

영상 복잡도에 기반한 GOP구조의 동적 제어 알고리즘

論 文

53D-4-7

Dynamic Control Algorithm of GOP Structure based on Picture Complexity

文榮得*·催金洙**

(Young-Deuk Moon · Kum-Su Choi)

Abstract - This paper propose a method that GOP structure based on the picture complexity change realtime adaptive without pre-analysis or time delay. Proposed algorithm calculates the complexity of pictures at first, and the ratio of the complexity(X_p/X_i) between P picture and I picture is calculated. The suitable M value for the three picture select by comparing with predetermined threshold. Used bit and vbv_delay the value of GOP is calculated according to selected M.

Experimental results show that the prediction error is reduce than the fixed GOP structure. Since the complexity distribution of the sequence is different, applied limits of threshold value is changed, also.

Key Words : MPEG, GOP, Dynamic Control Algorithm, Picture Complexity

1. 서 론

MPEG-1, 2는 비디오CD, DVD, 그리고 위성방송과 지상파 고화질 TV 시스템 등에 많이 응용되고 있는 영상 전송과 저장을 위한 매우 중요한 기술이다. MPEG 표준은 단지 압축된 영상의 비트스트림만을 정의하기 때문에 베이스라인 MPEG-2 TM(Test Model)5의 영상화질을 향상시키려는 알고리즘들이 많이 제안되어 왔다. MPEG에서 GOP(Group of Picture)의 길이가 고정되어 있다하더라도 부호화 효율 측면에서 최상의 화질을 제공하지는 않는다. GOP길이 N은 약 0.5초의 랜덤 접근을 위해 15나 12로 고정된다. 그러나 장면전환을 기대하지 않는 영역에서 장면전환이 효율적인 움직임 보상에 관계될 때 부호화 효율을 유지한다는 것은 어렵다. 이를 개선하기 위해 다이나믹 GOP 길이에 의한 부호화로 장면변화 전에 부호화 효율을 향상시키는 알고리즘들이 제안되어 왔다.

GOP길이를 가변시키기 위한 알고리즘으로는 Yoneyama와 Sugano는 MB(Macro Block)활동도와 MB활동 도메인상의 단순 움직임 예측에 대한 방법을 제안하였다.[1] 이 방법은 GOP 길이를 계산하기 위해 프레임 메모리에 15 ~ 30frame 정도를 미리 읽어들여 각 영상의 MB 활동도를 구해서 움직임 예측오차(Mean Absolute Error)값을 구한다. 그렇지만 GOP길이 계산을 위한 MB활동도와 움직임 예측 오차값을 구하기 위해 많은 프레임 메모리를 요구할 뿐만 아니라 사전분석을 하여야 하므로 계산량이 많아져 시간적

으로 지연이 생겨 실시간 적용에는 적합하지 못하다. GOP 길이를 가변 시키기 위한 또 다른 방법으로 Jiro Katto는 inter-frame 상관계수를 갖는 함수로 코딩 이득을 유도해서 주어진 최대 코딩 이득에 따른 최적의 GOP 구조를 탐색하는 것이다.[2] 그러나 이 방법은 inter-frame 상관을 계산하기 위해 사전분석이 요구되고 전체 GOP의 정보를 얻은 후에 GOP 구조를 결정하기 때문에 역시 지연이 필연적이라서 low-delay 실시간 응용에는 적합하지 않다. 그리고 J. W. Lee등은 Lagrange multiplier 방법을 사용하여 최소화 문제를 해결하므로 선택 가능한 후보로부터 GOP에서 P-picture의 최적수와 위치를 탐색할 수 있게 하였다.[3] 여기서는 GOP 내의 임의의 위치에 P-picture를 있게 하여 최적의 GOP구조가 선택할 수 있다. 그러나 이 방법은 많은 계산량을 요구하며 최적 수와 위치를 탐색하기 위해 지연이 필수적으로 수반된다. 그 외 여러 가지 측면에서 제안된 알고리즘에서도 GOP구조를 결정하기 위해서 일정량의 영상을 미리 읽어들여 사전분석을 하여야 하기 때문에 프레임 메모리를 많이 요구하게 되며 계산량도 많아져 지연이 필수적으로 수반되어 실시간적으로 적용하기에는 적합하지 못하였다.

본 논문에서는 사전분석이나 지연 현상을 없애기 위해 영상 복잡도에 근거하여 실시간 적으로 GOP구조를 변화시키는 적응적 알고리즘을 제안한다. 각 영상 부호화 방법을 위해 영상의 복잡성 변수를 비트 묶음의 예측기로 사용하며 영상에서 복잡도의 평균분포를 구하여 제안한 알고리즘의 문턱값으로 한다. 이 문턱값을 사용하여 최적의 GOP구조를 선택하며 GOP구조가 변경되면 사용 비트량이 달라져서 GOP 내의 사용 비트량과 vbv_delay값이 M=3일 때와 M=1일 때로 구분하여 조절하게 된다.

본 논문의 구성은 2장에서 MPEG의 picture와 GOP구조에 대하여 서술하고 3장에서는 제안한 동적 제어 알고리즘에 대하여 설명한다. 4장에서는 본 논문에서 제안한 방법과

* 正 會 員 : 釜山外大 情報科學大 컴퓨터電子工學科 教授

**學生會員 : 釜山外大 情報科學大 컴퓨터電子工學科

接受日字 : 2003年 4月 11日

最終完了 : 2004年 2月 6日

TM5와의 성능을 비교하여 그 결과와 함께 고찰한 내용을 기술하였으며 5장에서 결론으로 맺고 있다.

2. MPEG의 picture 와 GOP 구조

2.1 Picture 와 GOP구조

MPEG에서 규정하는 picture 형태에는 I, P, B-picture 등 세 가지가 있다. I(Intra coded)-picture는 움직임 예측과 같은 시간적인 압축을 사용하지 않고 DCT변환에 의한 공간적인 압축만을 사용한 화면이다. 시간적인 정보를 사용하지 않기 때문에 시간상으로 누적되는 오류를 방지할 수 있어 다른 화면을 압축 부호화할 때 기준이 되는 참조 picture의 역할을 한다. 그러나 시간적인 정보를 사용하지 않기 때문에 다른 picture들 보다 부호화 할 때 더 많은 비트수를 필요로 하며 단독으로 부호화가 가능하여 임의 접근할 수 있는 기준 영상이기도 하다. P(Predicted coded)-picture는 예측화면으로 직전의 I와 P-picture로 부터 원화면과 가장 가깝게 예측해 낸 화면이다. 즉 현재 화면과 이전 화면의 움직임을 비교하여 변경된 부분의 정보를 움직임 벡터로 표현하여 구성한다. B(Bidirectionally predicted coded)-picture 역시 예측화면이긴 하지만 이전 I, P picture에서만 예측하여 화면을 구성하는 것이 아니라, 그 다음 I, P-picture에서도 역 방향으로 예측한 후 이들 두 가지 방향에 대하여 평균한 화면을 만들어 이 세 가지 화면에서 원화면과 가장 비슷하게 예측되는 화면을 선택하는 다소 복잡한 화면이다. 이런 예측방법을 그림 1에서 보여주고 있다. I 와 P-picture 는 inter-frame 예측을 위해 다른 영상에 참조되므로 core picture라고도 한다.

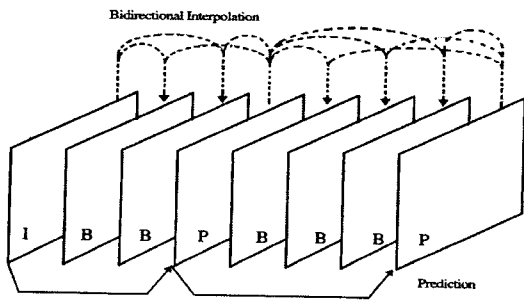


그림 1 MPEG 에서의 Picture 구조
Fig. 1 Picture Structure in MPEG

B-picture는 역 방향 예측에 의해 앞으로 발생할 영상을 참조하기 때문에 picture의 디스플레이 순서와 압축되어 저장되는 순서가 일치하지 않아 프레임의 재 순서화가 필요하기 때문에 이로 인한 지연이 발생하는 단점이 있다. B-picture를 많이 사용하면 발생 비트 수를 줄일 수 있지만 이는 한 방향 움직임 보상이기 때문에 P-picture에 비해 많은 오류를 포함하게 된다. 이러한 오류를 줄이기 위해서 B-picture를 적절하게 선택하여야 한다. 전통적인 방법에서 I, P, B 등 세 가지 영상의 배열은 주기적인 형식으로 정의 된다. core picture 사이의 거리가 3으로 고정되고 GOP내의

영상의 수가 15로 고정되어 있다면 GOP구조는 코딩순서에서 "IBBPBBPBBPBBPBB"로 된다. 영상화질을 향상시키기 위해서 입력영상에 따라 GOP구조가 동적으로 변화되어 진다면 고정 GOP구조 보다 효율적일 것이다.

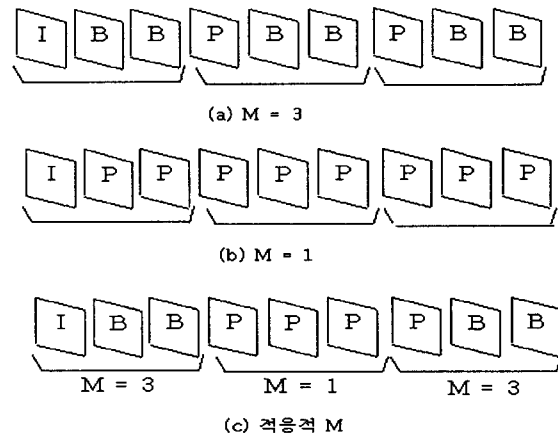


그림 2 MPEG 에서의 GOP 구조
Fig. 2 GOP Structure in MPEG

그림 2는 GOP 구조를 보여주고 있다. M은 core picture 사이의 거리이다. 그림2(a)와 (b)는 각각 M=3과 M=1인 것을 나타내고 있다. MPEG 부호화에서 B-picture의 사용은 높은 부호화 성능을 제공하지만 P-picture의 예측성능은 참조영상으로 부터의 거리가 커지면 감소되며 이에 따라 core picture사이의 B-picture에 대한 최적수가 결정된다. 적응적 방법은 그림 2(c)에서와 같이 MPEG 부호화에 일반적으로 사용되는 M값을 M=1에서 M=3 사이에서 변화시키고 있다.

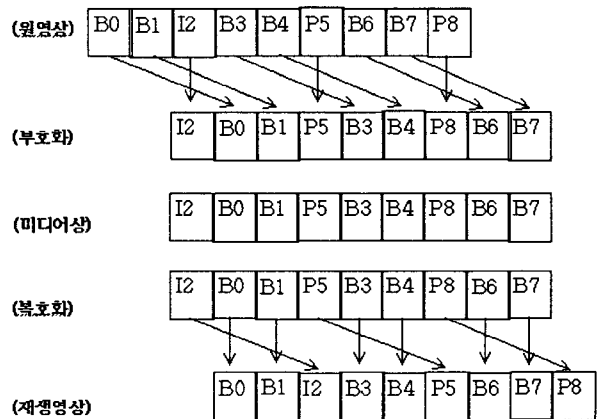


그림 3 부호화 과정의 화면 순서
Fig. 3 Display Order in Coding Processing

그림 3은 부호화를 위한 프레임 순서와 복호화 과정을 보여주고 있다. 부호화에서 I 나 P-picture는 M=3이 사용되는 동안 영상이 입력될 때 즉시 부호화되고 B 영상은 3프레임 뒤에 부호화된다. M이 3에서 1로 변화할 때 모든 프레임은 비록 B영상이 사용되지 않더라도 2개의 프레임을 위해 지연 된다. 부호화 과정은 이러한 순서제어에 의해서 연속적으로

동작한다. 만약 $M=1$ 이 처음에 사용된다면 부호화 과정은 지연된 상태로 시작된다. 부호화 과정을 위해서 부호화된 비트 스트림은 부호화 순서로 복호화된다. 복호화된 영상이 화면에 나타날 때 B영상은 영상이 복호화된 후 바로 나타나게 된다. 그리고 core picture는 다음 core picture가 복호화될 때까지 지연된다. 이러한 디스플레이 규칙에 따라서 복호화된 영상은 연속적으로 디스플레이될 수 있다. 이에 따라 적응적 M선택 제어를 효율적으로 수행한다면 실시간 MPEG-2 코딩에 적용할 수 있을 것이다.

2.2 데이터 전송과 버퍼제어

한 picture당 발생 비트 수는 가변적으로 버퍼의 충만도에 따라서 제어되는데 이 버퍼가 underflow나 overflow가 나지 않게 하기 위해서 인코더에서는 버퍼상태를 항상 감시하여 다음 picture의 비트 발생량을 조절해 주어야 한다. 인코더에서 버퍼를 1000분의 1초마다 조사하든지 아니면 매 picture마다 조사하든지 그것은 인코더에 따라 유연하게 할 수 있다. 즉 주어진 버퍼량에 대하여 underflow나 overflow가 발생하지 않게 하여야 한다. 그러나 이 작용이 최적화되지 않으면 정보량이 많은 picture에서 비트 할당을 많이 해주지 못하는 경우가 생기며 또 그 반대의 경우도 발생할 수 있다.

MPEG에서는 실제의 버퍼체계와는 관계없이 하나의 가상 버퍼 시스템에 대해서 언급하고 있는데 이것이 바로 VBV(video buffering verifier)이다. 이 VBV 버퍼란 인코더가 실제로 가지고 있지는 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. 버퍼의 동작을 나타내면 그림 4와 같다.

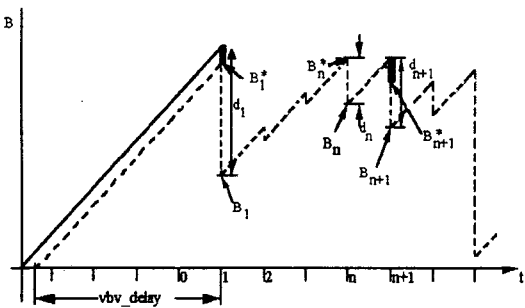


그림 4 VBV 버퍼의 동작
Fig. 4 VBV Buffer occupancy

그림4에서 picture의 데이터를 전송하는 시간간격은 $t_{n+1} - t_n$ 이며 progressive sequence의 경우에는 frame rate의 역수인 T 로 표시한다. B 는 VBV 버퍼의 크기이며 각 시간간격 내에서 직선의 기울기는 비트율인 R 이 된다. 그러므로 버퍼의 입력 단에는 매 초당 R bit씩의 데이터가 계속 들어오고 출력 단에는 T 간격으로 매 picture의 데이

터를 한꺼번에 뽑아가서 디코딩하고 디스플레이까지 한다. 그러나 디코딩을 맨 처음 시작할 때는 버퍼 내에 데이터가 없기 때문에 버퍼가 어느 정도 차는 것을 기다려야 하며 디코더가 기다려야 하는 시간을 vbv_delay 라 한다. 이 vbv_delay 는 MPEG-1의 시스템 클럭인 90kHz의 주기로 나타나며 $vbv_delay = (90000 \times B^*) / R$ 가 된다. 여기서 B^* 이란 $t=n$ 일 때 VBV 버퍼가 비트수로 얼마나 차 있는가를 나타낸다.

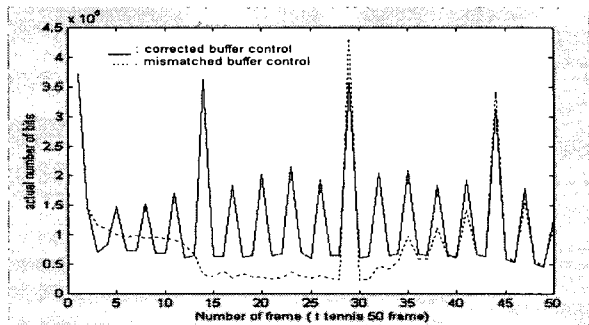


그림 5 영상에서의 비트율
Fig. 5 bitrate in MPEG sequence

그림 5에서는 버퍼 제어가 잘된 경우와 그렇지 않은 경우를 나타내고 있다. 실선은 정확한 버퍼제어에 의해서 사용 비트가 일정한 수준에서 진동하는 것을 볼 수 있다. 점선은 15frame에서 부터 낮은 비트율을 나타내며 underflow를 일으키고 29 frame은 비트 할당이 많아짐을 알 수 있다. 그러나 40 frame에서 부터는 안정되고 있음을 보여주고 있다. 그림5에서 처럼 underflow가 발생할 경우에 대하여는 버퍼 제어를 통해 GOP내의 사용 비트를 증가시키고 vbv_delay 값도 조절해 주어야 한다.

3. 제안한 GOP구조의 동적 제어 알고리즘

최적의 GOP 구조는 inter-frame 상관성에 의존하게 됨에 따라 inter-frame 상관성 계산을 위해 사전분석이 필요하다.[2] 상관성은 inter-frame 예측 부호화된 영상의 비트 카운터에 연관되므로 제안한 알고리즘에서는 TM5의 복잡성 변수 X_i , X_p , 그리고 X_b 를 각 영상 부호화방법을 위한 비트 묶음의 예측기로 사용한다.[4] 이러한 변수는 식(1)에서 보여주는 것과 같이 부호화 방법에서 이전 영상의 정보를 사용하여 나타낸다

$$X_i = S_i Q P_i, \quad X_p = S_p Q P_p, \quad X_b = S_b Q P_b \quad (1)$$

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 I 영상 부호화에 의해 영향을 받기 때문에 바로 inter-frame 상관성을 나타내지는 않는다. I 영상 부호화의 영향을 측정

하기 위해서는 I 영상의 복잡도 X_i 에 의해 나누어진 X_b 의 값이 사용된다. 버퍼제어에서 현재 영상의 비트 사용량으로 다음 영상의 양자화 값인 mquant를 조절하게 된다. 제안한 알고리즘 상에서 M=3에서 M=1로 변할 경우 GOP내의 B-picture의 개수가 변하고 low_delay 모드와 같은 상황이 일어난다. 이 경우 버퍼의 underflow가 발생하게 되는데 이를 방지하기 위하여 vbv_buffer의 크기나 vbv_delay값을 조절해 주어야 한다. M=3에서 M=1로 변하는 경우와 M=1에서 M=3으로 변하는 경우에 다시 계산된 GOP내의 사용 비트량과 vbv_delay값으로 underflow가 발생하는 경우를 방지할 수 있다.

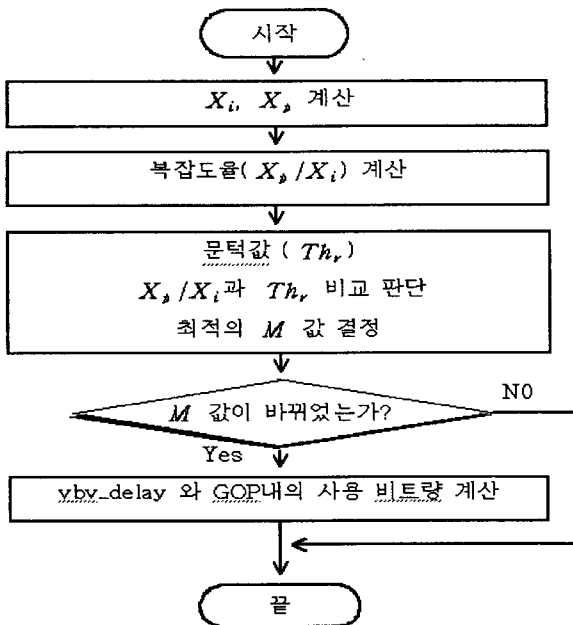


그림 6 제안한 알고리즘의 흐름도
Fig. 6 Flowchart of Proposed Algorithm

그림6은 제안한 알고리즘의 흐름도이며 I 와 P-picture의 부호화로 영상을 위한 부호화의 복잡도가 계산되고 P picture 와 I picture 사이의 복잡도율 X_b/X_i 는 그 다음 계산된다. 그리고 이것은 M값에 따라서 사전에 결정된 문턱값과 비교된다. X_b/X_i 가 M=3 상의 문턱값 보다 더 클 때는 현재의 세 영상을 위해 M은 1로 정해지고 그렇지 않을 경우에는 3을 유지한다. 한편 X_b/X_i 가 M=1상의 문턱값 보다 작으면 현재의 세 개의 영상에 대한 M 값은 3으로 되고 그렇지 않으면 1을 유지한다. 또한 M=3에서 M=1로 변하는 경우 GOP내의 사용 비트량과 vbv_delay값이 TM5상에서의 M=1일 경우로 계산되고 M=1에서 M=3으로 변하는 경우 GOP내의 사용 비트량과 vbv_delay값을 TM5상에서의 M=3일 경우로 계산한다.

4. 실험결과 및 고찰

제안한 알고리즘의 성능을 평가하기 위해 MPEG-2 TM5와 고정 GOP구조인 경우와 예측오차를 평가하였으며

사용된 영상은 704×480 "flower garden", "susie", "table tennis"로 MPEG-2 Main Profile, High Level (MP@HL)을 사용하였다.[5][6] 전송 비트율은 3Mbps 이고 각각 100frame을 사용하였다. 코딩 성능을 측정하기 위해 일반적으로 예측오차를 평가한다. 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 의해서는 영향을 받지 않는다.[4] 예측 오차는 식(2)와 같이 나타낼 수 있다.

$$MAE = \frac{1}{H \times V} \sum_{i=0}^{H-1} \sum_{j=0}^{V-1} |X(i, j) - X'(i, j)| \quad (2)$$

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

우선 영상열의 복잡도의 분포를 알아보기 위하여 각 영상의 복잡도 변수 X_b/X_i 를 조사하여 M=3일 경우와 M=1일 경우의 X_b/X_i 가 어떤 분포를 보이는가를 알아보았다. 이를 그래프로 나타낸 것이 그림 7과 8이다. 그림 7과 8은 각각 M=3일 때와 M=1일 때의 각 영상의 복잡도 변수 X_b/X_i 의 분포를 보여주고 있다. 그래프에서 알 수 있듯이 일반적으로 영상내의 복잡도 변수 X_b/X_i 는 M=3일 경우 0.45 ~ 0.62에 집중되고 M=1일 경우 0.27 ~ 0.42에 집중된다. 또한 복잡도 변수 X_b/X_i 의 구간을 정하기 위해서 각 영상열에서 X_b/X_i 의 평균을 구하였다.

<표-1>에 실험에 사용된 세 영상의 복잡도 변수 X_b/X_i 의 평균값을 나타내었다. 복잡성 변수의 분포와 평균값을 바탕으로 M=3일 경우에 임계값을 0.5에서 0.6정도로 하는 것이 적당하고 M=1일 경우에 0.3에서 0.4 정도로 정하는 것이 적당하다는 것을 알 수 있었다. 이를 참조하여 적응적으로 GOP구조를 변화시키기 위해 문턱값을 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.

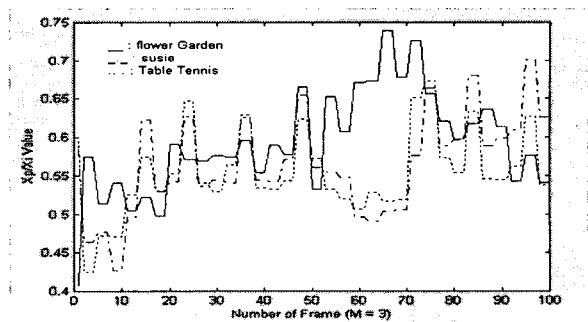


그림 7 M=3일 경우의 각 영상의 복잡도율(X_b/X_i)
Fig. 7 X_b/X_i value of each sequence in M=3

표 2. TM5와 제안한 방법의 예측오차(MAE)

Table 2. MAE Value of TM5 and Proposed Algorithm

(frame = 100)

		0.33~0.53	0.34~0.54	0.35~0.55	0.36~0.56	0.37~0.57
Flower Garden	M=3	13.1242	13.1242	13.1242	13.1242	13.1242
	M=1	14.0994	14.0994	14.0994	14.0994	14.0994
	proposed	12.6088	12.6088	12.6088	15.3910	14.2736
Susie	M=3	7.5040	7.5040	7.5040	7.5040	7.5040
	M=1	5.8210	5.8210	5.8210	5.8210	5.8210
	proposed	6.2786	5.6940	5.6940	5.6940	5.6940
Table Tennis	M=3	8.3518	8.3518	8.3518	8.3518	8.3518
	M=1	8.5944	8.5944	8.5944	8.5944	8.5944
	proposed	7.4120	7.8900	7.8834	7.0952	7.4497

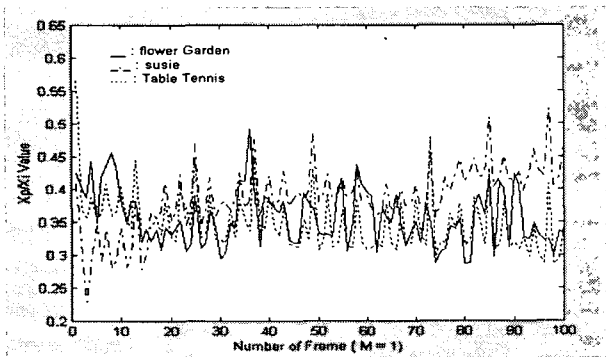


그림 8 M=1일 경우의 각 영상의 복잡도(X_p/X_i)

Fig. 8 X_p/X_i value of each sequence in M=1

표 1 각 영상에서의 복잡도(X_p/X_i) 평균값

Table 1 X_p/X_i mean value of each sequence

	Flower Garden	Susie	Table Tennis
M = 3	0.5951	0.5650	0.5548
M = 1	0.3578	0.3884	0.3497

<표-2>에 각 영상의 X_p/X_i 의 문턱값에 따른 고정 GOP구조인 TM5와 제안한 알고리즘에 의해 구해진 예측 오차값을 보여준다. 각 영상에 따라 X_p/X_i 의 문턱값이 변화할 때 예측오차의 차이값이 변화함을 볼 수 있으며, 각 영상마다 가장 좋은 결과를 보여주는 문턱값이 서로 다를 수 있다. flower garden은 0.35 ~ 0.55 이하에서는 동일한 결과를 보여주며 0.36 ~ 0.56 이상에서는 예측 오차값

이 커지는 것을 볼 수 있다. M=3 보다는 4%정도의 향상을 보여주고 M=1인 경우보다는 11%정도의 향상을 보여준다. susie영상은 M=3이나 M=1보다 좋은 결과를 보여주고 있으나 0.33 ~ 0.53에서는 M=3보다는 적은 오차값을 보여주지만 M=1일 때보다는 큰 값을 보여준다. 이는 영상의 복잡도의 분포가 영상마다 다르기 때문에 문턱값이 변화에 따라 GOP구조가 민감하게 변화하기 때문이다. susie 영상에서 예측 오차값은 M=3일 경우보다는 24%정도 M=1일 경우보다는 2% 정도의 향상을 보여준다. table tennis영상은 모든 구간에서 M=3이나 M=1일 때 보다 적은 예측 오차값을 보여준다. 특히 0.36 ~ 0.37에서 가장 큰 차이를 보여주고 있으며 M=3일 경우보다는 15%정도 M=1일 경우보다는 17%정도의 향상을 보여준다. 이러한 결과로부터 제안한 알고리즘은 M=3일 경우와 M=1일 경우보다 예측 오차값이 향상되었으며 문턱값은 영상의 특성에 따라 다르게 나타나는 것을 알 수 있었다.

그림 9, 10, 11은 각각 flower garden, susie, table tennis 영상의 고정 GOP구조상의 M=3, M=1일 경우와 제안한 방법에 의한 예측오차(MAE)값을 보여주고 있다. 세 영상 모두 고정 GOP 구조인 경우 보다 예측오차가 감소하는 것을 볼 수 있다. 그림 9는 flower garden 영상으로 문턱값이 0.35 ~ 0.55일 때의 그래프로 10 ~ 20 frame 과 60 ~ 80 frame 근처에서 예측 오차값이 값이 두드러지게 감소하는 것을 볼 수 있다. 그림 10은 susie 영상의 문턱값이 0.34 ~ 0.54일 때의 그래프이다. 그래프 상에서 30 ~ 60 frame 사이에서 예측 오차값이 눈에 띄게 감소하였다. 특히 M=3일 경우 15까지 예측 오차값이 올라갔으나 제안한 알고리즘에서는 8정도에 머물고 있어서 안정적이 되었다는 것을 알 수 있다. 그림 11은 table tennis 영상의 문턱값이 0.35 ~ 0.55일 때의 그래프이다. 20 ~ 40 frame 사이에서 예측오차가 감소하였으며 88 frame 근처에서 예측 오차값이 2정도까지 내려갔다가 다시 높게 올라가는 것을 볼 수 있는데 이는 사용 비트량의 변화로 양자화 값이 감소했다가 다시 증가했기 때문이다.

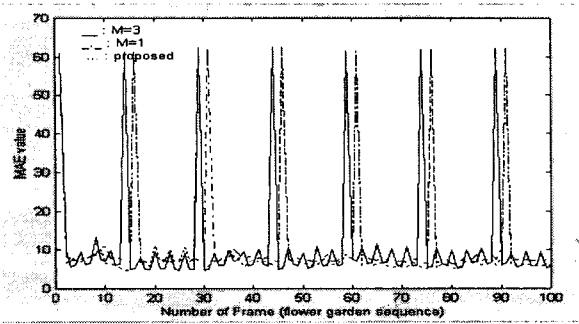


그림 9 Flower Garden 영상의 예측오차(MAE) 값
Fig. 9 MAE Value of Flower Garden sequence

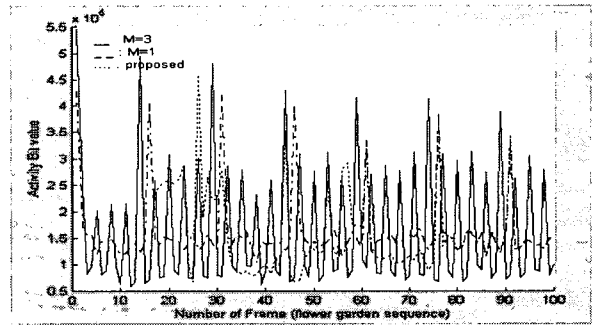


그림 12 Flower Garden 영상의 비트율
Fig. 12 bitrate of Flower Garden sequence

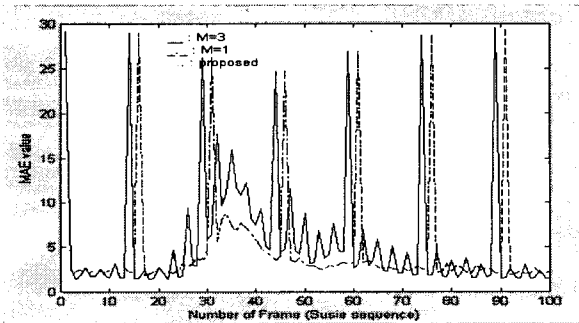


그림 10 Susie 영상의 예측오차(MAE) 값
Fig. 10 MAE Value of Susie sequence

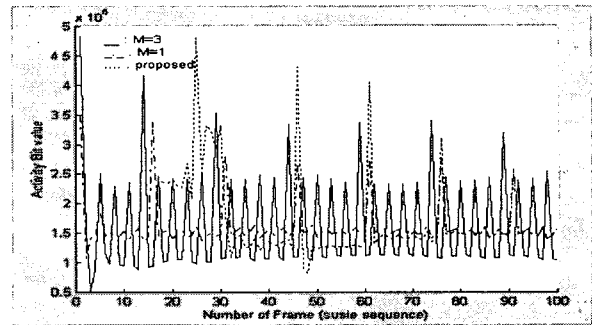


그림 13 Susie 영상의 비트율
Fig. 13 bitrate of Susie sequence

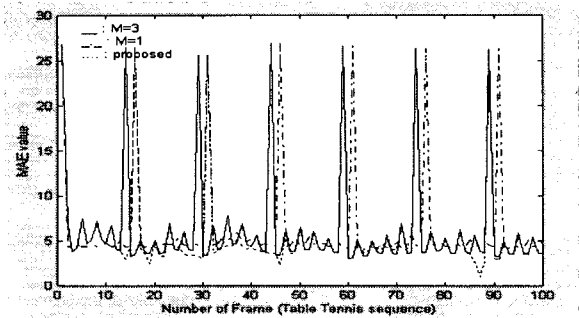


그림 11 Table Tennis 영상의 예측오차(MAE) 값
Fig. 11 MAE Value of Table Tennis sequence

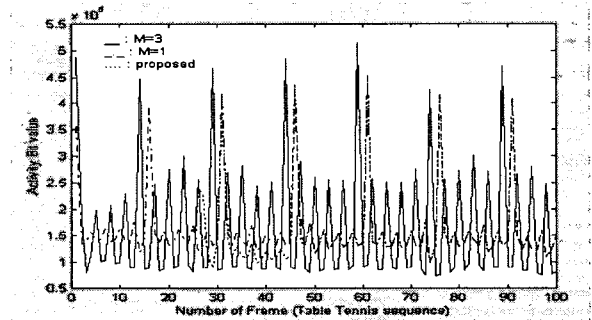


그림 14 Table Tennis 영상의 비트율
Fig. 14 bitrate of Table Tennis sequence

그림 12, 13, 14는 각각 flower garden, susie, table tennis 영상의 고정 GOP구조상의 M=3, M=1일 경우 과 제안한 방법에 의한 비트율을 보여주고 있다. M=3일 경우와 M=1일 경우에 B 영상의 차이로 진폭의 크기가 많은 차이가 나는 것을 알 수 있으며 제안한 방법은 M=3일 경우와 M=1일 경우의 중간 정도의 진폭을 나타냄을 알 수 있다. 이는 사용 비트량에서 고정 GOP구조보다 안정적이 되었다고 말할 수 있다. 그렇지만 flower garden영상의 16 ~ 25 frame 과 susie영상의 16 ~ 30 frame은 M=3일 경우보다 조금 많은 비트량을 보여주고 있다. 그러나 전체적으로 비트사용량이 안정되고 제안한 방법에 의한 비트율이 M=3일 경우의 비트 사용량을 벗어나는 경우는 거의 없었다.

5. 결 론

적용적 GOP 선택 알고리즘에 있어서 기존에 제안된 MB활동도와 MB활동 도메인상의 단순 움직임 예측을 사용하는 방법은 좋은 GOP 구조를 찾을 수 있으나 사전분석이 필요하기 때문에 실시간적으로 적용하는 GOP 구조 선택에는 적합하지 않을 뿐만 아니라 많은 계산량이 수반된다. 본 논문에서는 실시간 MPEG-2 비디오 코딩을 위해 영상 복잡도에 기반하여 실시간으로 적용할 수 있는 GOP구조 선택 방법을 제안하였다. 이 방법은 코딩 복잡도를 기반으로 하여 I 영상과 P 영상의 복잡도율을 계산하여 각 GOP에서 core picture의 길이 M을 예측 오차가 적은 M=3 과 M=1로 실시간에 적용적으로 선택하고 이에 따라 M=3으로 변화할 경우와 M=1로 변할 경우에 GOP 내의 사용 비트량과 vbv_

delay 값을 다시 계산해 주었다.

제안한 알고리즘은 주어진 문턱값으로 P 영상 자신을 포함한 세 개의 영상에 대하여 M=3과 M=1을 적용적으로 선택하였으며 고정 GOP 구조인 M=3일 경우와 M=1일 경우보다 2% ~ 15% 정도 예측오차가 감소함을 알 수 있었다. 복잡도 분포가 영상에 따라 다른 분포를 보여 주기 때문에 문턱값이 영상에 따라 적용범위가 달라졌으며 가장 좋은 예측오차로 나타나는 문턱값도 영상마다 달랐다. 평균적으로 문턱값이 0.35 ~ 0.55 로 향상된 성능을 보여주고 있으며 대부분의 영상에 M값을 적용할 수 있었다. 예측오차가 개선됨에 따라 영상의 PSNR도 향상되었다. 코딩시 P 영상에서 복잡도율을 기반으로 하여 M값이 실시간으로 적용하므로 시간지연을 해소할 수 있었으며 또한 사전분석이 필요하지 않을 뿐만 아니라 계산량도 현저히 줄일 수 있었다.

참 고 문 헌

[1] Akio Yoneyama, Y. Nakajima, H. Yanagihara, M. Sugano, "MPEG Encoding Algorithm with Scene adaptive Dynamic GOP Structure", IEEE, pp.832-835, 1999

[2] Jiro Katto, "mathematical analysis of MPEG compression capability and its application to rate control", ICIIP-95 Vol.II, pp.555-558(Oct. 1995)

[3] J. W. Lee et al., "Rate-Distortion Optimized Frame Type Selection for MPEG Encoding", IEEE Transaction on Circuit and its application to rate control", Vol.7, No.3 (jung 1997)

[4] 최금수, 정희태, 문영득, "복잡도를 기반한 적응적 GOP 구조 선택", 대한전기학회 하계학술대회 논문집, pp.2675-2677, 2001

[5] ISO/IEC-JTC1/SC29/WG11, "Test Model 5 (Draft)", "MPEG93/NO0400 (1993)

[6] J. L. Mitchell, W. B. Pennebaker, C. E. Fogg and D. J. LeGall, "MPEG Video Compression Standard", in Digital Multimedia Standards Series, Chapman & Hall, New York, NY, 1997. pp.179-181

[7] ISO/IEC 13818-1: "Information Technology Generic coding of Moving Picture and Associated Audio". Committee Draft, Mar.1994

[8] ISO/IEC 13818-2, Information technology - Generic coding of moving picture and associated audio information - Part 2: Video", International Standard (1995)

[9] Y. Yu, J. Zhou, Y. Wang, "A Fast Effective Scene Change Detection and AdaptiveRate Control Algorithm", Proc.ICIIP'98,ICIIP 1998.

[10] M. Lee, S. Kwon, J. Kim, "A Scene Adaptive Bitrate Control Method in MPEG Video Coding", SPIE VCIP, Vol.3024,pp.1416,1998.

[11] N. Zhang, C. Wu, "Augmenting STM5 Bit Rate Control Scheme for Improving Coding Quality of MPEG2 Video Encoders", Proc. ISCAS'97,Vol.2, pp.1297-1300, 1997.

[12] Ooi et al., "An MPEG-2 Encoder Architecture based on a Single-Chip Dedicated LSI with a Control MPU", ICASSP'97, Vol. 1, pp.599-602 (Apr.1997)

[13] B. G. Haskell, A. Puri, and A. N. Netravali, "Digital Video : An Introduction to MPEG2", Chapman & Hall, 1996.

[14] L. Teixeira, H. Rebeiro, "Analysis of a Two Step MPEG Video System", Proc. ICIIP'97, Vol.1, pp. 350-353, 1997

감사의 글

이 논문은 2004년도 brian busan21 사업에 의해 지원되었음

저 자 소 개

문 영 득 (文 榮 得)



1948년 8월 14일생. 1973년 동아대학교 전자공학과 졸업. 1977년 동 대학원 전자공학과 졸업(석사). 1987년 동 대학원 전자공학과 졸업(공학박). 1991년~1992년 영국 맨체스터대학교 객원교수. 2003년 캐나다 알버타대학교 객원교수. 1983년~현재 부산외국어대학교 정보과학대학 디지털공학부 교수
 관심분야 : Digital Circuit Design & Test, Digital Image Processing, MPEG-2 & 4, JPEG-2000, Multimedia Contents Design.
 Tel : 051-640-3175 / Fax : 051-640-3575
 E-mail : ydm@pufs.ac.kr

최 금 수 (崔 金 洙)



1977년 3월 18일생. 2000년 부산외대 전자공학과 졸업. 2002년 동 대학원 전자컴퓨터공학과 졸업(석사). 2002년~현재 동 대학원 전자컴퓨터공학과 박사과정
 Tel : 051-640-3175
 E-mail : mclarenf11@korea.com