

# 레지듀얼 정보를 이용한 다시점 동영상 부호화의 가중치 예측

김지영 김용태 서정동 손광훈

연세대학교 전기전자공학과

tomtom@diml.yonsei.ac.kr

## Weighted Prediction Using Residual Information for Multi-view Video Coding

Kim, Jiyoung Kim, Youngtae Seo, Jungdong Sohn, Kwanghoon

School of Electrical & Electronic Engineering, Yonsei University

### 요약

다시점 동영상 부호화는 서로 다른 카메라에 의해 영상을 획득하므로 카메라 내부 파라미터의 차이나 조명의 차이 및 변화 등에 의한 시점 간 명도 성분의 불균형을 가지고 있다. 이로 인해 잘못된 변이 추정이 이루어질 수 있으며, 따라서 전체적인 다시점 동영상 부호화의 성능을 크게 저하시킬 수 있다. 본 논문에서는 레지듀얼이 가지고 있는 밝기 차 정보를 이용하여 시점 간의 불균형을 해소하는 가중치 예측 알고리즘을 제안한다. 주변의 인과적인 블록의 레지듀얼 정보를 이용하여 현재 블록과 참조 블록의 밝기 차를 예측하고, 이 값을 이용해 시점 간 불균형을 보정 한 후 변이 추정을 수행한다. 변이 보상 후 계산된 현재 블록의 레지듀얼 평균값을 앞에서 예측된 밝기 차의 값에 누적하여 다음 블록의 밝기 차 예측에 사용한다. 제안된 방법을 실험 영상에 적용한 결과 평균적으로 약 0.2dB의 이득을 얻었다.

### 1. 서론

보다 현실적이고 자연스러운 영상의 구현을 위해 3차원 영상 시스템에 대한 연구가 활발히 진행 중이다. 넓은 시야각과 다양한 시점을 허용하는 다시점 동영상은 일반 2D 영상에 비해 수 배의 데이터 양을 요구하기 때문에 이를 처리하기 위한 효과적인 압축 알고리즘을 다시점 동영상 부호화기(MVC)라 한다.

다시점 동영상 부호화기에 대한 관심 증대로 인해 MPEG(Moving Picture Expert Group) 3DAV가 ISO/IEC SC29WG11에 설립 되었고, 현재 ISO/IEC 와 ITU-T 공동으로 JVT(Joint Video Team)의 SVC(Scalable Video Coding) 기반의 참조 모델을 통하여 다시점 동영상 부호화에 대한 표준화가 진행 중이다. 다시점 동영상의 표준화 작업을 통해 교육, 훈련, 3차원 영화 산업, 의료 산업, 화상 회의 및 3차원 방송 등 여러 분야에서 더 자연스럽게 현실감 있는 영상을 전달할 수 있을 것으로 기대된다.[1]

다시점 동영상 부호화기에서는 시간상의 중복성을 제거하기 위한 움직임 추정 및 보상 뿐 아니라, 공간상의 중복성을 제거하기 위해 시점간 변이 추정 및 보상을 사용한다. 그러나 서로 다른 카메라에 의해 획득된 영상들은 카메라 내부 파라미터의 차이나 조명의 차이 및 변화 등에 의한 시점 간 명도 성분의 불균형을 가지고 있으며, 이로 인해 잘못된 변이 추정이 이루어질 수 있다.[2] 이는 전체적인 다시점 동영상 부호화의 성능을 크게 저하시킨다. 이러한 문제는 변이 추정 이전에 먼저 부호화 하고자 하는 블록과 참조 블록의 불균형을 보정하여 정확한 대응점을 찾음으로써 해결할 수 있다.

본 논문의 전체적인 구성은 다음과 같다. 제2장에서는 레지듀얼 영상과 명도 불균형의 관계에 대해서 자세히 설명하고, 3장에서는 제안된 알고리즘과 그 구현 방법에 대해 설명하였다. 제4장에서는 제안된 알고리즘을 참조 소프트웨어의 결과와 비교하여 제안 방법의 우수성을 검증하였다. 마지막으로 제5장에서 결론을 맺는다.

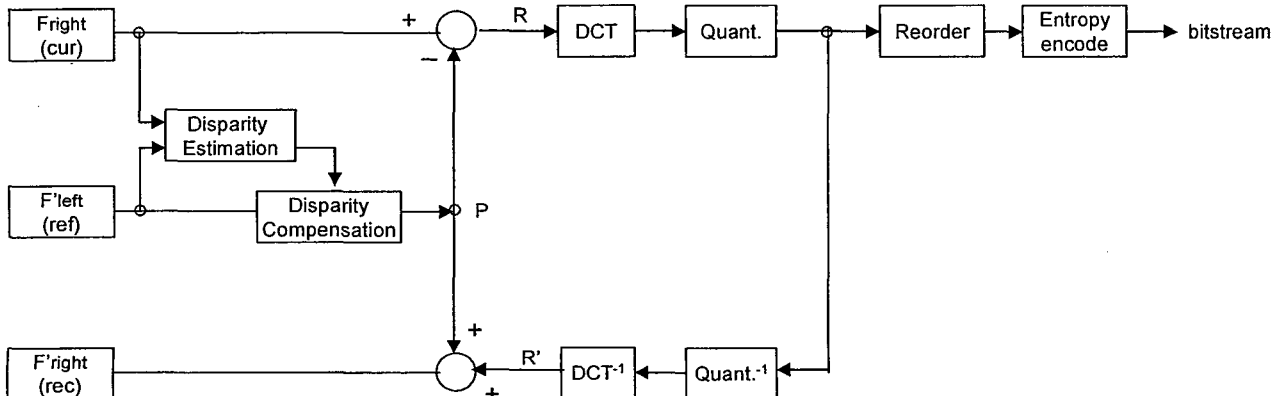


그림 1. 스테레오 동영상 부호화기

## 2. 레지듀얼 정보와 명도 불균형

그림 1은 일반적인 스테레오 동영상 부호화기의 간략한 블록도이다. 부호화하고자 하는 영상  $F_{right}$ 와 참조 영상  $F_{left}$ 는 같은 시간에 획득된 다른 시점의 영상이고, 이들 사이에 변이 추정과 변이 보상이 수행된다. 현재 부호화하는 블록이 가려진 영역이 아니고 변이 추정이 정확히 수행되었다면, 변이 보상 후 얻은 예측 영상(P)은 현재 블록과 유사한 값을 가지고 있어야 한다. 이 예측 영상과 현재 블록과의 차이인 레지듀얼(R) 값은 DCT와 양자화, 엔트로피 부호화를 거쳐 비트열에 쓰여진다. 그리고 양자화 된 레지듀얼을 다시 역양자화, 역DCT 거쳐 재구성(reconstruction) 하고, 이를 다음 영상의 참조 영상으로 사용한다.

현재 부호화하고자 하는 블록의 밝기 값이 1차원으로 나타내었을 때 그림 2.(a)와 같고, 참조 영상에서 대응되는 블록의 신호는 명도 성분의 불균형으로 인해 그림 2.(b)와 같다고 하자. 두 신호는 평균값의 차이가 있을 뿐 유사한 정보를 가지고 있으므로 변이 추정을 통해 두 블록이 대응되어야 한다. 그러나 변이 추정에 일반적으로 사용되는 비용 함수인 SAD(sum of absolute difference)는 다음과 같다.

$$SAD(u, v, d) = \sum_{(i,j)} |I_{cur}(u+j, v+i) - I_{ref}(u+d_u+j, v+d_v+i)| \quad (1)$$

따라서 블록의 크기가  $16 \times 16$ 이고 두 블록의 평균 밝기의 차를

$$M_{diff} = M_{cur} - M_{ref} \quad (2)$$

라 할 때 두 블록 사이의 에러는 약  $16 \times 16 \times M_{diff}$ 로 계산되므로, 이 참조 블록이 대응되는 블록으로 선택되지 않을 수 있다.

만약 그림 2.(b)의 참조 블록이 변이 추정을 통해 선택되었다면 이는 필터를 거쳐 변이 보정 후의 영상(그림 1에서 P) 신호가 된다. 따라서 그림 2.(a)와 그림 2.(b) 신호의 차이가 레지듀얼 영상과 유사하게 되며, 이는 그림 2.(c)와 같다. 즉 레지듀얼 영상의 평균 밝기  $M_{res}$ 이 부호화하고자 하는 블록과 참조 블록의 평균 밝기 차와 근사한 값을 갖게 된다.

$$M_{res} \approx M_{cur} - M_{ref} \quad (3)$$

따라서 레지듀얼 영상의 평균 밝기 정보를 이용해 두 영상의 명도 불균형 정도를 알아낼 수 있으며, 이를 오프셋으로 적용하여 변이 추정에 사용할 수 있다.

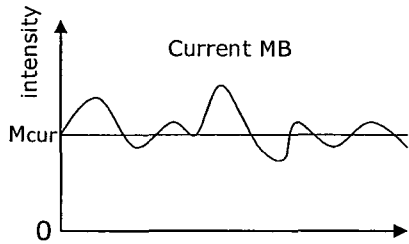
## 3. 가중치 예측을 이용한 변이 추정

앞 절에서 설명한 대로 현재 부호화하고자 하는 블록의 불균형 정보는 현재 블록의 레지듀얼 신호에 반영되어 있다. 하지만 레지듀얼 신호는 변이 추정과 보정 후에 계산되므로, 현재 블록의 변이 추정이 수행될 때 현재 블록의 레지듀얼 신호를 사용할 수 없다. 또한 영상에서 한 블록의 명도 불균형 정도는 주변 블록의 불균형 정도와 밀접한 관계가 있다. 따라서 현재 블록의 오프셋 정보는, 미리 부호화 된 블록의 레지듀얼 정보를 사용하여 예측해야 한다.

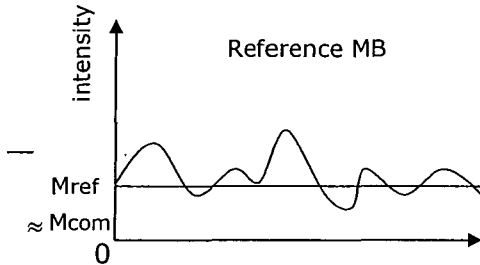
가장 먼저 부호화 되는 블록  $MB_0$ 는 예측에 사용할 블록이 없으며, 이 블록의 변이 추정이 비교적 정확하게 수행되어야 그 이후 부호화 될 블록들에서 명도 불균형 예측이 보다 정확하게 이루어진다. 따라서  $MB_0$ 는 명도 불균형에 강한 변이 추정이 필요하며, 이를 위해 다음과 같은 비용 함수를 사용한다.[3]

$$ZSAD(u, v, d) = \sum_{(i,j)} \left| \frac{I_{cur}(u+j, v+i) - M_{cur}}{I_{ref}(u+d_u+j, v+d_v+i) - M_{ref}} \right| \quad (4)$$

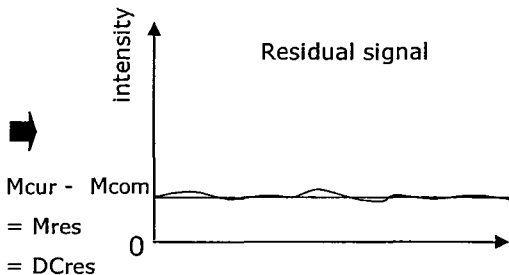
즉 가장 먼저 부호화되는 블록에서만 정확한 명도 불균형을 추정하여 변이 추정에 적용하며, 이 값은 복호화기로 전송되어야 한다. 그 이후 부호화되는 블록들은 이를 기반으로 예측된 오프셋이 적용되며, 레지듀얼 정보가 복호화기로 전송되기 때문에 불균형 오프셋은 전송하지 않아도 된다.



(a)



(b)



(c)

그림 2. (a) 부호화하고자 하는 블록의 신호 (b) 대응되는 참조 블록의 신호 (c) 레지듀얼 신호

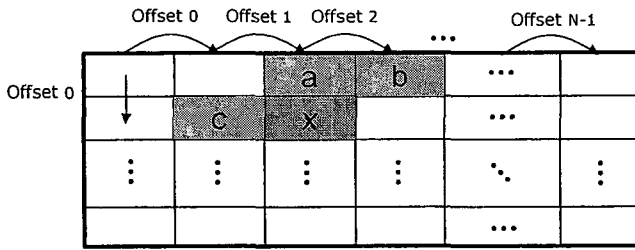


그림 3. 명도 불균형 예측 방법

$MB_0$ 를 변이 추정할 때 계산된  $offset_0$ 을 이용해 그 이후 블록에서 오프셋을 예측하는 방법이 그림 3에 나타나 있다. 블록의 부호화는 가로 방향(row)으로 진행되며 따라서  $MB_0$ 를 부호화한 후 그 블록의 오른쪽 블록  $MB_1$ 의 부호화가 진행된다. 이 때  $offset_0$  값을 미리  $MB_1$  각 픽셀의 밝기 값에서 빼주어 가중치 예측을 한 후 변이 추정을 수행하면 참조 블록과의 불균형이 해소된다. 그러나  $MB_0$ 와  $MB_1$ 의 명도 불균형 정도가 다르다면 레지듀얼 영상의 평균 밝기  $M_{res}$ 는 다시 0이 아닌 값을 가질 것이고, 이 값은 다음 블록에 누적되어 적용된다. 이 때 레지듀얼 영상의 평균 밝기는 레지듀얼이 DCT된 후의 DC 값과 같으므로, 추가적인 연산 없이 DCT 영상의 (0,0) 위치의 값을 사용하면 된다. 즉  $MB_n$ 의 DCT된 레지듀얼 영상의 DC 값  $DC_{res}^n$ 이 0이 아니라면 이 값은  $offset_{n-1}$ 에 누적되어  $offset_n$ 을 만들며, 이 값은 다음 블록  $MB_{n+1}$ 를 부호화할 때 적용된다.

$$offset_n = offset_{n-1} + DC_{res}^n \quad (5)$$

이와 같은 방법을 인과적인 주변 블록으로 확장시키면, 그림 3에서 블록  $MB_x$ 를 부호화 할 때  $MB_a$ ,  $MB_b$ , 그리고  $MB_c$ 의 오프셋  $offset_a$ ,  $offset_b$ , 그리고  $offset_c$  세 값의 메디안 값을 식 (6)과 같이 현재 블록  $MB_x$ 의 오프셋으로 예측하게 된다.

$$offset_x = median(offset_a, offset_b, offset_c) \quad (6)$$

현재 부호화 하고자 하는 프레임이 B-프레임이라면 두 개 이상의 다른 시점에서 참조 영상을 사용할 수도 있다. 이 때 각각의 시점에서 명도의 불균형 정도는 다르므로, 시점에 따라 오프셋을 따로 누적해야 한다. 한 블록에서 참조되지 않는 시점에 대해서는 예측된 오프셋을 그대로 저장하여 다음 블록의 오프셋 예측에 사용된다.

또한 부호화 과정에서 오프셋 값을 보상하여 영상을 재구성해야 하는데 부호화기에서 받은 레지듀얼 영상의 DC 값은 양자화 된 값이므로, 부호화기에서도 양자화 이후의 DC 값을 누적해야 추가적인 비트를 사용하지 않고도 정확히 부호화 할 수 있다.

그러나 이와 같은 예측 방법은 주변 블록과 명도 불균형의 정도가 비슷하다고 가정한 경우이며, 이를 가려진 영역이나 물체의 가장자리 부분에 그대로 적용할 경우 오히려 잘못된 변이 추정을 수행할 수 있다. 이를 방지하기 위해 각 블록마다 오프셋을 적용한 방법과 적용하지

않은 방법 중 더 좋은 결과를 내는 방법으로 부호화하며, 이러한 정보는 플래그를 통해 부호화기에 전달된다.

#### 4. 실험 결과

레지듀얼 정보를 이용한 명도 불균형 개선 정도를 확인하기 위해 다시점 동영상 부호화의 참조 소프트웨어인 JSVM(Joint Scalable Video Model) 6.5에 제안한 알고리즘을 구현하였다. 실험에 사용된 영상의 특성과 사용한 프레임 수는 표 1과 같다. 실험은 하나의 영상에 서로 다른 3개의 비트율에 대해 수행하였다. 비트율은 양자화 파라미터(QP)값으로 조절하고 각 영상에 사용한 양자화 파라미터는 표 2와 같다.[4]

제안된 알고리즘의 성능 검증에 위해 명도의 불균형을 고려하지 않은 참조 소프트웨어의 결과와 비교하였다. 실험 결과는 각 비트율에 대한 객관적 화질(PSNR)을 그래프로 도시하였으며, 그림 4-6와 같다.

표 1. 실험 영상의 특성과 사용된 프레임 수

실험 영상	영상 특성			사용된 프레임 수
	영상 크기	프레임율	시점 수	
Race1	640 x 480	30fps	8	248 frames = 8*2* GOP_15+8
Ballroom	640 x 480	25fps	8	200 frames = 8*2* GOP_12+8
Rena	640 x 480	30fps	16	496 frames = 16*2* GOP_15+16
Akko&Kayo	640 x 480	30fps	15	465 frames = 15*2* GOP_15+15

표 2. 실험 영상의 양자화 파라미터

실험 영상	각 비트율에 대한 기본 양자화 파라미터		
	저비트율	중간 비트율	고비트율
Race1	29	27	25
Ballroom	34	31	29
Rena	33	28	23
Akko&Kayo	37	31	25

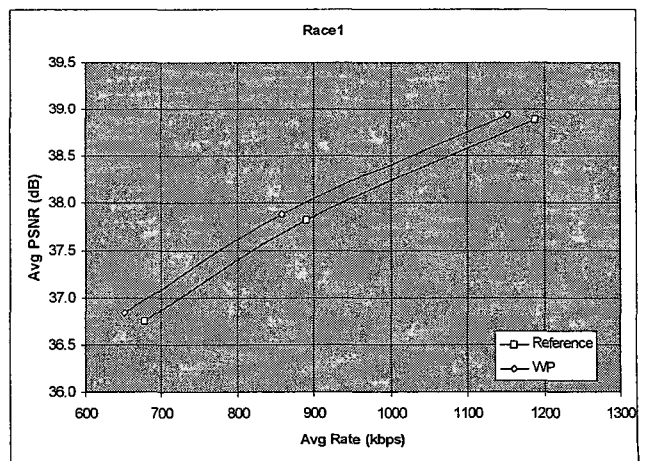


그림 4. "Race1" 영상의 실험 결과

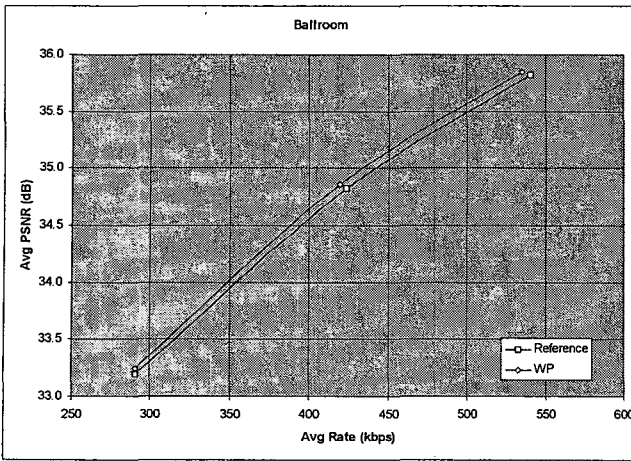


그림 5. "Ballroom" 영상의 실험 결과

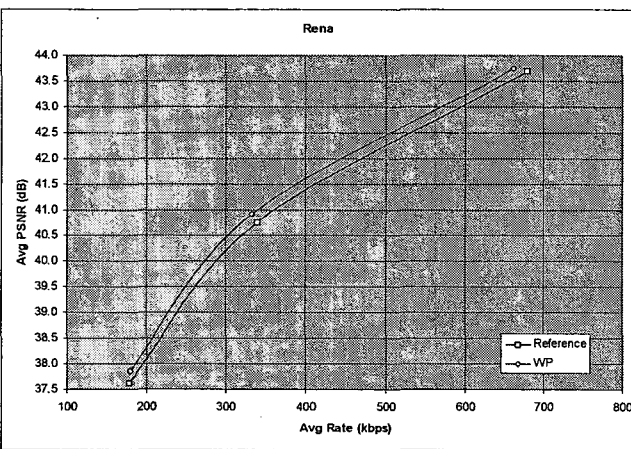


그림 6. "Rena" 영상의 실험 결과

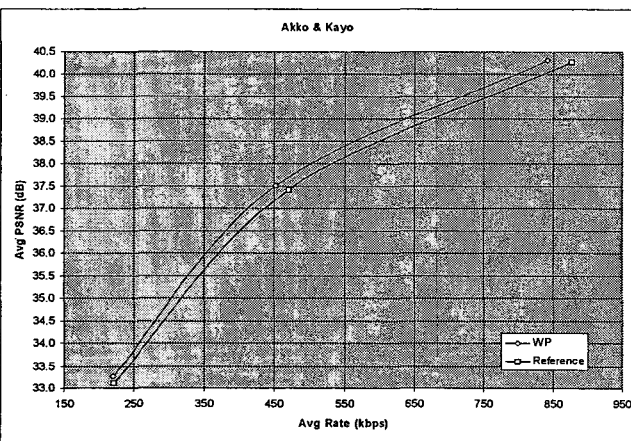


그림 7. "Akko & Kayo" 영상의 실험 결과

레이저듀얼을 이용해 시점 간의 명도 불균형을 보상하는 가중치 예측 알고리즘을 실험 영상에 적용한 결과 참조 소프트웨어 보다 최소 약 0.1dB에서 최대 약 0.3dB까지의 이득을 얻었다. 각 비트율의 양자화 파라미터가 다른 영상에 비해 작은 "Race1" 영상에서는 PSNR이 평균 0.1dB 상승함과 동시에 약 4%의 비트를 적게 사용하였다. "Ballroom" 영상에서는 평균 0.1dB의 PSNR 이득과 약 1%의 비트 감

소를 보였다. 고 비트율에서 저 비트율로 갈수록 성능 향상의 정도가 줄어들었다. 카메라 간 간격이 다른 영상에 비해 좁은 "Rena" 영상에 시점 간 명도 불균형을 보상해 준 결과 PSNR이 평균 0.2dB 상승하였고 비트율은 약 2% 감소하였다. 저 비트율의 양자화 파라미터가 비교적 큰 "Rena" 영상에서도 저 비트율로 갈수록 성능 향상 정도가 줄어들었다. 2D 카메라 배열을 가진 "Akko & Kayo" 영상은 고 비트율에서는 약 0.1dB PSNR 이득과 약 4%의 비트율 감소를 보였으나 저 비트율에서는 PSNR 이득이 약 0.2dB로 늘어난 반면 비트율 감소는 거의 없었다.

제안된 알고리즘은 다른 영상들에 비해 비교적 양자화 파라미터가 작아 전체적으로 비트율이 높은 "Race1" 영상에서 가장 좋은 성능을 보였으며 각 비트율에 대해서도 일정한 정도로 성능이 향상되었다. 그 외 영상에서는 저 비트율에서의 성능 향상의 정도가 고 비트율에 비해 조금 떨어지는 것을 볼 수 있다. 이는 변이 추정기 선택된 블록마다 한 비트의 플래그를 보내는 것이 고 비트율 보다 저 비트율에서 큰 영향을 끼치기 때문으로 보인다.

## 5. 결론

본 논문에서는 다시점 동영상 부호화에서 각 시점 간의 명도 불균형으로 인한 문제를 레이저듀얼 정보를 이용하여 보상하는 가중치 예측 알고리즘을 제안하였다. 제안한 알고리즘을 참조 소프트웨어에 구현하여 성능을 비교해 본 결과 평균적으로 약 0.2dB의 이득을 얻었다.

현 알고리즘은 명도 불균형을 측정하는 척도로 휘도 성분만을 사용하였으나 향후에는 색차 성분까지 고려해야 할 것이며, 가려진 영역이나 물체 가장자리 영역을 다루기 위한 좀 더 정교한 알고리즘이 연구되어야 할 것이다.

## 참고 문헌

- [1] A. Puri, R. V. Kollaritis and B. G. Haskell, "Basic of stereoscopic video, new compression results with MPEG-2 and a proposed for MPEG-4," *Signal Processing: Image Communication*, vol. 10, pp201-234, 1997.
- [2] A. Mancini, "Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video," *Master Thesis, McGill University, Feb. 1998*.
- [3] A. Giachetti, "Matching techniques to compute image motion," *Image Vision Comput.*, Vol. 18, pp247-260, 2000.
- [4] ISO/IEC JTC1/SC29/WG11, "Results of Core Experiment 1-D on Multiview Video Coding", M13228, Montreux, Swiss, Apr. 2006.