

H.264/AVC에서 비트율 제어와 시각 인지도를 고려한 영상화질 개선 기법

안수민*, 김강석*, 김재훈*

*아주대학교 대학원 지식정보공학과

e-mail : {wingjyn, kangskim, jaikim}@ajou.ac.kr

Improving Video Quality with Bitrate Control and Visual Recognition in H.264/AVC

Soomin Ahn*, Kangseok Kim*, Jai-Hoon Kim*

*Dept of Knowledge Information Engineering, Graduate School of Ajou University

요 약

H.264/AVC 영상 부호화기의 비트율(Bitrate) 제어에 사람의 눈에 잘 인식되는 부분의 화질을 향상하고 인식률이 낮은 부분의 화질을 낮추어 상대적으로 향상된 화질을 얻기 위한 방법을 연구하였다. 먼저 H.264/AVC 영상 부호화기에서, 장면 변화로 인한 참조 프레임의 비트 낭비를 막기 위해 GOP(Group of Pictures) 단위로 장면 변화를 검출한 후, 검출된 장면 변화에 대해 GOP를 적용 시킨다. 해당 GOP 내에서 시각 인지도에 기초하여 물체의 움직임으로 인한 인식률이 높은 부분을 검출하고, 인식률이 높은 부분에 대해 QP(Quantization Parameter)의 재분배로 비트율을 높임으로써 화질을 향상시키고, 인식률이 낮은 부분에 대해서는 비트율을 적게 분배한다. 그 결과 한정된 대역폭을 갖는 전송 환경에서 영상을 향상된 화질로 이용할 수 있는 방법을 제안한다.

1. 서론

인터넷에 접속 속도와 대역폭이 향상되면서 동영상 서비스가 가능해졌고, 많은 이용자들이 실시간 또는 그 외의 방법으로 모바일 기기를 통해 멀티미디어 영상을 접하고 있다. 이에 따라 모바일 기기로 접근하는 영상에 대해 높은 QoS(Quality of Service)를 보장하기 위해 다양한 영상 서비스 제공 방법이 제시되었다. 그 방법으로는 비디오 스트림의 전송에 있어서 대역폭, 버퍼 등의 네트워크상의 개선 방법뿐만 아니라 비트율 제어를 함으로써 한정되어 있는 대역폭에서 높은 화질을 보장하는 방법들이 있다 [2-6].

비트율 제어는 H.264/AVC 영상 압축 과정에서 중요한 부분이다. 비트율 제어의 주목표는 각각의 프레임의 이미지 왜곡을 최소화 시키고 현재 대역폭에 맞도록 비트들을 재분배 하는 것이다. 이는 실시간 데이터 교환에 있어서 통신량 증가에 따른 패킷 손실을 피하는 효과적인 방법이 다[4].

비트율 제어에서 고려할 수 있는 문제는 장면 변화이다 [3]. 동영상 부호화시에 각각의 프레임들은 한 그룹으로 묶여지게 되고, 그룹의 첫 장면은 키 프레임으로 설정되어 그룹 내부에서 이 후 프레임들이 참조하여 부호화가 진행

된다. 한 그룹 내부에서 장면의 변화, 즉, 키 프레임과 상관없는 장면으로의 변화가 일어나게 되면 참조할 프레임과 현재 프레임 사이에 많은 차이가 있기 때문에 큰 값을 저장하게 된다. 이와 같은 경우에는 부호화 출력 비트율이 심하게 변화하는데, 이럴 경우 일정 비트율 조건에서 비트 스트림을 출력한다면 화질 저하의 원인이 된다.

정보의 70% 정도가 눈으로 획득되고, 눈은 영상 정보의 최종 수신기이기 때문에 이미지의 높은 주관적 품질을 얻기 위해서는 시각 특성을 최대한 활용하는 것이 중요하다.

본 논문에서는 시각적 특성을 이용한 비트율 제어 알고리즘[2]에 장면 변화를 고려한 방법으로 같은 비트율에서 낭비되는 비트율을 줄이고 향상된 화질을 얻을 수 있는 방법을 제안한다.

2. 관련 연구

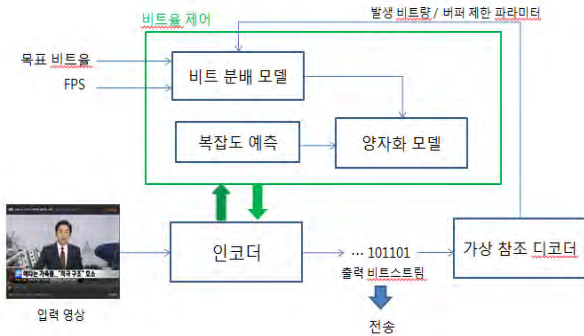
2.1 H.264/AVC 에서의 비트율 제어 방법

H.264/AVC의 참조 소프트웨어 JM(Joint Model)에서의 비트율 제어 알고리즘은 범용성을 고려하여 다양한 대역폭 조건을 만족시키는 것과 높은 압축 효율에 중점을 두고 개발되었다. JM은 비트분배모델, 비트율 양자화 모델, 복잡도 예측 모델을 사용하여, GOP, 프레임, 매크로 블록의 계층적인 순서로 비트율을 제어한다. (그림 1)은 H.264/AVC의 비트율 제어의 블록 다이어그램이다.

비트스트림 버퍼의 관리를 위해 부호화기에는 가상의

본 연구는 미래창조과학부 및 한국인터넷진흥원의 "고용계약형 지식정보보안 석사과정 지원사업"의 연구결과로 수행되었음. (과제번호 H2101-13-1001)

복호화기가 포함되는데, 이를 통하여 부호화기에서 비트 발생량에 따라 복호화기에 존재하는 비트스트림 저장 버퍼의 크기, 복호화 시작 시간까지의 지연, 복호화 과정에서 비트스트림 버퍼로부터 빠져나가는 비트율 등을 고려하여 표준에 정의된 비트스트림 버퍼의 오버플로우나 언더플로우 등의 문제가 발생되지 않는 비트율의 순응 여부를 알 수 있다. 이를 사용하여 발생 비트율을 적절히 조절한다.



(그림 1) H.264/AVC 비트율 제어 Block Diagram

2.2 기존 연구[2]

사람은 전체 프레임의 정보를 한 번에 받아들이는 것이 아니라, 어느 한 점에 대해 집중이 되거나 어떤 상황에 놓인 물체에 시각이 집중된다는 점에 기초하여 시각이 집중되는 부분은 색상 차이, 강도(휘도) 그리고 움직임 등에 의해 결정이 된다[1].

논문[2]에서는 매크로 블록 움직임에 따른 특성인 모션 벡터(MV: Motion Vector) 값에 따라 그 특성에 대해 분류를 하였다. 사람은 움직이는 물체나 정지해 있는 물체에서는 가장자리 변화에 민감한 것을 염두에 두고, 집중도를 두어 비트율 제어 알고리즘에서 움직임 특성과 가장자리 영역을 포함하여 계산하는 (식 2)를 비트율 제어 알고리즘(JVT-G012)에 추가하여 (식 1)과 같이 제안하였다.

$$\tilde{b}_l = T_r \times V(i, j) \times \frac{MAD_{cb}^2}{\sum_{k=1}^{N_{mit}} MAD_{cb}^2(k)} \quad (\text{식 1})$$

$$V(i, j) = motion(i, j) + edge(i, j) \quad (\text{식 2})$$

1) 움직임 특성에 대한 집중

실제 영상의 장면 내에서 움직이는 물체에 대해 그 특성에 따라 집중도에 대해 모션 벡터에 따른 장면 변화와 집중도를 예측하고 분류를 한다.

첫 번째로 다음 장면으로 갔을 때에 장면에 대한 변화는 적고 몇 개의 물체만 움직이거나 변화하는 부분이다. 이때 사람의 눈은 움직이거나 변화하는 부분에 집중이 된다. 이때 모션 벡터의 평균은 4.5 이하가 된다. 두 번째로 장면이 빠르게 변화할 때에는 모션 벡터 평균이 4.5가 넘게 되

며, 두 가지의 경우가 있다. 그 중 한 가지는 빠른 장면 변화 속에 움직임이 거의 없는 물체가 있을 때, 시각은 움직임이 적은 물체에 집중이 된다는 것이다. 이 때 프레임 내에서 물체의 모션 벡터 값은 2.5에서 6.5사이에 있고 배경 부분은 6.5 이상이다. 두 번째는 많은 물체들이 움직이고 있는 가운데, 몇 개의 특정 물체만이 더 빠르게 움직이는 장면의 경우에는 시각은 빠르게 움직이는 부분에 집중하게 된다. 이때 물체의 모션 벡터 값은 1.5이하이고 다른 부분은 1.5에서 2.5사이의 값을 갖는다. 이 두 가지를 고려하여 식으로 표현한 것이 (식 3)이다.

$$motion(i, j) = \begin{cases} 1 & MV < 1 \\ 2 - \frac{(6.5 - MV)}{8} & 2 < MV \leq 6.5 \\ 3 & MV > 6.5 \end{cases} \quad (\text{식 3})$$

2) 가장자리 특성에 대한 집중

이미지 특성에서 가장자리가 갖는 특성은 매크로 블록 값의 높은 변화량으로 보고 있다. 부·복호화 과정에서는 수치로만 해석하기 때문에, 매크로 블록이 갖는 값을 비교하여야 하는데 기존 논문에서는 그 수치가 매크로 블록이 갖는 변화량에 주목하고 있다. 여기서는 매크로 블록 내부에서 기본 단위가 급격하게 변화한다면 가장자리로 정의한다. 예를 들자면 과란색의 옷을 입고 있는 사람이 장면 내에 있고, 그 뒤로 보이는 배경이 과란색이 아닌 다른 색을 띄고 있다면 과란 색과 배경이 만나는 부분이 가장자리이고 이때의 두 값은 큰 차이를 보인다는 것이다.

$$edge(i, j) = \begin{cases} 1 - \frac{D(i, j)}{\max(D)} & D(i, j) < medium(D) \\ \frac{D(i, j)}{\max(D)} & other, \end{cases} \quad (\text{식 4})$$

3. 제안 방식

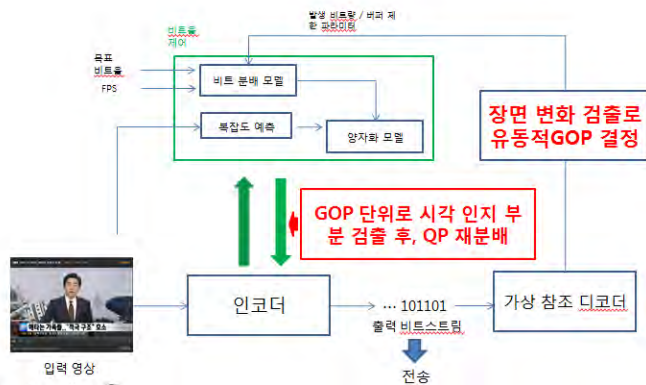
3.1 기존 개선 방법의 문제점

기존 연구는 프레임 변화에 따른 방식이다. 이것은 참조하는 프레임과 현재 프레임 간의 상관관계에 따른 방식이라고 할 수 있는데, 하나의 참조만 한다. 수치 분석에 있어서 단 하나의 참조만 있기 때문에 실험 결과를 분석하였을 때, 오차가 존재하였고, 특수한 상황, 즉, 전혀 다른 장면에서 집중도와 상관없는 부분에 비트율이 높게 분배되어 화질과 압축 용량에 있어서 큰 이득을 볼 수 없었다. 그리고 기존의 장면 전환에 따른 GOP 결정방법은 장면 전환들이 인접하여 발생할 경우 장면 전환 화면을 모두 I 프레임으로 처리하는데 따르는 급격한 비트율의 증가로 인한 버퍼 오버플로우의 위험을 갖고 있다. 그렇기 때문에 오차를 더 줄이며 비트 분배를 효율적이고, 또한 화질과 압축 용량을 개선하기 위해 다음과 같이 제안하였다.

3.2 제안 방법

본 논문에서는 여러 통계적 특성을 추출하여, 최종 QP 들을 결정하는 방법인 다중 경로(Multi-pass) 비트율 제어를 사용한다. 다중 경로 방식을 사용하는 것은 최적의 비트 분배를 위해 동영상 내에서 배경이 같은 부분을 위주로 GOP를 설정하기 위함이다. 상황이 변하기 전까지 같은 키 프레임이 써도 되는 프레임의 유동적 GOP를 결정하고 그에 속한 프레임의 모션 벡터를 이용한 움직임이 있는 부분을 측정한다. 그리고 모션 벡터 값들을 저장하고 비교하여, 초당 프레임레이트의 프레임 수 이상에서 움직임 변화와 시각 인지도에 관련이 있다고 판단이 되면, QP 값을 조절하여 화질을 향상 시킨다. 반대로 초당 프레임레이트의 프레임 수에 미치지 못한다고 판단이 되면 프레임 변화 중에 움직임 변화가 인지하지 못할 정도이기 때문에 화질을 향상시키지 않아도 되는 부분이므로 QP 값을 조절하여 비트율을 낭비할 필요가 없다. 또한 시각인지도와 관계없는 부분은 비트율을 낮추어 화질을 저하시켜 프레임의 비트율을 유지한다.

위와 같은 일련의 과정을 (그림 2)와 같이 표현하였다.



(그림 2) 제안한 H.264/AVC 비트율 제어의 Block Diagram

유동적 GOP 내에서 매크로 블록의 모션 벡터 값들을 저장하고 평균을 내어 그 값을 비교하고 (식 3)에 적용 한다. 여기에서 motion (i,j)의 값이 커진다는 것은 해당 매크로 블록의 목표 비트율인 b_t 값 (식 1)이 커짐을 의미하고 그에 따른 QP 값의 재배치로, 비트율을 높임으로써 화질을 향상시키고 인식률이 낮은 부분은 비트율을 낮게 분배한다. 프레임 측면에서 보면 프레임이 갖고 있는 비트율은 같지만 프레임 내부의 각 매크로 블록이 갖는 비트율이 조절되었기 때문에 인식률이 높은 부분은 화질이 향상된다.

4. 시뮬레이션 결과 및 고찰



(a)



(b)

(그림 3) Silent 영상에서 기존 방식(a)과 제안한 방식(b)의 비교

실험은 H.264/AVC의 레퍼런스 프로그램인 JM-18.6에서 진행하였다. 실험에 사용한 Highway, Bridge-close, Bridge-far, Silent 그리고 Container의 다섯가지 레퍼런스 영상을 사용하였다. 부호화 과정에서 옵션은 프레임 레이트를 25로 설정하였고 GOP는 제안방식이 유동적으로 변하기 때문에 제안방식에서 검출한 GOP와 JM을 동일하게 맞추어 주었다. 목표 비트율은 전체 프레임 수를 고려하여 유사하게 해주었다. 이외의 다른 옵션은 최대한 사용하지 않으면서 조건을 맞추어 주는데 집중하였다.

그림 3을 보면 기존방식(a)과 제안방식(b)에서 확연한 차이를 느낄 수 있다. 이 영상은 얼굴 표정과 손의 움직임이 뚜렷한 영상이며, 움직이는 부분으로 검출된 네모 상자 안의 부분의 화질이 향상되었음을 알 수 있다.

PSNR(Peak Signal to Noise Rate)는 화질을 평가할 때 사용하는 기준이다. 표 1을 보면 제안방식에

서 PSNR의 증가가 일어났음을 알 수 있다. 그리고 이것은 화질이 향상되었음을 나타내는 것이다.

표 2는 비트율을 비교한 것이다. 표 1과 표 2의 결과를 볼 때, 유사한 비트율에서 PSNR의 수치가 증가한 것을 알 수 있다.

따라서 본 논문의 제안 방식은 화질 향상에 효과적이다.

<표 1> 기존 방식과 제안 방식의 PSNR 비교

테스트 영상명	PSNR(dB)		
	기존 방식	제안 방식	차이
Highway	36.036	36.196	0.160
Bridge-close	35.716	35.986	0.270
Bridge-far	39.345	39.441	0.096
Silent	37.001	37.509	0.508
Container	35.095	37.883	2.788

<표 2> 기존 방식과 제안 방식의 비트율 비교

테스트 영상명	Output Bitrate(kbps)	
	기존 방식	제안 방식
Highway	34.88	35.61
Bridge-close	35.1	34.93
Bridge-far	32.09	31.46
Silent	63.43	63.5
Container	34.9	35.16

5. 결론

본 논문은 사람이 갖는 시각 특성에 기초하여 같은 비트율을 갖는 동영상을 더욱 향상된 화질로 접할 수 있도록 비트율 제어의 개선 방안을 제시하였다. 디지털 영상의 경우, 그 속에 포함된 정보는 매우 방대하고, 영상마다 정보가 다르기 때문에 출력 비트율의 심한 변화를 방지하고 화질의 급격한 저하를 최대한 방지하면서 일정 비트 스트림으로 접근하는 것이 필요하다. 특히, 비디오 시퀀스에 있어서 카메라 소스의 단절이나 비디오 편집 그리고 프로그램 삽입 등에 의해 발생하는 장면 전환은 인접하는 화면의 특성이 급작스럽게 변화되는 것을 의미하고 있는데, 이 부분이 한 GOP 내부에서 존재하게 된다면 통상의 방법으로 부호화될 경우 제한된 비트율에서 화질 저하의 문제점이 발생된다. 그러므로 장면 변화에 따른 화질 저하를 최대한 방지하기 위해 유동적 GOP와 시각 인지도에 따른 비트 분배로 같은 비트율에서 화질 면에서 향상을 시키고, 같은 화질일 경우에는 압축률 면에서 이득을 더 볼 수 있게 개선방안을 제시하였다.

향후 다양한 특징을 갖는 영상들과 다양한 전송 환경에

서 비트율 제약과 함께 화질 면에서 어느 정도의 향상이 있는지에 대한 연구를 진행할 것이다.

참고문헌

- [1] Z. Li, S. Qin and L. Itti, "Visual attention guided bit allocation in video compression," *Image and Vision Computing* 29.1 : 1-14, 2011.
- [2] X. Chen and H. Liu, "Basic Unit Layer Rate Control Algorithm for H.264 Based on Human Visual System," *Mathematical Problems in Engineering*, 2013.
- [3] X. Chen and F. Lu, "Adaptive Rate Control Algorithm for H.264/AVC Considering Scene Change," *Mathematical Problems in Engineering*, 2013.
- [4] JVT, "Adaptive basic Unit Layer Rate Control for JVT," *JVT 7th meeting document G012*, Mar. 2003.
- [5] P. Usach, J. Sastre, and J. M. López, "Variable frame rate and GOP size H. 264 rate control for mobile communications," *In Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pp. 1772-1775, 2009.
- [6] S. Park, "GOP Level Rate-Control for Real-Time Video Transmission", *International Journal of Multimedia and Ubiquitous Engineering*, Vol. 8, No. 4, pp. 207-215, 2013.
- [7] S. H. Park, "Initial QP modeling for GOP Layer Rate Control," *The Journal of the Korea Institute of Electronic Communication Sciences*, vol. 7, no. 6, pp. 1377-1383, 2013.