

논문 2007-44SD-11-18

저 비트율 장면 전환 영상에 대한 향상된 H.264 프레임 단위 데이터율 제어 알고리즘

(Frame-Layer H.264 Rate Control for Scene-Change Video at Low Bit Rate)

이 창 현*, 정 윤 호**, 김 재 석**

(Changhyun Lee, Yunho Jung, and Jaeseok Kim)

요 약

급 장면 전환된 영상 프레임은 이전 프레임들과의 상관관계가 거의 없다. 이러한 경우에 인터 코딩 (inter coding) 보다는 인트라 코딩이 (intra coding) 영상 왜곡이 작으므로 프레임 내 거의 모든 매크로블럭이 인트라코딩을 수행한다. 이는 기존의 데이터율 제어의 흐름을 깨뜨리고 코딩에 필요한 비트수를 증가시킨다. H.264 참조 소프트웨어는 장면 전환을 고려치 않았기 때문에 이차 율-왜곡 모델을 사용하여 이러한 문제를 해결하려는 여러 연구가 있었다. 하지만 이차 율-왜곡 모델은 인터 코딩에 적합하게 설계되어 급장면 전환된 인트라 프레임의 양자화 파라미터를 (quantization parameter) 계산하기에는 적절치 않다. 이 논문에서는 급장면 전환시 인트라 코딩의 특징을 고려한 데이터율 제어 알고리즘을 제안한다. 제안된 알고리즘은 범용적으로 많이 사용하는 16개의 테스트 영상에서 검증하였으며, 기존 H.264 알고리즘 대비 PSNR 측면에서는 평균 0.4-0.6 dB, 최대 1.1-1.6 dB 향상되었으며, 프레임에 따른 PSNR 변화 측면에서는 평균 18.6 % 향상되었다.

Abstract

An abrupt scene-change frame is one that is hardly correlated with the previous frames. In that case, because an intra-coded frame has less distortion than an inter-coded one, almost all macroblocks are encoded in intra mode. This breaks up the rate control flow and increases the number of bits used. Since the reference software for H.264 takes no special action for a scene-change frame, several studies have been conducted to solve the problem using the quadratic R-D model. However, since this model is more suitable for inter frames, the existing schemes are unsuitable for computing the QP of the scene-change intra frame. In this paper, an improved rate control scheme accounting for the characteristics of intra coding is proposed for scene-change frames. The proposed scheme was validated using 16 test sequences. The results showed that the proposed scheme performed better than the existing H.264 rate control schemes. The PSNR was improved by an average of 0.4-0.6 dB and a maximum of 1.1-1.6 dB. The PSNR fluctuation was also improved by an average of 18.6 %.

Keywords : Abrupt scene transition, H.264, Intra coding, Rate control.

I. 서 론

20세기 말부터 멀티미디어 트래픽, 특히 비디오 스트리밍에 대한 트래픽은 꾸준히 증가추세에 있다. 따라서

고품질, 저 비트율 비디오 통신은 일상 생활에서 중요시 되고 있다. 현존하는 최고의 비디오 코덱인 H.264/AVC (advanced video coding)는^[1] ITU-T 비디오 코딩 전문가 그룹 (VCEG: video coding experts group) 과 ISO/IEC 14496-10 AVC 동영상 전문가 그룹 (MPEG: moving picture experts group) 이 공동으로 만든 것이다. H.264/AVC 는 기존의 MPEG-2, H.263, MPEG-4 대비, 평균적으로 각각 64%, 49%, 39% 정도 비트율을 감소시키는 성능을 나타낸다.^[2] 이러한 향상은 주로 arithmetic entropy coding, multiple intra

* 학생회원, ** 정회원, 연세대학교 전기전자공학과
(Department of Electrical and Electronic Engineering Yonsei University)

※ 본 연구는 정보통신부 및 정보통신연구진흥원 대학 IT연구센터 육성·지원사업 및 삼성전자 정보통신총괄에 의해 지원되었음.

접수일자: 2007년9월12일, 수정완료일: 2007년11월1일

prediction modes, motion estimation (ME) with variable block sizes, quarter-pixel accuracy, multiple reference frames 등의 기능에 기인한다. H.26/AVC 코덱은 그 뛰어난 성능으로 전자업계의 관심을 받아 왔으며, 디지털 TV 시청이 가능한 이동 휴대 단말기등에 비디오 코덱으로 사용된다.^[3]

비디오 스트림이 유무선 네트워크 망을 통해 전송될 때, 가용할 수 있는 채널 대역폭과 버퍼 제약에 따른 데이터를 제어가 필요하다. 데이터율 제어는 각 프레임에서 사용할 비트수를 예상하여 그에 적합한 양자화 파라미터 (QP: quantization parameter)를 계산한다. 효율적인 데이터율 제어 알고리즘은 화질이 좋고, 프레임에 따른 화질의 변화가 적어야 하며 설정된 비트율과 인코딩된 비트율의 차이가 작아야 한다.

H.264/AVC 참조 소프트웨어는 인터 프레임에서 인트라 코딩을 하는 기능이 있다. 장면 전환 발생시, 해당 프레임은 이전 프레임들과 상관관계가 줄어든다. 특히 한 장면에서 다음 장면으로 갑작스럽게 전환되면 그 상관관계는 거의 없어지게 된다. (이후 이러한 장면전환을 급장면전환이라 칭한다.) 이러한 경우에 인트라 코딩된 프레임의 영상 왜곡이 인터 코딩된 프레임보다 작으므로 거의 모든 매크로블럭이 인트라 모드로 코딩된다. 급장면 전환 프레임이 GOP (group of picture) 내 첫 번째 프레임이 아니라면, 인터코딩이 인트라코딩으로 바뀐 것은 데이터율 제어 알고리즘의 흐름을 깨뜨리고 인코딩시 사용된 비트수를 증가시킨다. H.264/AVC 참조 소프트웨어는 급장면전환 프레임에 대해 어떠한 조치도 취하지 않으므로 영상의 화질을 열화시키고 프레임에 따른 영상화질의 변화를 심화시킨다. 특히, 비트율이 낮으면 프레임당 가용 비트수가 작아 비디오 성능은 더욱 나빠진다. 따라서 낮은 비트율이 사용될 때 (76.8 kbps 이하) 급장면 전환시 비디오 성능을 향상시키는 적절한 알고리즘이 필요하다.

데이터율 제어에 관한 연구의 관점은 대다수 프레임당 적절한 비트수 예측에 있다. (이를 목표 비트수라 칭한다.) 최근의 연구는 콘텐츠 복잡도 측면으로 접근하는 경향이다.^[4-8] 논문 [4]에서는 목표 비트수를 정확히 예측하기 위해 근사화된 헤더 비트를 이용했다. 하지만 사전 분석에 (pre-analysis) 추가적인 계산이 필요하다. 논문 [5], [6]에서는 미리 정해진 양자화 파라미터로 ME 과정을 통해 각 매크로블럭의 복잡도에 관한 정보를 얻은 뒤 정확한 양자화 파라미터를 계산한다. 큰 움직임과 급장면 전환이 있는 비디오 영상에 적용하기 위

해, 논문 [7]에서는 PSNR (peak signal to noise ratio)을 기반으로 프레임 복잡도를 측정하였으며 논문 [8]에서는 복잡도 예측 방법으로 정규화된 평균절대차 (MAD: mean absolute difference) 를 제안한다.

이러한 방식들은 양자화 파라미터를 얻기 위해 2차율-왜곡 모델 (quadratic rate-distortion model) 이 사용된다.^[9] 이 모델은 예측된 목표 비트수와 예측된 평균절대차가 필요한데, 급장면 전환 프레임의 아주 낮은 상관관계로 인해 예측된 평균 절대차는 실제 계산된 값과 다르다. 이로 인해 잘못된 양자화 파라미터를 얻게 된다. 논문 [7]의 방식이 급장면 전환이 있는 비디오 영상의 품질을 향상시키지만 급장면 전환 프레임에 대한 양자화 파라미터 계산시 정확치 않은 평균 절대차를 여전히 사용한다. 정확한 평균절대차를 제공하기 위해 논문 [8]에서는 PSNR 차이 비율과 (PSNR difference ratio) 정규화된 평균절대차를 사용한다. 하지만 2차율-왜곡 모델은 근본적으로 인터 프레임에 대한 양자화 파라미터 계산에 적합한 모델이다. 급장면전환 프레임에 있어서 거의 모든 매크로블럭이 인트라 모드로 코딩되므로 논문 [8]의 방식도 급장면전환 프레임의 양자화 파라미터 계산에는 적합하지 않다. 이 논문에서 우리는 인트라 코딩의 특성을 고려하여 급장면전환 프레임에서 적절한 양자화 파라미터를 얻는 향상된 프레임 레벨 데이터율 제어 알고리즘을 제안한다.

이 논문의 구성은 다음과 같다. II 장에서는 제안할 알고리즘의 전반적인 흐름을 설명한다. III 장에서는 급장면 전환 프레임에서 인트라 코딩의 특성을 추출하는 방법에 관해 설명한다. IV 장에서는 급장면 전환 프레임에서 목표 비트수를 할당하는 방법을 제시한다. V 장에서는 급장면전환 프레임에 대해 적절한 양자화 파라미터를 얻는 방법을 설명한다. VI 장에서는 제안한 방식에 대한 실험 결과를 보이고 마지막으로 VII 장에서 결론을 맺는다.

II. 제안 알고리즘의 전체 흐름

H.264/AVC 데이터율 제어에 있어서, 양자화 파라미터는 프레임 레벨 혹은 몇 개의 매크로블럭을 기본 단위 레벨로, 목표 비트수를 예측한 이후 계산된다.^[10] 이 논문은 computational load 가 낮은 프레임 레벨 데이터율 제어에 초점을 둔다. 양자화 파라미터의 계산은 예측된 목표 비트수의 부호에 따라 달라진다.^[11] 만약 목표 비트수가 양수이면 현재 프레임의 양자화 파라미터,

Q_i 는 2차 울-왜곡 모델을 사용하여 다음과 같이 계산된다. 여기서 $T_{tb,i}$ 는 i 번째 프레임에 대한 예측된 텍스처 비트 (texture-bit) 인데, 이는 예측된 목표비트수와 이전에 코딩된 헤더비트수의 차이를 나타낸다. 또한 $PMAD_i$ 는 i 번째 프레임에 대한 예측된 평균절대차이고, x_1 과 x_2 는 각각 1차, 2차수의 계수이다.

$$\frac{T_{tb,i}}{PMAD_i} = \frac{x_1}{Q_i} + \frac{x_2}{Q_i^2} \quad (1)$$

그러나 급장면전환 프레임에 있어서 수식 (1) 의 모든 변수들이 선형 회귀 방법 (linear regression technique) 에 사용되는 것이 적합하지 않기 때문에 계산된 양자화 파라미터는 적절한 값이 아니다. 만약 목표 비트수가 음수이면 양자화 파라미터는 수식 (1)을 통해 계산되지 않고 이전 양자화 파라미터 대비 2 가 늘어난 값이다.

앞서 설명한 것처럼 급장면 전환 프레임에 대한 기존의 데이터율 제어 알고리즘은 문제를 갖고 있다. 급장면전환시 해당 프레임은 인트라 코딩된다. 이때 최적의 양자화 파라미터를 얻기 위해 우리는 인트라 코딩의 특성을 고려한 데이터율 제어 알고리즘을 제안한다. 그림 1은 제안할 데이터율 제어 알고리즘의 전체 흐름도를 나타낸다. 제안할 알고리즘은 3단계로 나뉜다. 인트라 코딩 특성을 추출하는 단계, 목표 비트수를 예측하는 단계 (이후 비트 할당 이라 칭한다.), 마지막으로 양자화 파라미터를 계산하는 단계로 이뤄져 있다. 새로운 프레임이 입력되면 급장면 전환이 발생했는지를 확인한다. 급장면 전환을 검출하는 알고리즘은 다양하지만 H.264/AVC 의 복잡도와 인코딩 이전에 수행되어지는 데이터율 제어의 특성을 감안하여 계산상 부하가 적은 알고리즘이 적합하다. 본 논문에서는 MSE (mean square error) 방식의 급장면 전환 알고리즘을 활용했다.^[12] 만약 급장면 전환이 발생치 않았다면 양자화 파라미터 계산은 H.264/AVC 의 데이터율 제어를 따른다. 반면 급장면 전환 발생이 검출되면 해당 프레임에서 인트라 코딩의 복잡도를 예측하기 위해 픽셀간의 표준편차를 계산한다. 계산된 표준편차는 나중에 양자화 파라미터를 계산하는데 활용된다. 다음으로 $RatioIP$ 를 근사화한다. $RatioIP$ 는 인트라 코딩에 사용되는 평균 비트수에 대한 인트라 코딩에 사용되는 비트수의 비이다. $RatioIP$ 가 계산되면 비트 할당이 이뤄지고 마지막으로 양자화 파라미터가 계산된다. 만약 현재 계산된 양자화 파라미터와 이전에 계산된 양자화 파라미터간의 차이가 2 미만이면 현재 계산된 양자화 파라미터가 최

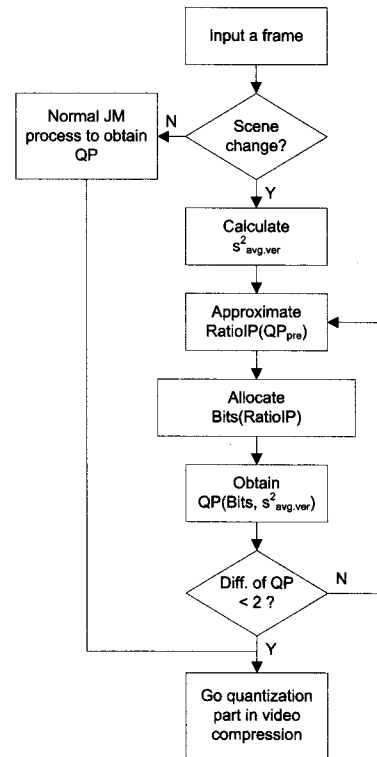


그림 1. 급장면 전환시 제안하는 데이터율 제어 알고리즘 흐름도

Fig. 1. Flow chart of the proposed rate control scheme for scene-change frames.

종값이 되고 인코딩 과정에 사용된다. 반면 차이가 2 이상이면 현재 계산된 양자화 파라미터로 $RatioIP$ 를 다시 얻어 비트 할당하고 양자화 파라미터를 다시 계산하게 된다. 수렴된 양자화 파라미터를 얻지 못하고 계산 루프가 무한정 도는 것을 막기 위해 루프 반복을 10 회로 제한한다.

III. 인트라 코딩 특성 추출

급장면 전환 프레임의 인코딩에 소요되는 비트수를 가능해 보기 위해 해당 프레임의 복잡도를 조사한다. 급장면 전환 프레임은 인트라 코딩을 할 것이므로 소요 비트수는 해당 프레임의 복잡도와 관계있다. H.264/AVC 인트라 코딩의 핵심은 에지 (edge) 방향을 찾는 데 있다. 이는 픽셀간 상관관계와 관련이 있다. 우리는 이러한 사실을 바탕으로 급장면 전환 프레임의 복잡도를 예측하기 위해 픽셀간 표준편차를 활용한다. 그림2는 프레임 사이즈가 QCIF 일 때 픽셀간 표준편차를 계산하는 방법이다.

픽셀간 표준편차는 두 가지 방향성을 갖는다. 수직 방향과 수평방향이 그것이다. 수직 방향의 픽셀간 표준편

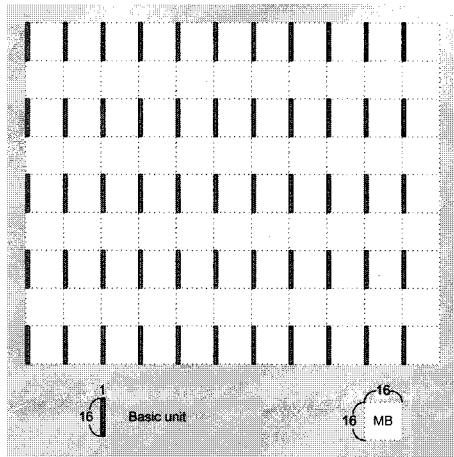


그림 2. 프레임 사이즈가 QCIF 일 때 픽셀간 표준편차를 계산하는 방법

Fig. 2. Architecture of basic units for calculating the standard deviation of pixels. QCIF frame size.

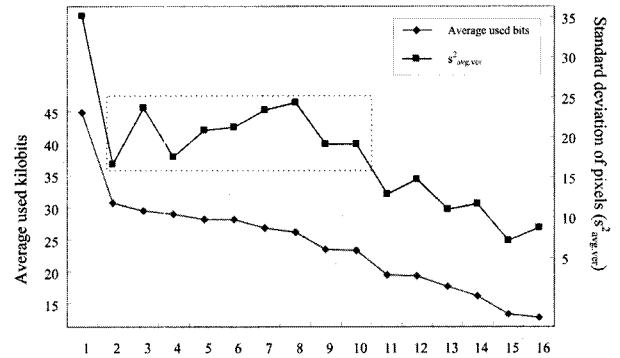
차값이 전체 표준편차를 대변할 만큼 지배적이다. 따라서 계산상의 부하를 감안하여 수직 방향의 픽셀간 표준편차만 계산한다. 추가적으로 계산상의 부하를 줄이기 위해 표준편차를 계산할 영역을 그림 2에서 보듯이 생성된 베이직 유닛 단위만 수행한다. 베이직 유닛은 가로 1 픽셀, 세로 16 픽셀을 기본 단위로 만들어 진다. 베이직 유닛의 생성은 수평방향으로 매 16 픽셀마다, 수직방향으로 매 2 베이직 유닛 마다 이뤄진다.

해당 프레임의 표준편차는 그림 2처럼 생성된 베이직 유닛에 대한 표준편차를 구하고 평균을 취한 값으로 다음과 같다. 여기서 s_{2i} 와 N_{BU} 는 각각 i 번째 베이직 유닛에 대한 표준편차와 베이직 유닛 개수를 나타낸다.

$$s_{avg.ver}^2 = \frac{1}{N_{BU}} \sum_{i=1}^{N_{BU}} s_i^2 \quad (2)$$

이러한 간략화로 계산 부하는 모든 픽셀에 대한 표준편차를 계산하는 것 대비 64배 만큼 감소하게 된다. 이때 간략화 된 수식 (2)의 결과값은 인트라 코딩의 복잡도를 인식하는데 오류를 범하지 않을 범위내에서 최대한 간략화를 수행하였다.

그림 3은 인트라 코딩의 복잡도를 가늠하는데 사용할 $s_{avg.ver}^2$ 의 결과를 나타낸다. 붉은색의 다이아몬드로 마크된 선은 테스트에 사용된 영상의 첫 번째 프레임에 인코딩하는데 사용된 비트수를 나타낸다. 이 값들은 다양한 양자화 파라미터 (15~45) 에서 얻어진 값들의 평균값이고 그림 3에서 내림차순으로 정렬되어 있다. 즉, 복잡도가 높은 순으로 정렬되어 있다. X 축상의 숫자는 각 테스트 영상을 나타내는데 그림에서 “Stefan” 영상



(1) Stefan, (2) Salesman, (3) Trevor, (4) Silent, (5) Coastguard, (6) News, (7) Foreman, (8) Container, (9) Carphone, (10) Hallmonitor, (11) Akiyo, (12) Suzie, (13) Motherdaughter, (14) Highway, (15) Claire, (16) Missamerica

그림 3. 각 테스트 영상에 대한 실제 복잡도를 나타내는 사용 비트수와 복잡도 예측에 사용될 픽셀간 평균 수직 표준편차값

Fig. 3. The average number of bits used over various quantization parameters for encoding the first frame and the average standard deviation of vertical basic units for the first frame.

의 복잡도가 최고이다. 반대로 검은색의 네모상자로 마크된 선은 해당 테스트 영상의 복잡도를 나타내는 $s_{avg.ver}^2$ 값이다. 비록 $s_{avg.ver}^2$ 값이 실제 복잡도에 대응되게 단조 감소 형태를 띄지는 않지만 네모 점선 박스를 중심으로 3단계로 나눌 수 있다.

그림 3을 바탕으로 인트라 코딩의 복잡도를 다음과 같이 3 단계로 나눈다.

$$\begin{aligned} \text{High} &: s_{avg.ver}^2 > 26 \\ \text{Medium} &: 15 < s_{avg.ver}^2 \leq 26 \\ \text{Low} &: s_{avg.ver}^2 \leq 15 \end{aligned} \quad (3)$$

이로써 인트라 코딩의 복잡도 정도는 $s_{avg.ver}^2$ 값에 의해 3단계로 결정된다.

IV. 급장면 전환 프레임 비트 할당

H.264/AVC 에 사용된 데이터율 제어방식에서 비트는 코딩 안된 프레임에 대해 남아 있는 비트들과 버퍼의 상태에 의해 할당된다. 엄밀히 말해 비트할당은 이전에 코딩된 비트를 참고하여 이뤄진다. 그래서 비트를 할당하는 수식은 인트라 프레임들인 P (forward predicted frame) 혹은 B (bidirectionally predicted frame) 프레임들에만 적합하다. 인트라 프레임을 위한 비트 할당 수식이 필요하므로 H.264/AVC 참조 소프트웨어에서 사용하고 있는 수식을 수식 (4)와 같이 변경한다.

$$T_{b,i} = \beta \frac{R_{b,i}}{N_{P,i}} \kappa_{IP} + (1 - \beta) \left[\frac{b_r}{f_r} - \Gamma(CBF_{i-1} - TBL_i) \right] \quad (4)$$

여기서 $R_{b,i}$ 와 $N_{P,i}$ 는 각각 i 번째 프레임을 인코딩 하 기전, 코딩안된 프레임들을 위해 남아 있는 비트수와 코딩 안된 P 프레임수를 의미한다. β 와 Γ 는 상수값이 고 각각 0.8 과 0.5 이다. 기존의 비트 할당 수식에서 달 라진 것은 변수 κ_{IP} 이다. 우리는 이것을 *RatioIP* 로 명 명한다. *RatioIP* 는 인터 코딩된 평균 비트수와 인트라 코딩된 비트수의 비를 나타내는데 다음 수식과 같다. 여기서 B_{tra} 와 B_{ter} 는 각각 인트라 코딩 후 생성된 비 트수와 인터 코딩후 생성된 비트수를 나타낸다. N_{ter} 는 GOP 내 인터 코딩되는 프레임수를 의미한다.

$$\kappa_{IP} = B_{tra} \frac{N_{ter}}{\sum_{i=1}^{N_{ter}} B_{ter,i}} \quad (5)$$

*RatioIP*를 사용하여 급장면 전환된 프레임에 좀더 효율적으로 비트를 할당 할수 있도록 기존 수식에 가중 치를 부여한다. 그러나 실질적으로 *RatioIP* 는 여러 변 수에 의해 변해서 분석적인 접근을 어렵게 한다. 따라 서 반복적인 실험값으로 접근하는 것이 더 타당하다. 표 1은 여러 테스트 영상에 여러 양자화 파라미터를 적 용했을 때 *RatioIP* 값을 나타낸다. GOP 변수값 세팅

표 1. 인터 코딩으로 생성되는 평균 비트수와 인트라 코딩으로 생성되는 비트수의 비

Table 1. Ratio of the number of intra-coded bits to the average number of inter-coded bits.

Sequences	Constant QP		
	20	30	40
Stefan	2.8	5.6	10.7
Container	11.1	39.1	47.3
Trevor	4.3	7.0	9.2
Foreman	4.1	10.5	13.8
News	10.9	18.4	25.1
Coastguard	2.7	5.7	11.3
Carphone	4.1	8.6	13.6
Hallmonitor	5.6	28.8	28.4
Silent	11.3	16.4	17.3
Salesman	11.9	20.1	25.5
Suzie	4.7	13.3	14.0
Akiyo	20.8	36.1	37.4
Highway	2.4	6.9	10.0
Motherdaughter	13.1	26.3	26.0
Missamerica	6.1	17.4	20.0
Claire	9.4	20.9	22.5

은, 인코딩 프레임 개수는 50개이고 인트라 프레임 주 기는 10, B 프레임은 사용되지 않았다. 움직임이 빠르 거나 인트라 코딩의 복잡도가 높을수록 *RatioIP* 는 작 아진다. 영상에 대한 움직임을 미리 예측하는 것은 거 의 불가능하므로 *RatioIP* 를 영상마다 구하는 것은 어 렵다. 따라서 모든 테스트 영상에 대해 *RatioIP* 평균값 을 구하여 *RatioIP* 를 주어진 양자화 파라미터의 함수 로 만들었다. II 장에서 설명했듯이 제안하는 데이터율 제어 알고리즘의 흐름도에서, 루프의 처음에는 *RatioIP* 는 이전 프레임에서 사용된 양자화 파라미터로 근사화 된다. 첫 번째 루프 이후에는 각 *RatioIP* 는 이전 루프 에서 얻어진 양자화 파라미터에 의해 계산된다. 이것은 계산된 양자화 파라미터가 특정한 값으로 수렴되도록 돕는다. 자세한 설명은 다음 장에서 한다.

V. 양자화 파라미터 계산

급장면 전환 프레임들에 대한 데이터율 제어의 궁극 적인 목적은 적절한 양자화 파라미터를 얻는데 있다. 이것을 하기 위해 우리는 III 장에서 인트라 코딩의 복 잡도를 결정하는 방법에 관해 설명했고, IV 장에서 급 장면 전환 프레임을 위한 비트 할당 방법에 관해 설명 했다. 이러한 것을 바탕으로 이번장에서 급장면 전환 프레임을 인코딩하는데 사용될 양자화 파라미터를 얻는 방식을 제안한다. 앞서 설명했듯이 양자화 파라미터는 2차 울-왜곡 모델^[9]에 의해 계산된다. 그러나 이 모델은 급장면 전환된 프레임에 적용하기는 적합치 않다. 그 래서 우리는 실험적인 결과를 바탕으로 새로운 모델을 제 안한다.

우선 주어진 양자화 파라미터와 인코딩에 사용될 비 트수와의 관계를 알아야 한다. 이러한 관계는 영상 마 다 변하게 된다. 달리 말해, 인트라 코딩의 복잡도에 따 라 변한다. 그림4는 이러한 관계를 보여준다. 모든 테스 트 영상을 사용했으며 인트라 모드로 코딩될 첫 번째 프레임을 코딩한 결과이다. 테스트 플랫폼으로 H.264/AVC 참조 소프트웨어 버전 JM9.8 을 사용했다. 실험을 위한 변수값 설정은, RDO (rate distortion optimization) 을 사용했으며, ME 을 위한 검색 영역은 16이며, 참조 프레임수는 한개이며, Hadamard 변환을 사용했으며, CABAC (context adaptive binary arithmetic coding) 을 엔트로피 코딩으로 사용했다.

입력 영상, 즉 인트라 코딩의 복잡도에 따라 여러 개 의 관계 커브가 있음을 알 수 있다. 이것은 계산상의 부

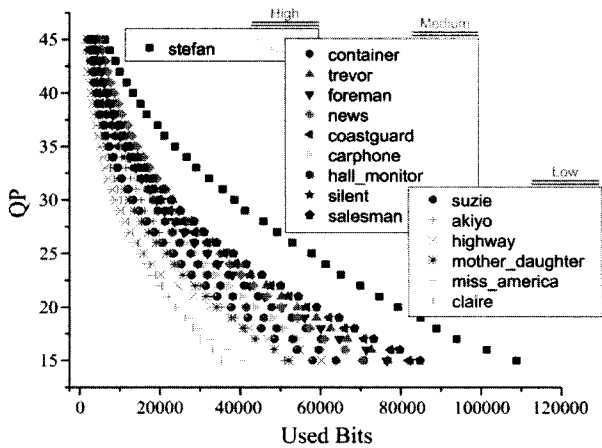


그림 4. 영상의 첫 번째 프레임을 인코딩 하는데 사용된 양자화 파라미터와 생성된 비트수의 관계
 Fig. 4. Relation between quantization parameter and number of bits used for the first frame.

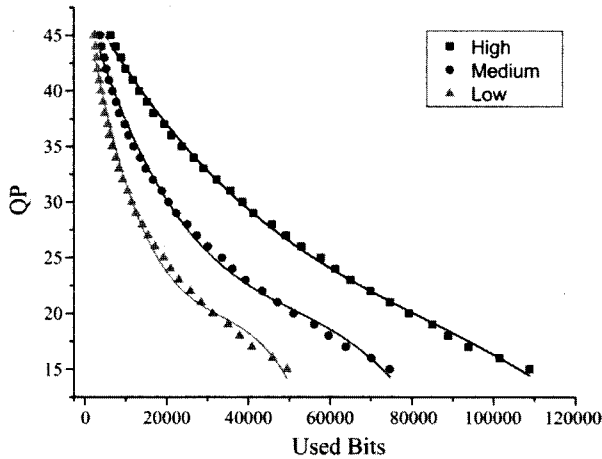


그림 5. 각 복잡도 단계마다, 영상의 첫 번째 프레임을 인코딩 하는데 사용된 양자화 파라미터와 생성된 비트수의 대표적 관계
 Fig. 5. Representative relation curves between QP and numbers of bits used for the three intra-coding complexity classes.

하를 증가시키기 때문에 우리는 인트라 코딩의 복잡도를 수식 (3)처럼 3단계로 나눴다. 그림 4에서 보듯이 Low 와 Medium 복잡도 단계에 여러 영상이 있다. 만약 동일한 인트라 코딩 복잡도 단계에 속하는 여러 관계 커브들이 하나의 대표적인 관계 커브로 표현될 수 있다면 급장면 전환 프레임의 양자화 파라미터는 2가지 변수에 의해 결정될 수 있다. 하나는 수식 (4)에 의한 비트 할당이고 다른 하나는 수식 (2)에 의한 인트라 코딩 복잡도 단계를 나타내는 픽셀간 수직 평균 표준편차 값이다. Low와 Medium 복잡도 단계에 대한 대표적인 관계 커브를 얻기 위해 소속된 영상에 대해 사용된 비트수의 평균을 취했다. 그림 5는 각 복잡도 단계에 대

한 양자화 파라미터와 사용된 비트수와의 대표적 관계 커브를 나타낸다. 네모박스, 원, 삼각형 도트들은 각기 High, Medium, Low 복잡도 단계에 대한 대표적 관계 커브를 나타낸다. 이러한 불연속 데이터에 대해 연속적인 양자화 파라미터를 부여하기 위해 3차 다항식으로 보간 (interpolation) 하였다. 각 도트를 연결하는 선이 보간된 값들을 의미한다.

인트라 코딩의 복잡도 단계를 결정한 이후, 할당된 비트를 바탕으로 양자화 파라미터를 결정할 수 있다. 그 관계식은 수식 (6)과 같다. 여기서 N_{AB} 는 할당된 비트수를 나타낸다. 함수 f_{md} 는 양자화 파라미터를 정수화 하기 위한 반올림 함수이다. 양자화 파라미터와 사용된 비트수와의 관계 커브는 변수 설정이 급격하지 않는 한 변화는 미미하다. 예를 들어, 엔트로피 코딩이 CAVLC (context adaptive variable length coding) 로 바뀌면 코딩에 사용된 비트수는 CABAC 대비 평균 2.5 % 정도 증가한다. 유사한 형태로 만약 RDO 를 사용하지 않는다면 코딩에 사용된 비트수는 RDO 사용시 대비 평균 7.5 % 범위에서 변한다.

정리해보면, 급장면 전환이 검출되면 s_{avgver}^2 와 $RatioIP$ 을 계산한다. $RatioIP$ 을 활용하여 비트를 할당하고 s_{avgver}^2 을 사용하여 대표적인 관계 커브를 선택한다. 마지막으로 할당된 비트에 대한 양자화 파라미터를 계산한다. 만약 양자화 파라미터의 차이가 2 이상이면 최대 10회 반복 계산한다. 루프를 반복할수록 양자화 파라미터는 특정값으로 수렴한다. 이는 다음과 같은 비례관계에 기인한다. $RatioIP$ 는 양자화 파라미터에 비례한다. 할당된 비트는 $RatioIP$ 에 비례하므로 양자화 파라미터에 비례한다고 할수 있다. 반대로 사용된 비트수와 양자화 파라미터는 서로 반비례한다. 따라서 각 루프에서 계산된 양자화 파라미터는 이전에 계산된 두 개의 양자화 파라미터값의 중간값인 어느 값을 갖게 되므로 특정값으로 수렴하게 된다.

VI. 실험

급장면 전환 영상에 대한 데이터율 제어 알고리즘의 성능을 검증하기 위해 테스트 영상은 16개를 사용하였다. 이는 범용적으로 사용하는 거의 모든 테스트 영상으로 다음과 같다. Akiyo, Carphone, Claire, Coastguard, Container, Foreman, Hallmonitor, Highway, Missamerica, Motherdaughter, News, Salesman, Silent, Stefan, Suzie, and Trevor. 모든 테

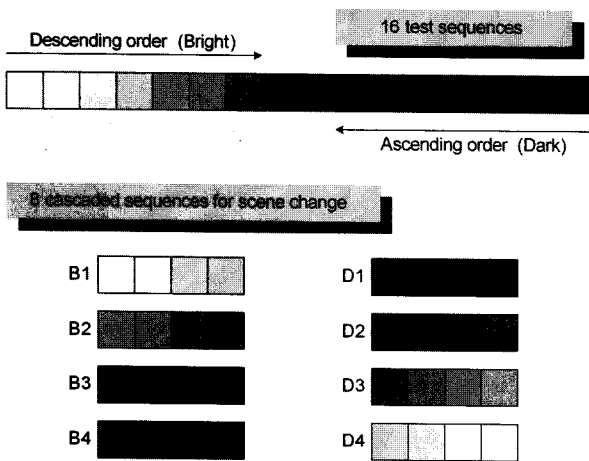


그림 6. 범용적으로 사용되는 16개의 테스트 영상을 휘도순으로 정렬한후 4개씩 중속 연결하여 만든 급장면 전환 테스트 영상

Fig. 6. Block diagram of construction of the eight cascaded sequences. Each rectangular box indicates a test sequence cut from the 0-th frame to the 49-th one.

스트 영상은 QCIF 4:2:0 포맷이다. 급장면 전환 검출은 휘도 (luminance)에 민감하므로 테스트의 신뢰성을 높이기 위해 16개의 테스트 영상에 대한 평균 휘도값을 내림차순과 오름차순으로 정렬한 후 4개씩 붙여 새로운 영상을 만들었다. (이를 중속접속 테스트영상이라 명명한다.) 각 중속접속 테스트 영상은 총 200 프레임으로 구성되어 있으며 매 50 프레임 마다 급장면 전환이 일어난다. 평균 휘도값을 내림차순으로 정렬한 후 만든 중속접속 테스트 영상을 B1, B2, B3, B4 라 명명하고 반대로 오름차순으로 정렬한 후 만든 중속접속 테스트 영상을 D1, D2, D3, D4 라 명명한다. 그림 6은 총 8개의 중속접속 테스트 영상을 만드는 방법을 나타낸다. 각각의 네모박스는 0번째 프레임에서 49번째 프레임까지의 테스트 영상을 나타낸다.

B1, B2, B3, B4 영상은 각기 Highway-Foreman-Container-Hallmonitor, Trevor-Stefan-Silent-Coastguard, Motherdaughter-Suzie-Carphone-Claire, Akiyo-News-Salesman-Missamerica이고 D1, D2, D3, D4 영상은 각기 Missamerica-Salesman-News-Akiyo,

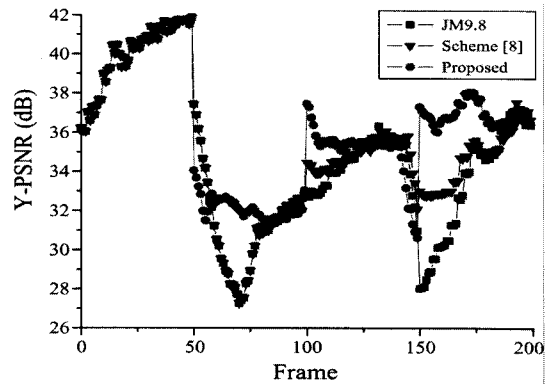


그림 7. 목표 비트율 38.4 kbps에서 D2 중속접속 테스트 영상에 대한 PSNR 커브

Fig. 7. PSNR curves of D2 cascaded sequence at 38.4 kbps. Scene transitions occur every 50 frames.

Claire-Carphone-Suzie-Motherdaughter, Coastguard-Silent-Stefan-Trevor, Hallmonitor-Container- Foreman-Highway이다.

제안하는 데이터율 제어 알고리즘은 JM9.8 을 기반으로 구현되었다.^[11] 모든 테스트에서 사용된 변수값은, 인코딩 프레임 개수는 200, GOP 내 첫 번째 프레임만 인트라 코딩수행, B 프레임은 없음, ME 검색영역은 16, 참조 프레임은 하나, RDO 는 사용하지 않음, 엔트로피 코딩은 CABAC, 버퍼 사이즈는 목표 비트율의 절반으로 설정, 프레임율은 30 fps. 실험에 사용한 목표 비트율은 19.2 kbps, 38.4 kbps, 76.8 kbps 이다.

좋은 데이터율 제어 알고리즘은 높은 비디오 품질과 프레임에 따른 비디오 품질의 변화가 적고 목표로 설정된 비트율과 인코딩 이후의 비트율간의 오차가 작아야 한다. 제안한 알고리즘의 성능은 JM9.8 과 논문 [8]의 방식과 비교한다. 이때 합리적인 비교를 위해, 논문 [8] 방식과 제안한 방식은 오직 급장면 전환 프레임에 한해서 양자화 파라미터를 각 방식에 따라 계산한다. 나머지 프레임은 JM9.8 의해 계산된다. 그림 7은 38.4 kbps 에서 D2 중속접속 테스트 영상에 대한 PSNR 커브를 나타낸다. 급장면 전환은 매 50번째 프레임에서 발생한다. 첫 번째 급장면 전환이 일어나기 전까지 3가지 데이터율 제어 방식의 결과는 동일하다. 첫 번째 급장면

$$QP = \left\{ \begin{array}{l} f_{rnd}(48.2 - 6.6 \times 10^{-4} N_{AB} + 5.7 \times 10^{-9} N_{AB}^2 - 2.3 \times 10^{-14} N_{AB}^3) : High \\ f_{rnd}(47.9 - 1.2 \times 10^{-3} N_{AB} + 2.1 \times 10^{-8} N_{AB}^2 - 1.3 \times 10^{-13} N_{AB}^3) : Medium \\ f_{rnd}(48.0 - 2.2 \times 10^{-3} N_{AB} + 5.9 \times 10^{-8} N_{AB}^2 - 5.8 \times 10^{-13} N_{AB}^3) : Low \end{array} \right\} \quad (6)$$

전환이 발생하는 50번째 프레임이후 제안한 데이터율 제어 알고리즘의 PSNR은 다른 두 가지 방식과 비교하여 거의 모든 프레임에서 높다. 프레임에 대한 PSNR의 변화 또한 제안한 알고리즘이 다른 두 가지 방식에 비해 낮다. 그림 7에서 보듯이 매 급장면 전환 프레임과 73번째 프레임에서 PSNR의 차이가 극명하게 나타난다. 이를 시각적으로 나타낸 것이 그림 8이다.

첫 번째 급장면 전환이후, 기존의 방식들은 첫 번째 급장면 전환 프레임에서 필요 이상으로 비트를 사용하여, 이후 사용할 비트수가 줄어, 양자화 파라미터를 크게 만들어 결론적으로 비디오 품질을 열화시킨다. 두 번째 급장면 전환 이후, 기존 방식들은 급장면 전환 프레임들에서 필요한 것보다 비트가 너무 작게 할당되었다. 이로 인해 열화된 비디오 품질은 이후 프레임에 전파되어 결론적으로 비디오 품질을 열화시켰다. 그러나 제안된 알고리즘은 기존 방식들보다 급장면 전환 프레임에서 양자화 파라미터를 효율적으로 계산하여 향상된

성능을 얻었다.

자세한 결과는 표 2에 정리되어 있다. 첫 번째 급장면 전환까지는 결과가 동일하므로 50번째 프레임부터 끝까지의 결과를 정리하였다. 색칠된 부분의 결과는 제안한 알고리즘이 기존의 방식들보다 성능적으로 우수함을 나타낸다. 제안한 알고리즘은 비디오 품질 측면에서 성능을 획기적으로 향상시켜, JM9.8 방식 대비 평균 0.6 dB, 최대 1.60 dB 성능이 개선되었으며, 논문 [8] 방식 대비 평균 0.41 dB, 최대 1.12 dB 향상되었다. PSNR 변동과 비트율 오차정도는 제안한 알고리즘과 기존 방식이 비슷한 결과를 보인다. 하지만 급장면 전환이 없는 장면끼리 비교하면 제안한 알고리즘의 PSNR 변동은 기존방식보다 우수함을 알수 있다. 이경우에 있어 제안한 알고리즘의 PSNR 변동은 JM9.8 대비 평균 18.6 % 향상되며, 논문 [8] 대비 평균 19.8 % 향상된다.

표 2. 제안한 데이터율 제어 알고리즘과 기존의 방식인 JM9.8, 논문 [8] 방식과의 성능비교
Table 2. Performance comparisons of the proposed rate control scheme with the existing schemes. (JM9.8 and scheme [8])

※ StDev 와 Ebps 은 각각 Y-PSNR 의 standard deviation 과 encoded bit rate 을 나타낸다.

Performance		B1	B2	B3	B4	D1	D2	D3	D4	
Target bit rate (19.2kbps)	JM9.8	Y-PSNR (dB)	27.96	24.25	30.66	30.45	29.38	29.52	23.31	29.03
		StDev	2.41	3.14	4.07	4.59	2.47	2.54	2.95	3.09
		Ebps (kbps)	19.70	19.49	19.28	19.28	19.36	19.43	19.31	19.41
	Scheme [8]	Y-PSNR (dB)	28.35	24.51	31.01	30.41	29.46	30.29	23.39	28.86
		StDev	3.00	3.36	2.92	4.97	2.83	2.38	3.57	3.06
		Ebps (kbps)	19.64	19.61	19.25	19.30	19.23	19.39	19.45	19.30
	Proposed	Y-PSNR (dB)	28.28	24.83	31.43	30.97	30.16	31.10	24.10	29.53
		StDev	2.08	3.59	3.27	4.32	2.84	2.06	3.46	3.15
		Ebps (kbps)	19.54	19.63	19.28	19.32	19.21	19.49	19.73	19.44
Target bit rate (38.4kbps)	JM9.8	Y-PSNR (dB)	33.30	26.95	34.52	34.54	33.78	33.06	26.65	33.07
		StDev	2.22	3.47	4.56	4.12	3.02	2.69	3.79	2.45
		Ebps (kbps)	38.41	38.66	38.47	38.58	38.48	38.44	38.50	38.45
	Scheme [8]	Y-PSNR (dB)	33.33	27.06	35.09	34.84	33.94	33.60	26.70	33.02
		StDev	3.04	3.81	3.70	4.27	3.33	2.39	4.10	3.05
		Ebps (kbps)	38.50	38.68	38.44	38.59	38.41	38.52	38.48	38.50
	Proposed	Y-PSNR (dB)	33.24	27.59	35.41	34.78	34.16	34.72	27.42	33.59
		StDev	1.63	4.12	3.83	4.20	2.90	2.17	3.97	2.47
		Ebps (kbps)	38.41	38.77	38.37	38.50	38.41	38.48	38.76	38.80
Target bit rate (76.8kbps)	JM9.8	Y-PSNR (dB)	36.43	29.64	38.17	38.08	37.95	36.42	29.76	35.90
		StDev	2.28	3.78	3.97	3.95	2.82	3.22	4.25	2.73
		Ebps (kbps)	76.87	77.12	76.81	76.88	76.61	76.72	76.92	76.78
	Scheme [8]	Y-PSNR (dB)	36.73	29.79	38.29	38.35	38.07	36.70	29.46	36.10
		StDev	2.38	4.12	3.82	3.73	2.44	3.17	5.14	2.37
		Ebps (kbps)	76.87	76.85	76.91	76.97	76.62	76.86	76.94	76.82
	Proposed	Y-PSNR (dB)	36.58	30.09	38.37	38.38	38.25	37.60	30.26	36.33
		StDev	2.33	4.67	3.89	3.33	2.74	2.68	4.73	2.12
		Ebps (kbps)	76.88	77.00	76.95	76.78	76.68	76.85	76.86	76.89

VIII. 결 론

낮은 비트율을 사용하고 급장면 전환이 있는 영상에 대한 향상된 H.264/AVC 프레임 레벨 데이터율 제어 알고리즘을 제안했다. 급장면 전환이 검출되면 향상된 비디오 품질을 위해 적절한 양자화 파라미터 계산이 필요하다. 그러나 급장면 전환된 인트라 프레임에 있어서 기존의 방식은 인트라 프레임에 적합한 2차 윌-왜곡 모델을 그대로 사용하여 양자화 파라미터를 계산하므로 적절치 못하다. 우리는 본 논문에서 인트라 코딩의 특성을 고려하여 최적의 양자화 파라미터를 계산하는 알고리즘을 제안했다. 이는 인트라 코딩의 복잡도를 계산하는 방법과 급장면 전환된 인트라 프레임을 위한 비트 할당하는 방법, 마지막으로 할당된 비트와 계산된 복잡도를 이용하여 양자화 파라미터를 계산하는 방법, 이렇게 3가지 블록으로 구성된다. 제안한 데이터율 제어 알고리즘은 비디오 품질과 프레임에 따른 비디오 품질의 변화정도에 있어서 기존 방식과 비교하여 우수한 성능을 보였다. 또한 목표한 비트율과 인코딩 후 비트율과의 오차정도는 제안 알고리즘과 기존방식이 비슷한 결과를 가졌다. 비디오 품질 측면에 있어서 제안 알고리즘은 평균 0.4 dB 에서 0.6 dB 향상되었으며 최대 1.1에서 1.6 dB 까지 향상되었다. 프레임에 따른 비디오 품질의 변화정도는 제안한 알고리즘이 평균 18.6 % 향상된 결과를 보였다. 따라서 제안하는 데이터율 제어 알고리즘은 낮은 비트율을 사용하고, 급장면 전환이 있는 영상에 있어서 기존 방식보다 우수한 성능을 보임을 알 수 있다.

참 고 문 헌

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, May 2003. Joint Video Team.
- [2] A. Joch, F. Kossentini, H. Schwarz, T. Wiegand, and G.J. Sullivan, "Performance comparison of video coding standards using Lagrangian coder control," *IEEE Int. Conf. Image Processing*, vol. II, pp. 501-504, Sept. 2002.
- [3] L. Gwangsoon, C. Sammo, Y. Kyutae, H. Youngkwon, and L. Sooin, "Development of terrestrial DMB transmission system based on Eureka-147 DAB system," *IEEE Trans. Consumer Electronics*, vol. 51, pp. 63-68, Feb. 2005.
- [4] L. Ping, X.K. Yang, and W.S. Lin, "Buffer-constrained R-D Model-Based Rate Control for H.264/AVC," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 2, pp. 321-324, Mar. 2005.
- [5] Y. Hongtao, L. Zhiping, and P. Feng, "An improved rate control algorithm for H.264," *IEEE Int. Sym. Circuits and Systems*, vol. 1, pp. 312-315, May 2005.
- [6] X. Jianfeng and H. Yun, "A novel rate control for H.264," *IEEE Int. Sym. Circuits and Systems*, vol. III, pp. 809-812, May 2004.
- [7] M. Jiang and N. Ling, "On enhancing H.264/AVC video rate control by PSNR-based frame complexity estimation," *IEEE Trans. Consumer Electronics*, vol. 51, pp. 281-286, Feb. 2005.
- [8] X. Yi and N. Ling, "Improved H.264 rate control by enhanced MAD-based frame complexity prediction," *Journal of Visual Communication and Image Representation*, vol. 17, pp. 407-424, April 2006.
- [9] L. Hung-Ju, C. Tihao, and Z. Ya-Qin, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 10, pp. 878-894, Sept. 2000.
- [10] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT," JVT-G012-r1, 7th Meeting, Pattaya II, Thailand, Mar. 2003.
- [11] The reference software version 9.8 of H.264/AVC is available on: http://ftp3.itu.ch/av-arch/jvt-site/reference_software/jm98.zip
- [12] W.A.C. Fernando, C.N. Canagarajah, and D.R. Bull, "A unified approach to scene change detection in uncompressed and compressed video," *IEEE Trans. Consumer Electronics*, vol. 46, pp. 769-779, Aug. 2000.

저 자 소 개



이 창 현(학생회원)
 1997년 경희대학교 전자공학과
 학사 졸업.
 1999년 연세대학교 전과공학과
 석사 졸업.
 2004년~현재 연세대 전기전자
 공학과 박사 과정.

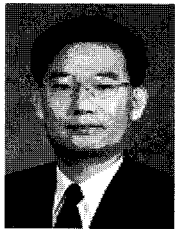
1999년~2004년 삼성전자 통신연구소 연구원
 <주관심분야 : H.264, Rate control, 영상통신>



정 윤 호(정회원)
 1998년 연세대 전자공학과 학사
 졸업.
 2000년 연세대 전자공학과 석사
 졸업.
 2005년 연세대 전기전자공학과
 박사 졸업.

2005년~2007년 삼성전자 통신연구소 책임연구원
 2007년~현재 연세대 TMS 정보기술사업단
 연구 교수

<주관심분야 : 초고속데이터 통신용 모뎀 및
 영상처리 프로세서의 알고리즘 및 SoC 설계>



김 재 석(정회원)
 1977년 연세대 전자공학과
 학사 졸업.
 1979년 KAIST 전기전자공학과
 석사 졸업.
 1988년 Rensselaer Polytechnic
 Institute, NY, 박사 졸업.

1993년~1995년 한국전자통신연구원 책임연구원
 1996년~현재 연세대학교 전기전자공학과 교수
 <주관심분야 : 통신 및 영상 시스템, VLSI 신호
 처리, 임베디드 S/W 및 SoC 구현>