

논문 2008-45SP-3-11

실시간 부호화를 위한 모델 파라미터 기반 일정 화질 비트율 제어 기법

(Model Parameter-based Rate Control Algorithm for Constant Quality Real-Time Video Coding)

정진우*, 조경민**, 최윤식**

(Jin-Woo Jeong, Kyung-Min Cho, and Yoonsik Choe)

요약

본 논문에서는 동영상의 화질을 일정하게 하기 위한 실시간 비트율 제어 기법을 제안한다. 일정 화질을 만족하기 위한 기존의 비트율 제어 알고리즘은 프레임의 부호화 복잡도를 잔여 신호인 MAD(mean absolute of difference)로 추정하여 비트 할당을 수행하였다. 그러나 MAD는 영상의 특성이나 부호화 파라미터에 따라 동일한 MAD라도 다른 비트를 생성하므로 영상의 부호화 복잡도를 적절히 나타내기 어렵다. 본 논문에서는 이 문제를 해결하기 위해 비트와 MAD사이의 기울기인 모델 파라미터를 프레임의 복잡도의 척도로 보고 이전 프레임과 현재 프레임의 모델 파라미터의 비율로 비트 할당을 수행한다. 또한 기존의 비트-복잡도 모델에서 구한 모델 파라미터는 양자화 파라미터가 변함에 따라 그 값이 크게 변하여 영상의 내재적 복잡도를 나타내기 어렵다. 따라서 본 논문에서는 비트-복잡도 모델에 양자화 파라미터를 추가하여 양자화 파라미터가 변하더라도 영상의 복잡도의 척도인 모델 파라미터는 변하지 않게 하였다. 광범위한 실험결과를 제안한 알고리즘이 기존의 알고리즘에 비해 비슷한 평균 화질을 유지하면서 화질의 변동을 큰 폭으로 줄였음을 보여준다.

Abstract

In this paper, we propose a rate control algorithm for constant quality real time video coding. To achieve constant quality, previous algorithm exploit mean absolute of difference(MAD) as measure of frame complexity. However, if scene is abruptly changed or if quantization parameter is not constant, encoder produces various output bits with same MAD. Therefore we know that MAD does not appropriately reflect characteristic of frame. To solve this problem, we exploit model parameter as measure of frame complexity. Because model parameter means slope between output bits and MAD, it reflects correctly complexity of frame. And because previous model, R-MAD model, is not considered quantization parameter, as quantization parameter increases or decreases, model parameter of frame also vary. So model parameter obtained using previous model cannot reflect internal characteristic of video. We solve this problem using proposed model, which is considered quantization parameter. Experiment results show that our algorithm provide better performance, in terms of quality smoothness than previous algorithm. Especially, when scene is abruptly changed, our algorithm alleviates quality drop.

Keywords : Rate control, constant Quality, model parameter, H.264

I. 서론

비트율 제어 기법의 궁극적인 목표는 주어진 채널이

나 저장 용량 아래에서 시각적으로 가장 좋은 화질을 제공하는 것이다. 시각적으로 가장 뛰어난 화질을 제공하기 위해서는 각각의 비디오 프레임의 화질이 높을 뿐만 아니라 프레임과 프레임 사이의 화질의 변동이 크지 않아야 한다. 그 이유는 프레임과 프레임 사이의 화질의 변동이 크면 “깜박임” 같은 현상이 일어나 인간의 시각 체계가 파악하는 비디오의 품질이 크게 훼손되어 낮은 품질의 동영상으로 인지되기 때문이다. 따라서 높은 품질의 동영상 화질을 제공하기 위해서는 일정 화질

* 학생회원, ** 정회원, 연세대학교 전기전자공학과
(Department of Electrical & Electronics Yonsei University)

※ 본 연구는 지식 경제부 및 정보통신 연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음 (IITA-2008-(C1090-0801-0011))

접수일자: 2008년1월24일, 수정완료일: 2008년4월11일

을 만족할 수 있는 비트율 제어 기법이 필요하다.

비트율 제어 기법은 비트 할당과 비트 소비라는 두 단계를 거쳐 수행된다. 비트 할당은 주어진 채널이나 저장 용량 아래에서 압축 대상의 부호화 복잡도와 복호기 버퍼의 상태를 고려하여 부호화하고자 하는 대상, 즉 GOP(group of picture), 프레임, 매크로블록 등에 목표 비트를 효과적으로 할당하는 단계를 의미하고, 두 번째 단계인 비트 소비 단계는 할당된 비트에 일치하게 적절한 양자화 파라미터를 선택하고 이에 따라 부호화를 수행하는 단계이다. 이중 일정한 화질을 만족하기 위해서는 압축 대상의 복잡도에 따라 비트를 할당하는 것이 중요하므로, 본 논문에서는 일정 화질을 만족하기 위한 효율적인 비트 할당기법의 연구에 초점을 맞춘다.

일정 화질을 달성하기 위한 비트 할당을 수행하기 위해서는 영상의 부호화 복잡도를 정확하게 측정할 필요가 있다. 이것을 달성하기 위하여 기존의 여러 연구에서는 두 번의 부호화 과정을 수행하였다. 즉 첫 번째 부호화 과정에서는 시퀀스 전체에 대하여 부호화를 수행하여 영상의 부호화 복잡도에 대한 정보를 얻은 후, 두 번째 과정에서 얻은 정보에 근거해서 비트를 할당하고 양자화 파라미터를 결정하여 실제 부호화를 수행한다^[1~3]. 이들 기법은 실제 부호화 복잡도에 근거하여 비트 할당을 수행하였기 때문에 비디오 품질의 변동 폭을 크게 줄일 수 있었지만, 여러 번의 부호화 과정이 필요하기 때문에 그 계산 복잡도가 매우 커 실시간 또는 저-지연(low-delay)응용에 그 활용이 제한적일 수 있다는 단점을 지닌다.

실시간 또는 저-지연 응용에 적합하면서 일정 화질을 위한 기존의 대표적인 비트 할당 기법으로는 Xie와 Zeng이 제안한 방법이 있다^[4]. 이 논문은 기존의 GOP 기반 비트 할당 기법^[5~7]들이 영상의 복잡도가 일정하다는 가정하에 GOP나 프레임에 균등한 비트를 할당하여 화질의 변동 폭이 커지는 것을 문제로 삼았다. 이것을 해결하기 위해 그는 잔여신호를 나타내는 MAD를 영상의 복잡도의 척도로 사용하여 복잡도와 비트 (R) 사이의 관계를 모델화하였다. 이러한 모델에 기반 하여 복잡도가 큰 프레임에는 많은 비트를 할당하고 복잡도가 작은 프레임에는 적은 비트를 할당하여 화질의 변동 폭을 완화시키고 평균 화질을 향상시키고자 하였다. 또한 Jiang등은 위의 방법과 같이 프레임의 복잡도로 MAD를 그대로 사용하여 H.264^[10] 부호기의 프레임 비트 할당을 수행하였다^[8]. 이들 기법은 프레임간의 복잡

도를 이용하여 비트를 할당함으로써 기존의 비트 할당 기법에 비해 프레임 간 화질의 변화를 적절히 완화시켰다. 그러나 이들 기법은 MAD만을 프레임 잔여 신호의 부호화 복잡도로 삼았기 때문에, 잔여 신호의 고차원 통계 특성에 기반 한 실제 부호화 복잡도 추정에 그 성능이 제한적이게 되고, 따라서 움직임이 다양하거나 복잡도 변화가 큰 비디오에서는 화질의 변동에 의한 품질 열화를 적절히 완화하기 어려운 한계를 가진다.

본 논문에서는 이와 같은 기존 연구의 한계를 극복하기 위한 새로운 비트 할당 기법을 제안한다. 제안된 비트 할당 기법은 프레임 잔여 신호에 대한 R-MAD 사이의 비례 관계가 비디오 시퀀스 별로 크게 다른 기울기를 가진다는 사실에 기반 한다. 이는 동일한 MAD를 지니더라도 시퀀스 별로 발생하는 비트는 큰 차이가 있음을 의미한다. 이와 같은 기울기 특성을 모델 파라미터라고 정의할 수 있고, 따라서 본 논문에서는 모델 파라미터와 MAD를 함께 반영하여 일정 화질을 위한 비트 할당 기법을 제안한다.

본 논문의 구성은 다음과 같다. II장에서는 기존 연구의 문제점을 살펴보기 위해 R-MAD 모델 기반의 비트 할당 기법을 소개한다. III장에서는 이를 해결하기 위해 본 논문에서 제안한 모델 파라미터 기반 비트 할당 기법을 설명하고 IV장에서는 양자화 파라미터 결정 방법을 간단히 살펴본다. 마지막으로 V장에서는 항등 전송을 채널 환경에서의 제안 기법의 성능을 모의실험을 통해 평가, 검증하고 VI장에서 본 논문의 결론을 맺도록 한다.

II. 일정 화질을 위한 기존의 비트 할당 기법

본 장에서는 기존 연구의 문제점을 구체적으로 살펴보기 위해 실시간, 저-지연 응용에서 일정 화질 비트 할당에 대한 대표적인 기법인 R-MAD 모델 기반의 비트 할당 기법을 살펴본다. 또한, 비디오 시퀀스 별, 양자화 단계 별로 R-MAD 모델의 타당성을 검증하고 그 내재적 문제점을 도출한다.

1. R-MAD 모델에 기반한 비트 할당 기법

Xie 등^[4]은 일정 화질 부호화를 위해 R-MAD 모델에 기반 하여 영상의 특성을 고려한 프레임 레이어 비트 할당 기법을 제안하였다. 즉, 프레임 단위의 잔여 영상의 절대값인 MAD를 그 프레임의 부호화 복잡도 척도로 활용하여, 복잡도가 큰 영상에 많은 비트를 할당

함으로써 프레임 간 화질 변동의 폭을 줄일 수 있도록 한 것이다. 그는 MPEG-4^[9] 부호화 코덱을 이용하여 R과 MAD의 관계를 확인하고 식 (1)과 같은 관계식으로 정의하였다.

$$R = K \times (MAD - \alpha)^\beta \quad (1)$$

고정된 양자화 크기를 사용하여 모든 시퀀스를 부호화한 후, MAD와 R의 값을 구한 후 얻어진 R과 MAD의 관계에 따라 α , β 를 0과 0.5로 각각 결정하여 식 (2)와 같이 R-MAD 모델을 정립하였다.

$$R = K \sqrt{MAD} \quad (2)$$

Xie는 식 (2)의 R-MAD 모델을 통해 식 (3)과 같은 비트 할당 기법을 제안하였다. 그는 식 (3)을 통해 이전 프레임의 복잡도와 현재 프레임의 복잡도의 비를 이용하여, 좀 더 복잡하고 변화가 큰 프레임에 많은 비트를 할당하고 복잡도가 적고 움직임의 변화가 적은 비교적 평탄한 프레임에는 적은 비트를 할당하여 일정 화질을 달성 하도록 하였다.

$$T_n = C \sqrt{\frac{MAD_n}{MAD_{n-1}}} \quad (3)$$

여기서 C 는 비트 율과 프레임 율의 비율(bit rate / frame rate)을 의미하며, $\overline{MAD_{n-1}}$ 은 이전 프레임까지의 모든 프레임 MAD 값의 평균을 의미한다. 또한 식 (2)의 모델 파라미터 K 는 $C / \sqrt{\overline{MAD_{n-1}}}$ 로 간략하게 표현됨을 알 수 있다.

2. 기존의 R-MAD 모델의 한계

R-MAD 모델은 Xie의 논문에서도 언급됐듯이 장면 변화가 크지 않은 영상에서는 화질의 변동을 줄이고 평균 화질을 향상시킬 수 있으나, 복잡도가 크고 장면 전환이 빈번하게 일어나는 영상에서는 그와 같은 효과를 나타낼 수 없으며, 특히 장면 전환에서의 화질의 급격한 변동을 막을 수 없다는 한계를 갖는다.^[4] 본 논문에서는 위와 같은 한계가 영상의 복잡도를 MAD만으로는 정확하게 반영할 수 없고 따라서 현재 프레임과 이전 프레임들의 MAD의 비율로만으로는 효과적인 비트 할당을 할 수 없기 때문이라고 판단하였다. 즉 동일한 MAD를 지니더라도 시퀀스나 양자화 파라미터가 다르면 R의 값은 크게 다를 수 있다는 의미이다. 우리는 이 문제를 실제 확인하기 위해 다양한 부호화 조건아래에

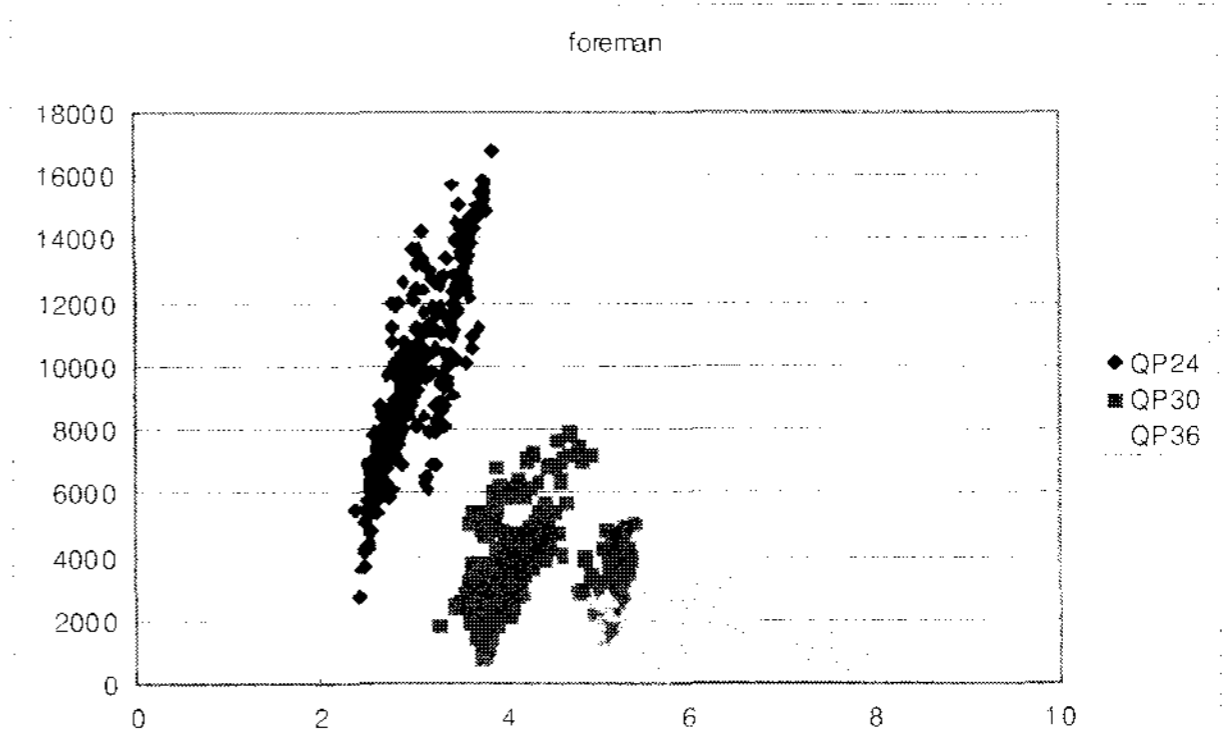


그림 1. Foreman(QCIF)영상의 QP(24,30,36)별 R-MAD 관계

Fig. 1. R-MAD relationship of Foreman (QCIF) per QP.

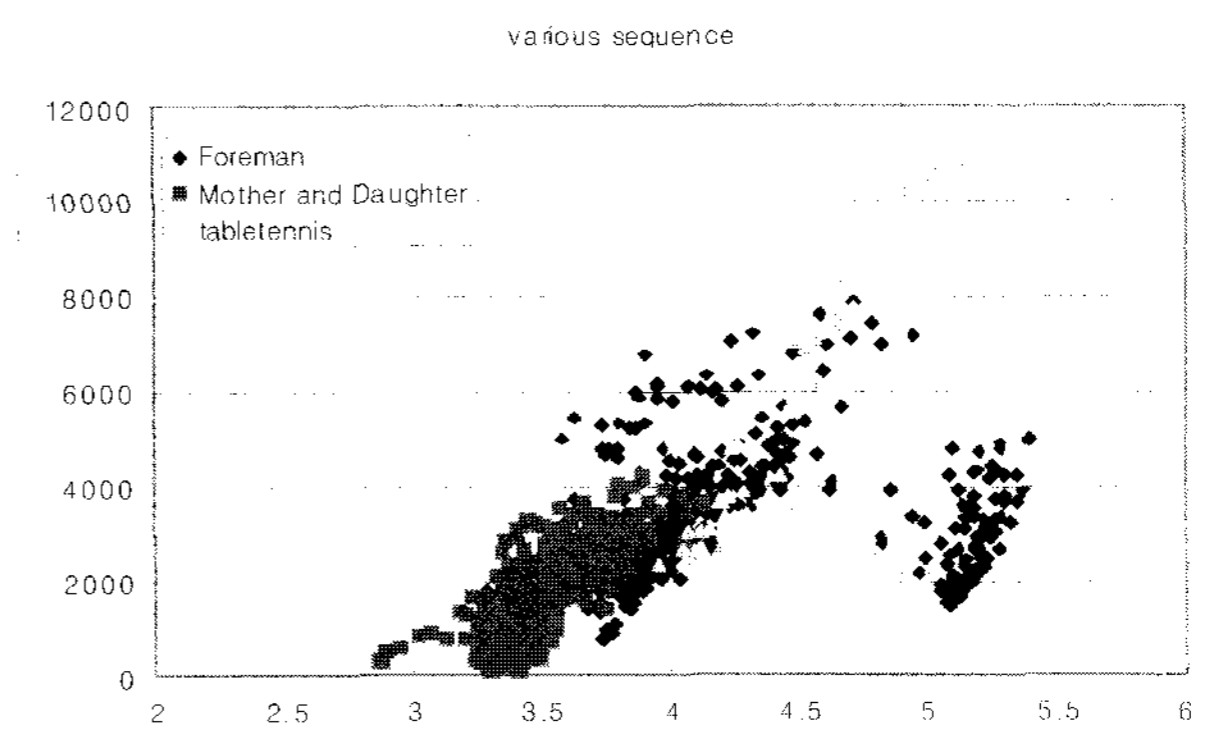


그림 2. QP 24일 때 영상별 R-MAD 관계

Fig. 2. R-MAD relationship of QP.24 per images.

서 실험을 수행하였다. 실험은 H.264의 레퍼런스 소프트웨어인 JM10.2^[12]에서 수행되었으며, 자세한 실험 결과는 아래와 같다.

가. 양자화 파라미터(QP)에 따른 R-MAD 관계

양자화 파라미터는 영상의 품질과 출력 비트(R)를 결정하는 주요한 파라미터이다. 따라서 다양한 양자화 파라미터에서 R-MAD의 관계를 확인할 필요가 있다. 우리는 Foreman 영상에 대하여 양자화 파라미터를 각각 24,30,36으로 설정한 후, 부호화를 수행하여 R-MAD 관계를 산출하고 그 결과를 그림 1에 나타내었다. 그림 1에서 볼 수 있듯이 양자화 파라미터의 변화에 따라 R-MAD의 분포가 크게 다르며 동일한 MAD라도 발생되는 비트의 양은 크게 차이가 남을 알 수 있다. 일반적으로 영상을 부호화함에 있어 버퍼나 채널 용량의 제한 때문에 한 개의 양자화 파라미터로만 시퀀스 전체를 부호화할 수 없다. 따라서 양자화 파라미터가 변하면 그림에서 보는 것과 같이 일관된 R-MAD의 관계를 얻을 수 없으므로 MAD만으로 영상의 부호화 복잡도를 판단

하기에는 한계가 있다.

나. 영상의 종류에 따른 R-MAD 관계

양자화 파라미터를 한 개의 값으로 고정하고, 영상의 종류를 변화시킴으로써 영상의 특성에 따른 R-MAD의 관계를 확인할 수 있다. 그림 2는 양자화 파라미터는 일정하게 한 후 여러 영상에서 R과 MAD의 관계를 보여주고 있다. 결과를 살펴보면 영상의 종류에 따라 동일한 MAD를 가지더라도 생성되는 비트의 양은 매우 큰 차이를 지님을 알 수 있다. 일반적으로 동영상은 여러 개의 장면으로 구성되어 있으므로 MAD만으로 부호화 복잡도를 판단하기에는 무리가 있음을 알 수 있다.

위와 같이 우리는 양자화 파라미터나 영상의 종류가 달라지면 R-MAD의 분포가 크게 바뀌어 동일한 MAD를 지니더라도 비트(R)는 크게 달라짐을 확인할 수 있었다. 동일한 MAD를 지니더라도 비트가 차이가 나는 이유는 식 (2)의 모델 파라미터 K의 차이로 이해될 수 있다. 즉 출력 비트는 단순히 MAD에만 비례하는 것이 아니며 영상의 특성을 고려해야 하며 이것은 모델 파라미터에 그 특성이 나타난다고 할 수 있다. 그러나 기존의 연구는 모델 파라미터를 직접적으로 고려하지 않고 단순히 MAD의 비만으로 모델 파라미터를 간략화하였고 따라서 효과적인 비트 할당을 수행할 수 없어 움직임이 심하거나 순간적인 장면 변환이 일어날 때 화질의 열화가 발생하였다.

III. 모델 파라미터 기반 비트 할당 기법

이 장에서는 위에서 언급한 모델 파라미터를 고려하지 못하여 효과적으로 비트 할당을 수행하지 못하는 기존 알고리즘의 문제점을 해결하기 위하여 모델 파라미터를 이용하여 비트 할당을 수행하였다. 또한 기존의 R-MAD 모델로 산출한 모델 파라미터는 양자화 파라미터에 따라 그 값이 크게 변하여 그 값으로 영상의 내재적 복잡도를 파악하는 것에는 어려움이 있으므로, 본 논문에서는 기존의 R-MAD 모델 식에 양자화 파라미터를 추가하여 양자화 파라미터에 관계없이 영상의 내재적 복잡도를 측정할 수 있게 하는 모델 파라미터를 산출하였다. 우리는 이와 같은 방법으로 산출될 모델 파라미터 K가 영상의 복잡도를 잘 대표할 수 있는 측도로 보고, 이 산출된 K 값을 바탕으로 현재 프레임의 비트 할당을 수행하여 일정 화질을 얻고자 하였다. 또한 모델 파라미터는 모든 부호화를 거쳐야 얻을 수 있

는 값이므로 저-지연, 실시간 응용에 제약이 따르므로 한 번의 부호화로 모든 과정을 수행할 수 있게 하기 위해 모델 파라미터를 예측하는 알고리즘을 제안하였다.

1. 비트 할당을 위한 모델

양자화 파라미터에 따른 R-MAD 관계는 그림 1에서 확인할 수 있듯이 양자화 파라미터에 따라서 그 기울기가 크게 변함을 알 수 있다. 그 기울기는 식 (2)의 모델 파라미터를 의미하므로 동일한 영상임에도 불구하고 양자화 파라미터에 따라 모델 파라미터가 크게 달라진다. 그러므로 이것은 식 (2)로 산출한 모델 파라미터가 영상의 고유한 특성을 대변할 수 없다는 것을 의미한다. 따라서 우리는 양자화 파라미터에 관계없이 영상의 고유한 특성, 즉 내재적 특성을 반영하기 위하여 양자화 파라미터를 고려한 R-MAD의 모델을 세웠다. 우선 양자화 파라미터가 일정하다고 가정하면 R과 MAD의 관계는 식 (1)과 같이 나타낼 수 있다. Xie는 식 (1)에서 MPEG-4에서 한 실험 결과를 바탕으로 식 (2)와 같은 모델을 만들었다. 그러나 본 논문에서 H.264를 대상으로 R과 MAD의 관계를 확인해 본 결과, 그림 1과 그림 2에서 볼 수 있듯이 선형 모델에 더욱 잘 맞음을 알 수 있다. 따라서 H.264에서 R과 MAD는 선형 관계를 갖는다고 할 수 있다. 그리고 일반적으로 R과 양자화 파라미터의 관계에서 R은 양자화 파라미터에 반비례하는 것으로 나타내어진다^[13]. 따라서 이들 R과 MAD와 양자화 파라미터의 관계를 하나의 모델로 표현하면 식 (4)와 같이 나타낼 수 있다.

$$R = K \times \frac{MAD}{Q^2} \quad (4)$$

여기서 Q는 양자화 파라미터를 나타내고 Q가 고정되어 있다면 식 (1)로 표현 될 수 있다.

식 (4)에서 모델 파라미터 K는 R과 MAD/Q^2 사이의 기울기로 볼 수 있으며 이것은 영상의 복잡도의 측도로 이해할 수 있다. 우리는 실험을 통해 기존의 R-MAD 모델에 기반한 모델 파라미터와 제안한 식 (4)에 기반한 모델 파라미터를 실험을 통해 비교하였다. 표 1은 R-MAD 모델과 제안한 모델인 식(4)로 산출한 모델 파라미터의 평균을 각각 나타낸다. R-MAD의 모델의 경우 표 1에서 볼 수 있듯이 양자화 파라미터가 커짐에 따라 모델 파라미터의 값이 점점 작아짐을 알 수 있다. 그에 비해 식 (4)에 기반하여 산출된 모델 파라미터의 값은 양자화 파라미터에 따라 그 값이 크게

표 1. 모델 파라미터의 비교 (식 (4)에서 Q는 실제 양자화 크기를 사용)

Table 1. Comparison of model parameter.

| | | QP24 | QP30 | QP36 |
|-------------|----------|--------|--------|--------|
| R-MAD 모델 | foreman | 3065 | 791 | 228 |
| | salesman | 1250 | 332 | 86 |
| | susie | 2153 | 574 | 186 |
| 식(4) | foreman | 306117 | 315969 | 365451 |
| | salesman | 124194 | 131999 | 137829 |
| | susie | 213912 | 228113 | 294873 |

변하지 않음을 알 수 있다. 따라서 이것은 제안한 모델로 산출된 모델 파라미터가 영상의 내재적 복잡도를 표현할 수 있음을 알 수 있다.

2. 모델 파라미터에 기반한 비트 할당 기법

위에서 우리는 제안한 모델로 모델 파라미터를 구해야 하는 이유를 설명하였다. 이 절에서는 제안한 모델을 사용하여 구한 모델 파라미터를 기반으로 프레임 비트 할당 기법을 제안한다. 모델 파라미터는 현재 프레임의 복잡도를 나타내는 파라미터이므로 모델 파라미터가 크다면 많은 비트를 할당하여야 하고 작다면 적은 비트를 할당해야만 한다. 우리는 모델 파라미터의 크고 작음을 판단하기 위해 이전 프레임의 모델 파라미터와 현재 프레임의 모델 파라미터의 비로써 현재 프레임의 복잡도를 판단하고 그 비율로 비트 할당을 수행한다. 마찬가지로 MAD도 프레임의 복잡도를 판단할 수 있는 하나의 측도이므로 MAD까지 고려하여 비트 할당을 수행하였다. 따라서 우리는 식 (5)와 같이 모델 파라미터 K와 MAD에 기반하여 비트 할당을 수행하는 기법을 제안한다.

$$T_n = \frac{K_n}{K_{n-1}} \times \frac{MAD_n}{MAD_{n-1}} \times C \quad (5)$$

여기서 n 은 n 번째 프레임을, T_n 은 n 번째 프레임에 할당되어진 비트를 나타낸다. 그리고 K_n 과 MAD_n 은 각각 n 번째 프레임에서의 모델파라미터와 MAD를 나타낸다. 또한 $\overline{K_{n-1}}$ 과 $\overline{MAD_{n-1}}$ 은 첫 번째 프레임부터 $n-1$ 번째 프레임까지의 평균 모델 파라미터와 평균 MAD를 의미한다. C 는 비트 율과 프레임 율의 비율(bit rate / frame rate)이며 모든 프레임에 균등한 비트를 분배했을 때 한 프레임에 할당할 수 있는 비트를 의미한다.

식 (5)에서 이전 프레임들의 평균 비트인 $\overline{R_{n-1}}$ 대신

C 를 사용한 이유는 잘못된 비트 할당이나 양자화 파라미터의 선택으로 인해 이전 프레임에서 적절한 비트로 부호화를 하지 못했을 때, 현재 프레임까지 그 오류가 전파되는 것을 방지하기 위해서이다. 또한 버퍼의 제한 때문에 이전 프레임들의 평균 비트는 결국 C 로 수렴한다. 따라서 우리는 $\overline{R_{n-1}}$ 대신 C 를 사용하였으며 이것은 안정적인 비트 할당을 수행하게 할 수 있다.

식 (5)에서 보는 것과 같이 제안한 비트 할당 기법은 현재 프레임과 이전 프레임들만의 모델 파라미터와 MAD의 비로써만 이루어졌다. 이전 프레임들과 현재 프레임의 복잡도가 매우 크거나 작다면 이 비율들은 매우 크거나 작아질 것이고, 따라서 매우 큰 비트나, 매우 적은 비트가 할당되어질 수 있다. 복호화기의 버퍼를 고려하지 않고 일정 화질만을 위해 비트 할당을 한다면 프레임에 매우 많은 비트나 적은 비트를 할당하여도 상관이 없겠으나 실제로는 버퍼를 고려해야 하므로 버퍼의 넘침이나 고갈을 방지하기 위해 매우 큰 비트나, 매우 적은 비트의 발생을 제한해야한다. 비트 할당을 제한하기 위해서는 모델 파라미터 K와 MAD가 생성될 수 있는 범위를 관찰해야한다. 우리는 실험을 통해 여러 영상에서 모델 파라미터와 MAD값의 최대값과 최소값을 확인하여 그 비율의 최대치를 표 2와 같이 계산하였다. 결과를 살펴보면 MAD의 경우 편차가 작지만, 모델 파라미터의 경우 그 편차가 매우 큼을 알 수 있다. 따라서 만약 모델 파라미터의 비를 식 (4)와 같이 사용할 경우 복잡한 영상에서 과도한 비트를 할당해 버퍼를 침해할 수 있는 경우가 생길 수 있다. 따라서 우리는 그런 과도한 비트 할당을 막기 위해 식 (6)과 같이 모델 파라미터의 비에 로그(log) 형태의 함수를 취하였다. 이와 같은 형태는 모델 파라미터로 인해 너무 적은 비트를 할당하여 버퍼가 고갈되는 현상을 방지하여 준다.

$$T_n = \sqrt{\frac{K_n}{K_{n-1}} \times \frac{MAD_n}{MAD_{n-1}}} \times C \quad (6)$$

식 (6)을 사용하여 현재 프레임의 비트 할당을 수행하기 위해서는 현재 프레임의 모델 파라미터가 필요하다. 그러나 현재 프레임의 모델 파라미터를 구하기 위해서는 실제 R, MAD와 Q의 값이 필요하고 이 중 R의 값은 최소한 한 번의 부호화를 수행해야만 그 값을 얻을 수 있다. 이것은 최소한 두 번의 부호화를 수행해야 하기 때문에 실시간 부호화를 어렵게 한다. 이 문제를

표 2. 영상 별 K, MAD값의 최대, 최소와 그 비율 (QP = 30)

Table 2. minimum/maximum value of K and MAD. (QP = 30)

| | | 최대값 | 최소값 | 최대값/ 최소값 |
|----------|-----|---------|-------|-------------|
| Foreman | MAD | 7.124 | 3.276 | 2.175 |
| | K | 1135439 | 87936 | 12.912 |
| Salseman | MAD | 8.526 | 3.574 | 2.386 |
| | K | 1102312 | 25200 | 43.743 |
| Susie | MAD | 6.409 | 2.757 | 2.325 |
| | K | 855258 | 54639 | 15.653 |

해결하기 위해 본 논문에서는 현재 프레임의 모델 파라미터를 이전 프레임의 모델 파라미터를 사용하여 선형적으로 예측하였다. 이것은 식 (7)과 같이 나타낼 수 있다.

$$\hat{K}_n = a \times K_{n-1} + b \quad (7)$$

여기서 \hat{K}_n 은 n 번째 프레임에서의 예측된 모델 파라미터이고 K_{n-1} 은 $n-1$ 번째 프레임의 모델 파라미터이다. 일반적으로 현재 프레임은 이전 프레임과 가장 상관도가 높다고 예상할 수 있으므로 식 (7)과 같이 모델 파라미터를 예측하였다. 그리고 a 와 b 는 계산의 편의를 위해 1과 0으로 각각 결정하였다.

위의 예측 모델을 식 (6)에 적용시키면 최종적으로 식 (8)과 같은 비트 할당 모델을 만들 수 있다. 식 (8)은 일정 화질 부호화를 위하여 본 논문에서 제안한 최종

비트 할당식이다.

$$T_n = \sqrt{\frac{\hat{K}_{n-1}}{K_{n-1}}} \times \frac{MAD_n}{MAD_{n-1}} \times C \quad (8)$$

위와 같이 우리는 프레임 비트 할당 기법을 제안하였다. 다음으로 우리는 버퍼의 넘침이나 고갈을 방지하기 위해 비트 할당을 제한하는 과정이 필요하다. 우리는 식 (8)를 통해 현재 부호화 하고자 하는 프레임에 비트 할당을 수행한 후, 식 (9)을 사용하여 버퍼 제어를 수행한다.

$$\begin{aligned} & \text{If}(B_c + T_n - C \geq B_s - B_1) \\ & \quad T_n = C \times 0.9; \\ & \text{else if}(B_c + T_n - C \geq B_s - B_2) \\ & \quad T_n = T \times 0.9; \\ & \text{else if}(B_c + T_n - C < 0) \\ & \quad T_n = C - B_c \end{aligned} \quad (9)$$

여기서 B_c 는 현재 버퍼의 상태를, B_s 는 복호기 버퍼의 크기를 의미한다. B_1 과 B_2 는 각각 버퍼의 넘침을 두 단계로 제어하기 위한 파라미터로 식 (10)과 같이 계산할 수 있다.

$$\begin{aligned} B_1 &= \min(0.5 \times B_r, B_s \times 0.05) \\ B_2 &= \min(B_r, B_s \times 0.1) \end{aligned} \quad (10)$$

여기서 B_r 은 초당 목표 비트 율을 의미한다. 위와 같은 버퍼 제어 방법을 제안한 이유는 H.264의 버퍼가 기존의 비디오 압축 표준인 MPEG-4나 H.263에 비해

표 3. 제안 알고리즘과 기존 알고리즘의 성능 비교 (PSNR, PNSRmdev)

Table 3. Comparison of performance between previous algorithms and proposed algorithm (PSNR, PNSRmdev).

| 영상 종류 | 목표 비트 율(Kbps) | JVT-H014 ^[11] | | | R-MAD 모델 기반 ^[4] | | | 제안 알고리즘 | | |
|----------------|---------------|--------------------------|-------|-----------|----------------------------|-------|-----------|----------|-------|-----------|
| | | bit-rate | PSNR | PSNR mdev | bit-rate | PSNR | PSNR mdev | bit-rate | PSNR | PSNR mdev |
| Foreman (QCIF) | 64 | 64.14 | 31.73 | 1.16 | 67.59 | 31.93 | 0.97 | 67.44 | 32.01 | 0.96 |
| M & D (QCIF) | 64 | 64.00 | 37.06 | 1.62 | 62.07 | 36.82 | 1.29 | 64.06 | 36.95 | 0.92 |
| Susie (QCIF) | 48 | 47.97 | 34.86 | 1.76 | 48.54 | 34.98 | 1.30 | 48.19 | 34.89 | 0.82 |
| Football (CIF) | 384 | 383.90 | 28.58 | 1.04 | 403.08 | 28.67 | 0.84 | 386.31 | 28.56 | 0.49 |
| Glasgow (CIF) | 512 | 511.88 | 33.60 | 3.44 | 484.04 | 32.95 | 2.79 | 513.88 | 33.24 | 1.72 |

복호기 버퍼의 크기가 매우 크기 때문이다. 따라서 더 유연한 버퍼 제어가 가능하므로 우리는 위와 같이 두 단계로 나누어 버퍼 제어를 수행하였다.

IV. 양자화 파라미터 결정

비트 할당을 수행한 후, 목표한 비트에 맞는 양자화 파라미터를 선택하여야 한다. 본 논문에서는 양자화 파라미터를 결정하기 위해 식 (4)의 모델을 사용하였다. 식 (4)에서 R을 목표 비트 T_n 으로 가정하면 식 (4)는 다음과 같이 양자화 크기에 관한 식으로 변환할 수 있다.

$$Q_{step}(n) = \sqrt{\frac{K_n \times MAD_n}{T_n}} \quad (11)$$

여기서 $Q_{step}(n)$ 은 n 번째 프레임의 실제 양자화 값을 나타낸다. 앞서도 언급했지만 식 (4)를 이용하여 K를 구할 때 실제의 양자화 값을 사용하여 K의 값을 산출하였다. 따라서 식 (11)에서 구한 양자화 값인 $Q_{step}(n)$ 을 H.264의 양자화 파라미터인 $QP(n)$ 로 변환해야 한다. Q_{step} 과 QP 의 관계는 식 (12)에 나타나 있다. 최종적으로 식 (12)를 통하여 구하여진 $QP(n)$ 을 사용하여 n 번째 프레임을 부호화한다.

$$Q_{step}(n) = 2^{(QP(n) - 4)/6} \quad (12)$$

IV. 실험 및 결과

본 논문에서 제안한 비트 할당 기법의 성능을 검증하기 위해서 JVT-H014^[11]에서 제안된 비트율 제어 기법 및 Xie의 R-MAD에 기반한 비트율 제어 기법과 비교, 검증하였다. JVT-H014는 현재 JM10.2 레퍼런스 소프트웨어에서 채택한 비트율 제어기법으로 일정 화질을 고려하지 않고 각각의 프레임에 거의 균등한 비트를 할당한다. Xie의 방법은 II장에서 언급했듯이 MAD를 프레임의 복잡도로 판단하고 MAD의 비로 프레임에 비트 할당을 수행하여 일정 화질을 달성하고자 하였다.

실험은 H.264의 레퍼런스 소프트웨어인 JM10.2^[12]에서 프로파일은 baseline 프로파일, 레벨은 1.3으로 설정하여 수행하였다. 시퀀스 구조는 I프레임과 연속된 P프레임으로만 이루어진 구조를 사용하였으며 한 프레임을 한 개의 양자화 파라미터로 부호화하였다. 실험영상으

로는 QCIF(176x144) 크기의 영상과 CIF(352x288)크기의 영상을 선택하였으며, QCIF 영상으로는 Foreman (400 프레임), Susie (150 프레임), Mother & Daughter (950 프레임)을 사용하였고 CIF 영상으로는 Glasgow (700 프레임)을 사용하였다.

제안한 알고리즘의 성능을 객관적으로 평가하기 위해 화질의 변동 정도를 나타낼 수 있는 PSNRmdev (PSNR mean absolute deviation)를 평가 측도로 사용하였다. PSNRmdev란 식 (13)과 같이 각각의 프레임의 화질과 평균 화질과의 차이의 정도를 나타낸다. 이 값이 클수록 화질의 변동이 큼을 의미하고, 작을수록 화질의 변동이 적어 일정 화질을 더욱 잘 구현하는 것으로 판단할 수 있다.

$$PSNRmdev = \frac{1}{N} \sum_{i=1}^N |PSNR_i - \overline{PSNR}| \quad (13)$$

여기서 N 은 총 프레임수를 \overline{PSNR} 은 평균 PSNR을 나타낸다.

표 3은 기존 알고리즘과 제안한 알고리즘의 총체적인 실험 결과를 보여준다. 실험 결과를 살펴보면 모든 영상에 걸쳐 제안한 알고리즘은 기존 알고리즘에 비해 더 낮은 PSNRmdev를 보임을 알 수 있다. 즉 제안한 알고리즘이 기존 알고리즘에 비해 더욱 일정한 화질을 보이는 것이다. 특히 Glasgow와 같이 장면의 변화가 심한 영상에서는 2배 이상 PSNRmdev가 감소함을 알 수 있다. 평균 화질 측면에서는 기존 알고리즘과 비슷한 화질을 보인다.

각각의 영상에 대한 프레임 별 화질은 그림 3에 나타나 있다. 그림 3에서 볼 수 있듯이 모든 실험 영상에서 제안한 알고리즘이 기존의 알고리즘에 비해 더욱 일정한 화질을 보여줌을 알 수 있다. 특히 기존 알고리즘에서 급격한 화질의 저하가 일어나는 프레임에서 제안 방법은 그 열화 정도가 덜함을 알 수 있다. 그림 4의 (d)는 Glasgow의 실험 결과를 보여주며 이 영상은 화질의 변동이 매우 심함을 알 수 있다. 이것은 Glasgow가 장면 변환이 심하기 때문에 장면 변환이 일어날 때마다 화질의 변동이 매우 큰 것이다. 그러나 제안 알고리즘에서는 기존 알고리즘에 비해 장면 변환이 일어나는 부분에서도 화질의 변동이 훨씬 작음을 알 수 있다. 특히 Xie의 알고리즘은 장면 변환 부분에서 PSNR의 갑작스러운 저하를 보인다. 이것은 보통 장면 전환에서는 인트라 모드로 부호화가 수행되며 일정한 화질을 유지하기 위해서는 매우 많은 비트를 필요로 하지만, Xie의

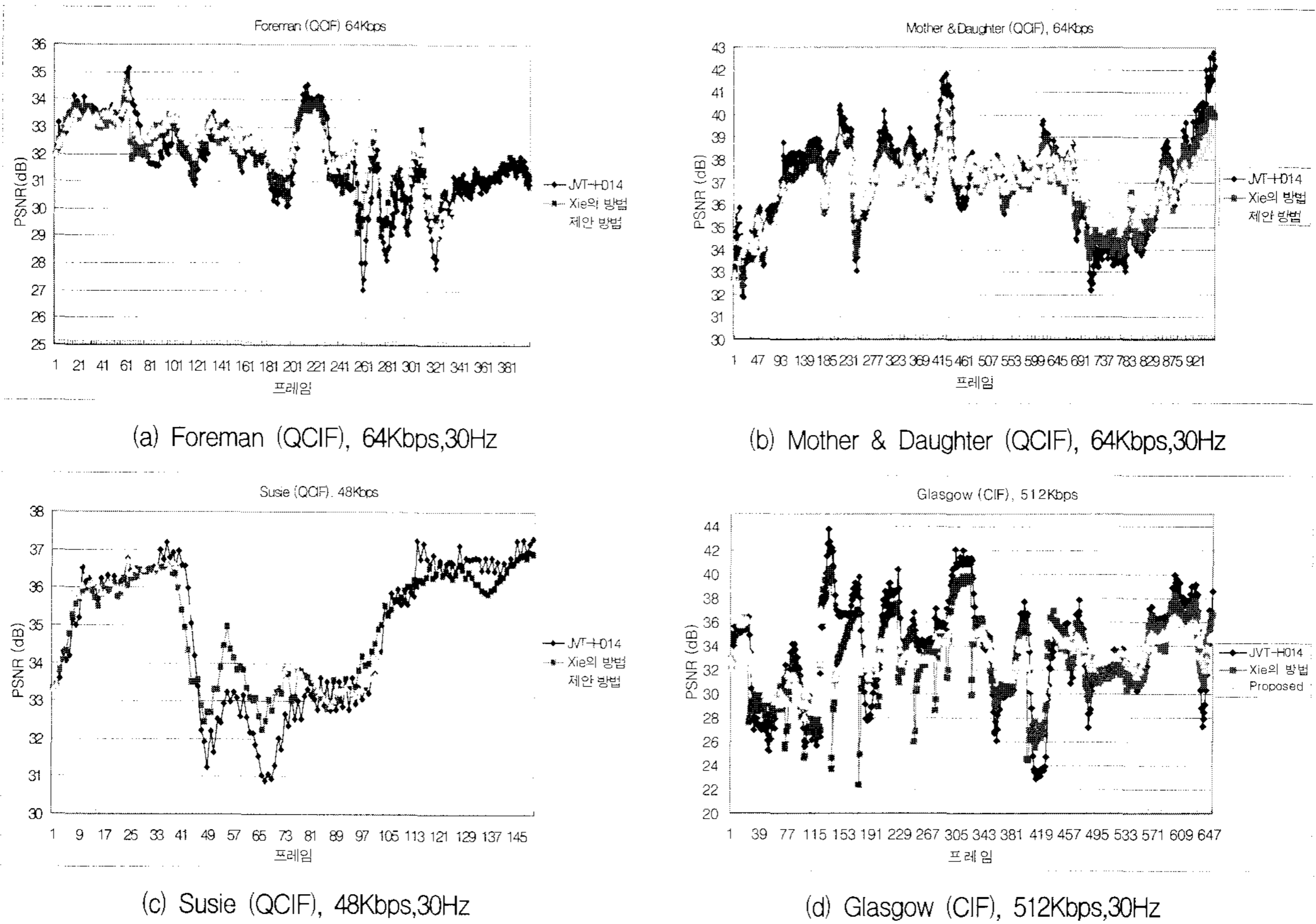


그림 3. 다양한 영상에서 프레임별 PSNR
 Fig. 3. PSNR variation for various sequences.

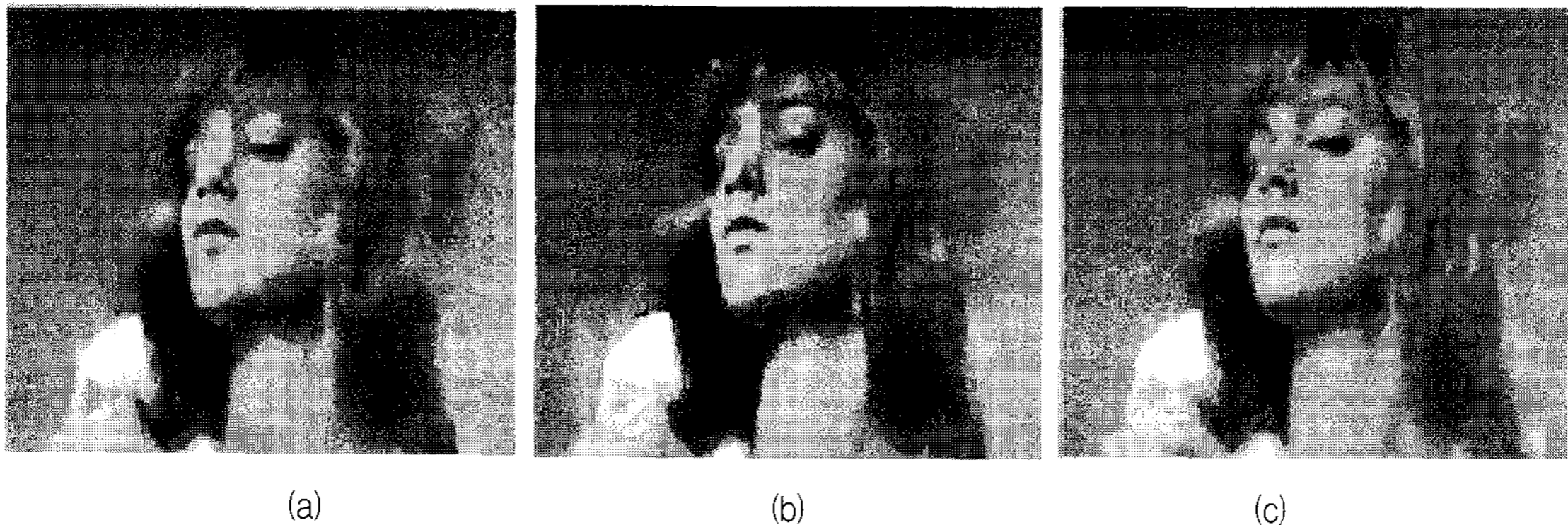


그림 4. Susie 영상의 49번째 프레임 (a) JVT-H014를 이용 (b) R-MAD 모델 이용 (c) 제안 알고리즘을 이용
 Fig. 4. 49th frame of Susie (a) using JVT-H014 (b) using R-MAD model (c) using proposed algorithm.

알고리즘에서는 그렇게 하지 못하고 있음을 보여준다. 이는 MAD로만으로는 영상의 복잡도를 적절하게 판단할 수 없음을 보여준다. 또한 일반적으로 사람은 영상의 화질을 판별할 때 가장 나쁜 화질일 때의 화질로 그 시퀀스 전체의 화질을 평가하는 경향이 있는 것으로 알려져 있다. 제안한 기법에서는 전체 시퀀스에서 최저 화질이 다른 알고리즘에 비해 최소 2dB 이상 높고, 이 사실은 제안 알고리즘이 기존 알고리즘에 비해 더욱 높은 주관적 화질을 제공한다는 것을 보여준다.

그림 4는 기존 알고리즘과 제안한 알고리즘으로 부호화 했을 때의 영상을 보여준다. 이 부분은 그림 3 (c)의 susie 영상에서 40에서 49프레임에 걸쳐 화질이 급격히 저하되는 부분에서 화질이 가장 저하됐을 때를 비교한 그림이다. 살펴보면 기존 알고리즘에 비해 제안한 알고리즘은 얼굴이나 머리카락의 뭉개짐이 훨씬 덜 함을 알 수 있다. 이것은 주관적으로도 더 좋은 화질을 제공한다.

그림 5는 Glasgow 영상에서 기존 알고리즘과 제안한

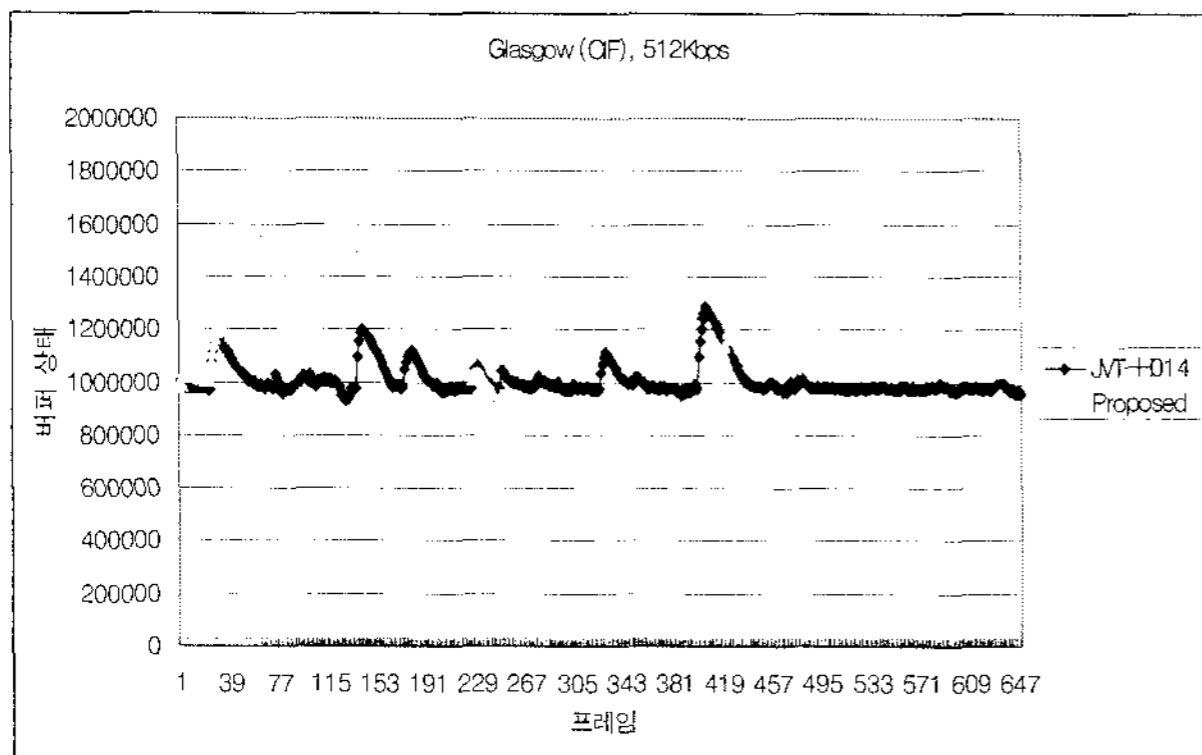


그림 5. Glasgow 영상의 버퍼 상태 비교
Fig. 5. Comparison of buffer level variation of Glasgow.

알고리즘의 버퍼 상태를 보여준다. 버퍼의 크기는 H.264의 레벨 1.3에서 정의된 2 Mbit로 삼았으며 초기 지연은 2초로 설정하였다. Glasgow는 많은 장면 변화가 있어 버퍼 제어가 어려우므로 제안한 알고리즘의 성능을 적절히 검증할 수 있다. Glasgow의 버퍼 상태를 보면 제안한 알고리즘의 버퍼의 변동 폭이 매우 크음을 알 수 있다. 이것은 기존의 알고리즘이 프레임마다 비트를 일정하게 소모하는 것에 비해, 제안 알고리즘은 화질을 일정하게 하기 위하여 프레임마다 소모하는 비트의 편차가 크기 때문이다. 제안한 알고리즘은 비록 버퍼의 변동 폭이 클지라도 버퍼의 고갈이나 넘침은 발생하지 않는다. 오히려 H.264가 가지고 있는 큰 버퍼를 효율적으로 사용할 수 있음을 알 수 있다. 그러나 화상 통화 같이 버퍼의 크기가 극단적으로 작은 환경에서는 제안한 알고리즘의 성능이 제한적일 수 있다.

IV. 결 론

본 논문에서는 모델 파라미터에 기반하여 실시간 부호화가 가능하며 일정화질을 만족하는 비트율 제어 기법을 제안하였다. 제안한 알고리즘은 다음과 같은 두 가지 기여가 있다. 첫 번째는 모델 파라미터는 영상의 복잡도를 판단할 수 있는 중요한 측도가 될 수 있다는 것을 보여주었다. 따라서 우리는 모델 파라미터의 비로 프레임에 비트를 할당하여 기존의 알고리즘에 비해 일정 화질을 달성할 수 있었다. 두 번째 기여는 모델 파라미터가 영상의 내재적 복잡도를 표현할 수 있게 하는 모델을 찾아내었다. 이것은 기존의 비트-복잡도 모델이 양자화 파라미터를 고려하지 못하여 기존 모델로 산출한 모델 파라미터가 양자화 파라미터가 변함에 따라 크게 바뀌어 영상의 내재적 특성을 반영하지 못함을 해결

하였다. 제안한 알고리즘은 이와 같은 두 가지 생각을 바탕으로 발전되었으며 실험결과는 제안한 알고리즘이 기존 알고리즘에 비해 화질의 변동폭을 크게 감소시켰음을 보여준다. 특히 제안한 알고리즘은 갑작스러운 장면의 변화에서도 화질의 저하를 기존의 알고리즘에 비해 크게 감소시켜 높은 주관적 화질을 제공하였다.

참 고 문 헌

- [1] Y.Yu, J. Zhou, Y.Wang, and C.Wen Chen, "A Novel Two-Pass VBR Coding Algorithm for Fixed-Size Storage Application", *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 11, no. 3, pp.345-356, March 2001.
- [2] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics.", *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 8, no. 4, pp.446-459, Aug. 1988.
- [3] N. Cherniavsky, G. Shavit, M. F. Ringenburg, R. E. Ladner, and E. A. Riskin, "MultiStage: A MINMAX bit allocation algorithm for video coders", *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 17, No.1, pp.59-67, Jan. 2007.
- [4] Bo Xie and W.Zeng, "A Sequence-based Rate Control Framework for Consistent Quality Real-Time Video", *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 16, No.1, pp.56-71, Jan.2006
- [5] Video codec test model, near-term, TMN8, in ITU Study Group 16, Video Coding Experts Group, Portland, OR, Jun. 1997.
- [6] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 9, no. 1, pp. 172-185, Feb. 1999.
- [7] Coding of Moving Pictures and Associated Audio, ISO/IEC 14 496-2, Oct. 1998.
- [8] M.Jiang, X.Yi and N.Ling, "Frame Layer Bit Allocation Scheme for Constant Quality Video", *IEEE Conf. on ICME*, 2004., pp. 1055-1058
- [9] ISO/IEC, Information Technology - Coding of audio-visual objects - Part 2 : Visual, ISO/IEC 14496-2, Dec.1999
- [10] ISO/IEC 14496-10: "Coding of Audiovisual Objects-Part 10: Advanced Video Coding" 2003, also ITU-T Recommendation H.264 "Advanced video coding for generic audiovisual services"
- [11] IZ. G. Li, W. Gao, F. Pan, S. W. Ma, K. P. Lim, G. N. Feng, X. Lin, S. Rahardja, H. Q. Lu, and

- Y. Lu, "Adaptive Rate Control with HRD Consideration", 8th JVT Meeting, JVT-H014, Geneva, May 2003.
- [12] Joint Video Team(JVT) of ISO/IEC MPEG and ITU-T VCEG, Joint Model Reference Software Version 10.2 (JM10.2)
- [13] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 7 ,no. 2, pp. 246-250, Feb. 1997.

저 자 소 개



정진우(학생회원)
 2004년 연세대학교 전기전자
 공학과 학사 졸업
 2006년 연세대학교 전기전자
 공학과 석사 졸업
 2006년~현재 연세대학교 전기
 전자공학과 박사과정 재학
 <주관심분야 : 영상/동영상 부호화, 웨이블릿, 비
 트율 제어 기법, 비디오 스트리밍>



조경민(정회원)
 2005년 연세대학교 전기전자
 공학과 학사 졸업
 2007년 연세대학교 전기전자
 공학과 석사 졸업
 <주관심분야 : 영상/동영상 부호
 화, 비트율 제어 기법, 비디오 스
 트리밍>



최윤식(정회원)
 1979년 연세대학교 전기공학과
 학사 졸업
 1984년 Case Western Reserve
 Univ. 시스템공학과 졸업
 1987년 Pennsylvania Stat Univ.
 전기공학과 석사 졸업
 1990년 Perdue Univ. 전기공학부 박사 졸업
 1990년~1993년 (주) 현대전자 산업전자 연구소
 책임연구원
 1993년~현재 연세대학교 전기전자공학부 정교수
 <주관심분야 : 영상신호처리, HDTV, 동영상 압
 축>