

Macro Block 기반의 효과적인 GOP구조 선택 알고리즘

Algorithm of effective GOP structure select by Macro Block base

전민정, 문영득, 정희태
부산외국어대학교

Jeon Min-Jeong, Moon Young-Deuk,
Chung Hee-Tae
Pusan University of Foreign Studies

요약

GOP 구조를 제어하여 양질의 영상을 제공하기 위한 방법들이 많이 제안되어 있다. 그러나 기존의 알고리즘들은 GOP 구조를 결정하기 위해 일정량의 영상을 사전 검토하여야 하는 과정을 거쳐야 하므로 계산 량이 증가하게 되어 대용량 프레임 메모리가 필요하고 시간적으로도 지연되므로 실시간 적용에는 적합하지 못하다. 본 논문에서는 매크로 블록을 기반으로 한 실시간 적용적인 알고리즘을 제안한다. 제안한 매크로 블록 기반의 GOP 구조를 이용하여 효과적인 영상처리 결과를 얻었다.

Abstract

There are many suggestion of methods which get good quality images by controlling the GOP structure. But the existing algorithm should have not good to deal the real time adaptive processing determine the GOP structure because which check the fixed quantity of image in advance and require high capacity of frame memories and have delay time by increasing calculation quantity. This paper propose the algorithm which adapt the real time by the basic of macro block. We show the good simulation results using the proposed GOP structure based on macro block.

I. 서론

MPEG-1,2는 영상 전송과 저장을 위한 표준 기술로 비디오, CD, DVD, 그리고 위성방송과 지상파 고화질 TV 시스템 등에 많이 응용되고 있다. MPEG 표준은 단지 압축된 영상의 비트스트림만을 정의하기 때문에 많은 베이스라인 MPEG-2 TM5의 영상 화질을 향상시키려는 알고리즘들이 제안되어 왔다. Teixeira와 Ribeiro는 MPEG 비디오 시스템에서 두 단계 접근을 사용하고 있다. 영상 복압도는 pre-coding 사용에 의해 첫 번째 단계에서 얻어진다. 그리고 GOP 길이에 관련하여도 몇몇 제안들이 있다. GOP길이 N은 약 0.5초의 랜덤접근을 위해 15나 12

로 고정된다. 그러나 장면변화가 장면변화를 기대하지 않는 영역에서 효율적인 움직임보상에 관계될 때 부호화 효율을 유지하는 것은 어렵다. 따라서 다이나믹한 GOP 길이에 의한 부호화는 장면 변화 전에 부호화 효율을 향상시키기 위해서 필요하다. 이것은 MPEG 편집의 관점에서 유용하고 사용자는 어느 경계에서나 용이하게 내용을 자르고 결합할 수 있다. GOP길이를 가변시키기 위해서 MB활동도와 MB활동도메인상의 단순 움직임 예측을 사용한다. 이 방법은 GOP 길이를 계산하기 위해서 프레임 메모리에 30프레임 정도를 미리 읽어 들이므로 각 영상의 MB 활동도를 구하고 움직임 예측 오차 값을 구하고 있다.

이 방법은 MB활동도와 움직임 예측 오차 값을 구하기 위해 많은 프레임 메모리를 요구하며 사전 분석을 위해 계산 량이 많아지며 또한 시간적 지연이 불가피하여 실시간 적용에는 적합하지 못하다. 다른 방법으로는 최소화 문제를 해결하여 선택 가능한 후보로부터 GOP에서 P-picture의 최적수와 위치를 탐색하는 것으로 이 방법은 GOP 내의 임의의 위치에 P-picture가 위치하게 되고 최적의 GOP구조가 선택된다. 그렇지만 이 방법도 많은 계산 량을 요구하며 최적수와 위치를 탐색하기 위한 시간적 지연이 문제이다.

본 논문에서는 사전 분석이나 지연이 없는 영상 복잡도에 기반하여 실시간 적용적으로 GOP구조를 변화시키는 알고리즘을 제안한다. 영상의 복잡성 변수를 각 영상 부호화 방법을 위한 비트 묶음의 예측기로 사용하여 먼저 영상에서 복잡도를 계산하고 그 복잡도의 평균분포를 구하여 제한한 알고리즘의 경계값으로 한다. 이 경계값을 사용하여 최적의 GOP구조를 선택한다.

II. MPEG의 picture 형태와 GOP 구조

2.1 picture형태와 GOP구조

MPEG의 picture 형태에는 I, P, B-picture 등 세 가지가 있다. I(Intra coded)-picture는 움직임 예측 같은 시간적인 압축을 사용하지 않고 DCT 변환에 의한 공간적인 압축만을 사용한 화면이다. 시간적인 정보를 사용하지 않기 때문에 시간상으로 누적되는 오류를 방지할 수 있어 다른 화면을 압축 부호화 할 때 기준이 되는 참조 picture의 역할을 한다. 그러나 시간적인 정보를 사용하지 않기 때문에 다른 picture들 보다 부호화시 더 많은 비트수를 필요로 한다. 또한 단독으로 부호화가 가능하고 이로 인해 임의의 접근을 할 수 있는 기준 영상이 되기도 한다.

P(Predicted coded)-picture는 예측 화면이라는 의미로서 가장 최근의 I와 P-picture로부터 원 화면

과 가장 가깝게 예측해 낸 화면이다. 즉 현재 화면과 이전 화면의 움직임을 비교하여 변경된 부분의 정보들을 움직임 벡터로 표현하여 구성된다. B(Bidirectionally predicted coded)-picture 역시 예측화면이긴 하지만 이전 I, P picture에서만 예측하여 화면을 구성하는 것이 아니라, 그 다음 I, P-picture에서도 역 방향으로 예측해 보고, 또 이들 두가지방법(순방향, 역방향)을 평균하는 화면을 만들어 본 뒤 이 세 가지 화면에서 원 화면과 가장 비슷하게 예측되는 화면을 선택하는 다소 복잡한 화면이다. 이런 예측 방법을 그림 1에서 보여주고 있다. I와 P-picture는 inter-frame 예측을 위해 다른 영상에 참조되므로 core picture라고도 한다. B-picture는 역방향 예측에 의해 앞으로 발생할 영상을 참조하기 때문에 picture의 디스플레이 순서와 압축되어 저장되는 순서가 일치하지 않아 프레임의 재순서화가 필요하여 이로 인한 지연이 발생하는 단점이 있다. B-picture를 많이 사용하면 발생 비트수를 줄일 수 있지만 그 만큼 많은 오류들이 누적되므로 B-picture의 적절한 선택이 이루어져야한다. I, P, B 등 세 가지 영상의 배열은 주기적인 형식으로 정의된다. core picture 사이의 거리가 3으로 고정되고 GOP 내의 영상의 수가 15로 고정되어 있다면 GOP 구조는 코딩 순서에서 "IBBPBBPBBPBBPBB"로 된다. 이것은 영상화질을 향상시키기 위해서 입력영상에 따라 동적으로 변화되어야 한다. MPEG 부호화에서 B-picture의 사용은 높은 부호화 성능을 제공하지만 P-picture의 예측성능은 참조영상으로부터의 거리가 커지면 감소되며 이에 따라 core picture사이의 B-picture에 대한 최적수가 결정된다. 제한한 방법은 MPEG 부호화에 일반적으로 사용되는 M값을 M=1에서 M=3 사이에서 변화시킨다.

2.2 영상 데이터 전송과 버퍼제어

picture당 발생하는 비트수는 가변적으로 버퍼의 충만도에 따라서 제어되는데 이 버퍼가 underflow

나 overflow가 나지 않게 하기 위해서 인코더에서는 버퍼상태를 항상 감시하고 있어야 하며 이에 따라서 다음 picture의 비트 발생량을 조절해 주어야 한다. 이러한 발생량 조절에 대해서는 MPEG-2에 기술되어 있지 않다. 인코더에서 버퍼를 1000분의 1초마다 검색하든지 아니면 매 picture마다 검색하든지 그것은 인코더의 자유이며 제작하기 나름이다. 즉 주어진 버퍼량에 대하여 underflow나 overflow가 나지 않게 해주어야 한다. 그러나 이 작용이 최적화되지 않으면 정보량이 많은 picture에 비트할당을 많이 해주지 못하는 경우가 생기며 또 그 반대의 경우도 발생할 수 있다.

MPEG에서는 실제의 버퍼체계와는 관계없이 하나의 가상 버퍼 시스템에 대해서 언급하고 있는데 VBV(video buffering verifier)이다. VBV 버퍼란 인코더가 실제로 갖고 있는 않지만 인코딩을 할 때 디코더 버퍼의 상태를 똑같이 따라가기 위해서 인코더가 생각하고 있는 가상의 버퍼 시스템을 말한다. 물론 디코더는 실제의 버퍼를 가지고 있으며 버퍼의 상태는 가상의 인코더 버퍼상태와 동일하게 유지된다. 여기서 매 picture 데이터를 뽑아가는 시간간격 ($t_{n+1} - t_n$)은 보통 progressive sequence의 경우 frame rate의 역수이며 T 로 표시한다. B 는 VBV 버퍼의 크기이며 각 시간간격 내에서의 직선의 기울기는 bit rate인 R 이 된다. 그러므로 버퍼의 입력 단에는 매 초당 R bit씩의 데이터가 계속 들어오고 출력 단에는 T 간격으로 매 picture의 데이터를 한꺼번에 뽑아가서 디코딩하고 디스플레이까지 한다. 그러나 디코딩을 맨 처음 시작할 때는 버퍼 내에 아무 데이터도 없기 때문에 버퍼가 어느 정도 차는 것을 기다려야 하며 이렇게 버퍼를 채우는 데까지 디코더가 기다려야 하는 시간을 vbv_delay라 한다. 이 vbv_delay는 MPEG-1의 시스템 클럭인 90kHz의 주기로 나타나며 $vbv_delay = 90000B * n/R$ 로 표현된다. 여기서 $B * n$ 이란 $t=n$ 일 때 VBV 버퍼가 얼마의 비트 수로 차있는 가를 나타낸다.

III. 제안한 방법에 의한 효과적인 GOP 구조 선택 알고리즘

효과적인 GOP 구조는 inter-frame의 상관성에 따라 결정되므로 사전분석으로 inter-frame의 상관성을 먼저 계산한다. 상관성은 inter-frame 예측 부호화된 영상의 비트 카운터에 연관되므로 제안한 알고리즘에서는 TM5의 복잡성 변수 X_i , X_p , 그리고 X_b 를 각 영상 부호화 방법을 위한 비트 묶음의 예측기로 사용한다. 이러한 변수는 식(1)에서와 같이 부호화 방법에서 이전 영상의 정보를 사용하여 나타낸다.

$$X_i = S_i QP_i, \quad X_p = S_p QP_s, \quad X_b = S_b QP_b$$

여기서 S_i , S_p , 그리고 S_b 는 이전 영상에서 사용한 비트 수이고 QP_i , QP_p , 그리고 QP_b 는 같은 부호화 방법에서 이전 영상의 평균 QP (Quantization Parameter)값이다. P영상의 복잡도 X_p 는 P영상의 예측오차를 반영하지만 그것은 I 영상 부호화에 의해 영향을 받기 때문에 바로 inter-frame 상관성을 나타내지는 않는다. I 영상 부호화의 영향을 측정하기 위해서는 I 영상의 복잡도 X_i 에 의해 나누어진 X_p 의 값이 사용된다. I와 P-picture의 부호화로 영상을 위한 부호화의 복잡도가 계산되고 P picture 와 I picture 사이의 복잡도를 X_p/X_i 는 그 다음 계산된다. 그리고 이것은 비트율과 M값에 따라서 사전에 결정된 경계 값과 비교된다. X_p/X_i 가 $M=3$ 상의 경계 값 보다 더 클 때는 현재의 세 영상을 위해 M은 1로 정해지고 그렇지 않을 경우에는 3을 유지한다. 반대로 X_p/X_i 가 $M=1$ 상의 경계 값 보다 작으면 현재의 세 개의 영상에 대한 M 값은 3으로 되고 그렇지 않으면 1을 유지한다. 버퍼 제어에서 현재 영상의 비트 사용량으로 다음 영상의 양자화 값인 m-quant를 조절하게 된다. 제안한 알고리즘 상에서 $M=3$ 에서 $M=1$ 로 변할 경우 GOP내

의 B-picture의 개수가 변하고 low_delay 모드와 같은 상황이 일어난다. 이 경우 버퍼의 underflow가 발생한다. 버퍼의 underflow가 발생하지 않도록 vbv_buffer의 크기나 vbv_delay값을 조절해 주어야 한다. 만약 M=3에서 M=1로 변하는 경우 TM5상에서의 M=1일 경우의 GOP내의 사용 비트 량과 vbv_delay값을 계산해주어야 한다, M=1에서 M=3으로 변하는 경우 TM5상에서의 M=3일 경우의 GOP내의 사용 비트 량과 vbv_delay값을 계산해주어야 한다.

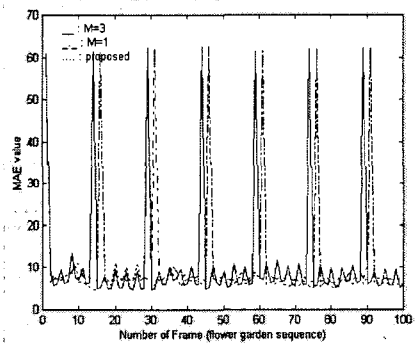
IV. 실험결과 및 고찰

제안한 방법으로 MPEG-2 TM5와 비교하여 예측 오차를 평가하였으며 평가에는 MPEG-2 MP@ML을 사용하였다. 전송 비트율은 3Mbps 이고 사용된 영상은 704×480으로 "flower garden", "susie", "table tennis"로 각각 100frame을 사용하였다. 코딩 성능을 측정하기 위해 일반적으로 예측오차를 평가한다. 부호화된 영상의 PSNR은 B영상을 위한 양자화 크기가 일반적으로 K_b 로 TM5을 제어방법으로 조절하기 때문에 사용하지 않는다. PSNR의 전체 평균은 B영상의 수가 서로 다르므로 각기 다른 GOP구조 사이에서는 비교하기가 적당하지 않다. 그러나 예측오차는 영상에서 inter-frame 예측에 의해 참조되지 않기 때문에 B영상의 양자화에 의해서 영향을 받지 않는다. 예측 오차는 다음과 같이 표현된다.

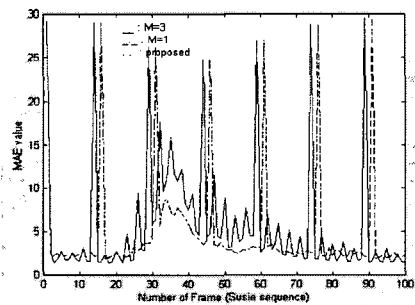
$$MAE = \frac{1}{H \times V} \sum_{H=0}^{H-1} \sum_{V=0}^{V-1} |X(i, j) - X'(i, j)|$$

여기서 $X(i, j)$ 는 현재 프레임의 (i, j) 위치의 화소 값을, $X'(i, j)$ 는 움직임 예측에 의해 얻어진 프레임의 (i, j) 위치의 화소 값을 나타내며 H 와 V 는 각각 수평과 수직 화소수를 나타낸다.

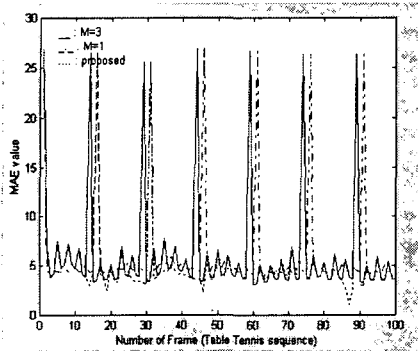
일반적으로 영상내의 복잡도 변수 X_b / X_i 는 M=3일 경우 0.45~0.62에 분포되고 M=1일 경우 0.27~0.42에 분포된다. 적응적으로 GOP구조를 변화시키기 위해 경계 값을 각각 0.33~0.53, 0.34~0.54, 0.35~0.55, 0.36~0.56, 0.47~0.57로 하였으며 이를 각 영상에 대입하여 실험하였다. 그림 1, 2, 3은 각각 Flower garden, Susie, Table Tennis 영상의 M=3, M=1 과 제안한 방법의 MAE 값을 보여주고 있다. 그림 1은 경계 값이 0.35~0.55일 때의 그래프로 10~20 frame과 60~80 frame 근처에서 예측 오차 값이 두드러지게 감소하는 것을 볼 수 있다.



▶▶ 그림 1. Flower Garden 영상의 MAE 값



▶▶ 그림 2. Susie 영상의 MAE



▶▶ 그림 3. Table Tennis 영상의 MAE

그림 2는 susie 영상의 경계 값이 0.34 ~ 0.54일 때의 그래프이다. 그래프 상에서 30frame~60frame 사이에서 예측 오차 값이 눈에 띄게 감소하였다. M=3일 경우 15까지 예측 오차 값이 올라갔으나 제안한 알고리즘에서는 8정도에 머물고 있어서 안정적이 되었다는 것을 알 수 있다. 그림 3은 susie 영상의 경계 값이 0.35~0.55일 때의 그래프이다. 20frame ~ 40frame 사이에서 예측오차가 감소하였으며 88frame 근처에서 예측 오차 값이 2정도까지 내려갔다가 다시 높게 올라가는 것을 볼 수 있는데 이는 vbv_delay 값의 조절로 양자화 값이 감소했다가 다시 증가했기 때문이다.

V. 결론

기존에 제안된 GOP 구조 제어 알고리즘들은 MB 활동도와 MB활동 도메인상의 단순 움직임 예측을 사용하는 방법은 사전분석이 필요하여 실시간 적응적 GOP 구조 선택에는 적합하지 못하였다. 본 논문에서는 실시간 MPEG-2 비디오 코딩을 위한 영상 복잡도에 기반한 적응적 GOP구조 선택 방법을 제안하였다. 이 방법은 코딩 복잡도를 기반으로 하여 I 영상과 P 영상의 복잡도 값을 계산하여 각 GOP에서 core picture의 길이 M을 작은 예측 오차를 갖는 M=3 과 M=1로 실시간에 적응적으로 선택하고 이에 따라 M=3으로 변화할 경우와 M=1일 변할 경우

GOP내의 사용 비트 량과 vbv_delay 값을 다시 계산해 주었다. 실험에서 제안한 알고리즘은 주어진 경계 값으로 P 영상 자신을 포함한 세 개의 영상에 대하여 M=3과 M=1을 적응적으로 선택하였으며 고정 GOP 구조인 M=3일 경우와 M=1일 경우보다 2%~15% 정도 작은 예측 오차를 보여주었다. 복잡도 분포가 영상에 따라 다른 분포를 보여 주기 때문에 경계 값이 영상에 따라 적용범위가 달라졌으며 가장 좋은 예측오차를 갖는 경계 값도 영상마다 달랐다. 코딩 시에 P 영상에서 복잡도 값을 조사하므로 지연이 발생하지 않았으며 사전분석이나 불필요하고 계산 량도 많지 않았다.

참고 문헌

- [1] ISO/IEC 13818-2, Information technology- Generic coding of Moving Picture and Associated Audio Information-Part2: Video, International Standard(1995)
- [2] M. Lee, S> Kwon, J. Kim, "A scene Adaptive Bitrate Control Method in MPEG Video Coding", SPIE VCIP, Vol.3024, pp.1416, 1998.
- [3] Akio Yoneyama, Y. Nakajima, H. Yanagihara, M. Sugano, "MPEG Encoding Algorithm with Scene adaptive Dynamic GOP Structure", IEEE, pp.832-835, 1999.
- [4] K. S. Choi, Y. D. Moon, "Image Improvement in Virture of Block Coefficient Sampling by usingDCT", 대한전기학회 하계학술대회 논문지, pp.2915-2917, 2000.
- [5] 최금수, 정희태, 문영득"복잡도를 기반한 적응적 GOP 구조 선택", 대한 전기학회 하계학술대회 논문집, pp.2675-2677, 2001.
- [6] Young-Deuk Moon, Kum-Su Choi"Dynamic Control Algorithm of GPO Structure based on Picture Complexity" Trans. KIEE, Vol.53D. No.4, APR. 2004.
- [7] Kum-Su Choi, Young-Deuk Moon" A effective algorithm for the scene change detection on compressed video", ITC-CSCC2004, 6E3L-4_1, July5-8, 2004.