

다시점 비디오 부호화기를 위한 효율적인 참조 영상 선택 알고리즘

류승철, 서정동, 김동현, 손광훈

연세대학교

khsohn@yonsei.ac.kr

An Efficient Reference Picture Selection Method for MVC

Seungchul Ryu, Jungdong Seo, Donghyun Kim, Kwanghoon Sohn

Department of Electrical and Electronic Engineering, Yonsei University

요약

다시점 비디오 부호화기(MVC)는 다양한 블록 크기 기반 움직임 추정과 변이 추정을 수행한다. 또한 2개 이상의 다중 참조 영상 움직임 추정 기술을 사용한다. 이 기술들을 통해 MVC는 높은 부호화 효율을 얻을 수 있지만 실제 적용하기에는 너무 높은 부호화 복잡도가 걸림돌로 작용한다. 본 논문에서는 MVC의 부호화 복잡도를 감소시키기 위하여 효율적인 참조 영상 선택 알고리즘을 제안한다. 부호화에 사용된 참조 영상들은 인접한 블록들 간에 높은 상호 연관성을 가지므로, 부호화된 이웃 블록들의 참조 영상 정보를 기반으로 현재 블록의 참조 영상을 효율적으로 선택할 수 있다. 실험을 통해 제안된 알고리즘이 부호화 시간을 기존의 MVC에 비해 최대 73.3%, 평균 57.3% 감소시키며 부호화 효율의 감소는 무시할 만한 수준임을 확인하였다.

1. 서론

최근 3DTV는 HDTV의 다음 세대 방송매체로 각광받고 있다. 3DTV에 적합한 비디오 신호를 부호화하기 위해서는 기존의 2D 비디오 부호화기와 다른 새로운 부호화기가 필요하다. 이에 새로운 부호화기를 위한 많은 연구가 선행되었고, 그 결과로 여러 시점의 카메라로부터 동시에 획득한 영상을 부호화 하는 다시점 부호화기(MVC)가 제안되었다.

MVC의 주요 부호화 기술은 다양한 크기의 블록 기반 움직임 추정 기술과 변이 추정 기술이다. 움직임 추정 기술은 같은 카메라로부터 다른 시간에 얻은 영상들의 시간적 중복성을, 변이 추정 기술은 다른 시점의 카메라로부터 같은 시간에 얻은 영상들의 시점 간 중복성을 제거하는 기술이다. 또한 MVC에서는 시간적, 시점 간 중복성을 효율적으로 줄이기 위해 계층적 B 픽처 구조가 사용된다[1]. MVC에서 널리 사용되는 계층적 B 픽처 구조의 예시를 그림 1에 나타내었다. B 픽처 구조에서 각 픽처들은 2개 이상의 다중 참조 영상 움직임 추정과 변이 추정을 수행한다. 이러한 기술들 외에도 전 방향, 후 방향 그리고 양방향 예측 기술, RD 최적화 방법 등의 사용으로 MVC는 높은 부호화 효율을 가진다.

하지만 이 기술들은 실제로 적용하기에는 너무 높은 부호화 복잡도를 가진다. 따라서 실제 적용을 위해서는 부호화 복잡도를 줄이기 위한 연구가 필수적이다. MVC의 바탕이 된 2D 비디오 부호화기인 H.264/AVC의 부호화 복잡도를 줄이기 위한 몇 가지 방법들이 제안되었다[2-3]. MVC가 H.264/AVC를 기본으로 했음에도 불구하고 이러한 방법들은 MVC의 복잡한 구조 때문에 적용하기 어렵거나 효율적이지 못하다. 최근에 MVC의 부호화 복잡도를 줄이기 위한 몇 개의 논문들이 제안되었다 [4-6]. [4]에서는 탐색 범위와 참조 영상의 개수를 줄이기 위해 카메라세트와 움직임 벡터의 기하학적 관계에 기반한 방법이 제안되었다. 그 외에 Inter-모드를 RD 최적화 방법을 거치지 않고 결정하는 방법들이 [5][6]에 제안되었다. 하지만 움직임 추정과 변이 추정에 사용되는 참조 영상의 개수는 전체 부호화 복잡도에 큰 영향을 미침에도 불구하고 MVC를 위한 참조 영상을 효율적으로 선택 알고리즘은 제안된 바가 없다. 따라서 이를 효율적으로 선택하는 방법을 제안한다.

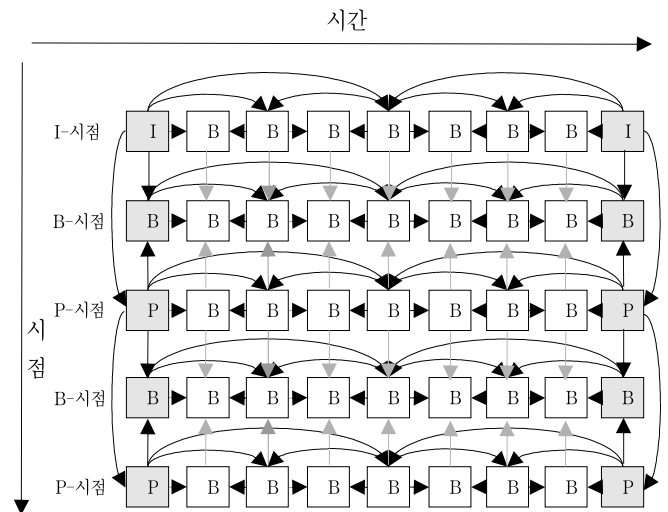


그림 1. 다시점 영상 부호화 계층적 B 픽처 구조

본 논문의 구성은 다음과 같다. 2장에서 효율적인 참조 영상 선택 알고리즘을 제안하고 제안된 기법의 실험 결과를 3장에서 기술한다. 그리고 마지막으로 4장에서 결론을 맺는다.

2. 제안된 효율적 참조 영상 선택 알고리즘

본 논문에서 제안된 알고리즘은 이웃 블록의 참조영상과 이웃블록과의 유사도를 기반으로 움직임 추정 및 변이 추정에 사용될 참조 영상을 선택한다. 2.1절에서 이웃 블록의 참조 영상 RPN, 2.2절에서 이웃 블록과의 유사도 SN_k 와 최소 이웃 유사도 MSN을 설명하고, 2.3절에서 전체적인 알고리즘에 대해서 설명한다.

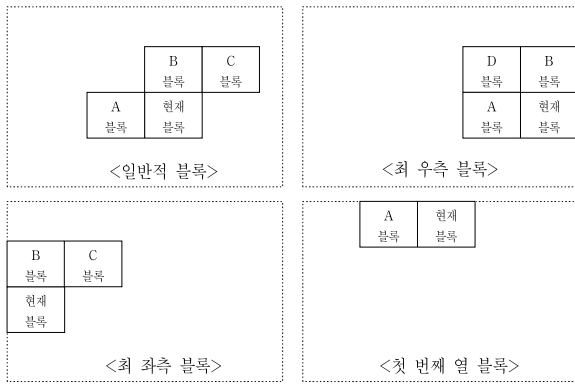


그림 2. 이웃 블록의 경우의 수

2.1 이웃 블록의 참조 영상 (RPN)

이웃 블록을 부호화 하는데 사용한 참조 영상의 집합을 RPN으로 정의하고 이를 수식으로 표현하면 다음과 같다.

$$RPN = \{x | x \text{는 이웃 블록들이 참조한 영상의 인덱스}\} \quad (1)$$

현재 블록의 위치에 따라 참조 가능한 이웃 블록은 그림 2에 나타난 것과 같이 4가지의 경우가 있다. 각각 경우 RPN은 2개, 3개 혹은 1개의 원소를 가진다.

2.2 이웃 블록과의 유사도 (SN_k)

현재 블록과 k번째 이웃 블록과의 유사도 SN_k를 다음과 같이 정의한다.

$$SN_k = \frac{1}{\frac{1}{16 \times 16} \sum_i \sum_j |p_c(i, j) - p_{n_k}(i, j)|} \quad (2)$$

여기에서 p_c(i, j)는 현재 블록의 픽셀 p_{n_k}(i, j)는 k번째 이웃 블록의 픽셀을 의미한다. 이웃 블록과의 유사도는 현재 블록과 이웃 블록의 평균 절대 차이 값(MAD)에 반비례한다. 사용되는 이웃 블록은 그림 2에 나타난 것처럼 4가지의 경우가 있다. SN_k 중 가장 작은 값을 최소 이웃 유사도(MSN)라고 정의하고 효율적인 참조 영상을 선택하는데 사용한다.

$$MSN = \min \{SN_1, \dots, SN_k\} \quad (3)$$

2.3 효율적 참조 영상 선택 알고리즘

본 논문에서 제안된 효율적인 참조 영상 선택 알고리즘은 2.1과 2.2에서 설명한 이웃 블록의 참조 영상 RPN과 최소 이웃 유사도 MSN을 이용한다. 움직임 추정 혹은 변이 추정 과정에서 사용가능한 모든 참조 영상을 사용하는 기존의 MVC와 달리 제안된 방식에서

표 1 실험 영상

| 실험 영상 | | |
|-------------|----------|-----------|
| 영상 | 해상도 | 시점(I-B-P) |
| Newspaper | 1024x768 | 2-4-6 |
| Lovebird1 | 1024x768 | 4-6-8 |
| Bookarrival | 1024x768 | 10-8-6 |
| Champagne | 1280x960 | 37-39-41 |

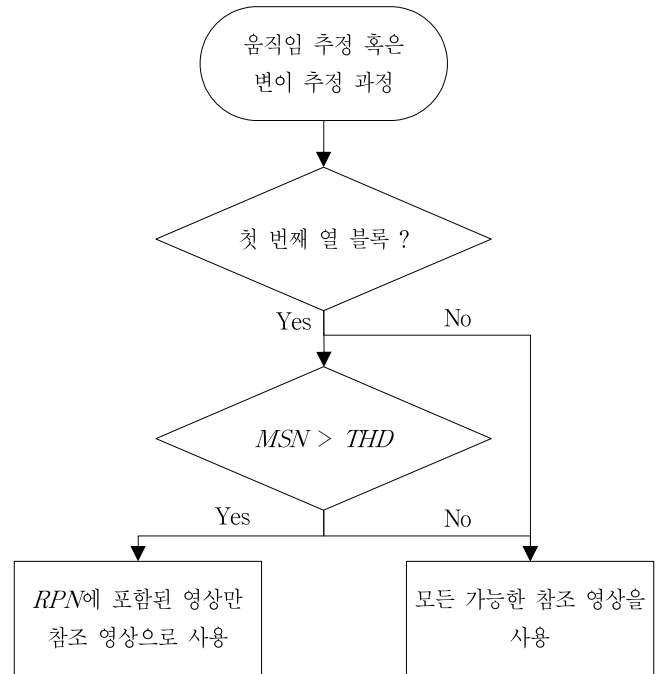


그림 3. 효율적인 참조 영상 선택 알고리즘

는 사용되는 참조 영상을 RPN의 원소로 제한한다. 즉, RPN에 포함되지 않은 참조 영상은 움직임 추정 혹은 변이 추정 과정에서 사용하지 않는다. 이 방법으로 참조 할 필요가 없는 영상을 제외할 수 있기 때문에 부호화 복잡도를 크게 줄일 수 있다.

하지만 만약 RPN이 적합하지 않는 참조 영상들만을 포함한다면 부호화 효율이 떨어지는 결과를 가져 오게 되고, 이것은 이후에 부호화 될 블록에도 연쇄적으로 영향을 끼치게 된다. 이를 방지하기 위해 MSN이 특정 값(THD) 보다 작거나 첫 번째 열의 블록의 경우 (그림 2의 우측 하단)에는 모든 참조 영상을 사용한다. THD는 실험적으로 선택되었으며 전체적인 알고리즘을 그림 3에 도시하였다.

3. 실험 결과

제안된 효율적 참조 영상 선택 알고리즘의 성능을 테스트하기 위해 4개의 영상(Newspaper, Champagne, Lovebird1, Bookarrival)을 사용하였다 [7]. 각 영상의 특성을 표 1에 나타내었다.

제안된 알고리즘은 JMVC 6.0 [8] 참조 소프트웨어에 구현되었다. 실험 조건은 표 2에 나타난 바대로 22.27, 32.37의 양자화 파라미터, 96 픽셀의 탐색 범위 그리고 GOP 크기는 15 혹은 8 (Bookarrival)의 조건에서 실험하였다. 탐색 모드는 고속 탐색을 사용하였으며, 각 영상마다 100개의 프레임을 사용하였다. 그 외의 조건은 [7]과 같다. 제안된 알고리즘의 성능을 비교하기 위해 JMVC 6.0과 그 결과를 비교 분석하였다.

표 2 실험 조건

| 실험 조건 | |
|----------|------------------------|
| GOP 길이 | 15 (8 for Bookarrival) |
| 탐색 범위 | 96 픽셀 |
| 탐색 모드 | 고속 탐색 |
| 총 프레임 수 | 100 |
| 참조 영상 갯수 | 2 |

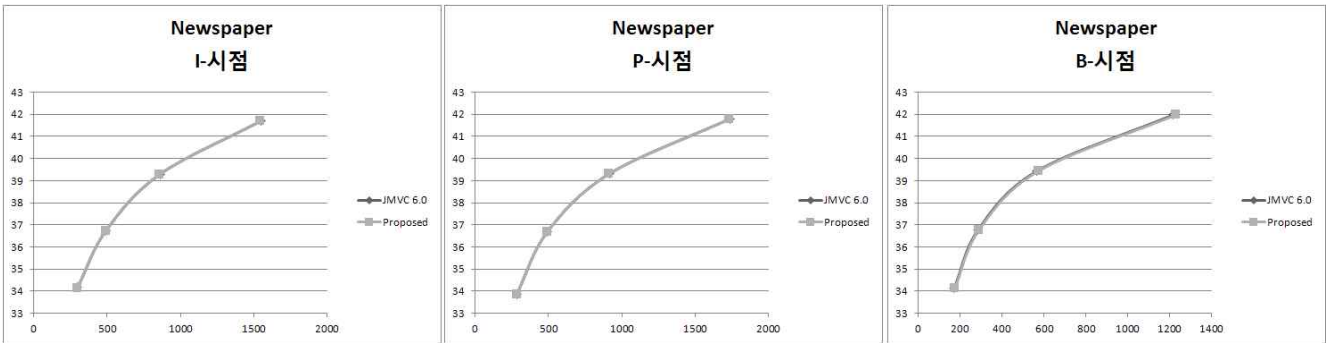


그림 3 Newspaper 영상의 I-,P-,B-시점의 RD-curve

표 3 효율적인 참조 영상 선택 알고리즘의 성능

| 영상 | 시점 | Δ Time (%) | BDPSNR (dB) | BDBR (%) |
|-------------|-------|-------------------|-------------|----------|
| Newspaper | I-시점 | -35.1% | 0.0104 | -0.2% |
| | P-시점 | -67.9% | 0.0043 | -0.1% |
| | B-시점 | -73.3% | -0.0552 | 1.4% |
| | Total | -66.1% | -0.0135 | 0.3% |
| Champagne | I-시점 | -31.9% | -0.0018 | 0.0% |
| | P-시점 | -59.5% | -0.0251 | 0.6% |
| | B-시점 | -64.7% | -0.0532 | 1.3% |
| Lovebird1 | I-시점 | -27.5% | 0.0049 | -0.1% |
| | P-시점 | -64.5% | -0.0486 | 1.4% |
| | B-시점 | -72.8% | -0.0407 | 1.0% |
| Bookarrival | I-시점 | -18.6% | 0.0028 | -0.1% |
| | P-시점 | -42.4% | -0.0173 | 0.6% |
| | B-시점 | -48.6% | -0.0529 | 1.7% |
| Total | | -41.9% | -0.0225 | 0.7% |

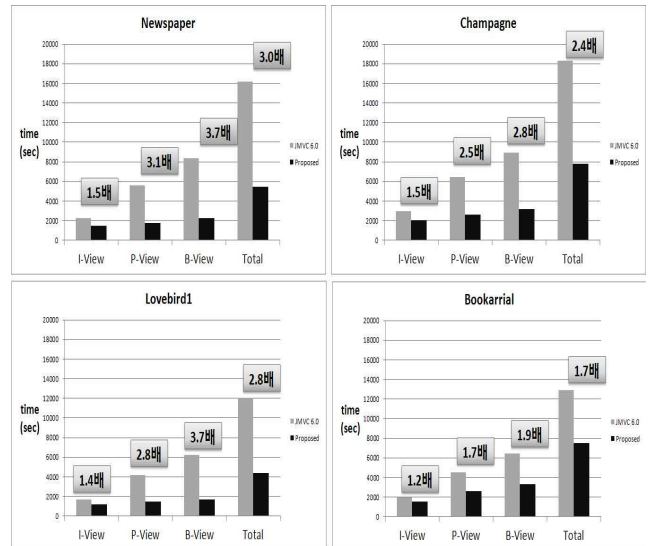


그림 4 각 영상의 부호화 시간 비교

제안된 효율적 참조 영상 선택 알고리즘이 기존의 JMVC 6.0에 비해 부호화 시간을 평균적으로 57.3%, 최대 73.3% (Newspaper 영상, B-시점) 줄임을 확인하였다. 그 손실로 평균 0.02dB의 BDPSNR[9]이 감소하였으나 그림 3의 RD-곡선에서도 확인 할 수 있듯이 이는 무시할만한 정도의 양으로, 제안된 알고리즘이 큰 부호화 효율의 손실 없이 부호화 시간을 크게 단축 할 수 있음을 의미한다. 표 3에 그 결과를 도시하였다. 그림 4는 각 영상의 부호화 시간을 나타낸다. 영상마다 차이는 있지만 제안된 알고리즘의 부호화 시간이 JMVC 6.0보다 평균적으로 2.3배 빠름을 확인 할 수 있다.

4. 결론

본 논문에서는 기존의 MVC의 실제 적용에 있어 걸림돌이 되는 높은 부호화 복잡도를 줄이기 위한 효율적 참조 영상 선택 알고리즘을 제안하였다. 제안된 알고리즘에서는 이웃 블록의 참조 영상(RPM)과 최소 이웃 유사도(MSN)를 이용하여 효율적으로 참조 영상을 선택 한다. 실험 결과 제안된 알고리즘은 평균 0.02dB BDPSNR의 무시할 만한 화질저하를 보였지만, 부호화 시간이 JMVC 6.0보다 평균적으로 57.3%, 최대 73.3% 감소하였다.

5. 참고문헌

- [1] H. Schwarz, D. Marpe and T. Wiegand, "Hierarchical B pictures," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-P014, Poznan, Poland, July 2005.
- [2] D. Wu, F. Pan, K. P. Lim, S. Wu, Z. G. Li, X. Lin, S. Rahardja and C.C. Ko, "Fast intermode decision in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 953-958, July, 2005.
- [3] L. Shen, Z. Liu, Z. Zhang, G. Wang, "Novel intermode decision algorithm in H.264/AVC," *J. Electron. Imaging*, Vol. 17, No.1, 013006, 2008.
- [4] X. Li, D. Zhao, X. Ji, Q. Wang, W. Gao, "A fast inter frame prediction algorithm for multi-view video coding," *ICIP*, Texas, USA, September 2007.
- [5] S. Y. Lee, K. M. Shin, and K. D. Chung, "An object-based mode decision algorithm for multi-view video coding," *ISM*, pp. 74-81, 2008.
- [6] L. Shen, T. Yan, Z. Liu, Z. Zhang, P. An, Lei. Yang, "Fast mode decision for multiview video coding," *ICIP*, Egypt, November, 2009.

- [7] H. Schwarz, D. Marpe and T. Wiegand, "Description of exploration experiments in 3D video coding," ISO/IEC JTC1/SC29/WG11 MPEG2010/N11274, Dresden, Germany, April 2010.
- [8] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. WD1 Reference software for MVC (JMVC) 6.0, Doc. JVT-AF14, November, 2009.
- [9] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-T Video Coding Experts Group document VCEG-M33, Mar. 2001.