed of pairs of QP value and bit rate as shown in Table 1. The amount of bits in QP-Bitrate table is changed with the most recently used QP and the related amount of bit generated during the video coding, but QP is fixed as initial. This table gives more adaptable QP than those of linear, non-linear methods which consider only QP and buffer status. That is, QP-Bitrate table is adaptively changed by consideration of the characteristics of video sequence.

Table 1. An example of QP-Bitrate table

QP value	Amount of Bits
1	545
2	480
⋮	⋮
31	20

Since the source coder generates different amount of bits according to coding modes such as intra, inter in motion compensated hybrid video coding methods, it is necessary to use separate QP-Bitrate table for each ones. With the same rate control scheme written in Equation (6), we can make QP-Bitrate tables for different coding modes to predict accurately bit amounts to be generated. The formula described above says that the QP for current macroblock is determined by considering the prediction of relationship between QP and bit rates and buffer utilization whose range varies from 0 to 1. As buffer utilization comes close to 1, the source coder has lower QP but buffer overflow would be occurred more probably.

To get more adaptivity to physical buffer size, we use the block classification information(or coding type of block) COD which tells whether the macroblock will be coded or not. The COD is used for only inter-mode macroblock. In very low bit rate video coding, more than three quarters of macroblocks must be coded by inter-mode to meet the bit rate. So if we use the COD well we can get good video quality without great loss of degradation. In presented scheme, only buffer fullness is considered to decide COD. COD is determined as Equation (7).

$$COD = \begin{cases} 0 & \text{if } (\mathit{buffer\ fullness} < \mathit{threshold}) \\ 1 & \mathit{otherwise} \end{cases} \tag{7}$$

where buffer fullness is defined as $\frac{remaining\ fuffer\ size}{total\ buffer\ size}$,

value 0 means that the macroblock must be coded and value 1 means that only COD value must be transmitted without coded information of macroblock. If a macroblock is coded with 'not coded' mode, the content of the macroblock at the same position in the last picture is copied into the present macroblock. So we can reduce the amount of bit generated by source coder. The block classification information is useful to avoid the buffer overflow in case of wrong prediction of QP at QP-Bitrate table.

Ⅳ. Simulation Results

In this section, we show the experimental results to evaluate our rate control scheme which is implemented on H.263. As a distortion measure, we take the PSNR, which is defined as

$$PSNR = 10 \cdot \log\left(\frac{255^2}{MSE}\right) \tag{8}$$

where MSE is the mean square error of the compressed image[4].

In simulation, the Miss America, Carphone, Foreman and Sales man video sequences are used to show the effectiveness of the suggested rate control scheme. The format of test sequences is 4:2:0 QCIF which consists of Y, Cb, Cr with size of 176 pels/line, 144 lines/frame for Y, a quarter of Y for Cb, Cr. The frame rate of test sequences is 5 frames/sec. The simulation is performed with the 2kbits, 4kbits size of physical buffer and 16kbit/s, 32kbit/s transmission rate in three method, respectively. And we take buffer utilization factor $\mu = 0.6$ and threshold value $threshold = 0.8$ for COD in proposed scheme. To make use of the characteristics of each coding modes we use two QP-Bitrate table for intra and inter modes. At each table, the QP and Amount of Bitrate are changed based on generated bitrate at the corresponding modes.

The physical buffer size is closely related the coding delay. If we use the unlimited buffer, we can not guarantee the delay. In simulation the delay is less than (buffer size/transmission rate) at worst case. Because we set the average buffer utilization at 0.6, the average coding delay for each macroblock is $\frac{3}{40}$ at each simulation conditions.

The simulation results are given in Table 2. At each test sequence, the proposed scheme has better PSNR performance by 2dB than those of any other two methods at 16kbit/s and 32kbit/s. The buffer utilizations of linear and non-linear methods are dependent on the physical buffer size and the characteristics of test sequences. But in proposed scheme, buffer utilization is near to initial buffer utilization factor μ. That is to say, the proposed scheme has good adaptability to video sequences. Table 3 shows the relation between buffer utilization and buffer overflow. We find out as

Table 2. Average PSNR and buffer utilization(B-utl) for test video sequences when we adapt the conventional linear rate control, non-linear rate control suggested by [9] and proposed rate control schemes at the rate of 16kbit/s, 32kbit/s, respectively

Bit Rate	Schemes	Miss America		Carphone		Foreman		Salesman	
		PSNR	B-utl	PSNR	B-utl	PSNR	B-utl	PSNR	B-utl
16kbit/s	linear	38.48	0.20	30.94	0.52	30.48	0.53	32.32	0.28
	nonlinear	38.11	0.37	30.80	0.53	29.46	0.51	32.12	0.40
	proposed	42.59	0.58	32.92	0.61	32.34	0.63	34.91	0.60
32kbit/s	linear	40.43	0.13	32.93	0.38	31.41	0.46	34.92	0.18
	nonlinear	40.56	0.31	32.50	0.45	31.86	0.49	34.79	0.34
	proposed	43.80	0.59	34.18	0.59	34.53	0.60	36.80	0.57

Table 3. Trade-off between buffer utilization(B-utl) and number of overflowed macroblock(NOVM)

Bit rate	B-utl	Miss America		Carphone		Foreman		Salesman	
		PSNR	NOVM	PSNR	NOVM	PSNR	NOVM	PSNR	NOVM
16kbit/s	0.1	37.07	0	27.79	0	27.75	205	30.77	0
	0.2	35.86	0	27.86	0	27.99	247	31.35	0
	0.3	35.95	0	27.86	0	28.07	247	31.61	0
	0.4	36.55	0	28.06	0	28.31	245	31.76	0
	0.5	37.22	4	28.36	0	31.32	246	32.70	4
	0.6	42.59	43	32.92	0	32.34	247	34.91	43
	0.7	44.76	362	33.14	105	35.85	385	33.47	362
	0.8	43.74	462	32.03	1121	34.11	1485	33.82	462
	0.9	41.74	2590	31.78	2531	34.13	1382	34.38	2590
32kbit/s	0.1	36.94	0	29.15	0	27.21	0	31.48	0
	0.2	37.18	0	29.22	0	27.21	0	33.20	0
	0.3	36.83	0	29.19	0	27.20	0	32.91	0
	0.4	37.58	0	29.18	0	28.51	0	32.83	0
	0.5	42.13	0	29.24	0	30.32	0	32.81	0
	0.6	43.80	0	34.18	0	34.53	0	36.80	0
	0.7	44.69	7	35.51	15	36.58	58	37.82	7
	0.8	43.07	218	34.18	501	37.42	531	36.04	218
	0.9	42.57	328	33.37	488	36.56	764	36.67	328

the buffer utilization factor comes close to 1, buffer overflow will be occur more frequently. In case of buffer overflowing, we discard the macroblock and we copy a macroblock from the previous frame at the same position to reconstructthe broken one.

In Figure 3, 4, and 5, we depict PSNR and amount of bits generated at each frame of Carphone video sequence to show some more detailed results of the schemes at 16kbit/s, 2kbits buffer size. We find that the PSNR performance is much higher and data generation is more regular with the suggested strategy. In other test video sequences, transmission bit rates and buffer sizes, similar results are obtained. Figure 5 shows the buffer status of three rate control schemes sampled from block number 100 to 500. The plotted points in below 0 and above 1 of buffer_ fullness means that buffer underflow and buffer overflow have occurred, respectively. In the figure, the proposed rate control scheme has better buffer utilization without buffer overflow and underflow than those of any other schemes. In other words, the proposed strategy predicts the amount of bits to be generated accurately. Thus, a good video quality can be obtained by taking appropriate QP.

With the proposed scheme, we can get encouraging results for very low bit rate video coding algorithm with video telephony image sequences and control the bit rate to be generated more adaptively.

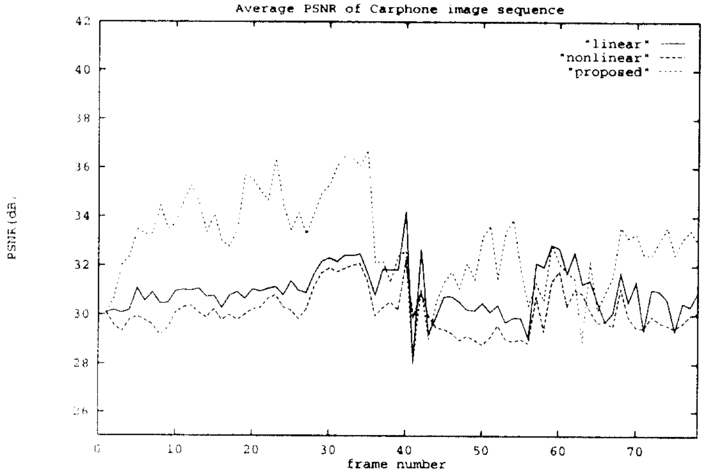

Figure. 3 The PSNR of Carphone sequence along the frame number at 16kbit/s transmission rate

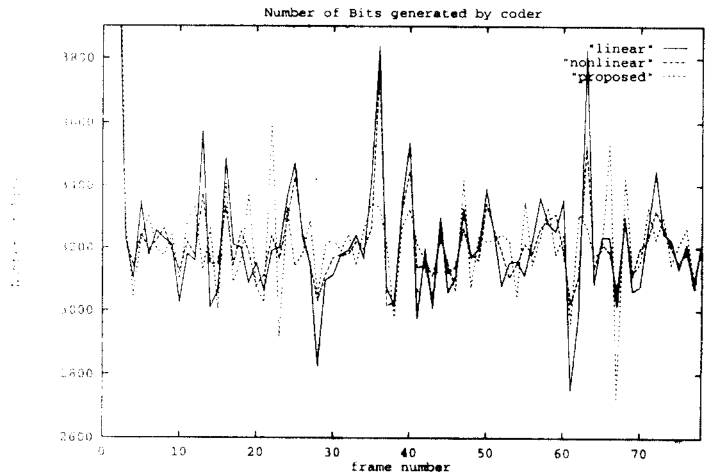

Figure. 4 The number of generated bits for Carphone sequence along the frame number

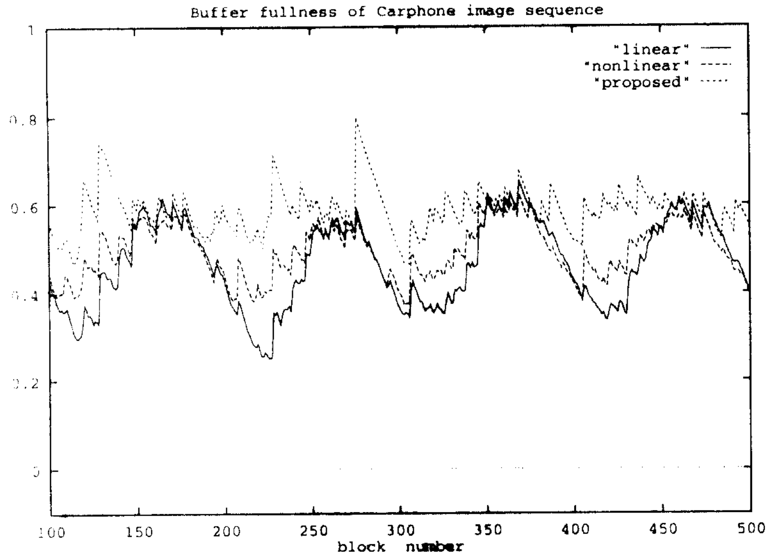

Figure. 5 The Comparison of buffer status between conventional linear, non-linear and proposed schemes at 16kbit/s, 2kbits buffer size

V. Conclusion

In this paper, we present a new adaptive rate control method for H.263. To regulate the amount of bit generated by source coder with constant image quality, we use the statistics of macroblock, buffer_fullness, and buffer utilization factor. The new method results in more acceptable picture quality of test sequences such as Miss America, Carphone, Foreman and Salesman with more regular generation of data compared with the results of linear, non-linear rate control schemes. The potential future work in development of more adaptable rate control method is by using of the n^{th}-order markov model to find out scene change in video sequence.

References

1. H.Gharavi and M. Mills, "Block matching motion estimation algorithm-new results," *IEEE Trans. on Circuits and Systems*, vol. 37, pp. 649-651, May 1990.
2. D.~L. Gall, "MPEG:A video compression standard for multimedia application," *Communication of the ACM*, vol. 34, pp. 46-58, Apr. 1991.
3. M. Liou, "Overview of the p×64 kbit/s video coding standard," *Communication of the ACM*, vol. 34, pp. 59-63, Apr. 1991.
4. M. Rabbani and P. W. jones, *Digital Image Compression Techniques*, SPIE-Optical Engineering Press, 1991.
5. W. Tawbi, F. Horn, E. Horlait, and J.-B. Stefani, "Video Compression Standards and Quality of Service," *The Computer Journal*, vol. 36, no. 1, pp. 43-54, Jan. 1993.
6. J. Kim, J. Nam, J. Jeon, L. Kim, D. Yoon, and J. Jeong, *ISO/IEC MPEG and CCITT ATM Video Coding Algorithms*, KAIST, 1993.
7. Y. Takishima, M. Wada, and H. Murakami, "An Analysis of Optimal Frame Rate is Low Bit Rate Video Coding," *IEICE Trans. on Communicaiton*, vol. E76-B, no. 11, pp. 1389-1396, Nov. 1993.

8. H. Li, A. Lundmark, and R. Forchheimer, "Image Sequence Coding at Very Low Bitrate: A Review," *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 589-609, Sep. 1994.
9. Y. Kim and H. Kim, "A New Buffer Control Strategy for Image Data Compression," *IEEE Trans. on Consumer Electronics*, vol. 40, pp. 932-937, Nov. 1994.
10. R.Schaphorst and Rapporteur, "Draft Recommendation H.263(Video coding for low bitrate communication)," *ITU-T SG15 WP15/1*, Jul. 1995.

오 황 석(Hwang-Seok Oh) 정회원
1969년 11월 17일생
1992년 2월: 경북대학교 자연과학대학 전자계산학과 졸업(학사)
1994년 2월: 한국과학기술원 전산학과 졸업(공학석사)
1994년 3월~현재: 한국과학기술원 전산학과 박사과정
주관심분야: Video coding, Parallel processing for video coding, Multimedia service system

이 흥 규(Heung-Kyu Lee) 정회원
1955년 10월 3일생
1978년 2월: 서울대학교 전자공학과 졸업(학사)
1981년 2월: 한국과학원 전산학과 졸업(이학석사)
1984년 8월: 한국과학기술원 전산학과 졸업(공학박사)
1985년 3월~1986년 8월: U. of Michigan, U.S.A (Research Scientist)
1986년 9월~현재: 한국과학기술원 전산학과 교수(부교수)
1991년~현재: 공업진흥청 JTC1/SC29(전문위원)
1994년~현재: ITU-T SG9(VOD 연구그룹 의장)
주관심분야: Real-time processing, Fault-tolerant system, Multimedia service system, Real-time operating system

전 준 현(Joon Hyeon Jeon) 정회원
1984년 2월: 동국대학교 전자공학과 학사
1986년 2월: 한국과학기술원 전기 및 전자공학과 석사
1991년 8월: 한국과학기술원 전기 및 전자공학과 박사
1991년 9월~현재: 한국통신연구개발본부 멀티미디어연구소(선임연구원)