

스케일러블 다시점 비디오 부호화를 위한 효율적인 움직임 예측 구조

김지훈, 이홍래, 정태준, 서광덕, *이하현, *강정원
 연세대학교, *한국전자통신연구원
 dwarfkjh@naver.com, *jungwon@etri.re.kr

Efficient Motion Estimation Architecture for Scalable Multi-view Video Coding

Ji-hoon Kim, Hong-rae Lee, Tae-jun Jung, Kwang-deok Seo, *Hahyun Lee, and *Jung won Kang
 Yonsei University, *ETRI

요약

본 논문에서는 SVC와 MVC를 통합하여 스케일러블 다시점 비디오 부호화를 구현할 때, SVC 부호화의 계층간 움직임 추정을 위해서 다른 시점의 비디오 정보를 활용함으로써 압축의 효율을 증대시킬 수 있는 움직임 예측 구조를 제안한다. 제안된 예측 구조의 적용으로 스케일러블 다시점 비디오 부호화의 부호화 효율을 1% 가량 향상시킬 수 있다.

1. 서론

H.264/AVC[1]의 표준화가 완료된 이후, ISO/IEC MPEG에서는 다계층의 비디오 부호화를 위한 스케일러블 비디오 부호화(SVC: Scalable Video Coding)[2]와 다시점의 비디오 부호화를 위한 MVC(Multi-view Video Coding)[3]에 대한 표준화를 진행하였고, JVT(Joint Video Team)에 의해서 표준이 완성되었다.

향후 비디오 콘텐츠 서비스가 다양한 종류의 네트워크와 다양한 성능을 갖춘 단말기를 바탕으로 확산되는 추세임을 고려할 때 다시점 비디오도 계층화된 형태로 부호화하여 서비스할 필요가 있다. 이렇게 할 경우 다시점 비디오 서비스를 주어진 전송 및 서비스 조건에 적응적으로 최적화하여 서비스할 수 있다. 따라서, 기존에 이미 표준 기술 개발이 완료된 계층화된 비디오 부호화 기술인 SVC와 다시점 비디오 부호화 기술인 MVC를 통합한 스케일러블 다시점 비디오 부호화 구조에 대해 연구하는 것은 매우 중요하다.

그림 1은 본 논문에서 구현한 SVC와 MVC를 통합한 스케일러블 다시점 비디오 부호화의 구조를 나타낸다. 이 그림에서는 2개의 카메라를 사용하여 얻어진 dual-view 비디오(시점 0+시점 1)를 입력 받아서 각각의 view를 2개의 공간적 계층으로 부호화하는 구조를 나타낸다.

본 논문에서는 SVC와 MVC를 통합하여 스케일러블 다시점 비디오 부호화를 구현할 때, SVC 부호화 과정에서의 계층간 움직임 추정을 위해서 다른 시점의 비디오 정보를 활용함으로써 압축의 효율을 증대시킬 수 있는 움직임 예측 구조를 제안한다.

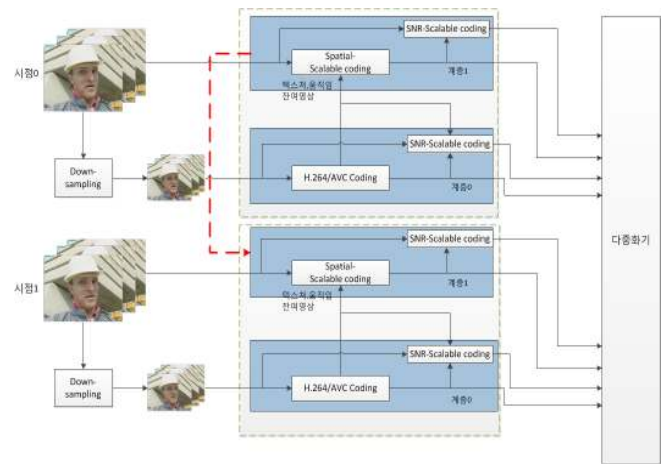


그림 1. 스케일러블 다시점 비디오 부호화의 구조.

2. 제안하는 움직임 예측 구조

스케일러블 비디오 부호화의 예측 방법은 동일한 계층에서 인접한 프레임(Frame) 간의 중복성을 제거하는 방법과 다른 계층 간의 상관 관계를 이용하여 부호화하는 방법이 있다. 같은 계층의 영상에서 인접하는 프레임은 거의 유사하므로, 이 유사함을 이용하여 다음 프레임의 움직임을 예측하고 보상하는 방식으로 시간적 중복성을 제거한다.

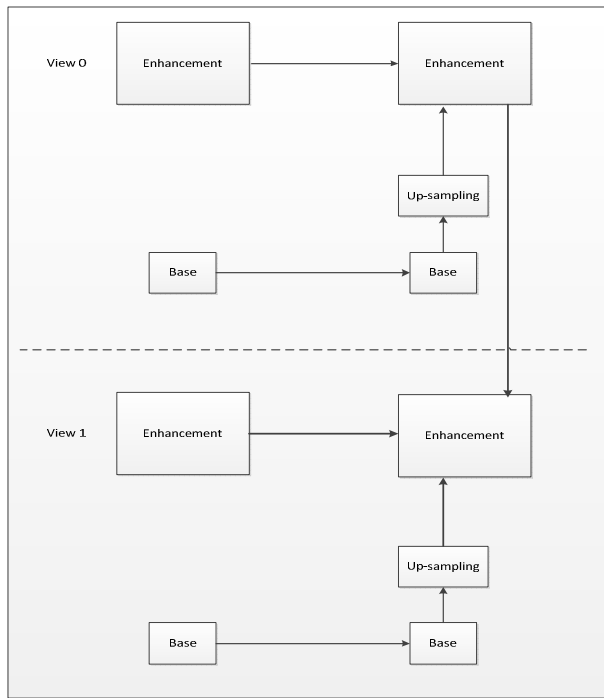


그림 2. 두개의 view 간의 Enhancement layer 예측 구조.

계층간(inter layer) 상관 관계를 이용하여 부호화 효율을 높이는 기본적인 방법은 하위 공간적 계층의 프레임을 복호화 한 후에, 이를 상위 계층의 프레임 크기에 알맞게 업 샘플링 한 영상을 상위 계층의 참조 프레임(Reference frame)으로 이용하는 방법이다. 동일 시간 내에 있는 하위 계층의 프레임과 상위 계층의 프레임은 높은 공간적 중복성을 갖기 때문에 부호화 효율이 향상 시킬 수 있다. 이처럼 하위 계층 프레임을 참조 프레임으로 쓸 경우, 업 샘플링 과정으로 인해 참조 프레임의 화질 저하가 발생한다. 참조 프레임의 화질 저하는 영상의 복호화 성능을 저하 시키는 큰 요인으로 작용하기 때문에, 본 논문에서 제안하는 방법은 동일 시간을 가지는 다른 view 의 프레임을 참조 프레임으로 이용하여 부호화 효율을 높이는 방법이다. 다른 view 일 지라도 동일한 시간을 가지는 프레임 간에는 높은 공간적 중복성을 가진다. 추가적으로 다른 view 의 상위 프레임을 참조영상으로 사용 할 경우, 인접한 프레임과 하위계층에서 업 샘플링된 프레임보다 더 효율적인 움직임 정보와 예측에 의한 잔여 정보를 얻을 수 있다.

3. 성능 평가

본 절에서는 앞에서 제안하는 스케일러블 다시점 부호화를 위한 움직임 예측 구조의 성능을 알아보기 위해 스케일러블 비디오 부호화(SVC) 방법과 성능 비교 실험을 수행하였다. 부호화 방법은 스케일러블 비디오 부호화의 참조 소프트웨어인 JSVM 9.19.14 와 본 논문의 제안 방법을 적용하여 구현된 소프트웨어를 사용하였다. 실험을 위해 참조 비디오 시퀀스로 좌영상 우영상 두 개로 이루어진 BMX, Tunnel, Trapeze 시퀀스를 사용하였다. 성능 평가를 위한 실험 조건으로 GOP size 는 8, Intra period 는 16, Frame rate 은 30 frame 으로 설정한 후, 각 영상에 따라 QP 값을 15, 25, 35 로 변경하면서 실험을 하였다. 제안된 예측구조가 적용 되었을 때, 스케일러블 비디오 부호화와의 성능 비교 결과는 그림 3에 나타나 있다.

QP 값을 35 로 설정하여 Tunnel 시퀀스를 부호화한 비교

결과에서는 제안된 예측구조가 적용되었을 때, 스케일러블 비디오 부호화 방법보다 비트율 측면에서 약 1.3%의 데이터를 절약 할 수 있음을 확인할 수 있다. PSNR 측면에서는 약 0.02dB 성능이 향상 되었다. QP 값이 15 로 낮아지면 비트율 측면에서의 성능이 약 0.4% 정도로 감소되는 것을 확인할 수 있는데, 이는 base layer 의 화질이 좋기 때문에 제안된 예측구조에서 추가된 picture 가 reference 로 선택되는 빈도가 감소 했기 때문에 성능이 떨어진 것으로 볼 수 있다. 따라서 QP 값이 점점 증가 할수록, 제안된 움직임 예측 구조 방법이 SVC 의 inter layer 예측 구조보다 비트율 측면과 PSNR 측면에서 성능이 점차 향상 되는 것을 확인할 수 있다.

제안된 예측구조는 기존의 SVC 의 inter layer 예측구조보다 비트율과 PSNR 측면에서 이득이 있지만, 이 이득을 얻기 위해서는 약간의 복잡도가 증가한다.

| | QP=15 | QP=25 | QP=35 |
|----------------|--------|--------|--------|
| 데이터 size 이득(%) | 0.378 | 0.903 | 0.911 |
| PSNR 이득(dB) | 0.0046 | 0.0044 | 0.0246 |

(a) BMX left sequence

| | QP=15 | QP=25 | QP=35 |
|----------------|--------|--------|--------|
| 데이터 size 이득(%) | 0.485 | 1.115 | 1.336 |
| PSNR 이득(dB) | 0.0103 | 0.0040 | 0.0277 |

(b) Tunnel left sequence

| | QP=15 | QP=25 | QP=35 |
|----------------|--------|--------|--------|
| 데이터 size 이득(%) | 0.318 | 0.733 | 0.539 |
| PSNR 이득(dB) | 0.0066 | 0.0244 | 0.0390 |

(c) Trapeze left sequence

그림 3. 제안된 예측구조와 SVC 기반의 예측구조의 성능 비교

4. 결론

본 논문에서는 스케일러블 다시점 비디오 부호화를 위한 효율적인 움직임 예측 구조를 제안하였다. 제안된 예측구조의 적용으로 스케일러블 다시점 비디오 부호화의 부호화 효율을 1% 가량 향상시킬 수 있었다.

감사의 글

본 연구는 방송통신위원회의 무안경 다시점 3D 지원 UHDTV 방송기술 개발사업의 연구결과로 수행되었음 [KCA-2012-11921-02001].

참고 문헌

- [1] ISO/IEC ITU-T Rec. H.264, Advanced Video Coding, June 2011.
- [2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits and Syst. for Video Technol.*, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.
- [3] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient Prediction Structures for Multiview Video Coding", *IEEE Trans. Circuits and Syst. for Video Technol.*, vol. 17, no. 11, pp. 1461-1473, Nov. 2007.