
GOP 레이어 비트율 제어를 위한 초기 QP 모델링

박상현*

Initial QP Modeling for GOP Layer Rate Control

Sang-Hyun Park*

요약

비디오 코딩에서 GOP의 첫 번째 프레임은 많은 비트를 발생시키는 인트라 모드로 압축되고 다음 프레임의 인터 모드 압축에 사용되기 때문에 첫 프레임을 위한 초기 QP 값은 첫 프레임뿐만 아니라 이후 프레임에도 영향을 주게 된다. 일반적으로 초기 QP 값은 bpp 값에 따라 4가지 값 중에 하나로 설정된다. 이렇게 설정하는 것은 간단한 반면 부정확한 문제가 있다. 정확한 초기 QP 값 예측을 위해서는 bpp 뿐만 아니라 영상의 복잡도와 전송률도 함께 고려하여야 한다. 본 논문에서는 GOP의 인코딩 특성을 분석하여 전체 GOP의 PSNR 값을 최대로 하는 초기 QP 값을 찾기 위한 트래픽 모델과 실시간 영상 압축에서 모델 파라미터를 실시간으로 구하는 방법을 제안한다. 실험 결과는 제안하는 모델이 초기 QP 계산에 필요한 트래픽 특성을 잘 반영하고 있으며 또한 실시간으로 모델 파라미터를 구하는 방법도 효과적으로 작동함을 보여준다.

ABSTRACT

The first frame of a GOP is encoded in intra mode which generates a larger number of bits. In addition, the first frame is used for the inter mode encoding of the following frames. Thus the initial QP for the first frame affects the first frame as well as the following frames. Traditionally, the initial QP is determined among four constant values only depending on the bpp. Although this initialization scheme is simple, yet it is not accurate enough. An accurate initial QP prediction scheme should not only depends on bpp but also on the complexity of the video sequence and the output bandwidth. In this paper, we propose a traffic model for finding the optimal initial QP which maximizes the PSNR of the GOP. We also propose a method to find model parameters for real-time video encoding. It is shown by experimental results that the proposed traffic model captures initial QP characteristics effectively and the proposed method for model parameters accurately estimates the real values.

키워드

Rate Control, Video Compression, H.264/AVC, Initial QP, Traffic Model
비트율 제어, 영상 압축, H.264/AVC, 초기 QP, 트래픽 모델

1. 서론

비디오 코딩에서 비트율 제어는 MPEG, H.263, H.264와 같은 다양한 표준에서 다양한 응용에 맞게

연구가 진행되고 있다 [1, 2]. 비디오 인코더를 구현하기 위해서는 매크로블록 모드, 양자화 파라미터(QP : Quantization Parameter), 프레임 타입, 프레임 레이트와 같은 코딩 파라미터들이 결정되어야 한다. 이러한

* 순천대학교 멀티미디어공학과(shark@snu.ac.kr)

접수일자 : 2012. 06. 29

심사(수정)일자 : 2012. 11. 23

게재확정일자 : 2012. 12. 10

코딩 파라미터들을 어떻게 결정하는지는 비디오 표준에 명시되어 있지 않지만 코딩 파라미터의 설정은 비디오 압축에서 성능에 많은 영향을 주게 된다. 특히 비트율 제어에서 QP 값의 설정은 결과 영상의 비트율을 목표 비트율에 맞추게 하거나 버퍼의 오버플로우와 언더플로우를 방지하는데 중요한 역할을 수행한다 [2, 3].

H.264 비디오 표준은 이전의 표준들에 비해 아주 뛰어난 압축 효율 때문에 최근에 많은 관심을 받고 있다 [4]. H.264 비디오 표준은 다양한 영상 압축 기법들이 사용되었다. 특히 RDO (Rate Distortion Optimization)라고 불리는 비트율-왜곡 (R-D: Rate-Distortion) 최적화 움직임 예측과 압축 모드 선택은 비디오 데이터의 압축율을 크게 향상시킨다. 또한 다양한 화면내 예측과 화면간 예측 모드 및 복수개의 참조 영상 사용 등으로 압축율을 극대화한다 [5].

기존의 많은 비트율 제어 알고리즘들은 최적의 QP를 찾기 위해 RDO와 영상의 통계치 간의 상호 의존적인 관계를 해결하기 위한 모델 개발에 초점을 맞추고 있다. 비트율 제어에서 모델은 중요한 요소이지만 그 외에는 다양한 요소들을 고려하여야 한다. 초기 QP (Initial QP)의 설정도 비트율 제어에서 중요한 요소이다. 많은 H.264/AVC의 비트율 제어 알고리즘에서 초기 QP는 JVT-W057에서 제안된 bpp (Bits Per Pixel)에 따라 설정하는 방법을 그대로 사용하고 있다 [6]. bpp 에 따라 설정하는 방법은 간단하지만 정확하지 않다는 치명적인 단점이 있다. 초기 QP 값은 다음 GOP (Group of Pictures)의 초기 QP 값에 영향을 미치기 때문에 처음 잘못 설정되면 그 효과가 다음 GOP까지 파급되어 전체 영상에 걸쳐 손실을 초래한다.

초기 QP 설정 문제는 영상의 처음뿐만 아니라 장면 전환에서도 적용된다. H.264/AVC는 인트라 코딩 프레임도 인트라 모드로 압축할 수 있다. 즉, 장면 전환이 발생하면 장면 전환된 첫 프레임은 이전 프레임과 연관성이 매우 낮기 때문에 이전 프레임을 이용한 영상 압축이 불가능해지고 따라서 인트라 모드로 압축되게 된다. 이때도 정확한 초기 QP 값이 설정되지 않으면 그 효과가 다음 GOP까지 파급되어 전체 영상의 품질을 저하시킨다 [7].

본 논문에서는 저전송률로 영상을 압축할 때 초기 QP에 따른 GOP의 PSNR 특성을 분석하고 이를 바

탕으로 효과적으로 QP를 결정하기 위한 트래픽 모델을 제안한다. 이를 위해 JVT에서 제공되는 참조 소프트웨어 중 최신 버전인 JM 18.4를 사용하였다. JM 18.4에서도 초기 QP를 결정하는 방법은 JVT-W057에서 제안된 방법을 그대로 사용한다. 먼저, 기존 방법의 문제점을 분석하고 이를 바탕으로 초기 QP 모델을 제안한다. 그리고 제안하는 모델의 파라미터를 효과적으로 결정하는 알고리즘을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 JVT 알고리즘에서 사용되는 초기 QP 값 결정 알고리즘에 대해서 설명한다. 3장에서는 제안하는 초기 QP 모델과 실시간으로 모델 파라미터를 결정하는 알고리즘을 설명하고, 4장에서는 제안하는 알고리즘의 성능을 분석한다. 마지막으로 5장에서 결론을 맺는다.

II. 연구 배경

H.264/AVC의 비트율 제어는 JVT-W057에 제시되어 있다 [6]. 비트율 제어는 레이어별로 구성되며 본 논문에서는 GOP 레이어 비트율 제어에 대해서 설명한다. JVT-W057에서는 각 GOP의 첫 프레임과 두 번째 프레임은 비트율 제어를 적용하지 않고 고정 QP 값을 사용한다. 세 번째 프레임부터 비트율 제어가 적용된다.

첫 GOP의 경우 초기 QP 값을 결정하는 방법은 다음과 같다. 영상 시퀀스의 시작인 첫 프레임은 이용할 수 있는 이전 프레임의 정보가 없기 때문에 단순히 프레임 레이트, 화면 크기, 목표 대역폭에 따라서 식 (1)과 같이 결정된다 [6].

$$bpp = \frac{Target\ bandwidth}{Frame\ rate \times Frame\ size} \quad (1)$$

여기서 bpp 는 픽셀당 비트 수를 나타내고, $Target\ bandwidth$ 는 목표 대역폭을, $Frame\ size$ 는 한 프레임의 화소 수를 나타낸다. 이렇게 구해진 bpp 에 따라 식(2)와 같이 첫 GOP의 초기 QP 값 ($QP_1(1)$)이 결정된다.

표 1. PSNR이 최대일 때 QP 값 비교
Table 1. Comparison of QP values when PSNR is maximum

영상	전송속도	$PSNR_{max}$	ΔQP
Akiyo	60kbps	42.18	7.14
	80kbps	43.80	7.45
	100kbps	44.89	7.38
Carphone	60kbps	35.19	2.55
	80kbps	36.44	4.03
	100kbps	37.49	4.66
Salesman	60kbps	35.16	4.79
	80kbps	36.97	5.62
	100kbps	38.25	6.10

$$QP_1(1) = \begin{cases} 40 & bpps \leq l_1 \\ 30 & l_1 < bpps \leq l_2 \\ 20 & l_2 < bpps \leq l_3 \\ 10 & bpps > l_3 \end{cases} \quad (2)$$

영상의 크기에 따라에 권장하는 l_1, l_2, l_3 값은 표준 안에 정의되어 있다 [6].

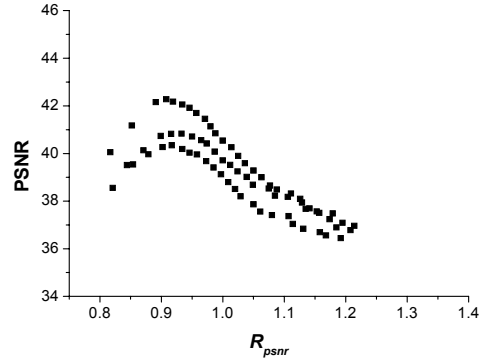
두 번째 GOP 부터는 첫 번째 프레임과 두 번째 프레임의 QP 값이 식 (3)과 같이 결정된다.

$$\begin{aligned} QP_1(i) &= \overline{PQP}(i-1) - \min\left(2, \frac{N(i-1)}{15}\right), \\ QP_1(i) &= \min\{QP_1(i-1) + 2, QP_1(i)\}, \\ QP_1(i) &= \max\{QP_1(i-1) - 2, QP_1(i)\}. \end{aligned} \quad (3)$$

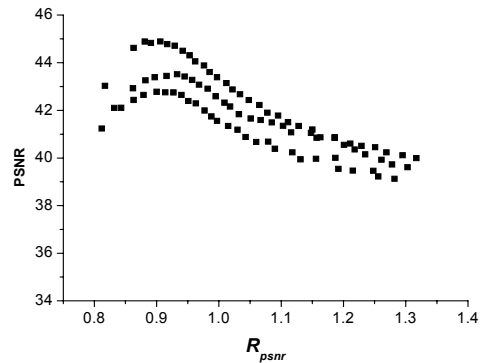
여기서 $QP_1(i)$ i 번째 GOP의 초기 QP이고 $N(i)$ 는 i 번째 GOP의 크기를 나타내며, $\overline{PQP}(i)$ 는 i 번째 GOP의 P 프레임 압축에 사용된 QP의 평균이다. 초기 QP는 이전 GOP의 P 프레임 압축에 사용된 QP의 평균값에서 GOP 크기를 15로 나눈 값을 뺀 값으로 설정된다. 이때 GOP 크기를 15로 나눈 값이 2보다 클 수는 없다. 이 값은 다시 이전 GOP의 초기 QP와의 차이가 2보다 크지 않고 수정된다. 수정된 값을 현재 GOP의 첫 번째 프레임과 두 번째 프레임을 압축하기 위한 QP 값으로 사용한다.

JVT에서 사용하는 GOP 레이어 비트율 제어에서는 식 (1)에 의해 초기 QP가 이전 GOP의 P 프레임

에 사용된 QP들의 평균 (\overline{PQP})과의 차이가 2 이하가 되도록 설정된다. 이 방법이 적절함을 평가하기 위해서 다양한 영상에 대해서 여러 초기 QP를 적용하여 실험을 수행하였다. 표 1은 각 영상이 60kbps, 80kbps, 100kbps로 압축했을 때 PSNR이 최대일 때의 값 ($PSNR_{max}$)과 그때의 초기 QP와 \overline{PQP} 의 차이 (ΔQP)를 보여준다. 표 1에서 알 수 있듯이 PSNR이 최대가 될 때의 초기 QP와 \overline{PQP} 의 차이는 2를 넘어 간다. 그리고 영상에 따라 전송 속도에 따라 그 차이가 다양함을 알 수 있다.



(a)



(b)

그림 1. R_{psnr} 비교 (a) 64kbps (b) 100kbps
Fig. 1 Comparison of R_{psnr} (a) 64kbps (b) 100kbps

III. 제안하는 트래픽 모델

본 논문에서는 다양한 영상에 대해서 여러 초기 QP 값을 이용하여 코딩한 결과를 바탕으로 I 프레임과 P 프레임의 PSNR 비율 (R_{psnr})에 대해서 분석을 수행하였다.

한 GOP에 대해서 R_{psnr} 는 식 (4)와 같이 정의된다.

$$R_{psnr} = \frac{PSNR_P}{PSNR_I} \quad (4)$$

여기서 $PSNR_I$ 는 I 프레임의 PSNR이고 $PSNR_P$ 는 GOP의 P 프레임들의 평균 PSNR이다.

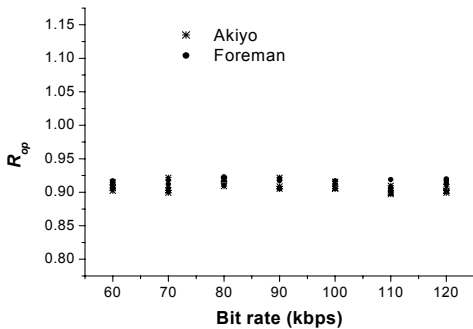


그림 2. R_{op} 비교
Fig. 2 Comparison of R_{op}

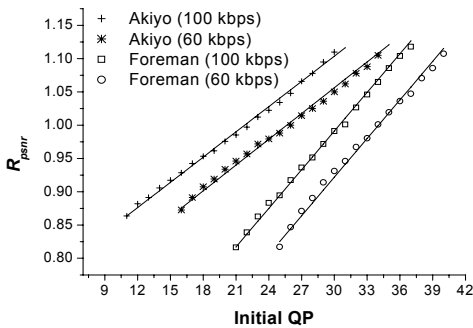


그림 3. R_{psnr} 와 초기 QP 사이의 관계
Fig. 3 Relations between R_{psnr} and initial QP

그림 1은 Akiyo 영상에 대해서 60kbps와 100kbps로 압축했을 때 R_{psnr} 에 따른 GOP의 PSNR을 보여준다. 전송 속도별로 세 개의 GOP를 선택하여 실험을 수행하였다. 그림에서 볼 수 있듯이 영상에 따라

PSNR 값을 달라진다. 같은 영상일 때도 GOP가 다른 PSNR 값은 역시 다른 값을 가진다. 하지만 GOP의 PSNR이 최대가 될 때의 R_{psnr} 값 (R_{op})은 비슷함을 알 수 있다. 이 결과를 바탕으로 전송 속도별 R_{op} 값을 실험하였다. 그림 2는 Akiyo 영상과 Foreman 영상에 대해서 전송 속도를 60kbps에서 120kbps까지 변화시키며 측정된 R_{op} 를 보여준다. 그림에서 R_{op} 는 영상과 전송속도에 상관없이 0.91에서 0.95 사이의 값을 가지는 것을 알 수 있다. 즉, R_{op} 는 영상과 전송속도에 상관없이 비교적 일정한 값을 가짐을 알 수 있다. 따라서 R_{op} 와 초기 QP 사이의 관계를 모델링 할 수 있으면 쉽게 최적의 초기 QP 값을 구할 수 있다.

그림 1과 그림 2의 결과를 바탕으로 R_{psnr} 와 초기 QP와의 관계를 분석하였다. 그림 3은 Akiyo 영상과 Foreman 영상에 대한 R_{psnr} 과 초기 QP 사이의 관계를 보여준다. 그림에서 볼 수 있듯이 영상에 따라 다른 특성을 보여주고 또 전송 속도에 따라서도 다른 특성을 보여준다. 하지만 각각의 그룹이 선형의 형태를 띠는 것을 알 수 있다. 따라서 본 논문에서는 R_{psnr} 과 초기 QP 사이의 관계에 대해서 식 (5)와 같은 선형 모델을 제안한다.

$$QP_1(i) = a \cdot R_{psnr} + b \quad (5)$$

여기서 a 와 b 는 모델 파라미터이다.

이 모델을 이용하면 최적의 초기 QP를 쉽게 구할 수 있다. 그림 2에서 R_{op} 는 거의 일정함을 알 수 있었다. 예를 들어, 0.91에서 0.95의 범위에서 I 프레임과 P 프레임의 PSNR 차이를 적게 하기 위해 0.95로 설정할 수 있을 것이다. R_{op} 결정되고 모델 파라미터를 알면 최적의 초기 QP를 식 (5)를 이용하여 구할 수 있게 된다.

식 (5)의 선형 모델에서 중요한 것은 모델 파라미터를 어떻게 결정하느냐이다. 실시간 압축에서는 데이터를 미리 분석할 수 없기 때문에 실시간으로 적용할 수 있는 적절한 알고리즘이 필요하다. 본 논문에서는 Recursive Least Squares Estimation (RLSE) 방법 [8]을 이용하였다. RLSE를 식 (5)에 적용하기 위해서는 다음 식에 맞게 상태를 정의하여야 한다.

$$y_i = H_i \cdot x_i + v_i, \quad (6)$$

제안하는 알고리즘에서는 식 (7)과 같이 상태를 정의하고 정의된 상태를 바탕으로 필요한 파라미터도 정의하였다.

$$x_i = \begin{pmatrix} a \\ b \end{pmatrix},$$

$$y_i = QP_1(i), \quad (7)$$

$$H_i = (1, R_{psnr}(i)),$$

$$E(v_i^2) = R_i.$$

여기서 값을 예측해야 하는 것은 x_i 이다. x_i 에 대한 예측값을 \hat{x}_i 라고 하면 예측값은 식 (8)과 같이 재귀적으로 구할 수 있다.

$$K_i = P_{i-1}H_i^T(H_iP_{i-1}H_i^T + R_i)^{-1},$$

$$\hat{x}_i = \hat{x}_{i-1} + K_i(y_i - H_i\hat{x}_{i-1}), \quad (8)$$

$$P_i = (I - K_iH_i)P_{i-1}(I - K_iH_i)^T + K_iR_iK_i^T.$$

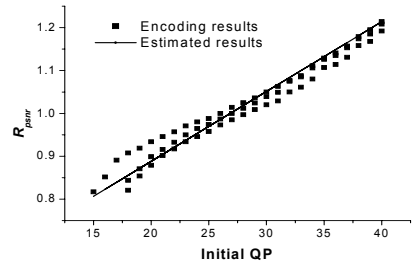
여기서 K_i 은 이전 예측값에 어느 정도 측정값을 반영하여 현재 예측값을 결정할지를 나타내는 가중치이고 P_i 는 예측 에러에 대한 공분산이다.

IV. 실험 결과

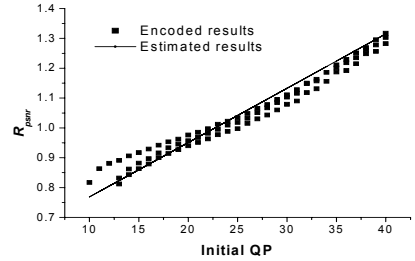
제안하는 트래픽 모델의 성능을 분석하기 위해서 다양한 영상과 다양한 전송률에 대해서 실험을 수행하여 제안하는 파라미터 예측 방법이 실시간으로 모델 파라미터를 정확하게 예측하는지를 분석하였다. 모델 파라미터를 예측하기 위한 RLSE에서는 초기값을 설정하여야 한다. 다양한 영상에 대한 실험을 통해 a 의 경우 40의 값을 초기값으로 설정하였다. b 의 경우 첫 GOP를 압축한 후 결과값을 식 (5)에 대입하여 결정하였다.

R_{psnr} 와 초기 QP 사이의 관계를 확인하기 위해서

각 실험 영상에서 세 개의 GOP를 선택하여 다양한 QP로 압축하여 결과를 분석하였다. 파라미터 예측에 대한 실험에서는 JVT 알고리즘을 그대로 적용하여 초기 QP 값을 계산하였고, 그 값을 이용하여 매 GOP를 압축하고 압축된 결과를 바탕으로 제안하는 모델의 파라미터를 예측하였다. 본 논문에서는 최적의 초기 QP를 계산하기 위한 초기 QP 모델에 대한 것으로 실험은 제안하는 모델에 대한 파라미터 계산 방법이 타당함을 검증하는데 초점을 두고 있다.



(a)



(b)

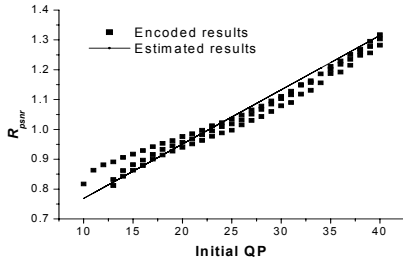
그림 4. Akiyo 영상에 대한 실험 결과 (a) 64kbps (b) 100kbps

Fig. 4 Experimental results for Akiyo sequence (a) 64kbps (b) 100kbps.

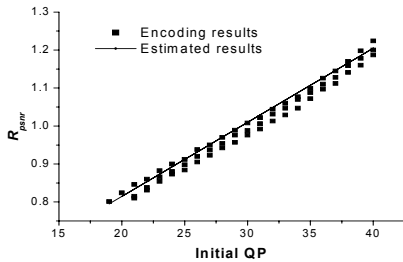
그림 4는 Akiyo 영상에 대해서 전송속도가 60kbps와 100kbps일 때의 결과이다. 각 점들은 다양한 초기 QP에 대해서 압축한 결과의 R_{psnr} 값을 나타낸다. 실선의 경우 제안한 파라미터 예측 알고리즘을 이용하여 예측한 결과를 식 (5)에 적용하여 그린 것이다. 파라미터는 10개의 GOP를 압축한 후의 값을 사용하였다. 그림 5, 6, 7은 그림 4와 동일한 조건으로 Carphone, Salesman, Container 영상에 대해서 실험

한 결과를 보여준다.

모든 실험 결과는 초기 QP와 R_{psnr} 사이에 선형 관계가 있음을 보여주고 있다. 또한, 제안하는 모델 파라미터 예측 방법이 영상이 가지는 선형 특성을 정확하게 예측함을 보여준다. 실험 결과에 제시된 결과 외에도 다양한 영상에 대해서 실험을 수행하였고 모든 결과에서 유사한 결과를 확인할 수 있었다.

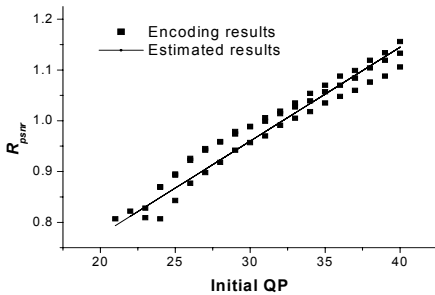


(a)

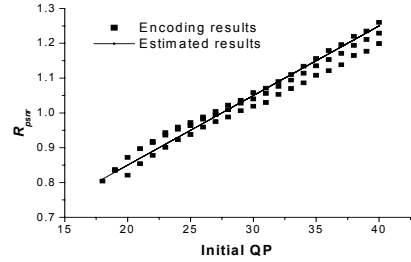


(b)

그림 5. Carphone 영상에 대한 실험 결과 (a) 64kbps (b) 100kbps
Fig. 5 Experimental results for Carphone sequence (a) 64kbps (b) 100kbps

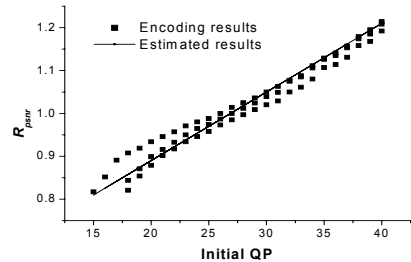


(a)

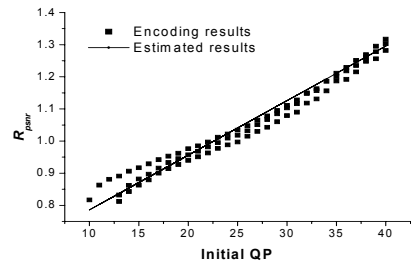


(b)

그림 6. Salesman 영상에 대한 실험 결과 (a) 64kbps (b) 100kbps
Fig. 6 Experimental results for Salesman sequence (a) 64kbps (b) 100kbps.



(a)



(b)

그림 7. Container 영상에 대한 실험 결과 (a) 64kbps (b) 100kbps
Fig. 7 Experimental results for Container sequence (a) 64kbps (b) 100kbps

V. 결론

본 논문에서는 H.264 비트율 제어에서 GOP 레이 어 비트율 제어에 해당하는 초기 QP 값 결정에 필요

한 초기 QP 모델과 모델 파라미터를 실시간으로 예측하기 위한 알고리즘을 제안하였다. JVT에서 사용하는 초기 QP 값 결정 알고리즘은 영상의 복잡도를 고려하지 않으며, 영상의 크기와 프레임 레이트, 대역폭만을 고려하여 네 개의 범주로 나누고 각 범주마다 하나의 값을 할당하기 때문에 최적의 초기 QP 값을 찾기가 어렵다. 제안하는 모델은 여러 GOP에 대해서 다양한 초기 QP를 이용하여 압축한 결과를 바탕으로 제안되었으며, 영상의 특성 및 전송 속도에 대한 특성을 효과적으로 반영할 수 있다. 실험결과에서 제안하는 모델이 다양한 영상의 초기 QP 특성을 잘 반영하고 있음을 확인하였고, 파라미터 예측 알고리즘도 정확하게 해당 영상의 특성을 예측함을 알 수 있었다.

제안하는 모델은 R_{op} 값이 결정되면 간단하게 최적의 초기 QP값을 구해준다. R_{op} 의 경우 단순히 GOP의 PSNR을 최대로 하는 R_{op} 로 정의할 수 있지만 응용에 따라서는 I 프레임과 P 프레임의 PSNR 차이를 줄이는 방향으로 설정할 수 있다. 응용에 맞게 R_{op} 를 설정한 후 제안하는 모델을 이용하여 목적에 맞게 초기 QP를 결정할 수 있기 때문에 제안하는 모델은 다양한 응용에 활용될 것으로 기대된다.

참고 문헌

[1] Z. Chen and K.N. Ngan, "Recent advances in rate control for video coding," *Signal Process.: Image Commun.*, Vol 22, pp. 19-38, 2007.
 [2] 신윤정, 손남례, N.D. Toan, 이귀상, "움직임벡터 군집화를 이용한 H.264/AVC에서 MPEG-2로의 트랜스코딩", *한국전자통신학회논문지*, 5권, 1호, pp. 23-30, 2010.
 [3] 광성우, 최홍, 양정민, "HSPDA 기반 실시간 영상 전송 및 위치 인식 시스템", *한국전자통신학회논문지*, 7권, 1호, pp. 21-26, 2012.
 [4] 윤홍준, 서영호, 최현준, 김동욱, "H.264/A VC의 고속 인트라 예측 방법," *한국해양정보통신학회논문지*, 10권, 7호, pp. 1172-1179, 2006.
 [5] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 7, No. 7, pp. 1-19, 2003.

[6] K. Lim, G. Sullivan, and T. Wiegand, "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods," *JVT-W057*, San Jose, CA, 2007.
 [7] H. Wang and S. Kwong, "Rate-Distortion Optimization of Rate Control for H.264 With Adaptive Initial Quantization Parameter Determination," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 18, No. 1, pp. 140-144, 2008.
 [8] P.C. Young, "Recursive Estimation and Time-Series Analysis," Springer-Verlag Heidelberg, pp. 29-45, 2011.

저자 소개

박상현(Sang-Hyun Park)



1995년 고려대학교 전자공학과 졸업 (공학사)

1997년 고려대학교 대학원 전자공학과 졸업(공학석사)

2002년 고려대학교 대학원 전자공학과 졸업(공학박사)

2004년 순천대학교 컴퓨터공학과 교수

※ 관심분야 : 영상처리, 패턴인식