

MPEG-2에서의 영상에 적합한 GOP 할당 기법

전승홍, 조남익

서울대학교 전기컴퓨터공학부

e-mail : iplita@ispl.snu.ac.kr, nicho@snu.ac.kr

Scene Adaptive GOP Allocation in MPEG-2

Seung-Hong Jeon, Nam-Ik Cho

School of Electrical Engineering & Computer Science

Seoul National University

Abstract

Fixed GOP allocation in MPEG-2 cannot cope with scene change and amount of motion, which results in degradation picture quality. By finding suitable N and M and allocating dynamic GOP, the improvement of picture quality can be achieved. In this paper, N and M are determined by scene change detection and estimation of amount of motion using color histogram per each macroblock. The simulation results show that the average PSNR is improved, especially around the shot boundaries.

I. 서론

DVD와 디지털 TV 등의 표준 형식으로 자리잡아 현재 널리 쓰이는 MPEG-2는 Test Model 5(TM5)[5]를 기본으로 하여 화질을 향상시키려는 여러 가지의 시도가 있다. 그런 시도의 대부분은 TM5의 고정 비트율(CBR) 대신 가변 비트율(VBR)을 적용하는 것과 고정된 GOP(Group Of Pictures)를 할당하는 대신 장면 전환이나 움직임의 크기에 따라 그에 맞는 GOP를 할당하고자 하는 것으로 나눌 수 있다.

장면 전환에 맞게 GOP를 할당하는 방법은 [1-3] 등에서 제시되었다. TM5는 GOP의 길이인 N을 미리 고정한 값(주로 12 또는 15)으로 할당한다. 그러나 장면 전환이 있을 경우에는 고정된 N으로는 이에 대하여 제대로 대처할 수 없다. 하나의 GOP안에 서로 다른 두 장면이 포함되었을 경우에 서로 다른 장면 사이에서는 움직임 예측이 제대로 수행되지 못한다. 즉 새로운 장면이 시작된 직후의 픽처가 전방 예측을 수행할 때는 그 참조 프레임이 다른 장면에 속하기 때문에 전방 예측이 제대로 수행되지 못한다. 이로 인하여 화질의 열화가 일어나게 되고 이후의 인코딩에서도 이와 같이 나쁘게 인코딩된 픽처를 참조하게 되므로 화질의 열화가 생긴다. 그림 1은 장면 전환이 있는 영상을 고정된 GOP를 할당하는 TM5로 인코딩한 결과를 보여 주고 있다. 화살표로 표시된 장면 전환 부분에서 화질

이 떨어지는 것을 확인할 수 있다. 이러한 상황을 막기 위해서는 서로 다른 장면에는 서로 다른 GOP를 할당하여 인코딩 시에 아무 관련이 없도록 해야 한다. 즉 새로운 장면이 시작하는 픽처를 I 픽처로 하여 인트라 코딩을 하면 움직임 예측 없이 새로운 장면을 올바르게 인코딩 할 수 있다[1]. 그리고 여기서 장면 전환을 올바르게 검출하는 것 역시 중요한데, 이웃하는 두 픽처의 픽셀별 밝기의 차이를 이용하는 방법[1]이나 평균 밝기 및 평균 분산 등 픽처의 특징을 비교하는 방법[2-3] 등이 제시되었다.

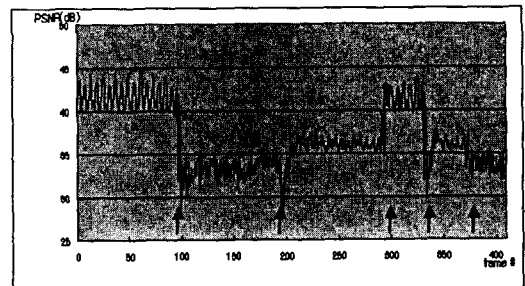


그림 1 장면 전환 부분에서의 화질저하

Fig. 1 Quality degradation at scene change

그리고 GOP의 길이와 더불어 GOP내의 I-P/P-P 간격인 M을 조정하여 화질을 향상시킬 수 있다[2]. 일반

적으로 TM5에서는 $M = 3$ 이 널리 쓰인다. 그러나 움직임이 매우 클 경우에는 I-P/P-P 사이의 간격이 넓어서 전망 예측이 잘 수행되지 않을 가능성이 있다. 이런 경우에 M 을 3보다 작게 하면 움직임 예측이 잘 이루어질 수 있고 이에 따른 화질의 향상을 기대할 수 있다. 반대로 움직임이 거의 없는 경우에는 M 이 약간 커도 움직임 예측이 잘 수행되므로 3보다 큰 값을 할당해준다. [2]에서는 간단한 움직임 예측 알고리즘을 제시하고 이를 통하여 움직임의 크기를 미리 계산하여 이에 맞는 M 을 할당하였다.

본 논문에서는 매크로블록별로 컬러 히스토그램을 만들어 이를 이용하여 GOP를 할당하고자 한다. 우선 매크로블록별 컬러 히스토그램의 합으로 한 픽처 전체의 컬러 히스토그램을 만들어 이를 이전 픽처의 것과 비교하여 장면 전환을 검출하고 N 을 결정한다. 그리고 이에 따라 정해진 N 픽처에 대하여 이웃하는 픽처의 같은 위치에 있는 매크로블록의 히스토그램을 비교하고 이를 이용하여 움직임의 크기를 예측하고 그에 따른 적절한 M 을 결정하여 최종적으로 GOP를 할당한다.

II. 적합한 GOP 할당기법

2.1 매크로블록별 컬러 히스토그램

본 논문에서 제시하는 알고리즘은 전체적으로 매크로블록별 컬러 히스토그램에 기반하고 있다. 이는 각 매크로블록마다 YCbCr의 컬러 성분별로 히스토그램을 작성하는 것이다. 히스토그램의 계급은 3가지 컬러 각각 0~255의 값을 가지므로 $256 \cdot 3$ 개이다. 일반적으로 각 프레임의 히스토그램은 화면 내부의 움직임에 큰 영향을 받지 않는 특징이 있는데, 이를 각 매크로블록별로 나누어서 작성하여 움직임을 반영할 수 있는 지표로 삼고자 하였다. 그림 2에서 보듯이 두 그림 전체에 대한 컬러 히스토그램을 작성하면 완전히 동일하지만 이를 매크로블록별로 작성할 경우 움직임이 있는 블록일 때는 히스토그램이 달라지게 되어 움직임이 있는지 파악할 수 있다.

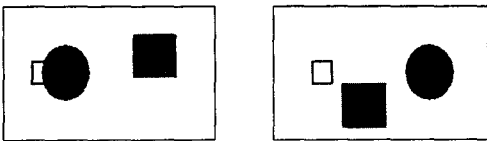


그림 2 매크로블록별 히스토그램
Fig. 2 Histogram per macroblock

2.2 장면 전환 검출 및 N 결정

N 을 결정하기 위해서는 장면 전환을 올바르게 검출할 수 있는 방법이 제시되어야 한다. 이는 기본적으로 장면이 전환되었을 때와 전환되지 않고 같은 장면이 계속되는 경우 이웃하는 프레임 사이에 어떤 차이가 있는지 알아내고 이를 기준으로 삼는 문제가 된다. 이에 대한 내용은 이미 동영상에 대한 인덱싱을 위하여 오래전부터 연구되어 왔다. 그 중 특히 본 논문에서는 인코딩에 크게 지장을 주지 않는, 단순하고 빠른 알고리즘이 필요한데 이에 적합한 방법으로는 두 프레임 간의 픽셀별 밝기 차이의 합(밝기의 SAD)을 이용하는 방법[1]과 프레임의 평균 밝기, 블록별 분산의 평균과 분산 등 프레임의 특성값을 이용하는 방법[2-3], 컬러 히스토그램의 공통부분의 비율을 이용하는 방법[4] 등이 있다. 밝기의 SAD를 이용하는 방법은 픽처 내의 움직임에 매우 취약한 약점이 있고, 이후 움직임의 크기를 측정할 때 작성하는 매크로블록별 컬러 히스토그램을 쉽게 이용할 수 있으므로 본 논문에서는 컬러 히스토그램을 이용하는 방법을 적용하였다. 이는 픽처 전체에 대한 컬러 히스토그램을 그리고 이를 이전 프레임의 히스토그램과 비교하는 방법으로 장면 전환인 곳에서는 두 히스토그램의 공통부분의 비율이 작아지는 것을 이용한다. 우선 이전 픽처의 히스토그램을 O , 현재 픽처의 것은 C 라 할 때 공통부분의 비율 H 는 다음과 같다.

$$H(O, C) = \frac{\sum_j \min(O_j, C_j)}{(\text{전체픽셀수} \times \text{컬러수})} \dots\dots(1)$$

j 는 히스토그램의 계급의 인덱스이고 n 은 전체 계급수를 나타낸다. H 는 0과 1사이의 값을 가지는데 같은 장면일 경우 1에 가깝다. H 가 문턱값보다 작으면 장면전환으로 판단하는데 문턱값 Th 는 다음 (2)와 같이 결정한다.

$$Th = m_H - w \times \sigma_H \dots\dots(2)$$

GOP가 시작될 픽처로부터 N_{\max} 장의 픽처에 대해 장면 전환을 조사하여 장면 전환이 일어나기 직전의 픽처까지를 하나의 GOP로 하여 N 을 결정한다. 장면 전환이 없을 경우는 미리 주어진 기본값 N_{def} 로 결정한다.

2.3 움직임의 크기 측정 및 M 결정

정해진 N 장의 픽처에 대하여 움직임의 크기를 측정하여 M 을 정한다. 움직임의 크기를 측정하는 가장 좋은 방법은 실제로 움직임을 해보고 실제 움직임의 크기 및 다른 정보들을 참고하는 것이다. 그러나 M 은 해당 GOP가 인코딩되기 전에 결정되어야 하므로

이러한 방법은 많은 계산량 및 시간을 필요로 한다. 따라서 움직임의 크기를 측정할 간단한 알고리즘이 필요하다. [2]에서는 매크로블록별 분산으로 이루어진 영역에서 간단한 움직임 예측을 통하여 움직임의 크기를 측정하는 방법을 적용하였다. 본 논문에서는 이를 매크로블록별 컬러 히스토그램을 이용하여 측정하는 방법을 제시하였다. 움직임의 크기는 이웃하는 픽처의 같은 위치의 매크로블록의 히스토그램을 비교하여 공통부분의 비율을 구한 후 이들의 평균 및 표준편차를 이용한다. 움직임이 있는 부분에서는 비율이 작으므로 평균이 작아지고, 비율이 큰 것과 작은 것이 섞여 있으므로 표준편차는 커지게 된다. 이를 통하여 M을 결정한다. 각 픽처별로 이들의 평균과 표준편차를 구한 후 이들을 N장의 픽처에 대해 평균하여 움직임의 크기의 지표로 삼는다. 주어진 픽처가 X개의 매크로블록으로 이루어지고 이전 픽처의 x번째 블록의 히스토그램을 O_x , 현재 픽처의 것을 C_x 라 할 때, 각 블록에 대한 평균 μ_x 및 표준편차 σ_x 는 다음과 같다.

$$I_x(O_x, C_x) = \frac{\sum_{i=1}^N \min(O_x(i), C_x(i))}{16^2 \times 3} \quad \dots(3)$$

$$\mu_i = \frac{\sum_x I_x(O_x, C_x)}{X} \quad \dots\dots\dots(4)$$

$$\sigma_i = \sqrt{\frac{\sum_x I_x(O_x, C_x)^2}{X} - \mu_i^2} \quad \dots\dots\dots(5)$$

이 값들을 N장의 픽처에 대해 계산하므로 각각 N-1개의 값을 얻는다. M을 구하기 위해 이들의 평균을 얻는다. 이를 각각 μ 과 σ 라 할 때 다음과 같이 정한다.

$$\mu = \frac{\sum_i \mu_i}{N-1} \quad \dots(6), \quad \sigma = \frac{\sum_i \sigma_i}{N-1} \quad \dots(7)$$

이로부터 M을 결정한다. μ 는 대체로 움직임의 크기를 잘 반영하나 복잡한 영상일 경우 작은 움직임에도 큰 영향을 받을 수 있다. 이 때 블록들의 공통부분의 비율은 전체적으로 작아지므로 σ 는 비교적 작은 값을 가지게 되어 이를 보완할 수 있다. 따라서 M은 다음과 같이 정한다.

$$M = \max(1, \min(M_{\max} - v(1 - \mu)\sigma)) \quad \dots(8)$$

v는 비례상수이다.

2.4 GOP 할당 및 인코딩

결정된 N, M에 따라 GOP를 할당한다. 만일 장면 전환이 있을 시에는 그림 3과 같이 장면 전환 직후의 GOP는 시작 부분을 닫힌 형태로 하여 다른 장면의 픽처간에 참조가 일어나지 않도록 한다. 그림 3은 장면

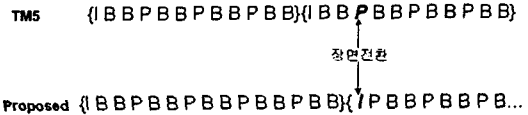


그림 3 최종적인 GOP 할당

Fig. 3 Final GOP allocation
 전환에 따라 N=15로 결정된 경우를 나타내고 있는데 직후의 GOP가 시작부분이 닫힌 형태여서 앞의 GOP와 무관하게 인코딩되는 것을 알 수 있다.

그리고 새로운 GOP에는 결정된 N에 따라 비트를 할당한다. 이는 TM5와 같은 방법으로 다음과 같이 이루어진다.

$$R = N \times \frac{\text{비트율}}{\text{프레임율}} \quad \dots\dots(9)$$

III. 실험 결과

제안된 알고리즘에 대한 실험은 $N_{\max}=20$, $N_{\text{ref}}=15$, $w=2$, $v=60$ 의 조건하에서 수행되었다. 실험 결과 측정된 평균 PSNR과 [2]에서 제시된 방법과 비교한 결과를 표 1, 2에 수록하였다. 표 3은 720 x 480 기준으로 3Mbps로, 표 4는 5Mbps로 하여 실험한 결과이다. 그리고 그림 4에서 PSNR의 변화를 직접 비교하였다.

표 2 3Mbps에서의 실험결과

Table 3 Experimental result at 3Mbps

3Mbps	Proposed	[2]	TM5
Ad1	35.1dB	34.8dB	34.7dB
Ad2	36.2dB	36.1dB	35.8dB
Ad3	34.5dB	34.1dB	33.4dB
Drama1	37.3dB	37.3dB	36.9dB
Drama2	37.4dB	37.5dB	37.2dB

표 3 5Mbps에서의 실험결과

Table 4 Experimental result at 5Mbps

5Mbps	Proposed	[2]	TM5
Ad1	39.1dB	38.8dB	38.6dB
Ad2	39.6dB	39.5dB	39.1dB
Ad3	37.1dB	36.6dB	36.9dB
Drama1	40.8dB	40.8dB	40.4dB
Drama2	40.8dB	41.1dB	40.3dB

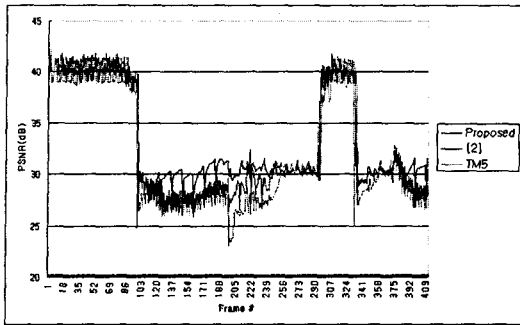


그림 4 3Mbps에서의 PSNR

Fig. 4 PSNR at 3Mbps

표 3, 4 및 그림 4의 결과를 볼 때 제시된 방법이 TM5에 비해 향상된 PSNR을 보여 주었고, [2]와의 비교에서도 우수하거나 비슷한 성능을 보이는 것을 알 수 있다. 따라서 N과 M을 적합하게 변경한 것이 화질을 향상시켰음을 알 수 있다.

IV. 결론

본 논문에서는 매크로블록별 컬러 히스토그램을 이용하여 장면 전환과 움직임의 크기에 맞게 N과 M을 결정하고 그에 따라 GOP를 할당하는 기법을 제안하였다. 제안된 방법은 TM5에 비하여 향상된 PSNR을 나타내고, 기존 방법과도 비슷하거나 나은 화질을 보여주었다.

그러나 제시된 기법의 영향은 상당히 제한적인 것으로, N의 조절은 장면 전환 부근에서만 화질 향상에 기여하고 M을 조절하는 것 또한 화질 향상에 어느 정도의 한계를 지니고 있다. 따라서 이러한 기법에 R-D 최적화를 고려한 가변 비트율(VBR) 등의 방법이 추가될 경우 더욱 좋은 화질을 얻을 수 있을 것이다.

참고문헌

- [1] Yue Yu, Jian Zhou and Yiliang Wang, "A Fast Effective Scene Change Detection and Adaptive Rate Control Algorithm", ICIP, Vol. 2, p.379-382, October, 1998.
- [2] Akiyo Yoneyama, Yasuyuki Nakajima, Hirosama Yanagihara and Masaru Sugano, "MPEG Encoding Algorithm with Scene Adaptive Dynamic GOP Structure", 3rd workshop on MMSP, p.297-302, 1999.
- [3] Akiyo Yoneyama, Hirosama Yanagihara, Yasuyuki Nakajima, "One-pass VBR MPEG Encoder using Scene Adaptive Dynamic GOP Structure",

ICCE, p.174-175, June, 2001.

[4] Michael J. Swain and Dana H. Ballard, "Indexing Via Color Histograms", 3rd ICCV, p.390-393, December, 1990.

[5] ISO/IEC/JTC1/SC29/WG11, "Test Model 5", April, 1993.