

모바일 비디오 응용을 위한 적응적 공간 해상도 제어 인코딩 기법

정회원 이희정*, 이용희*, 이정훈*, 종신회원 신현식*

A Video Encoding Scheme using Adaptive Spatial Resolution Control for Mobile Video Applications

Heejung Lee*, Yonghee Lee*, Jonghun Lee* *Regular Members*,
Heonshik Shin* *Lifelong Member*

요약

모바일 비디오 스트리밍을 위해 비디오 스트림들을 사용가능한 대역폭에 맞추면서 인코딩을 하기 위해서 시간적 해상도, 공간적 해상도, 화면 화질의 세 가지 요소를 가지고 조절할 수 있는데, 그 중에서도 양자화 파라미터 (quantization parameter)를 가지고 화면 화질을 조절하는 것이 가장 널리 사용되고 있는 방법이다. 그러나 본 논문에서는 낮은 대역폭 환경에서는 적응적으로 공간적 해상도를 변경하는 것이 화면 화질이나 에너지 소모량 면에서 더 효율적이라는 것을 보이고, 사용 가능한 대역폭에 대해서 최적의 공간적 해상도를 찾기 위한 모델을 제시한다. 적응적 공간 해상도 제어 기법은 비디오 서버와 모바일 기기 간의 대역폭이 심하게 변화거나, 모바일 기기가 에너지 소모에 민감한 상황에서 특히 효과적이다. 본 기법은 낮은 비트율 상황에서 기존의 방법에 비해 화질을 약 0.5dB 만큼 개선하고, 에너지 소모를 50% 이상 줄일 수 있다.

Key Words : spatial resolution, wireless, video streaming, H.264/AVC

ABSTRACT

Video streams for mobile video streaming can be encoded to fit the available network bandwidth by controlling three factors: temporal resolution, spatial resolution, and picture quality. The controlling of picture quality by modifying the quantization parameter (QP) is most widely used. In this paper, we demonstrate that reducing the spatial resolution adaptively can be more efficient in terms of picture quality and energy consumption in low bit-rate environment, and present a model to find the optimal spatial resolution for the available bandwidth. Adaptive spatial resolution control scheme is especially effective when the bandwidth between the video server and the mobile device varies considerably with time, and when the mobile device is sensitive to energy consumption. Our scheme can improve the picture quality by approximately 0.5dB and reduce energy consumption by more than 50% compared to the conventional video coding in low bit-rate environment.

I. 서론

휴대용 기기 상에서 비디오 응용을 재생하기 위

해서는 많은 계산량과 에너지 소모로 인해서 어려움을 겪는다. 그러나 계속되는 기술 발전으로 인하여 이러한 문제들이 점점 극복되어 나가고 있으며,

※ 이 논문은 정부(교육과학기술부)의 재원으로 BK21프로젝트와 한국과학재단의 지원을 받아 수행된 연구임 (No. 2009-0080823).

* 서울대학교 전기컴퓨터공학부 컴퓨터시스템 연구실 (hjlee ; yhlee ; terminal@cslab.snu.ac.kr, shinhs@snu.ac.kr)

논문번호 : KICS2008-12-564, 접수일자 : 2008년 12월 23일, 최종논문접수일자 : 2009년 6월 26일

갈수록 많은 비디오 서비스들이 휴대용 기기 상에서 사용 가능해 지고 있다. 또한 HSDPA와 WiBro와 같은 무선 통신 기술의 발전은 휴대용 미디어 재생기가 단순히 저장된 내용에 국한되지 않고 스트리밍 비디오를 재생하는 것을 가능하게 만들었다. 그러나 무선 통신망을 통하여 모바일 비디오 스트리밍 서비스를 제공하기 위해서는 유선 통신망에 비해서 대역폭이 낮고 불안정하다는 문제에 대응해야만 한다.

모바일 비디오 스트리밍을 위해 비디오 스트림들을 사용가능한 대역폭에 맞추면서 인코딩을 하기 위해서 시간적 해상도, 공간적 해상도, 화면 화질의 세 가지 요소를 조절할 수 있다. 시간적 화질의 척도로는 일반적으로 초당 프레임 수 (frames per second) 가 사용되는데, 각 시간 간격 내에서 더 많은 프레임이 보여 질수록 장면 내의 움직임이 더 부드럽게 보이게 된다. 공간적 화질은 주로 화면 내의 화소 수로 표현될 수 있다. 일반적으로 시간적, 공간적 해상도는 미리 결정되어 주어지고 인코딩 과정에서 동적으로 결정되지 않는다. 비디오 인코더는 시간적, 공간적 해상도가 주어졌을 때 사용가능한 비트율에 맞추기 위해서 양자화 파라미터 (quantization parameter)를 수정함으로써 화면 화질을 조절할 수 있는데, 이 방법은 MPEG1, MPEG2, MPEG4, H.264 등의 널리 사용되는 비디오 코딩 표준에 사용되었다^{1, 2}. 더 큰 양자화 파라미터를 적용할수록 더 많은 데이터가 압축 과정에서 버려지게 되고, 결과적으로 더 많이 압축되지만 더 낮은 화면 품질을 보여준다.

사용 가능한 대역폭의 변경이 상당히 심한 환경에서 화면 품질을 너무 떨어뜨리는 대신에 초당 프레임 수를 동적으로 조절하는 방법들이 연구되어 왔다³⁻⁵. 그러나 비교적 낮은 비트율 환경에서 초당 프레임 수가 너무 떨어지게 되면 움직임이 매우 부자연스럽게 보이게 되고 주관적 화질에서 안 좋은 결과를 낳는다. 반면에, 동적인 공간 해상도 조절은 비교적 덜 주목을 받아왔는데, 화면의 해상도 변화는 사용자에게 시각적으로 매우 거슬리기 때문이다. 그러나 사용자가 화면 크기를 계속 고정하여 본다면, 효율적인 전송을 위해서 화면의 해상도를 줄여서 인코딩 할 수 있다. 모바일 휴대용 기기는 일반적으로 디스플레이 크기가 매우 작기 때문에, 보통 원본 화면 크기에 관계없이 고정된 크기로 비디오 콘텐츠를 보게 된다. 이러한 상황에서는 낮은 해상도로 비디오를 인코딩 하여 전송하고, 모바일 기기

에서 디코딩 후 원래 해상도로 확대함으로써, 사용자가 인지하는 화질 저하가 더 적으면서 대역폭의 효율적인 사용이 가능하다.

또한 휴대용 기기는 보통 외부 전원 대신에 제한된 용량을 가진 배터리를 사용하기 때문에, 에너지 소모가 많은 멀티미디어 응용을 동작시키기 위해서는 여러 가지 에너지 절감 기법들이 적용된다. 화질의 저하가 없는 기법 중에서 DVS (Digital Voltage Scaling) 가 대표적이인데, 실행 시간에 프로세서의 동작 전압을 변화시켜서 수행 속도와 에너지 소모를 조절한다 [6]. 화질의 저하를 감수한다면, 재양자화 (Requantization)이나 낮은 QoS 파라미터를 할당하는 방법 등이 제안되었다^{7, 8}.

본 논문에서는 사용 가능한 대역폭이 낮은 환경에서 높은 양자화 파라미터 값을 사용하는 대신에, 적응적으로 공간적 해상도를 조절하는 비디오 인코딩 스킴을 제안한다. 이를 위해 화질과 에너지 소모를 예측할 수 있는 모델을 제시하고, 사용가능한 비트율과 사용자의 에너지 관심도에 따라 최적의 공간 해상도를 결정하는 방법을 제시한다. 제시된 적응적 공간 해상도 조절을 이용하는 모바일 비디오 스트리밍 시스템은 대역폭을 효율적으로 사용함으로써 화질을 개선할 수 있고, 디코더의 연산량을 크게 줄임으로써 모바일 기기의 에너지 소모를 줄일 수 있다.

본 논문의 제 2장에서는 적응적 공간 해상도 조절 스킴과 예측 모델을 제시하고, 제 3장에서는 모바일 비디오 스트리밍 시스템에의 적용을 설명한다. 제 4장에서는 실험을 통해 제시된 스킴을 평가하고, 제 5장에서 결론을 맺는다.

II. 적응적 공간 해상도 조절

공간적 해상도 조절을 통해서 대역폭이 부족한 상황에 대처하고자 하는 노력이 이전 비디오 코딩 표준에서 반영되어 왔다. MPEG4의 Advanced Simple 프로파일에서 제공되는 DRC (dynamic resolution conversion)는 VOP (video object plane) 의 해상도를 줄여서 인코딩 할 수 있게 한다⁹. 줄어든 해상도의 VOP 내의 매크로블록은 디코더에서 복호된 후에 업샘플된다. H.263 Annex Q에서 제공되는 RRU (Reduce Resolution Update) 도 비슷한 기능을 한다¹⁰. 그러나 이런 코딩 툴들은 현재 가장 코딩 효율이 좋아서 널리 사용되고 있는 H.264/AVC 표준에는 포함되어 있지 않다.

최근에 H.264/AVC의 Annex G로 발표된 SVC (scalable video coding) 확장^[11]은 하나의 스트림으로부터 시간적, 공간적, 화질의 3가지 측면의 스케일러빌리티를 제공한다. SVC 스트림은 하나의 기본 계층 (base layer)와 여러 개의 강화 계층(enhancement layer)으로 구성되는데, 강화 계층을 버리고 추출함으로써 손쉽게 원하는 부스트림 (substream)을 만들어낼 수 있다. 따라서 하나의 스트림을 가지고 재인코딩이나 변환코딩 (transcoding) 할 필요 없이 다양한 기기와 가변적인 네트워크 상황에 쉽게 적응하는 것이 가능하지만, 스케일러빌리티를 지원하기 위해서 코딩 효율이 저하된다는 부담을 가지게 된다.

본 논문에서는 현재 압축 효율이 좋기 때문에 가장 널리 사용되고 있는 H.264/AVC 코덱^[12]을 변경하지 않고 이용하는 모바일 비디오 스트리밍 시스템을 목표로 하였다. 본 논문에서 제시하는 적응적 해상도 조절 인코딩 스킴을 이용하는 모바일 비디오 스트리밍 시스템 모델을 그림 1에서 간략하게 제시하고 있다. 비디오 서버 내의 해상도 결정부에서는 현재 비트율과 에너지 상태에 대해서 최적의 화질과 에너지 소모를 가지는 공간적 해상도를 결정하고, 이에 맞추어 입력 이미지의 해상도를 다운 샘플링 한 후 인코딩한다. 디코더에서는 디코딩 결과 이미지를 가지고 목표 해상도로 업샘플링 하여 디스플레이 한다.

이 장에서는 먼저 화질과 에너지 소모 측면에서 최적의 공간적 해상도를 찾기 위한 모델을 제시한다.

2.1 화면 화질과 비트율

비디오 신호의 비트율과 왜곡 간의 관계를 분석적으로 모델링하기 위한 연구는 폭넓게 계속되어 왔다. 이미지의 양자화 (Quantization) 처리 과정에서 손실되는 화질과 비트율 간의 관계를 분석적으로 모델링하기 위해서는 DCT 변환 후의 계수의 분

포에 대한 가정이 필요하다.

초기 연구에서는 DCT 변환 후의 계수 분포를 가우시안 (Gaussian) 형태로 모델링 했으며, 이 경우 율-왜곡 관계식은

$$D = \sigma_x^2 2^{-2R} \tag{1}$$

으로 나타낼 수 있다. 여기에서 D 는 평균 제곱 오차 (MSE), 즉 원본과 복원된 픽셀 데이터 간의 차이의 제곱으로 측정된 값이다. R 은 인코딩 된 비트 스트림의 비트율이고, σ_x^2 은 입력 신호의 분산을 나타낸다^[13].

그러나 일반적으로 DCT 변환 후의 계수 분포는 라플라시안 (Laplacian) 분포로 모델링 되고, 분포와 율-왜곡 관계는 각각

$$R = -\log(\alpha D_M) \tag{2}$$

로 나타낼 수 있다. 여기에서 테일러 확장을 이용하여 더 적용하기 쉬운 형태로 변형하여

$$R = aD^{-1} + bD^{-2} \tag{3}$$

로 많이 사용되고 있다^[14].

최근에는 DCT 변환 후의 계수는 라플라시안 분포보다 다음과 같이 코시 (Cauchy) 분포를 따른다고 주장이 제시되었다^[15]. 이 경우에 율-왜곡 관계는 다음과 같이

$$D = cR^{-\gamma} \tag{4}$$

으로 근사화 할 수 있다.

그러나 화면 화질에 적용되는 더 실질적인 척도

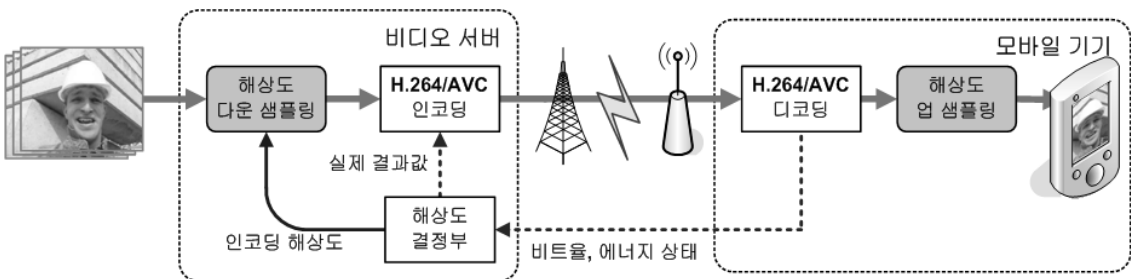


그림 1. 제한된 전체 시스템 구조

는 최대 신호 잡음비 (PSNR) 이고, 다음과 같이 왜곡과 연관되어 있다.

$$PSNR = 10 \log_{10} \frac{255^2}{D} \quad (5)$$

$$= 20 \log_{10} 255 - 10 \log_{10} D$$

여기에 앞에서의 율-왜곡 관계식을 대입하면

$$PSNR = q_1 \log R + q_2 \quad (6)$$

이다. 여기에서 R 은 비트율이고, q_1, q_2 는 실험적으로 얻을 수 있는 상수 값인데, 코딩 방법과 비디오의 내용에 따라 달라진다.

그러나 적응적으로 공간 해상도를 조절하는 경우에는 양자화에서 발생하는 율-왜곡 관계뿐만 아니라, 해상도 축소와 확대에 인해서 발생하는 화질 열화까지도 고려해야 한다. 실험을 통해서 공간적 왜곡 효과는 비트율과 해상도 확대 비율에 따라 거의 비례하는 것으로 나타났다. 따라서 최종적으로 화면 화질은 다음과 같이

$$PSNR = q_1 \log R + q_2 - q_3 (S_r - 1) R \quad (7)$$

로 나타낼 수 있고, 여기서 S_r 은 원본과 축소본 간의 크기의 비율이다. 식 (7)이 실제 결과와 잘 일치한다는 것을 실험을 통해 확인하였다.

2.2 공간적 해상도와 디코딩 시간

비디오의 한 프레임을 복호 하는 데 필요한 시간은 다음과 같이 매크로블록의 수 N_{MB} 와 한 매크로블록의 평균 복호 시간 T_{MB} 의 곱으로 해서

$$T = N_{MB} \times T_{MB} \quad (8)$$

나타낼 수 있다. T_{MB} 는 다음과 같이 두 가지 부분으로 나눌 수 있는데, 매크로블록 헤더 처리와 움직임 벡터복호와 움직임 보상 등을 포함한 고정적인 처리 시간인 M_{MB} , 그리고 가변 길이 DCT 계수 복호, 역 양자화, 역 DCT 와 같이 가변적인 처리 시간인 O_{MB} 이다. 일반적으로 화면 화질은 O_{MB} 에 비례하는데, 이것은 각 매크로블록에 할당되는 비트율에 비례한다. 그래서 한 프레임의 복호 시간은

$$T = N_{MB}(O_{MB} + M_{MB}) \quad (9)$$

$$= N_{MB} \times a \frac{R_f}{N_{MB}} + N_{MB} \times M_{MB}$$

으로 나타낼 수 있고, 여기에서 a 는 상수 값이다. 이 식은 다시 정리하면,

$$T = aR_f + N_{MB}M_{MB} \quad (10)$$

$$= aR_f + b(S_x \times S_y)M_{MB}$$

이고, 여기에서 b 는 상수 값이고, S_x 와 S_y 를 각각 수평 방향과 수직 방향의 픽셀 수라고 했을 때, $N_{MB} = (S_x \times S_y) / (16 \times 16)$ 의 관계를 가진다. 더 일반적으로 표현하기 위해 상수 계수들을 정리 하면

$$T = t_1R + t_2S \quad (11)$$

와 같이 나타낼 수 있다.

그러나 적응적으로 공간 해상도를 조절하는 경우에는 화면에 표시하기 위해 복호 된 이미지를 확대 하는 데 걸리는 부담까지 고려해야만 한다. 일반적으로 확대 알고리즘에서는 결과 영상의 픽셀에서 입력 영상의 픽셀 위치를 구한 후에 그 위치에서의 픽셀값을 참조하는 역방향 매핑을 한다. 따라서 확대 알고리즘의 시간 복잡도는 결과 이미지의 해상도에 비례하고, 입력 이미지의 해상도와의 연관성은 거의 없게 된다. 따라서 확대 연산에 들어가는 계산 복잡도는 비트율과 해상도 비율에 상대적으로 상수 값으로 추가되며, 최종적으로 복호와 화면 표시에 드는 시간 복잡도는 다음과 같다.

$$T = t_1R + t_2S + t_3 \quad (12)$$

이 관계식은 실험을 통해서 실제와 잘 들어맞는다는 것이 입증되었고, 이전의 연구 결과들과도 부합한다^{16, 17}.

2.3 통합 모델

화질과 에너지에 대한 요구 사항은 각 모바일 기기의 사용자마다 중요도가 달라질 수 있다. 따라서 본 논문에서는 비디오 화질과 에너지 소모를 결합해서 고려했고, 다음과 같이 나타낼 수 있다.

$$\Phi = \frac{Q}{E^\omega} \quad (13)$$

여기에서 Φ 는 합계 성능 척도이고, Q 는 화면 화질이고 E 는 디코딩을 포함한 처리 시간이다. 에너지 소모에 대한 사용자가 느끼는 중요도는 상황에 따라 달라지기 때문에 추가적인 변수 ω 를 도입하였다. ω 가 0이면 오직 화면 화질만이 고려 대상이 되고, 이것은 외부 전원을 공급받는 상황이 될 것이다. ω 값이 클수록 에너지 소모를 줄이기 위해서 화면 화질을 희생하는 것을 의미한다.

이 성능 척도에 앞에서의 관계식들을 대입하면,

$$\Phi = \frac{q_1 \log R + q_2 - q_3 (S_r - 1)R}{(t_1 R + t_2 S + t_3)^\omega} \quad (14)$$

가 된다. 이 관계식은 결국 다음과 같은 형태의 함수가 된다.

$$\Phi = f(R, S, \omega) \quad (15)$$

일단 비트율과 에너지 중요도가 결정될 경우에, 최종 성능 척도는 공간적 해상도에만 의존하게 된다.

III. 제한된 비디오 스트리밍 시스템

II장에서 제시된 모델을 가지고, 현재 사용가능한 비트율과 사용자의 에너지 관심도를 반영하여

1. 네트워크 상태를 관측하면서 사용가능한 대역폭을 탐지하고, 사용자의 에너지 관심도를 전달받는다.
2. 대역폭이나 에너지 관심도에 한도 이상의 변화가 발생하면, GOP 경계점에서 각 공간적 해상도에 대해 성능 척도를 계산한다.
 - 2.1. 식 (7)에 따라서 화면 품질을 계산한다.
 - 2.2. 식 (12)에 따라서 복호와 확대에 필요한 에너지를 예측한다.
 - 2.3. 에너지 중요도 ω 를 포함하여 식 (13)에 따라 합계 성능 척도를 계산한다.
 - 2.4. 가장 높은 성능 척도 값을 보여주는 공간적 해상도를 선택한다.
3. 선택된 해상도에 따라 인코딩 또는 변환코딩을 수행한다.
4. 실제 생성된 비트스트림을 가지고 선형 회귀 분석(linear regression analysis)을 통하여 식 (7), (12)의 상수 계수들을 업데이트한다.
5. 1단계로 돌아가서 반복한다.

화질과 에너지 측면을 함께 고려할 때 최적의 공간적 해상도를 구할 수 있다. 공간적 해상도가 결정된 다음에는 기존의 비트율 제어(rate control) 방법을 이용하여 매 프레임의 QP를 결정하여 비트율에 맞춘다.

해상도가 변경되는 상황에서 화면간 예측(inter prediction)을 수행하기 위해서는 이전 프레임의 해상도를 같은 크기로 변경해야 하고, 이것은 H.264/AVC 코덱의 변경을 필요로 한다. 그러나 매 프레임마다 수행되는 QP 조절과 달리, 공간적 해상도 결정은 네트워크 대역폭이나 에너지 관심도가 기준치 이상으로 변동이 있을 때만 수행한다. 따라서 제안된 시스템에서 적응적 공간 해상도 제어는 각 GOP(group of pictures) 경계점에서만 수행하도록 하고, 결과적으로 해상도가 변경되었을 경우에는 I 프레임으로 인코딩이 시작된다. 제안된 비디오 스트리밍 시스템의 동작은 다음과 같이 정리된다.

IV. 실험 결과 및 분석

4.1 실험 환경과 확대 알고리즘

최적의 공간적 해상도를 결정하는 과정을 시뮬레이션 하기 위해서 H.264/AVC 참조 소프트웨어 JM 14.0을 사용하여 실험하였다. CIF (352×288) 과 4CIF (704×576) 크기의 원본 이미지들을 각각 더 작은 해상도로 축소해서 입력으로 사용하여 인코딩하였다.

인코딩 된 비디오 스트림들을 디코딩 한 후 원래 해상도로 확대하는 과정에서 사용되는 알고리즘은 매우 중요한 역할을 한다. 영상의 크기 변환에 관해서는 처리 속도와 결과 품질에 따라서 다양한 종류의 것들이 연구되어 왔다. 본 논문에서는 근래에 발표된 SVC의 표준 확장 방법을 적용하였다. 이 확장 알고리즘은 AVC의 움직임 내삽 필터(motion interpolation filter)에 적용되었던 랑쥬쉬(Lanczos-3) 함수에 기반한 6탭 필터를 계산량을 줄이기 위해 4탭 필터로 대체하였다^[11].

4.2 공간 해상도에 따른 물-왜곡

그림 2에서 원본보다 낮은 해상도로 인코딩 된 비디오들을 디코딩 한 후의 결과 이미지들을 원래 해상도로 확대한 후에 PSNR 값을 비교하였다. 실험 결과 곡선들을 분석해 보면 최상의 화면 화질을 보여 주는 공간적 해상도는 비트율에 따라 달라진다. 공간적 해상도를 줄여서 인코딩 하는 것은 비트

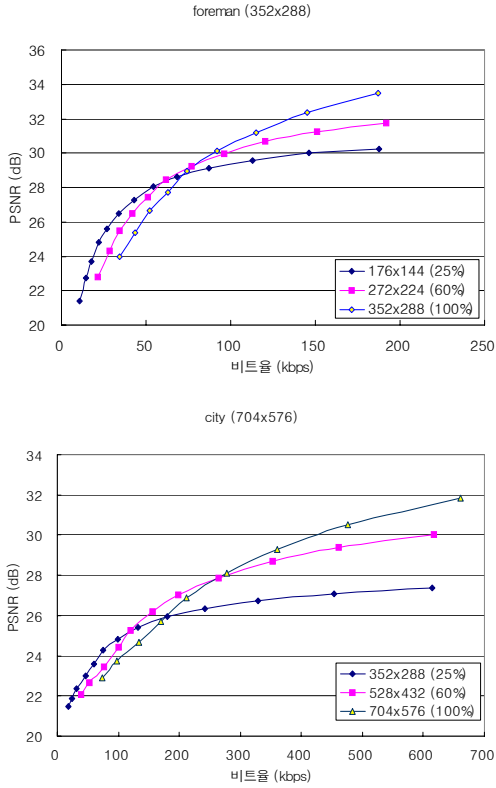


그림 2. 해상도별 율-왜곡 성능: “foreman” (위), “city” (아래)

율이 매우 낮은 경우에 특히 효과적이다. 여기서 사용된 성능 척도는 가장 널리 사용되는 객관적 척도인 PSNR이지만, 실제 사용자가 겪는 주관적 품질과 정확히 대응되지 않는 경우가 많다. 실제 결과 이미지들을 비교해 보면 거의 비슷한 PSNR 값을 가지지만, 작은 해상도의 이미지를 복호 한 후 확대한 이미지가 주관적인 화질에서 더 낫다는 것을 그림 3에서 쉽게 확인할 있다. 원본 크기대로 인코딩했을 경우에 PSNR 값이 약간 높음에도 불구하고, 블록킹 효과 (blocking effect)에 의해 블록 단위로 화질 저하가 두드러지게 보인다.

따라서 각 스트림별로 현재 비트율에 맞추어 가장 적합한 해상도로 적응적으로 변경하여 인코딩하여 전송하는 것이 객관적, 주관적 화질 측면에서 최적의 결과를 얻을 수 있음을 알 수 있다. 그러나 어느 정도 비트율 이상으로 올라가면 축소와 확대에 의한 화질 열화가 너무 심해지기 때문에 적응적으로 해상도를 변경할 여지가 없게 된다.



그림 3. 결과 화면의 주관적 화질 비교: 원본 CIF 해상도, 23.99 dB (위), QCIF 디코딩 후 CIF로 확대, 23.79 dB (아래)

4.3 에너지를 고려한 합계 성능

각 해상도 별로 디코딩 시간 및 확대시간을 더한 전체 수행 시간을 측정하여 그림 4에 표시하였다. 디코딩 시간은 비트율과 해상도에 비례하여 증가하며, 확대에 드는 시간은 상대적으로 일정하기 때문에 일정한 비율로 추가된다. 실험 결과는 식 (14)와 잘 일치하며, 본 논문에서 제시된 에너지 모델을 뒷받침하고 있다. 각 해상도별로 비트율에 따른 디코딩 시간의 증가 비율은 계속 유지되기 때문에, 작은 해상도 스트림을 디코딩 할수록 에너지 이득은 거의 일정한 비율로 얻어진다. 그러나 확대에 드는 고정적인 비용을 추가할 경우 전체적인 수행 시간의 이득은 줄어들게 된다.

에너지 중요도 ω 를 각각 0.05와 0.15로 적용한 결과가 그림 5에 나타나 있다. 에너지 소모를 줄

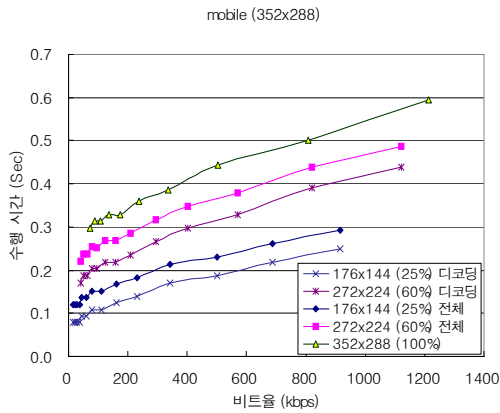


그림 4. 해상도별 디코딩 및 전체 실행 시간

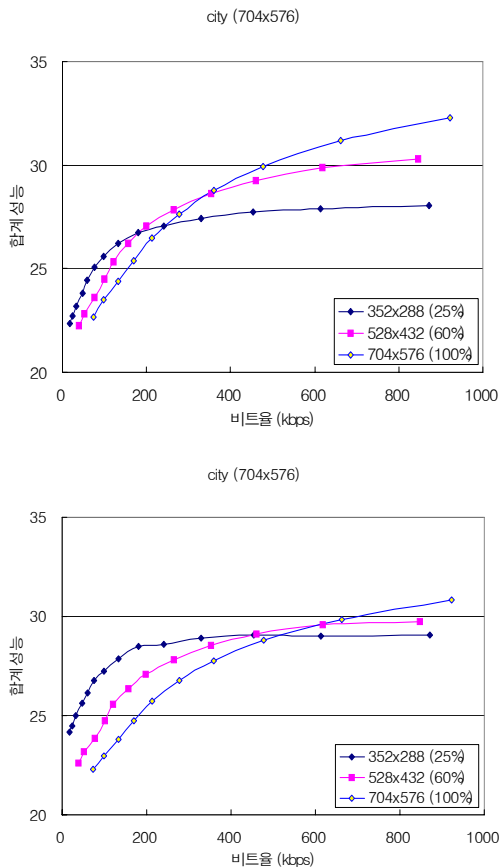


그림 5. 에너지 관심도와 비트율에 따른 합계 성능: $\omega=0.05$ (위), $\omega=0.15$ (아래)

이는 것이 얼마나 중요한 가를 나타내는 ω 가 작아 질수록 합계 성능 척도는 그림 2의 형태로 수렴한

다. 반대로 ω 가 커질수록 합계 성능 척도는 작은 해상도로 인코딩하는 것이 더 유리해지는 형태로 변하게 된다. 본 논문에서 제시된 적응적 해상도를 적용하게 되면 일반적으로 낮은 비트율에서는 낮은 해상도로 인코딩 하는 것이 유리하게 되고, 비트율이 증가함에 따라 더 높은 해상도로 인코딩 하는 것이 유리하다. 또한 에너지 관심도가 더 증가 할수록 낮은 해상도로 인코딩 하는 것이 유리해 지는 구간의 폭이 증가하게 된다.

표 1에서 기존의 H.264/AVC 비트율 제어만을 이용하는 것보다 제안된 적응적 공간 해상도 조절 방법을 이용했을 때의 화질과 처리 시간에서의 이익을 정리하였다. 제안된 방법은 낮은 비트율 환경에서 약 0.5 dB의 화질을 개선할 수 있고, 50% 이상의 처리 시간을 줄일 수 있으며, 에너지 관심도 ω 가 증가할수록 에너지 이익도 증가한다.

마지막으로 AVC 와 SVC 를 이용하여 인코딩 할 때 적응적 공간 해상도 조절 스킴을 적용했을 경우의 결과를 비교하였다. 공간적 스케일러빌리티를 지원하기 위해서 SVC 스트림의 기본 계층은 QCIF 해상도로, 강화 계층은 CIF 해상도로 인코딩 되었다. 또한 네트워크 대역폭에 적응을 위해 다양한 비트율에 맞추어 추출이 가능한 화질 스케일러빌리티를 지원하기 위해서, 각 계층에 MGS (medium-grain scalability) 가 적용되었다. AVC와 마찬가지로, 낮은 비트율 환경에서는 강화 계층 대신 기본 계층이 추출되어 CIF 해상도로 확장되어 보는 상황을 고려하였다.

그림 6과 같이 에너지 관심도 ω 가 0인 상황에서 Rate-distortion 성능만을 비교해 보았을 때, SVC는 scalable 특성을 지원하기 위해서 오버헤드가 들므로, AVC에 비해 상대적으로 코딩 효율이

표 1. PSNR과 처리 시간에서의 이익 ($\omega = 0$)

	50 kb/s		80 kb/s	
	PSNR (dB)	시간 (%)	PSNR (dB)	시간 (%)
foreman	0.24	22.9	0.00	0.00
stefan	0.60	60.6	0.32	23.6
bus	0.62	56.4	0.12	23.4
mobile	0.36	57.6	0.01	22.0
football	0.42	55.0	0.65	57.0
평균	0.45	50.5	0.27	26.0

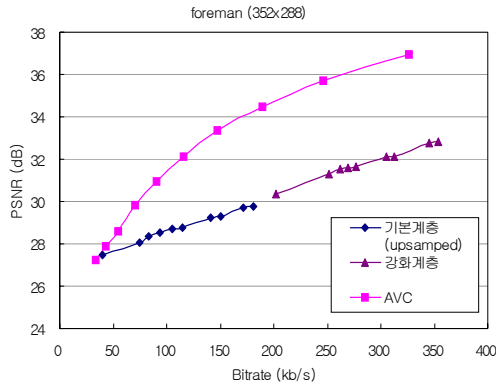


그림 6. 적응적 공간 해상도 조절을 적용할 경우 H.264/AVC 와 SVC의 압축 효율

떨어짐을 알 수 있다. 그러나 SVC를 이용할 경우에는 재인코딩 할 필요 없이 하나의 스트림을 가지고 상황에 맞게 추출만 해 주면 되기 때문에 더 많은 수의 클라이언트를 서비스 하는 환경에서 더 적합하다는 장점이 있다.

V. 결 론

모바일 비디오 스트리밍을 위해서 적응적으로 공간 해상도를 조절하는 비디오 인코딩 스킴을 제시하였고, 사용 가능한 대역폭에 맞추기 위해서 양자화 파라미터만을 조절하는 기존의 비트율 조절 방법보다 화면 화질과 에너지 소모 측면에서 더 효율적임을 보였다. 이를 위해 각 해상도와 비트율, 에너지 관심도에 따른 모델을 정의하고, 최적의 해상도 결정을 위한 성능 척도를 제시하였다. 본 기법은 낮은 비트율 상황에서 기존의 방법에 비해 화질을 약 0.5dB 만큼 개선하고, 에너지 소모를 50% 이상 줄일 수 있다.

참 고 문 헌

[1] J. Ribas-Corbera and S. Lei, "Rate Control in DCT video coding for low-delay communications," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 9, pp. 172-185, Feb. 1999.

[2] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 1, pp. 1533-1544,

Dec. 2005.

[3] J. Kim, Y.-G. Kim, H. Song, T.-Y. Kuo, Y. J. Chung, and C.-C. J. Kuo, "TCP-friendly Internet video streaming employing variable frame rate encoding and interpolation," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 10, pp. 1164-1177, Oct. 2000.

[4] H. Song and C.-C. J. Kuo, "Rate control for low-bit-rate video via variable-encoding frame rates," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 11, pp. 512-520, Apr. 2001.

[5] Y. Yuan, D. Feng and Y. Zhong, "Fast adaptive variable frame-rate coding," *IEEE Vehicular Technol. Conf.*, pp. 2734-2738, 2004.

[6] D. Son, C. Yu, and H. Kim, "Dynamic voltage scaling on MPEG decoding," *Int'l Conf. Parallel and Distributed Systems*, 2001.

[7] S. Park, Y. Lee, J. Lee, and H. Shin, "Quality-adaptive Requantization for low-energy MPEG-4 video decoding in mobile devices," *IEEE Trans. Consumer Electron.*, vol. 51, pp. 999-1005, Aug. 2005.

[8] M. Mesarian and Y. Turner, "Reduced energy decoding of MPEG streams," *ACM/Springer Multimedia Systems Journal*, vol. 9, pp. 202-213, Aug. 2003.

[9] "Information technology - Coding of Audio-visual Objects - Part 2: Visual," *International Organization for Standardization, ISO/IEC 14496-2:1999 /Amd.1:2000(E)*, Jul. 2000.

[10] G. Cote, B. Erol, M. Gallant, and F. Kossentini, "H.263+: video coding at low bit rates," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 8, no. 7, pp. 849-866, Nov. 1998.

[11] Joint Draft ITU-T Rec. H.264 | ISO/IEC 14496-10/Amd.3 Scalable video coding, *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6*, 2007.

[12] Draft ITU-Recommendation and FDIS of Joint Video Specification (ITU-T Rec. H.264 - ISO/IEC 14496-10 AVC), *ISO/IEC JTC1/SC29 /WG11 and ITU-T SG16 Q.6*, 2003.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

