

병합 방법에 의한 가변 블록 움직임 예측

이규호*, 손남례, 이귀상

*전남대학교 전산학과

e-mail:herolkh@cad.chonnam.ac.kr

Variable Block-size Motion Estimation based on Merging Procedure

Kyu-Ho Lee*, Nam-Rye Son, Guee-Sang Lee

*Dept of Computer Science, Chonnam University

요 약

본 논문에서는 가장 최근의 동영상 표준인 H.264에서 가변 블록 움직임 예측 시 인접한 블록과의 상관성을 분석하여 병합 절차를 추가함으로써 매크로블록의 최종 모드를 결정하는 시간을 줄이기 위한 알고리즘을 제안한다. H.264에서는 매크로블록의 모드를 결정하기 위하여 총 7가지 모드를 사용하여 움직임 예측을 실시함으로써 부호화 효율을 극대화시킨 반면 이러한 움직임 예측이 부호화기의 복잡도를 높이는 주요 요인으로 현재 커다란 단점으로 지적되고 있다. 본 논문에서는 8x8 움직임 예측이 끝난 후 인접한 두 블록 사이의 거리를 임계값(Threshold)과 비교하여 다음 모드의 움직임 예측의 실시 여부를 먼저 결정함으로써 불필요한 움직임 예측에 소비되는 시간을 단축시켰다. 여기서 실험 조건으로 명시하고 있는 것은 대표적인 단일모드 중에서 수행 성능이 가장 좋은 8x8 모드를 기본모드로 사용하고 병합 시 16x16 모드 쪽으로 상향식(bottom-up) 방법의 병합을 수행 해 나아간다. 모의실험을 통해 수행 성과 전체 부호화 시간 측면을 본 논문에서 제안한 방법과 4가지 모드인 16x16, 16x8, 8x16, 8x8 모드를 모두 사용한 경우, 8x8 단일모드를 사용한 경우를 비교하였다. 실험 결과 8x8 단일모드보다 수행 성능이 향상되었으며, 시간 단축 면에서 제안한 방법이 4가지 모드인 16x16, 16x8, 8x16, 8x8 모드를 모두 사용한 경우와 8x8 단일모드를 사용한 경우보다 계산 시간이 감소하였음을 확인하였다.

1. 서론

무선 통신의 발달로 인하여 동영상 서비스에 대한 요구가 높아지고 있는 가운데 최근에 표준화가 완성된 새로운 동영상 압축 방식인 H.264는 기존의 동영상 압축 방식인 H.261/H.263/MPEG-1/MPEG-4에 비하여 새로운 요소들과 방식들을 채택하면서 중요한 기술로 부각되고 있다. 그중 움직임 예측에 있어서 16x16, 8x8인 고정된 블록 단위로 처리를 했던 이전 압축 방식과 다르게 H.264에서는 매크로블록을 총 7가지 모드를 사용하여 움직임 추정하는 가변 블록 단위의 움직임 예측 방법이 채택되었다[1]. 이는 블록의 크기나 모양이 움직임 예측에 있어서 중요한 요소로 작용되고 있음을 말한다. 가변 블록 단위의 움직임 예측 방법이 높은 수행 성능을 보임은 여러 실험과 문서를 통하여 입증되었다[2]. 하지만 매크로블록의 최적 모드를 찾기 위하여 7가지 움직임 예측 모드를 사용하여 가변 블록 단위로 움직임 예측을

하는 과정에서 많은 계산 시간을 필요로 한다는 것이 커다란 단점으로 지적되고 있다.

그림 1에서 볼 수 있듯이 여러 가지 모드를 사용하여 움직임 예측이 이루어지는데 이러한 움직임 예측 과정을 효율적으로 줄이기 위한 방법을 여기서 제안하고자 한다. H.264 부호화기에서는 매크로블록의 최적의 모드를 결정하기 위하여 하나의 매크로블록을 먼저 4개의 8x8 블록으로 나누어서 세밀한 부분(4x4)부터 움직임 예측을 시작하여 가장 큰 부분인 매크로블록(16x16)까지 움직임 예측이 이루어진다. 대표적인 단일모드들의 수행 성능을 알아보기 위하여 7가지 모드를 모두 사용한 경우와 단일모드인 16x16, 8x8, 4x4 모드의 R-D(Rate-Distortion) 성능을 실험을 통해 비교해 본 결과 단일모드 중에서는 8x8 모드가 대부분의 영상에서 결과가 좋음을 알 수 있었다[그림 2].

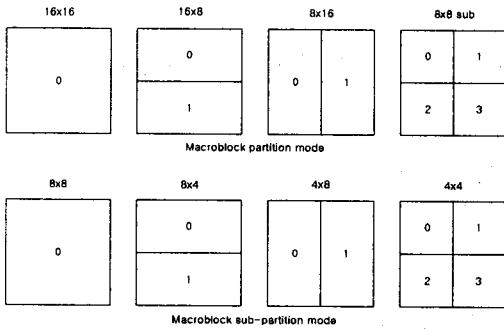


그림 1 H.264 움직임 예측모드

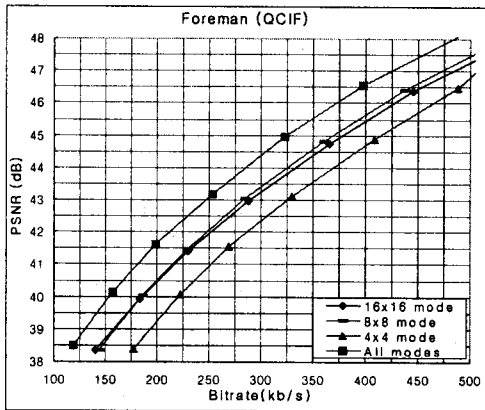


그림 2 Foreman의 비트율-왜곡 곡선

따라서 본 논문에서는 8x8 움직임 예측모드를 기본 모드로 사용하며, 제안 알고리즘은 8x8 움직임 예측 후, 인접한 블록과의 거리를 임계값과 비교하고, 병합 시 상향식 방법의 병합을 전개한다. 즉, 두 개의 블록이 병합될 때에는 8x8 블록에서 움직임이 거의 없는 16x16 블록 방향으로 병합한다. 본 논문은 16x16, 16x8, 8x16, 8x8의 4가지 모드를 사용했을 경우, 8x8 단일모드를 사용했을 경우 그리고 제안한 알고리즘을 비교하면서 결론을 맺는다.

2. 제안 알고리즘

2.1 임계값을 결정하는 방법

H.264에서는 매크로블록의 최적 모드를 찾기 위하여 총 7가지 예측모드를 사용하여 각각의 모드마다 움직임 예측을 수행한다.

H.264에서 지원하는 단일모드로는 가장 작은 블록인 4x4 모드를 포함하여 기존의 대표적인 단일모

드인 16x16, 8x8 모드가 있다. 3가지 단일모드의 수행 성능을 실험 해 본 결과 8x8이 대부분의 영상에서 우월함을 알 수 있었다. 따라서 병합을 위한 기본 예측모드로는 8x8 예측모드를 사용한다.

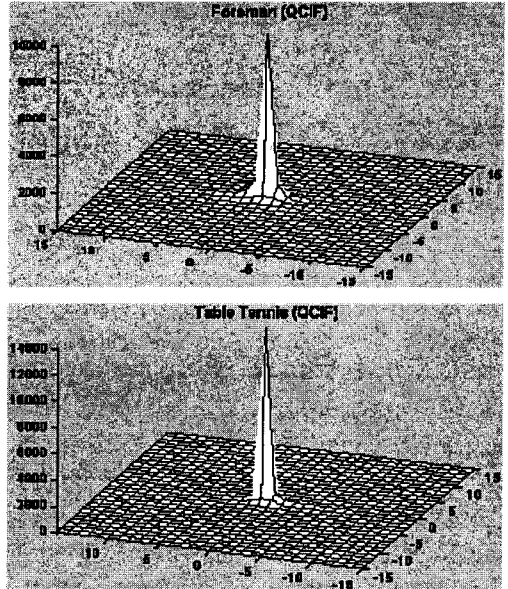


그림 3 움직임 벡터의 분포도

그리고 블록의 병합은 8x8 블록부터 이루어진다. 8x8 모드의 움직임 예측이 끝나면 각각의 블록마다 움직임 벡터값을 알 수 있다. 블록이 얼마나 인접해 있는지를 판단하기 위하여 4개의 8x8 블록에 대해 움직임 벡터를 이용하여 블록간의 거리(distance)를 구한다. (식1)은 움직임 벡터가 (a, c), (b, d)일 때 거리를 구하는 식이다.

$$distance = \sqrt{(a-b)^2 + (c-d)^2} \quad \text{---(식1)}$$

임계값은 블록간의 거리와 많은 관계가 있기 때문에 두 블록 사이의 거리를 구하기 위하여 사용되는 움직임 벡터의 분포를 실제 실험을 통하여 알아 보았다. 그림 3은 실험한 영상 중 두 가지 영상에 대한 움직임 벡터의 분포도를 나타내고 있다. 그림 3에서 확인해 볼 수 있듯이 움직임 벡터의 90% 이상이 ± 2 범위 안에 있다는 것을 알 수 있다. 즉, 움직임 벡터의 대부분이 이웃한 블록의 움직임 벡터와 거의 비슷한 위치에 있음을 알 수 있다. 여기서 병합을 위한 임계값은 ± 2 를 기준으로 실제 실험을 통

해 얻은 수치와 두 움직임 벡터 사이의 거리에 관한 기존의 연구를 토대로 하여 $\sqrt{2}$ 를 병합을 위한 임계값으로 결정하였다[3].

2.2 제안된 병합 방법

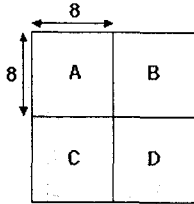


그림 4 매크로블록

매크로블록의 최적의 모드를 결정하는 과정에서 8x8 블록 단위의 움직임 예측이 끝나면 그림 3의 A, B, C, D 각각 4개의 블록에 대한 움직임 벡터가 구해진다. 각각의 움직임 벡터를 가지고 A와 B블록의 거리, C와 D블록의 거리, A와 C블록의 거리, B와 D블록의 거리를 구한다. 여기서 구한 거리를 인접한 블록과의 병합을 위한 정보로 사용하게 된다. A와 B블록의 거리가 임계값 안에 존재하고 C와 D블록의 거리 역시 임계값 안에 존재할 경우 A와 B블록 그리고 C와 D블록은 서로 병합이 되고 16x8 움직임 예측모드를 실행하게 된다. 마찬가지로 A와 C블록 사이의 거리가 임계값 안에 존재하고 B와 D블록 사이의 거리가 임계값 안에 존재할 경우 각각의 블록은 서로 병합되고 8x16 모드를 사용하여 움직임 예측을 하게 된다. 만약에 전자와 후자가 모두 참일 경우에는 다시 두 개의 블록이 병합되어서 마지막 움직임 예측 모드인 16x16 모드를 실행하여 움직임 예측을 하게 된다. 이것을 간단하게 순서도로 표현하면 그림 5와 같다. 순서도에서 A_B, C_D, A_C, B_D는 블록들간의 거리를 나타내고 Th는 임계값을 나타낸다. (1)의 결과가 참일 경우 16x8 모드를 실행하게 되고 (2)의 결과가 참일 경우에는 8x16 모드를 실행하게 된다. 마지막으로 (1)과 (2)의 경우를 모두 만족한다면 16x16로 병합되고 16x16 모드를 사용하여 움직임 예측이 이루어진다. 두 블록 사이의 거리가 임계값을 벗어난 경우에는 두 개의 블록의 움직임이 서로 다르다고 판단한다. 이런 경우에는 나머지 모드를 사용하지 않기 때문에 움직임 예측 시간을 단축할 수 있는 이점이 생긴다.

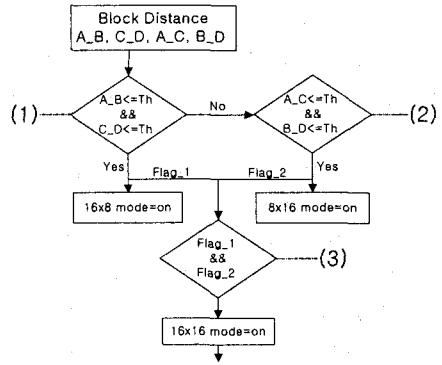


그림 5 병합 알고리즘의 순서도

3. 성능 평가 및 결과

제안된 알고리즘을 검증하기 위하여 본 논문에서는 JM6.1a 버전의 부호화기를 사용하여 실험하였다. 실험 영상은 Foreman(QCIF), Table Tennis(QCIF), Bus(CIF) 영상 100장으로 프레임율은 10Hz이다. 실험 조건으로는 7가지 움직임 예측모드 중 기본 예측모드로는 수행 성능이 평균적으로 가장 좋은 8x8 모드를 사용하고 상향식 방법으로 블록의 병합이 이루어진다. B 프레임은 사용하지 않았고, 실험한 영상에 대하여 IPPP의 순서로 부호화하였으며, 엔트로피 코딩 방법으로는 CAVLC를 사용하였다.

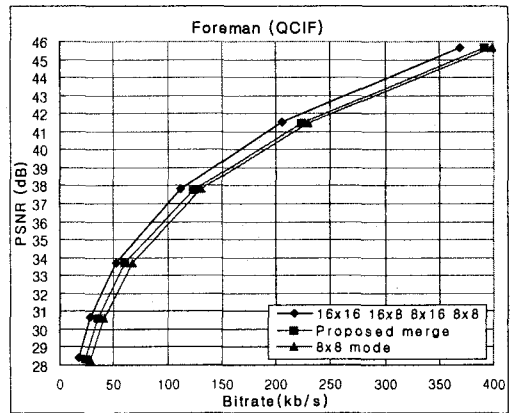


그림 6 Foreman의 비트율-왜곡 곡선

제안한 알고리즘의 비교는 4가지 모드인 16x16, 16x8, 8x16, 8x8 모드를 모두 사용한 경우와 8x8 단일모드를 사용했을 경우를 실험을 통하여 서로 비교하였다. 부호화 성능의 비교를 위하여 객관적인 지표로는 PSNR을 사용하였고, 3가지 모드를 비교하기

위하여 각각 동일한 양자화 값에 따라서 비트율이 일정할 때 전체 부호화 시간을 비교하였다.

그림 6은 제안 알고리즘과 4가지 모드(16x16, 16x8, 8x16, 8x8)를 모두 사용한 경우 그리고 8x8 단일 모드를 사용한 경우에 대하여 모의실험 결과를 그래프로 표현한 것이다. 그래프에서 알 수 있듯이 블록 병합을 실시한 제안 알고리즘이 8x8 단일모드를 사용했을 경우보다 수행 성능이 향상되었음을 확인할 수 있다. 나머지 다른 실험 영상의 경우에서도 Foreman 영상의 결과 그래프와 비슷한 결과를 보였다.

QP	4가지 모드	8x8 모드	제안 알고리즘
0	235.609	216.645	209.657
10	209.894	195.763	187.498
20	171.687	161.061	153.234
31	136.331	127.050	122.763
40	119.546	110.282	109.486

표 1 Foreman 영상의 부호화 시간 비교

QP	4가지 모드	8x8 모드	제안 알고리즘
20	168.782	159.001	153.110
24	153.907	143.642	139.342
28	142.658	132.731	130.061
32	133.565	124.081	122.391
36	126.344	117.154	116.422

표 2 Table Tennis 영상의 부호화 시간 비교

QP	4가지 모드	8x8 모드	제안 알고리즘
20	384.443	360.623	344.450
24	351.200	329.173	313.826
28	321.282	300.217	287.168
32	292.591	272.451	262.982
36	268.766	249.470	243.000

표 3 Bus 영상의 부호화 시간 비교

영상의 부호화에 따른 계산량을 비교를 위하여 전체 부호화 시간을 4가지 모드(16x16, 16x8, 8x16, 8x8)를 모두 사용한 경우와 8x8 단일 모드를 사용한 경우 그리고 본 논문에서 제안한 병합 알고리즘을 적용하여 실험한 결과를 표로 요약하였다. 실험 영상 중 Foreman 영상은 양자화 값의 변화를 크게 Table Tennis, Bus 영상의 경우 양자화 값의 변화를 적게 하였다. 위의 표는 3가지 실험 영상을 가지고 여러 가지 양자화 값에 따른 전체 부호화 시간

비교한 결과이다. 물론 부호화 시간의 비교는 동일한 양자화 값에 대하여 비교하였다. 실험 결과에서 알 수 있듯이 제안한 병합 알고리즘이 4가지 모드(16x16, 16x8, 8x16, 8x8)를 모두 사용한 경우와 8x8 단일모드를 사용한 경우보다 전체 부호화 시간이 단축되었음을 확인할 수 있다.

4. 결론

H.264의 가변 블록 움직임 예측 기법은 매크로블록을 7가지 모드로 정의하여 각각의 모드마다 움직임 예측을 실시함으로써 부호화기 전체의 복잡도를 높이는 요인으로 작용하고 있다. 특히 실시간 통신에 있어서 부호화 시간이 많이 걸린다는 것은 커다란 단점으로 지적될 수 있다. 따라서 본 논문에서는 이러한 움직임 예측으로 인하여 발생하는 부호화기의 복잡도를 줄일 수 있는 방법으로 움직임 예측 시 인접한 블록간의 상관성을 조사하여 블록을 병합함으로써 움직임 예측 시간을 단축할 수 있는 방법을 제시했다. 모의실험 결과 제안한 알고리즘이 4가지 모드(16x16, 16x8, 8x16, 8x8)를 모두 사용한 경우보다 전체 부호화 시간에 있어서 평균 10%~11%의 시간이 단축되었음을 확인할 수 있었다.

추후에 더 많은 실험을 통하여 블록간의 상관성을 결정하는 임계값에 대한 세밀한 분석이 따른다면 PSNR 측면에서도 더 향상된 결과를 얻을 수 있을 것이다.

참고문헌

- [1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Study of Final Committee Draft of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", JVT-100.Doc, 6th Meeting: Awaji, Island, JP, 5-13 December, 2002.
- [2] Thomas Wiegand, Heiko Schwarz, Anthony Joch, Faouzi Kossentini, and Gary J. Sullivan "Rate-Constrained Coder Control and Comparison of Video Coding Standards", IEEE Trans. Circuits Syst. Video Technol., vol. 13, NO. 7, July 2003.
- [3] Ming-gang Liu and Chao-huan Hou, "A Fast Block-Matching Motion Estimation Algorithm Based on Spatial-Temporal Motion Vector Correlation." Proc. 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, May 2001.