

특집논문 (Special Paper)

방송공학회논문지 제17권 제6호, 2012년 11월 (JBE Vol. 17, No. 6, November 2012)

<http://dx.doi.org/10.5909/JBE.2012.17.6.976>

ISSN 1226-7953(Print)

스케일러블 다시점 비디오 부호화를 위한 효율적인 움직임 예측구조와 DPB 설계

김지훈^{a)}, 정태준^{a)}, 이홍래^{a)}, 서광덕^{a)*}, 김진수^{b)}, 이하현^{c)}, 강정원^{c)}

Efficient Motion Prediction Architecture and Design of DPB for Scalable Multi-view Video Coding

Ji-Hoon Kim^{a)}, Tae-jun Jung^{a)}, Hong-rae Lee^{a)}, Kwang-deok Seo^{a)*}, Jin-soo Kim^{b)},
Hahyun Lee^{c)}, and Jung Won Kang^{c)}

요약

본 논문에서는 SVC와 MVC의 부호화 구조를 결합하여 구현된 스케일러블 다시점 비디오 부호화를 위한 효율적인 움직임 추정 기법과 DPB 설계 메커니즘에 대해 제안한다. 제안된 움직임 추정 기법에서는 부호화 과정에서 필요한 예측 부호화의 성능 향상을 위해 서로 다른 시점 (view)의 픽처 정보를 참조픽처의 후보로서 활용한다. 제안된 움직임 예측 구조에 의해서 압축된 비디오 데이터의 크기를 감소시켜 압축 효율을 증대시킬 수 있다. 또한, 스케일러블 다시점 비디오 부호화를 수행할 때 SVC와 MVC의 DPB (Decoded Picture Buffer)를 통합한 통합형 DPB 설계 메커니즘에 대해 제안한다. 다양한 실험을 통해서 제안된 예측 구조를 적용함으로써 스케일러블 다시점 비디오 부호화에서의 압축 효율의 향상을 얻어낼 수 있음을 확인하였다.

Abstract

In this paper, we propose an efficient motion prediction architecture and DPB design mechanism for scalable multi-view video coding which is implemented by integrating SVC and MVC coding algorithms. In the proposed motion prediction architecture, we employ pictures associated with other views as a candidate for reference picture for improved motion prediction performance. By the proposed prediction architecture, we could enormously reduce the size of compressed video data. When performing scalable multi-view video coding, an integrated DPB design mechanism is also proposed. It is shown by various simulations that the proposed motion prediction architecture for scalable multi-view video coding can result in reduced data size in the compressed bitstream.

Keyword : Motion prediction, Decoded Picture Buffer, SVC, MVC, Scalable Multi-view Video Coding

a) 연세대학교 (Yonsei University)

b) 한밭대학교 (Hanbat National University)

c) 한국전자통신연구원 (ETRI)

* Corresponding Author : 서광덕 (Kwang-deok Seo)

E-mail: kdseo@yonsei.ac.kr

Tel: +82-33-760-2788 Fax: +82-33-760-4323

※ 본 연구는 방송통신위원회의 “무안경 다시점 3D지원 UHDTV 방송기술 개발”사업 [KCA-2011-11921-02001] 및 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업의 지원을 받아 수행된 연구임 [No. 2010-0025268].

· Manuscript received September 10, 2012 Revised October 26, 2012 Accepted October 26, 2012

I. 서론

인터넷과 각종 기술의 급속한 발전에 따라, 동영상 정보를 포함한 각종 멀티미디어 서비스에 대한 요구가 커지고 있다. 최근에는 HDTV와 같은 고품질 방송 서비스가 등장하면서, 정보의 질적 향상에 더 많은 관심을 가지게 되었다. 특히, 3차원 정보를 이용한 다양한 멀티미디어 분야가 출현하여, 2차원 디지털 시대에서 3차원 디지털 시대로의 전환이 이루어지고 있다. 3차원 비디오는 이러한 3차원 디지털 시대의 영상 정보로서 그 중요성이 나날이 커지고 있다.

이러한 시대적 요청에 대응하여, 여러 가지 비디오 압축 방식이 개발되었다. H.264/AVC^{[1][2]}의 표준화가 완료된 이후, ISO/IEC MPEG에서는 다계층의 비디오 부호화를 위한 스케일러블 비디오 부호화 (SVC: Scalable Video Coding)^{[3][4]}와 다시점의 비디오 부호화를 위한 MVC (Multi-view Video Coding)^{[5][6]}에 대한 표준화를 진행하였고, JVT (Joint Video Team)에 의해서 표준이 완성되었다.

다시점 비디오 서비스를 다양한 네트워크 환경과 다양한 단말기에서 최적의 품질로 제공하기 위해서는 다시점 비디오 또한 계층화된 형태로 부호화할 필요성이 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. 그러나, 계층화 비디오 부호화 기술인 SVC와 다시점 비디오 부호화 기술인 MVC를 통합하여 부호화할 경우에는 참조픽처 관리에 대한 문제가 제기 된다. 비디오 부호화 및 복호화 과정에서 유연하게 화면 간 예측을 실현하기 위해서는 많은 참조픽처를 메모리에 저장해야 할 뿐 아니라, 비트열을 복호하는 순서와 출력장치에 표현하는 순서를 적절한 순서로 전환하기 위해서는 메모리가 필요하다. 이러한 과정을 수행하는 메모리들을 하나의 메모리로 통합하여 관리하는 메모리를 복호 픽처 버퍼(DPB: Decoded Picture Buffer)라고 부른다^{[7][8]}.

본 논문에서는 스케일러블 다시점 비디오 부호화 구조를 위해 필요한 DPB의 중요 매커니즘인 참조픽처의 재순서화를 위한 참조픽처 재순서화 (reference picture re-ordering), 참조픽처를 저장 및 개방하는 마킹 프로세스 (marking

process), 그리고 참조픽처 리스트를 구성하는 reference picture list construction 과정을 새롭게 제안하여 스케일러블 다시점 비디오 부호화를 효율적으로 수행할 수 있도록 한다.

그리고 다시점 스케일러블 비디오 부호화 과정에서 움직임 추정을 수행할 때, 계층 간 상관관계를 이용하여 부호화 효율을 높이는 계층 간(inter-layer) 예측구조와 시점 간 상관관계를 이용한 시점 간(inter-view) 예측구조를 결합하여 압축의 성능을 향상시킬 수 있는 IL(Inter layer)-IV(Inter view) 예측구조를 제안한다. IL-IV 예측구조는 계층 간 공간적 중복성과 시점 간 공간적 중복성을 모두 이용함으로써 기존의 SVC의 예측구조보다 효율적인 움직임 추정이 가능하다. 제안된 IL-IV 예측구조를 바탕으로 스케일러블 다시점 비디오 서비스를 주어진 전송 환경과 조건에 적응적이면서 유연하게 최적화 하여 서비스할 수 있다.

II. 관련된 기존 기술 분석

1. SVC의 예측 구조

SVC의 공간적 스케일러빌리티는 그림 1에서와 같이 작은 영상 크기의 비디오 시퀀스부터 큰 영상 크기의 시퀀스들을 각 공간적 계층에 입력하여 부호화 하는 다계층 부호화 기법에 기반한다^[3]. 이 때 각 공간적 계층들은 각각 부호화 할 수도 있지만, 하위 계층의 부호화 결과물을 상위 계층에서 이용하여 부호화 효율을 향상시킨다. 공간적 계층의 계층 간 상관관계를 이용해서 부호화 효율을 높이기 위한 계층 간 예측 기술의 가장 기본적인 방식은 하위 공간적 계층의 영상을 복호화한 후 상위 계층의 영상 크기에 맞게 화면을 업샘플링(spatial upsampling)하여 참조 화면으로 이용하는 것이다. 또 다른 계층 간 예측 방법으로는 계층 간 잔여 신호 참조 (inter-layer residual prediction)가 있다. 공간적 하위 계층의 양자화 계수 값이 크면, 잔여 정보를 예측하는 것은 부호화 효율에 도움이 될 수도 있다.

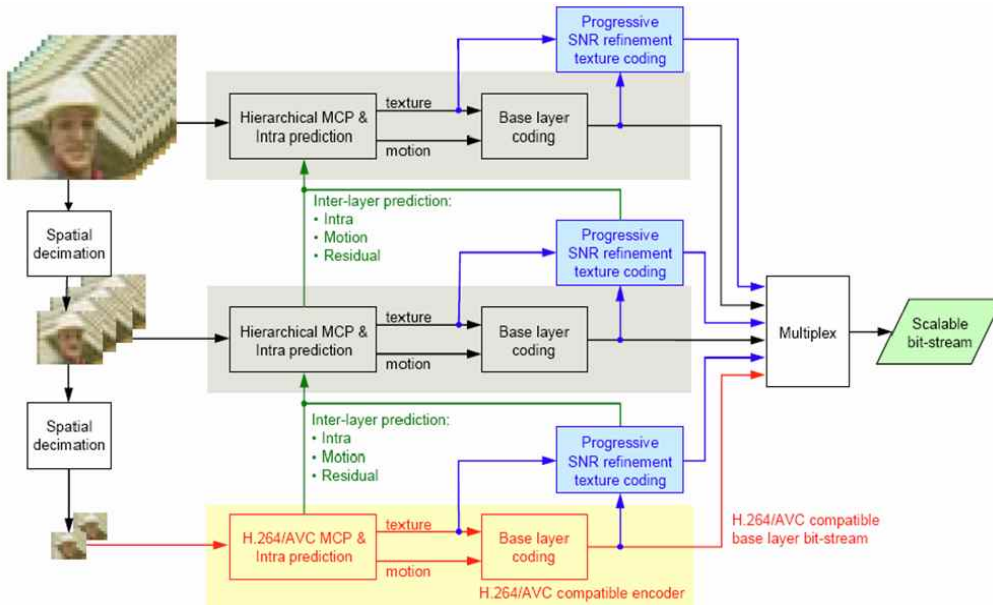


그림 1. 공간적으로 3 계층의 스케일러빌리티를 제공하는 SVC 부호화 예
 Fig. 1. Exemplary SVC encoding architecture supporting spatial scalability with three layers

2. MVC의 예측 구조

다시점 영상의 시점 간에 유사도가 높다는 사실을 기반으로 MVC에서는 기존 H.264에서 사용하는 단일 시점 코

딩 방법에서의 시간방향 예측과 함께 시점방향으로의 예측을 함께 수행함으로써 높은 부호화 효율을 성취하고 있다 [9][10].

그림 2는 다시점 비디오 부호화의 기본이 되는 예측 구조

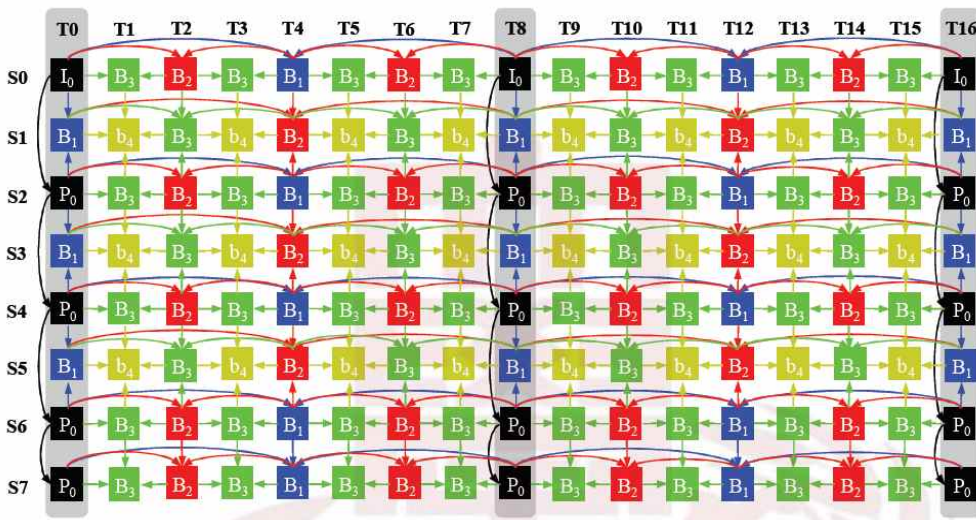


그림 2. MVC의 예측 구조의 예 [10]
 Table 2. Exemplary prediction architecture of MVC [10]

로써 8개의 시점을 포함하고 GOP의 크기를 8로 정하여 부호화를 수행 할 때의 구조를 보여준다. S0, S1, S2, ..., S7은 각각의 시점을 의미하고, T0, T1, T2, ..., T16은 시간 방향의 영상을 나타낸다. 그림 2에서 앵커픽처의 예측 구조를 살펴보면 S2 시점은 동일한 시간대의 S0 시점으로부터, S1 시점은 동일한 시간대의 S0 시점과 S2 시점으로부터, S4 시점은 동일한 시간대의 S2 시점으로부터, S3 시점은 동일한 시간대의 S2 시점과 S4 시점으로부터, S6 시점은 동일한 시간대의 S4 시점으로부터, S5 시점은 동일한 시간대의 S4 시점과 S6 시점으로부터, S7 시점은 동일한 시간대의 S6시점으로부터 예측을 수행하고, S7 시

점의 경우는 마지막 시점이기 때문에 S6시점으로부터 예측만을 수행한다. 다시점 비디오 부호화에서는 예측 부호화 설계가 상당히 유연하게 되어 있어서, 그림 2의 예측 구조 뿐만 아니라 응용에 따라서 다양한 예측 구조를 구성할 수 있다^[9].

III. 다시점 스케일러블 비디오 부호화를 위해 제안된 예측구조

본 연구를 위해 구현된 다시점 스케일러블 비디오 부호

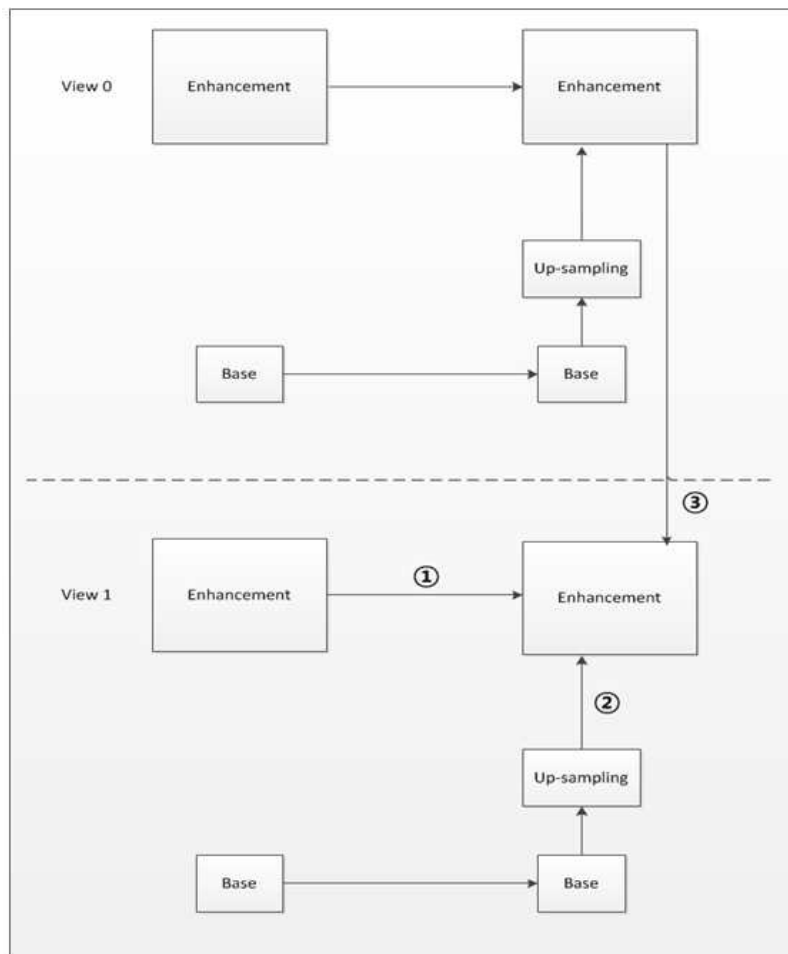


그림 3. 스케일러블 다시점 비디오 부호화를 위한 제안된 예측구조
 Fig. 3. Proposed prediction structure for scalable multi-view video coding

화의 예측구조는 단일 시점 부호화 방법에서의 시간방향의 예측 뿐만 아니라 서로 다른 시점 간의 예측을 함께 수행함으로써 부호화 효율의 향상을 목표로 한다. 다른 시점의 영상이라도 동일한 시간에 해당하는 영상 간에는 높은 공간적 중복성이 존재하기 때문에 다른 시점의 영상을 움직임 추정을 위한 참조픽처로 사용할 경우 보다 효율적인 움직임 추정이 가능하고 잔여신호 (residual signal)를 줄일 수 있다.

그림 3은 두 개의 시점을 가지는 영상의 공간적 상위 계층을 부호화할 때 제안된 예측구조를 적용한 경우이다. 시점 1의 상위계층의 예측방법에는 3가지 방법이 있다. 첫 번째로 단일 시점에서 동일한 계층 내의 시간방향으로 예측을 수행하는 것이다(그림 3에서 ①에 해당). 이는 시간상으로 인접한 영상을 참조픽처로 사용하는 것이다. 두 번째로 단일 시점에서 동일한 시간 내의 하위 계층으로부터 예측을 수행하는 방법이다(그림 3에서 ②에 해당). 하위의 공간

적 계층의 영상을 상위 공간적 계층에 알맞은 크기로 업-샘플링하여 참조픽처로 사용하는 것이다. 세 번째로는 동일 시간 내에서 다른 시점 (view)의 같은 공간적 계층의 픽처를 참조픽처로 사용함으로써 예측을 수행하는 방법이다(그림 3에서 ③에 해당). ①과 ②의 경우 기존의 SVC 예측구조에서 이미 활용하고 있는 방법이지만, ③의 방법을 추가함으로써 기존의 예측구조의 성능을 향상시키는 것이 본 논문의 목적이다. 따라서, 이 3가지 방법을 모두 고려하여 스케일러블 다시점 비디오 부호화를 수행함으로써 보다 더 최적의 움직임 추정 방법을 선택하여 영상 압축의 성능을 높일 수 있다.

그림 4는 상기의 움직임 추정 방법을 포함하여 본 논문에서 구현한 SVC 와 MVC를 통합한 스케일러블 다시점 비디오 부호기의 구조를 나타낸다. 이 그림에서는 2개의 카메라를 사용하여 얻어진 dual-view 비디오(시점0+시점1)를 입력 받아서 각각의 view를 2개의 공간적 계층으로 부호화하

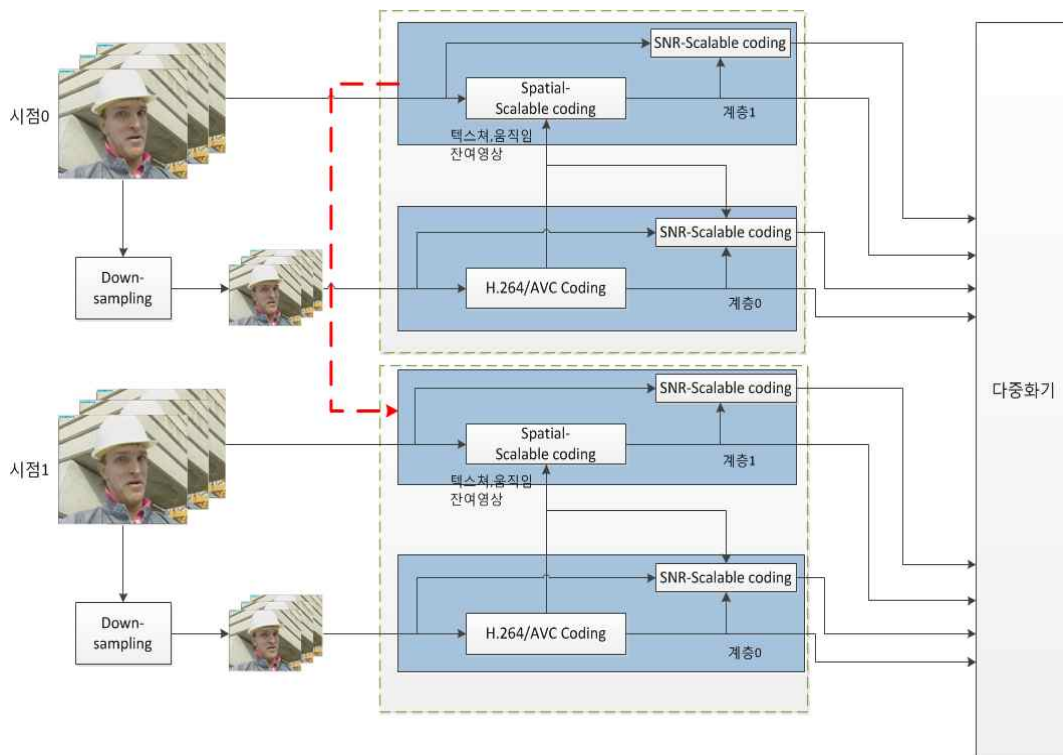


그림 4. 스케일러블 다시점 비디오 부호기의 구조
Fig. 4. Architecture of scalable multi-view video coding

기 위한 구조를 나타낸다. 각 시점 별로 원본 영상이 입력되면 시점 별로 원본 영상이 down-sampling 과정을 통해서 base layer 영상으로 만들어진다. 이후 각 시점 별로 SVC 부호화 과정을 통해 계층 간 예측이 이루어진다. 이와 동시에 시점 0의 Enhancement layer의 프레임을 동일 시간에 해당하는 시점 1의 Enhancement layer 프레임의 움직임 예측을 위한 참조픽처로 사용함으로써 시점방향으로의 예측을 동시에 고려하게 된다.

IV. 제안하는 통합형 DPB 설계 메카니즘

1. 통합형 DPB에서의 참조픽처 리스트 재순서화 (Reference Picture List Re-ordering) 방법

SVC의 부호화 구조와 MVC의 부호화 구조를 결합하여 그림 4의 스케일러블 다시점 비디오 부호화를 구현할 때 통합형 DPB에서 재순서화가 일어났을 경우 고려해야 할 사항들은 다음과 같다. 먼저, 각 식별번호를 구분하는 조건 연산자에서 SVC에서 사용되는 슬라이스인 EB/EP/EI (Enhancement layer B/P/I) 를 모두 고려해야 한다. 또한 MVC에서 사용되는 식별번호 4, 5 번 즉 view간 재순서화 과정을 관리하는 매커니즘이 고려되어야 한다. 재순서화 과정에서 참조픽처 번호의 할당은 할당 방법을 위한 ID(식별번호)와 그 할당 참조픽처를 지정하는 번호의 조합으로 구

성이 되는데 통합된 DPB에서의 재순서화 식별번호와 그 과정은 표 1과 같다. 표 1을 살펴보면 기존 MVC의 동작 과정을 기반으로 동작한다.

2. 통합형 DPB에서의 참조픽처 (Reference Picture) 리스트 구성

통합형 DPB에서 참조픽처 리스트를 구성하기 위한 매커니즘은 AVC/SVC/MVC가 모두 고려된 리스트 구성 방법이어야 한다. 즉, SVC에서 사용하는 dependency_id와 MVC에서 사용하는 view_id와 “only view reference picture”가 고려되어진 2중적 참조픽처 리스트 구성이 이루어져야 한다. 그림 5에서 시점 1의 상위계층의 EB1의 픽처를 복호화 한다고 가정을 하면 EB1은 시점 0에서의 EB0를 참조픽처로 사용하고 있다. 여기서 view_id가 고려된 참조픽처 리스트 구성이 이루어지게 되는데 EB0의 경우는 시점 간 예측에만 사용이 되기 때문에 시점 간 예측에서만 사용되는 참조픽처의 픽처로 마킹 (marking)이 이루어진다. EB0의 경우에는 EP0와 EI0에서 참조를 하고 있는데 이때는 SVC의 dependency_id가 고려되어져서 dependency_id가 1인 경우의 참조픽처 리스트만 가지게 된다. 그렇기 때문에 통합된 DPB에서 참조픽처 리스트 구성을 하는데 있어서 dependency_id, view_id, only-view reference picture가 모두 고려된 이중적 참조픽처 리스트 구성 과정이 고려되어야 한다.

표 1. 식별번호에 따른 재순서화의 동작 과정[5]
 Table 1. Operation of re-ordering according to ID number

ID (식별번호)	동작
0	지정한 단기참조픽처에 참조픽처 번호를 할당 지정한 단기 참조픽처의 참조픽처 번호는 직전에 할당한 참조픽처번호로부터 a를 뺀 값이다.
1	지정한 단기 참조픽처에 참조픽처 번호를 할당 지정한 단기 참조픽처의 참조픽처 번호는 직전에 할당한 참조픽처번호에 a를 더한 값이다.
2	지정한 장기 참조픽처에 참조픽처 번호를 할당
3	동작의 완료를 표시
4	지정한 단기 참조픽처의 참조픽처 번호는 직전에 할당한 view간 참조픽처 번호로부터 a를 뺀 값이다.
5	지정한 단기 참조픽처의 참조픽처 번호는 직전에 할당한 view간 참조픽처 번호로부터 a를 더한 값이다.

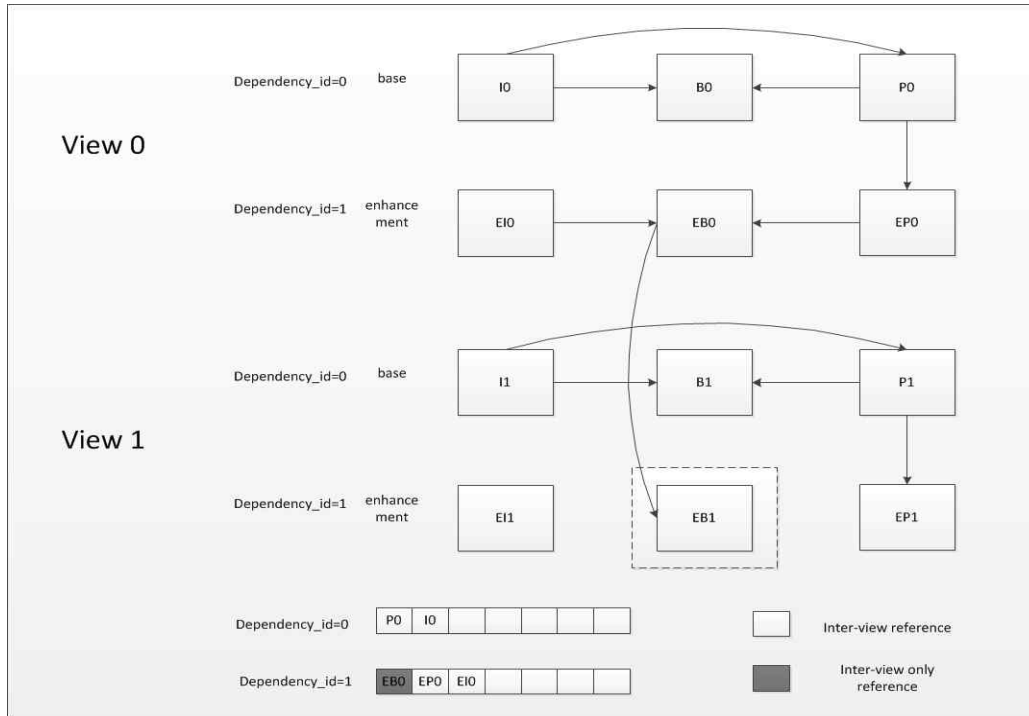


그림 5. 통합된 DPB에서 참조픽처 리스트 구성
 Fig. 5. Construction of reference picture list in unified DPB

3. 통합형 DPB에서의 참조픽처 마킹 프로세스

통합형 DPB에서의 참조픽처 마킹 프로세서는 SVC와 MVC의 dependency_id와 view_id가 모두 고려되어야 한다. SVC에서의 dependency_id를 고려해서 각각의 dependency_id별로 독립적으로 마킹 프로세스가 적용되는 점과 MVC에서 고려하고 있는 view_id를 적용해서 픽처 단위가 아닌 view component 단위로 마킹프로세스가 적용되는 것을 동시에 고려해야 한다.

4. 통합된 DPB의 순서도

통합된 DPB의 작동 순서는 그림 6과 같다. 처음에 슬라이스 판별 시 SVC에서의 Enhancement layer 계층에서의 EI/EB/EP와 MVC의 view_id에 따라 리스트 초기화 과정이 이루어진다. I/EI의 경우에는 각 계층별 종속성을 표현하는 dependency_id가 적용되어 Frame_num_wrap이 계산

되어지고 참조픽처 리스트 구성은 이루어질 필요가 없기 때문에 바로 재순서화 과정을 거치지 않고 마킹 프로세서로 넘어가게 된다. P/EP의 경우에는 순방향 예측만 이루어지기 때문에 List0이 구성되는데 view 간 예측이 일어나기 때문에 view_id도 고려되어 참조픽처 리스트가 구성되어야 한다. 다음은 B/EB의 경우인데 이때는 역방향의 예측 과정도 발생한다. 즉 List0 뿐만 아니라 List1도 생성된다. view_id와 dependency_id가 고려된 리스트가 구성되고 P/EP 그리고 B/EB의 경우는 두 가지 경우에 모두 참조픽처 리스트가 존재하기 때문에 재순서화 과정이 일어날 수 있다. flag 값이 0이면 재순서화는 일어나지 않고 마킹 프로세서로 넘어가게 되고 만약 flag가 1일 경우에는 재순서화 과정을 수행한다. 기존의 재순서화와 가장 큰 차이가 있는 부분은 상위계층인 Enhancement layer 부분의 슬라이스 EI/EB/EP가 고려되어야 하고 또한 view 간 예측에 의해서 MVC에서 존재하는 view 간 참조픽처의 재순서화 과정이 추가된다. 즉 식별번호 4, 5번에 따른 단기참조픽처

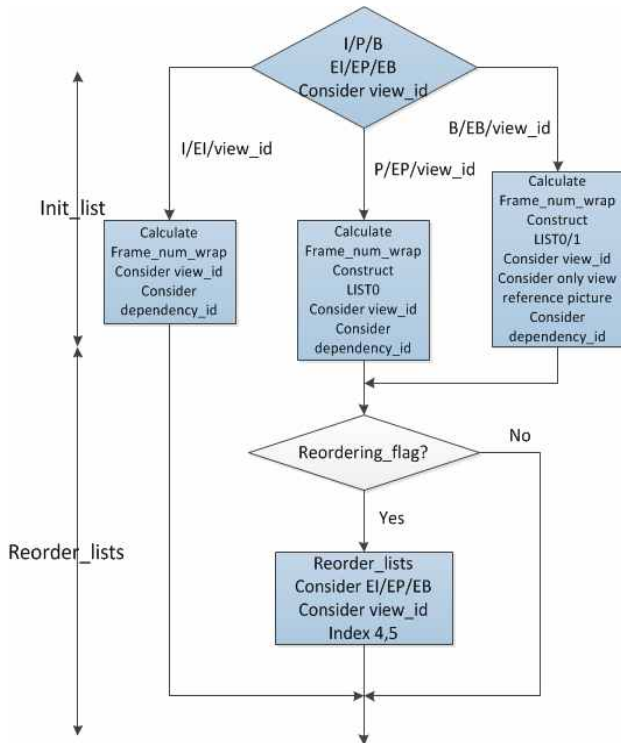


그림 6. 통합된 DPB의 참조 리스트 초기화와 재순서화 과정의 작동 순서도
 Fig. 6. Flow of reference list initialization and re-ordering in unified DPB

리스트 재순서화가 추가된다. 이후의 재순서화 과정은 기존의 참조픽처에 대한 번호 할당 방법대로 이루어지고 마킹프로세서로 넘어가게 된다.

마킹 프로세스 과정에서도 view_id와 dependency_id가 적용되어 동작한다. 그림 7은 통합된 DPB의 마킹 프로세스 순서도를 나타낸다. 기본적으로 기존의 AVC의 마킹프로세서를 따르지만, 계층 간 예측에서 종속성 표현인 dependency_id와 view 간 예측에 사용되는 view_id에 따라 동작을 수행한다. IDR 픽처의 여부를 먼저 확인하고, IDR 픽처의 경우에는 IDR 메모리 관리에 의해서 동작하고 DPB로의 저장이 이루어진다. IDR 픽처가 아닌 경우에는 adaptive flag에 의해서 단기 참조픽처만 존재하는 슬라이딩 윈도우 방식을 사용 할지 아니면 적응적 메모리 관리를 통해 장기 참조픽처와 단기 참조픽처를 모두 다 관리를 할 것인지 판별하게 된다^{[3][5]}. 이후 DPB에 저장이 되고 표시 순서를 맞추기 위해서 체크과정을 통해 출력이 이루어진다.

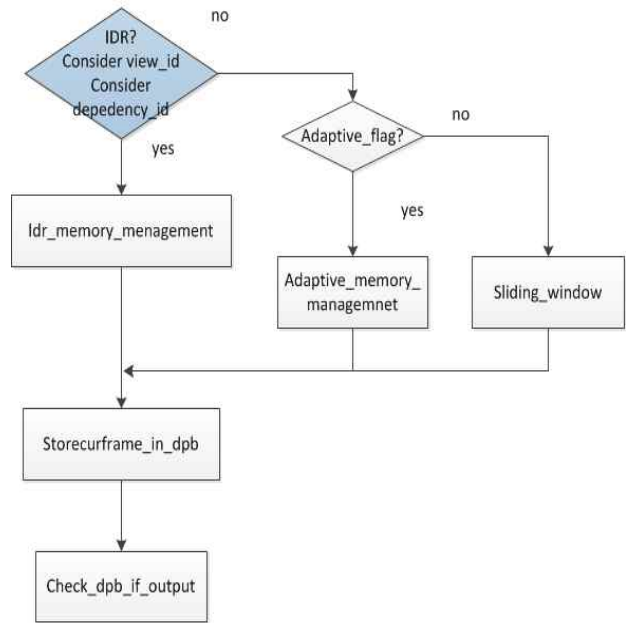


그림 7. 통합된 DPB의 마킹 프로세스 순서도
 Fig. 7. Flow of marking process in unified DPB

V. 실험 결과

실제적인 부호화 과정에서 제안된 예측구조가 적용 되었을 때의 영상의 압축률과 화질적 측면에서의 성능을 평가하기 위해서 JVT에서 주관하는 계위적 비디오 표준화 과정에서 개발된 공개 소프트웨어인 JSVM (Joint Scalable Video Model)^{[11][12]}을 기반으로 그림 4에 나타난 스케일러블 다시점 비디오 부호기를 구현하여 실험을 수행하였다.

성능평가를 위한 공통적인 실험 조건으로 GOP size=9, Intra period=16으로 설정하였으며, 초당 화면율 (frames/sec)은 30 으로 설정하였다. 참조 비디오 시퀀스로 좌영상, 우영상 두 개의 시점으로 구성된 BMX, Tunnel, Trapeze 시퀀스와 총 8개의 시점을 가지는 Exit 시퀀스를 사용하였다. 그리고 영상의 해상도는 Base layer는 QCIF 이고 Enhancement layer는 CIF 로 설정하여 두 개의 서로 다른 공간적 해상도를 지원한다. 움직임 추정에 사용되는 참조 영상의 최대 개수는 6장으로 설정하였다.

성능 평가 방법은 각각의 테스트 시퀀스에 대해 QP

(quantization parameter)값을 15, 25, 35로 변경해 가면서 기존의 SVC의 예측구조를 통해 부호화된 비트스트림과 제안하는 계층 간 및 시점 간 예측을 수행하는, 즉 IL-IV (inter layer-inter view) 예측구조가 적용되어 생성된 비트 스트림 간의 압축률과 평균 PSNR 을 비교하였다.

표 2는 두 개의 시점을 가지는 Tunnel 테스트 영상의 좌 시점 영상에 대한 실험 결과이다. 표 2에서 SVC prediction 은 각 시점 별로 독립적으로 SVC의 예측구조를 적용하여 부호화 한 결과이고 IL-IV는 제안된 예측구조를 적용하여 부호화 한 결과이다. 표 2의 실험 결과에 의하면, 제안하는 IL-IV 방식이 기존의 스케일러블 비디오 부호화 방식보다 대부분 압축률 측면과 PSNR 화질 측면에서 모두 이득을 얻을 수 있음을 알 수 있다. 표 2에서 QP 크기가 35일 때, 기존의 SVC의 예측 구조를 적용했을 경우 압축 데이터의

크기는 1141176 bits이고, IL-IV 방식을 적용 했을 경우에는 1105928 bits로 약 3.08%의 비트율이 감축되었다. 그리고 평균 PSNR 측면에서는 SVC의 예측구조를 적용 했을 경우에는 32.2930 dB이었지만, IL-IV 예측구조를 적용했을 경우에는 32.3207로 약 0.027 dB 만큼의 화질 향상을 가져왔다. 이는 동일 시간에 해당하는 다른 시점의 영상이 참조픽처로 사용될 경우 영상의 부호화시에 예측 성능을 향상 시켜 압축 데이터의 크기를 감소시키는 동시에 화질 측면에서도 이득을 얻을 수 있음을 의미한다.

QP 크기가 25일 경우에는, SVC의 예측구조를 적용 했을 경우에는 영상의 데이터 크기가 2873048 bits이고, IL-IV 방식을 적용했을 경우에는 2821000 bits로 1.81%만큼 데이터 크기가 감소하였다. 그러나 QP 크기가 35일 때 보다 데이터 압축률 향상은 조금 감소되었다. PSNR은 SVC의 예

표 2. QP 크기에 따른 Tunnel Left 영상의 압축 데이터의 크기와 평균 PSNR 비교
Table 2. Comparison of compressed data and average PSNR for Tunnel Left sequence

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
QP=15	8324368	8234016	1.0854	44.3549	44.3652	0.0103
QP=25	2873048	2821000	1.8116	38.1628	38.1668	0.004
QP=35	1141176	1105928	3.0887	32.2930	32.3207	0.0277

표 3. QP 크기에 따른 BMX Left 영상의 압축 데이터의 크기와 평균 PSNR 비교
Table 3. Comparison of compressed data and average PSNR for BMX Left sequence

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
QP=15	11615120	11541224	0.6362	44.8528	44.8482	0.0046
QP=25	4015672	3949408	1.6501	38.2574	38.2618	0.0044
QP=35	1472192	1428776	2.9490	33.0646	33.0892	0.0246

표 4. QP 크기에 따른 Trapeze Left 영상의 압축 데이터의 크기와 평균 PSNR 비교
Table 4. Comparison of compressed data and average PSNR for Trapeze Left sequence

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
QP=15	15715408	15465368	1.5910	45.0900	45.0966	0.0066
QP=25	4628016	4514080	2.4618	39.0228	39.0472	0.0244
QP=35	1774552	1714984	3.3568	34.2795	34.3185	0.039

측구조를 적용했을 경우 38.1628 dB이고 IL-IV 예측구조를 적용했을 경우에는 38.1668 dB로 나타나 0.004 dB의 이득이 있었지만 QP 크기가 35일 때 보다는 성능 향상이 조금 감소되었다.

QP 크기가 15일 때의 경우에는, SVC의 예측구조보다 제안한 IL-IV 예측구조가 데이터 압축률 측면에서 1.0854%만큼의 이득이 있었다.

전체적으로 IL-IV 예측구조를 적용 했을 경우에는 SVC의 예측구조보다 압축된 데이터 크기가 감소하는 동시에 평균 PSNR이 증가하는 성능 향상을 가져오고 이때의 성능 향상률은 QP 크기가 높아질수록 대체적으로 커진다. 그 이유는 QP 크기가 커지면 양자화 과정에서 영상의 화질이 크게 저하된다. 화질이 저하된 base layer를 up-sampling하여 계층 간 예측을 사용하는 것 보다 동일 시간 내의 시점 간 예측을 수행하는 것이 움직임 추정에 보다 더 효율적이기 때문이다. 시점 간 예측의 추가를 통해 계층 간 예측만을 적용할 때 보다 더 정확한 움직임 추정이 가능하여 압축률 측면과 화질적 측면 모두에서 이득을 얻을 수 있다.

표 2의 실험과 유사하게 표 3 및 표 4는 두 개의 시점으로 구성된 BMX 및 Trapeze 영상의 Left 시점 영상을 SVC의

예측구조와 IL-IV 예측구조를 적용하여 부호화한 결과를 비교한다. 실험 결과를 보면, BMX 영상은 QP 크기가 35일 때 데이터 크기는 약 2.94% 감소되었으며, 동시에 약 0.024 dB 만큼의 PSNR 향상이 있었다. QP 크기가 25, 15일 때는 데이터 크기는 각각 약 1.65%와 0.63% 감소되었으며, 평균 PSNR 성능에는 큰 차이가 없었다. Trapeze 영상의 경우에도 QP값을 35, 25, 15를 적용했을 경우 압축 데이터의 크기는 3.3~1.5%가량 감소하였고 PSNR은 약간 상승하였다. 표 2~4의 실험결과로부터 SVC 부호화의 예측구조보다 IL-IV 예측구조를 추가적으로 적용했을 경우에 영상 데이터의 압축률 측면에서 약 1~3% 가량의 향상이 있었고, PSNR 측면에서도 대체적으로 약간의 향상이 있었다.

표 5~6은 총 8개의 시점을 가지는 Exit 영상에 대한 실험 결과이다. 영상의 시점간의 거리가 IL-IV 예측구조 성능에 어떠한 영향을 미치는지 알아보기 위해 시점 간의 거리를 늘려가며 실험을 수행하였다. 참조 시퀀스로 사용된 Exit 시퀀스는 시점이 순서대로 0번부터 총 7번까지 8개의 시점으로 각 시점 별로 일정한 각도 차이를 두고 카메라로 촬영한 영상이다. 표 5는 시점간의 거리가 가장 가까운 시점 3과 시점 4를 그림 4의 부호기를 통해 함께 부호화 하였고,

표 5. Exit의 시점 3, 4 중 시점 3 영상에 대한 압축 데이터의 크기와 평균 PSNR 비교
 Table 5. Comparison of compressed data and average PSNR for view 3 among view 3 and 4 of Exit sequence

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
QP=15	7667464	7607896	0.7769	44.0936	44.0970	0.0034
QP=25	1480360	1450944	1.9870	39.8891	39.8969	0.0078
QP=35	600360	577952	3.7324	35.7868	35.8459	0.0591

표 6. Exit의 시점 1, 6 중 시점 1 영상에 대한 압축 데이터의 크기와 평균 PSNR 비교
 Table 6. Comparison of compressed data and average PSNR for view 1 among view 1 and 6 of Exit sequence

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
QP=15	8192344	8172008	0.2482	43.9721	43.9753	0.0032
QP=25	1416816	1407000	0.6928	39.8158	39.8170	0.0012
QP=35	566296	556376	1.7517	35.9549	35.9598	0.0049

시점 간의 거리가 가장 가까운 시점 3과 시점 4에 대한 실험 결과인 표 5를 살펴보면, QP 크기가 35일 때 SVC의 예측구조를 사용했을 경우에 압축 데이터 크기는 600360 bits 이고 IL-IV방식을 적용 했을 경우에는 577952 bits 로써 약 3.73%의 비트율이 감소되었다. 평균 PSNR 측면에서는 SVC의 예측구조를 적용했을 경우에는 35.7868 dB 이었고 IL-IV 예측구조를 적용했을 경우에는 35.8459 dB로써 약 0.0591 dB 만큼의 화질 향상을 가져왔다.

시점 간의 거리가 늘어난 Exit의 시점 2와 시점 5에 대해 동일하게 QP 크기가 35일 때의 결과를 살펴보면, SVC 부호화의 예측구조를 적용했을 경우에는 압축 데이터의 크기는 564504 bits이고 IL-IV방식을 적용 했을 경우에는 552352 bits로 약 2.15%의 비트율이 감소되었다. 그러나 시점의 거리가 가장 가까운 시점 3과 시점 4에 대한 부호화의 경우보다는 데이터 감소율이 축소되었다. 평균 PSNR 측면에서는 SVC의 예측구조를 사용했을 경우에는 36.0199 dB 이고, IL-IV 예측구조를 적용했을 경우에는 36.0271 dB로써 약 0.0072 dB 만큼의 화질 향상을 가져왔다. 그러나 PSNR 측면에서의 성능 향상률은 표 2의 경우보다 감소하였다.

시점 간의 거리가 더 늘어난 시점 1과 시점 6에 대해 QP 크기가 35 일 때 부호화 한 결과인 표 6을 살펴보면, 데이터 감소율은 SVC의 예측구조를 사용했을 경우보다 IL-IV 예측구조를 적용했을 때 데이터 압축률은 2.15%의 이득이 있었다.

표 5~6에서 다른 QP 값인 15, 25에 대해서도 유사한 결과를 얻을 수 있었다. 즉, 시점간 거리가 먼 시점에 해당하

는 영상을 압축할 경우 얻을 수 있는 데이터 압축률은 감소하였다.

결론적으로, IL-IV 예측구조는 시점간의 거리에 따라 그 성능 향상률이 달라지는데, 시점간의 거리가 가까운 영상에 적용 했을 경우, 높은 성능 향상을 보이고 시점 간의 거리가 멀어질수록 점차 성능 향상률은 떨어진다. 그 이유는 동일한 시간에서 다른 시점의 영상이 base layer를 up-sampling한 영상보다 화질적으로 우수하더라도 시점 간의 거리가 멀어지면 영상 간의 공간적 중복성이 감소하기 때문에 IL-IV 예측구조의 성능이 떨어지기 때문이다.

마지막으로 참조가 가능한 참조영상의 최대 개수에 따른 IL-IV 예측구조의 성능을 분석하기 위해서 참조영상의 개수를 4장으로 줄여서 실험을 수행하였다. 표 7은 QP 크기가 25이고 참조영상이 최대 4장인 각각의 테스트 영상에 대해 부호화한 결과이다. 표 2에서 Tunnel 시퀀스를 부호화할 때 최대 참조영상이 6장일 경우에는 데이터 감소율이 1.81% 이었다. 표 7에서 참조영상의 개수를 4장으로 줄일 경우에는 데이터 감소율이 1.94%로 성능이 향상되었다. Trapeze 시퀀스인 경우에도 데이터 감소율이 2.46%에서 2.67%로 참조영상이 4장인 경우에 데이터 감소율이 증가하여 성능이 향상되었다. BMX 영상의 경우에는 참조영상이 4 장일 때의 데이터 감소율이 6장인 경우의 데이터 감소율인 1.65% 보다 조금 향상되어 1.85%를 기록하였다. 그리고 Exit의 시점 3, 4 중 시점 3 영상을 부호화 했을 경우에 최대 참조영상 6장을 적용하면 데이터 감소율이 1.98% 이었는데, 참조영상의 개수를 4장으로 줄일 경우에는 데이터 감소율이 2.03%로써 성능이 조금 향상되었다.

표 7. 참조영상 4장을 적용한 경우 각각의 테스트 영상에 대한 압축률과 평균 PSNR 비교
Table 7. Comparison of compressed data and average PSNR for the case of four reference pictures

구분	압축된 영상 데이터 크기			평균 PSNR		
	SVC prediction (bits)	IL-IV (bits)	이득 (%)	SVC prediction (dB)	IL-IV (dB)	이득 (dB)
BMX	4015672	3941088	1.8573	38.2574	38.2596	0.0022
Tunnel	2863304	2807520	1.9482	38.1649	38.1688	0.0039
Trapeze	4603944	4486192	2.6737	39.0140	39.0384	0.0244
Exit(3,4)	1468200	1438344	2.0335	39.8982	39.9062	0.008

참조영상의 개수가 줄었을 경우, IL-IV 예측구조의 압축 성능이 대체로 향상되는 이유는 다음과 같이 분석된다. B 픽처를 부호화 할 경우에는 여러 영상을 참조영상으로 사용할 수 있다. 참조 영상이 충분히 많을 경우, 다른 시점의 영상을 참조영상으로 선택하는 것 보다 인접한 영상을 최적의 참조영상으로 선택할 확률이 증가한다. 그러나 참조영상의 개수가 작은 값으로 제한될 경우 동일 시간에 해당하는 다른 시점의 영상이 최적의 참조영상으로 선택될 확률이 증가하여 IL-IV 예측구조의 성능이 향상되게 된다.

한편, 제안된 IL-IV 예측 구조를 스케일러블 다시점 비디오 부호화에 적용할 경우에 기존의 SVC의 예측구조를 적용한 경우에 비해 시점 간 예측을 추가적으로 수행하기 때문에 부호화 시의 계산복잡도가 증가하게 된다. 상기의 모든 실험에 대한 계산 복잡도를 분석한 결과 IL-IV에 의해 추가되는 계산 복잡도는 대략 10~15% 범위로 측정되었다. 따라서, 실시간 부호화가 필요하지 않고 부호화 시간에 크게 제약이 없는 환경에서의 서비스의 경우 제안된 방법은 유용하게 활용될 수 있다.

VI. 결 론

본 논문에서는 스케일러블 다시점 비디오 부호화를 수행할 때, 부호화 과정에서 필요한 움직임 추정을 위해서 서로 다른 시점 (view)의 픽처 정보를 참조픽처로 활용하는 새로운 움직임 예측 구조를 제안하였다. 제안된 예측구조에서는 계층 간 움직임 추정을 위해서 다른 시점 (view)의 비디오 정보를 예측에 활용함으로써 압축의 효율을 증대시킬 수 있었다. 제안된 통합형 DPB는 dependency_id, view_id, 그리고 식별번호 등 SVC와 MVC에서 고려되는 부분들을 통합하여 설계되었다. 실험 결과로부터 제안된 움직임 예측 구조를 적용할 경우 SVC의 예측구조보다 압축 성능 측면에서 약 1~4%의 압축된 데이터의 크기를 절약 할 수 있었고, 화질적 측면에서도 대체적으로 약간의 PSNR이 개선

되었다. 제안된 방법은 실시간 부호화가 필요하지 않고 부호화 시간에 크게 제약이 없는 서비스 환경에서 효과적으로 활용될 수 있을 것으로 기대된다.

참 고 문 헌

- [1] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audio-visual services", ITU-T Recommendation H.264 and ISO/IEC 14496-10.
- [2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, July 2003.
- [3] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, and M. Wien, "Joint draft 11 of SVC amendment," Joint Video Team, Doc. JVT-X201, July 2007.
- [4] 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.
- [5] A. Vetro, P. Pandit, H. Kimata, A. Smolic, and Y.-K. Wang, eds., "Joint Draft 8 of Multiview Video Coding", Joint Video Team (JVT) Doc. JVT-AB204, Hannover, Germany, July 2008.
- [6] E. Martinian, A. Behrens, J. Xin, A. Vetro, and H. Sun, "Extensions of H.264/AVC for Multiview Video Compression", Proc. IEEE International Conf. on Image Processing, Atlanta, USA, Oct. 2006.
- [7] ISO/IEC JTC1/SC29/WG11, "SVC DPB management", JVT-P064, France, Oct. 2005.
- [8] ISO/IEC JTC1/SC29/WG11, "DPB management", JVT-F047, Japan, Dec. 2002.
- [9] M. Kitahara, H. Kimata, S. Shimizu, K. Kamikura, Y. Yashima, K. Yamamoto, T. Yendo, T. Fujii, and M. Tanimoto, "Multi-view video coding using view interpolation and reference picture selection", Proc. IEEE International Conference on Multimedia & Expo, Toronto, Canada, pp. 97-100, July 2006.
- [10] P. Merkle, A. Smolic, K. Mueller, and T. Wiegand, "Efficient Prediction Structures for Multiview Video Coding", IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 11, Nov. 2007.
- [11] "JSVM Software", Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, CVS tag : JSVM_9_19_14, June 2011.
- [12] "JMVC Software", Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, CVS tag : JMVC_8_5, March 2011.

저 자 소 개



김 지 훈

- 2011년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2011년 3월 ~ 현재 : 연세대학교 전산학과 석사과정
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, DASH, MMT



정 태 준

- 2010년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2012년 8월 : 연세대학교 전산학과 석사
- 2012년 9월 ~ 현재 : 연세대학교 전산학과 박사과정
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, DASH, MMT



이 흥 래

- 2010년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2012년 8월 : 연세대학교 전산학과 석사
- 2012년 9월 ~ 현재 : 연세대학교 전산학과 박사과정
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, DASH, MMT



서 광 덕

- 1996년 2월 : KAIST 전기및전자공학과 학사
- 1998년 2월 : KAIST 전기및전자공학과 석사
- 2002년 8월 : KAIST 전기및전자공학과 박사
- 2002년 8월 ~ 2005년 2월 : LG전자 단말연구소 선임연구원
- 2005년 3월 ~ 현재 : 연세대학교 컴퓨터정보통신공학부 부교수
- 2012년 10월 ~ 현재 : 미국 플로리다 대학교 방문교수
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, DASH, MMT



김 진 수

- 1991년 : 경북대학교 전자공학과 학사
- 1993년 : KAIST 전기및전자공학과 석사
- 1998년 : KAIST 전기및전자공학과 박사
- 1997년 ~ 2000년 : 삼성전자 선임연구원
- 2000년 ~ 현재 : 한밭대학교 정보통신컴퓨터공학부 교수
- 주관심분야 : 멀티미디어 스트리밍, SVC, Distributed Video Coding, 디지털 방송

저 자 소 개



강 정 원

- 1993년 2월 : 한국항공대학교 항공전자공학과 학사
- 1995년 2월 : 한국항공대학교 항공전자공학과 석사
- 2003년 8월 : Georgia Institute of Technology Department of ECE 박사
- 2003년 10월 ~ 현재 : 한국전자통신연구원(ETRI) 방송통신융합연구부문 선임연구원
- 2007년 3월 ~ 현재 : 과학기술연합대학원 대학교 이동통신 및 디지털방송공학점임 조교수
- 주관심분야 : 비디오 신호처리, 비디오 부호화, 비디오 적응변환, 디지털방송 미디어, 미디어 컨버전스



이 하 현

- 2002년 : 한국항공대학교 항공전자공학과 학사
- 2007년 : 과학기술연합대학원대학교(UST) 이동통신 및 디지털방송공학 석사
- 2003년 ~ 2005년 : LG전자 연구원
- 2008년 ~ 현재 : 한국전자통신연구원(ETRI) 실감미디어연구팀 선임연구원
- 주관심분야 : 영상 부호화, 영상 처리, 멀티미디어 시스템