

## 계층적 양자화를 이용한 고속 움직임 탐색 기법

박상욱, 이대현, \*심재영, 이상욱  
 서울대학교, \*울산과학기술대학교

supark05@snu.ac.kr, dlcjswol@snu.ac.kr, jysim@unist.ac.kr, sanguk@snu.ac.kr

### Fast motion estimation algorithm using hierarchically quantized image

Sang-Uk Park, Dae-Hyun Lee, \*Jae-Young Sim, and Sang-Uk Lee

Seoul National University, \*Ulsan National Institute of Science and Technology

### 요 약

본 논문에서는 H.264 부호기의 복잡도를 낮추기 위한 고속 움직임 탐색 기법을 제안한다. H.264 부호기는 움직임 탐색 알고리즘 및 다양한 모드 결정으로 인해 매우 높은 복잡도를 갖기 때문에 고해상도 영상의 압축에 시간적 부담을 준다. 따라서, 우리는 비디오 프레임에 계층적으로 양자화를 적용한 뒤, 각 단계마다 움직임 탐색 영역을 적절히 분배함으로써 최적의 움직임 벡터를 예측하는 기법을 제시한다. 제안 기법은 H.264 부호기의 한가지 예측 모듈로 제안되며, 부가적으로 움직임 벡터 예측자 주변을 탐색하는 모듈이 수행된다. 실험 결과에서 제안 알고리즘이 기존의 H.264 부호기에 비해 매우 빠르면서 안정된 비트율-왜곡 성능을 가짐을 보인다.

### 1. 서론

최근 들어 정보통신 기술의 급격한 발전에 힘입어, Full-HD 와 같은 고해상도 영상의 보급이 일반화 되고 있다. 더불어, 영상을 제공 받기만 하던 소비자가 능동적으로 영상을 부호화 하는 경우가 지속적으로 증가하고 있으며, 이는 H.264 [1]와 같이 높은 복잡도를 가지는 기존의 부호기로는 실제적인 제품 개발에 한계를 갖는다. 따라서, 낮은 복잡도를 가지는 효율적 동영상 부호기의 개발이 요구되고 있는 실정이다.

H.264 부호기의 복잡도는 80% 이상이 움직임 예측 기법에 의해 발생되므로, 이를 단순화하는 기법들이 제안되었다 [2-4]. [3]은 움직임 탐색을 육각형 모양 및 축 방향의 샘플 포인트에 수행함으로써 90% 가량의 시간적 이득을 갖는 기법을 제안하였다. [4]는 다이아몬드 형태의 샘플 포인트에 대하여 점진적으로 탐색하는 기법을 사용하여 부호기의 복잡도를 크게 낮추었다. [2]는 탐색 영역을 세가지 영역으로 세분화하고 각 영역의 샘플 포인트 및 모드를 효율적으로 분배함으로써 약 93%의 시간적 이득을 얻었다.

이러한 노력에도 불구하고, Full-HD 와 같이 고해상도 영상의 부호화는 여전히 그 복잡도가 문제점으로 지적된다. 특히, 하드웨어적인 측면에서 H.264 부호기의 높은 연산량은 반드시 해결되어야 할 과제이다.

본 논문에서는 이상의 문제점을 해결하기 위하여 계층적 양자화가 적용된 움직임 벡터 탐색 기법을 제안한다. 효율적인 탐색을 위해, [2]에서 제안된 예측자 주변에 대하여 완전 탐색하며, 추가적으로 256x256 의 넓은 영역에 대하여 세 단계에 걸쳐 샘플링 및 양자화 기법을 사용하여 움직임 벡터를 예측한다. 이때, 탐색 영역에서 최적의 포인트를 찾기 위한 매칭 연산자로 저화질 영역에서는 NNMP (number of non-

matching points)를 사용하며 [5], 나머지 고화질 영역에서는 SAD (sum of absolute difference)를 사용한다.

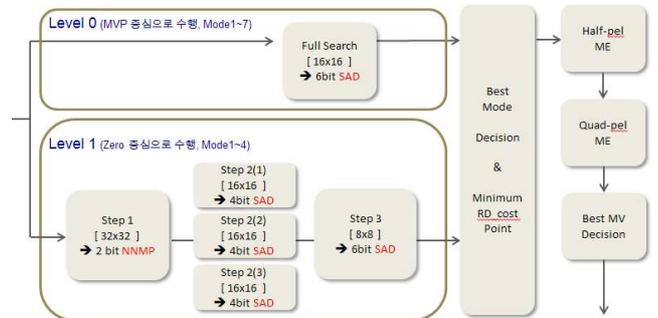


그림 1. 제안하는 움직임 벡터 예측 기법.

### 2. 제안 알고리즘

그림 1 은 제안하는 움직임 벡터 예측 기법을 보여준다. 우리는 예측 프레임에 대하여 두 단계로 구분한 뒤, 움직임 벡터 탐색을 진행한다. 먼저, Level 0 의 경우 움직임 벡터 예측자 (motion vector predictor, MVP)를 중심으로 16x16 영역에서 완전 탐색을 수행한다. 이때, 사용되는 픽셀 값은 6bit 로 양자화된 값을 사용하며, 매칭 연산자는 SAD 를 사용한다. 또한, 7 가지 모드로 이루어진 다양한 크기의 블록에 대하여 모두 수행한다. 블록 매칭 알고리즘에서 영상의 대부분 영역은 움직임 벡터 예측자와 동일한 움직임 벡터를 갖기 때문에, Level 0 만 수행하는 것으로도 비트율-왜곡 성능의 큰 손해 없이 시간적 이득을 얻을 수 있다.

Level 1 은 영상에 존재하는 물체의 불규칙하거나 빠른

움직임 및 가림 현상 등에 의해 움직임 예측이 어려운 경우를 보완해 주는 역할을 한다. [2]의 경우, 64x64 영역을 Level 1 으로, 256x256 영역을 Level 2 로 놓고 픽셀을 샘플링 하여 움직임 예측을 하였다. 그러나 이는 앞서 언급한 움직임 예측이 어려운 경우를 보완하지 못한다. 본 논문에서는 이러한 문제점을 해결하기 위하여, Level 1 을 256x256 전체 영역에서 8x8 영역까지 세 단계에 걸쳐 점진적으로 움직임을 탐색하도록 하였다.

첫 번째 단계는 (0,0)를 중심으로 256x256 영역에서 움직임을 탐색한다. 이때, 복잡도를 낮추기 위하여 8:1 로 샘플링 된 도메인에서 블록 연산을 수행한다. 이와 더불어 2bit 양자화 된 픽셀 값에 대하여 NNMP 매칭 연산값이 최소인 세 개의 위치를 저장한다. 두 번째 단계에서는 이렇게 넘겨 받은 세 위치에 대하여 32x32 영역에서 움직임을 재탐색 한다. 이때는 4:1 로 샘플링 된 도메인에서 2bit 가 추가된 4bit 픽셀 값으로 SAD 매칭 연산값이 최소인 한 위치를 저장한다. 마지막으로, 8x8 영역에서 완전 탐색을 수행한다. 이번 역시 2bit 추가된 6bit 픽셀 값으로 SAD 매칭 연산값이 최소인 위치 벡터를 Level 1 의 최종 결과로 저장한다. 결론적으로, Level 0 와 Level 1 의 매칭 연산값을 비교하여 최소인 위치 벡터 및 모드에 대하여 1 픽셀 이하의 탐색 모듈을 진행한다.

### 3. 실험 결과

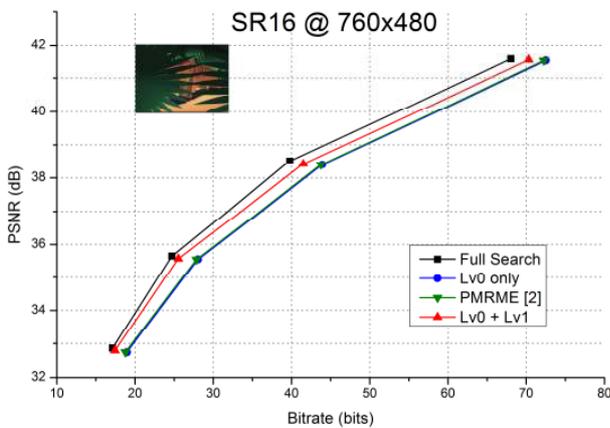


그림 2. Sr16 영상에 대한 비트율-왜곡 그래프.

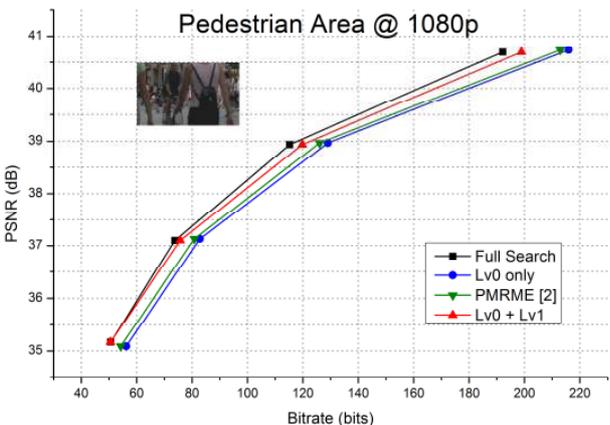


그림 3. Pedestrian 영상에 대한 비트율-왜곡 그래프.

		$\Delta$ PSNR	$\Delta$ bit rate	run time
Sr16	Lv0 only	-0.12	9.50%	1.45%
	PMRME	-0.07	9.38%	5.32%
	Proposed (Lv0+Lv1)	-0.02	3.59%	3.71%
PD	Lv0 only	-0.13	8.24%	1.13%
	PMRME	-0.05	7.81%	4.81%
	Proposed (Lv0+Lv1)	-0.04	3.32%	3.52%

표 1. H.264 대비 비트율-왜곡 변화량 및 수행 속도 비교

본 논문은 760x480 크기의 “Sr16” 영상 및 1080p 크기의 “Pedestrian Area” 두가지 영상에 대하여 실험하였다. 그림 2 와 그림 3 은 H.264 의 완전 탐색 부호화 비트율-왜곡 성능과 제안 알고리즘 및 비교 알고리즘 [2]의 비트율-왜곡 성능을 비교하여 보여준다. 두 영상 모두 움직임이 빠르고 불규칙적이기 때문에 완전 탐색에 비해 약간의 성능 저하가 있지만, 제안 알고리즘이 비교 알고리즘에 비해 보다 높은 성능을 보여준다. 표 1 은 평균적인 비트율-왜곡 차이 및 수행 시간을 나타낸다. 제안 알고리즘이 비트율-왜곡 성능 저하가 가장 적으면서도 어느 정도 빠른 수행 시간을 가짐을 보인다.

### 4. 결론

본 논문에서는 H.264 부호기의 복잡도를 낮추기 위한 고속 움직임 탐색 기법을 제안하였다. 예측 프레임에 계층적으로 양자화를 적용한 뒤, 각 단계마다 움직임 탐색 영역을 적절히 분배함으로써 최적의 움직임 벡터를 예측하는 기법을 제시하였다. 실험 결과에서 제안 알고리즘이 기존의 H.264 부호기에 비해 매우 빠르면서 안정된 비트율-왜곡 성능을 가짐을 보였다.

### 5. 참고 논문

[1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Transactions on Circuits System and Video Technology, vol. 13, no. 1, pp. 560-576, Jan. 2003

[2] C.-C. Lin, Y.-K. Lin, T.-S. Chang, "PMRME: a parallel multi-resolution motion estimation algorithm and architecture for HDTV sized H.264 video coding," IEEE International Conference on Acoustics, Speech and Signal Processing, 2007, pp. 385-388

[3] Z. Chen, , "Fast integer-pel and fractional-pel motion estimation for H.264/AVC," Journal of Visual Communication and Image Recognition, 2006, pp. 264-290

[4] A. M. Tourapis, "Enhanced predictive zonal search for single and multiple frame motion estimation," SPIE Visual Conference on Image Processing, 2002, pp. 263-266

[5] A. M. Tourapis, "Two-bit transform for binary block motion estimation," IEEE Transaction on Circuit System and Video Technology, vol. 13, no. 1, pp. 263-266, Jan. 2005