

# 실시간 스케일러빌리티 변환 SVC 비트스트림에 대한 연구

배태면\*, 김덕연\*, 노용만\*, 강정원\*\*, 김재곤\*\*

\*한국정보통신대학교 멀티미디어 그룹

\*\*한국전자통신연구원 방송미디어연구 그룹

{yro, heartles}@icu.ac.kr

## A Study of Realtime scalability converted bitstream extraction

Tae Meon Bae\*, Duck Yeon Kim \*, Yong Man Ro\*, Jung Won Kang \*\*, Jae Gon Kim\*\*

\*Multimedia Group, Information and Communication University

\*\*Broadcasting Media Research Group, Electrics and Telecommunication Research Institute

### 요 약

스케일러블 비디오 부호화 (Scalable Video coding)는 시간, 공간, 화질의 스케일러빌리티를 가지는 비트스트림을 제공하므로, 가변하는 네트워크 환경에서 QoS 를 만족시키는 비디오 서비스를 가능케 한다. 먼저 스케일러블 비디오 부호화에서 실시간으로 가변 가능한 스케일러빌리티에 대한 분석을 하였으며, 이를 바탕으로 본 논문에서는 실시간으로 스케일러블 비디오 비트스트림의 시간, 및 화질을 변화하며 추출할 수 있는 방법을 제안한다. 제안한 방법은 JSVM2.0 에 구현되었으며, 실험 결과는 제안한 방법이 유효함을 보여준다.

### I. 서 론

ISO/IEC MPEG 와 ITU-T VCEG 의 JVT 에서는 MCTF(Motion Compensated Temporal Filtering) 와 H.264 를 확장한 SVC(Scalable Video Coding)을 표준화하고 있다[1]. 현재 표준화 되고 있는 SVC 는 JSVM(Joint Scalable Video Model)으로 구현되어 있으며, 공간, 시간, 품질상의 스케일러빌리티를 가진 비트스트림을 제공하고 있다[2]. SVC 의 목적은 공간적, 시간적, 품질적 특성의 확장성을 제공하는 비트스트림을 생성하여, 다양한 이종의 네트워크 환경 및 단말에 대해 QoS(Quality of Service)가 보장되는 비디오 스트리밍 서비스를 사용자에게 제공하는 것이다. 이와 같은 서비스가 효율적으로 제공되기 위해서는 원본 비디오 비트스트림으로부터 네트워크 환경에 맞추어 비트율을 가변하면서 비디오 비트스트림을 추출하는 방법에 대한 연구를 수행 할 필요가 있다.

본 논문에서는 스케일러블 비디오 부호화 (scalable video coding)에 있어서, 실시간으로 스케일러블 비디오

비트스트림의 프레임율 및 화질을 변화하며 추출할 수 있는 방법을 제안한다.

### II. SVC 에서의 스케일러빌리티 비트스트림 추출

#### 2.1 SVC 에서의 스케일러빌리티 지원

SVC(Scalable Video Coding)에서의 공간 해상도에 대한 스케일러빌리티는 레이어 기반 부호화를 통해 지원된다. 그림 1은 SVC 에서의 스케일러빌리티를 보여준다. 하위 레이어에서 낮은 공간 해상도의 영상에 대한 부호화가 이루어 지고, 상위 레이어에서는 하위 레이어의 영상정보를 추가로 이용하여 영상부호화를 수행한다. 시간적인 스케일러빌리티는 MCTF(Motion Compensated Temporal Filtering)를 통해 지원이 되며, 여러 개의 시간적 레벨을 통해 다양한 프레임율을 제공할 수 있다. MCTF 는 움직임에 의한 화질 저하를 막기 위해 움직임 예측/보상을 고려한 시간 방향으로의 wavelet 필터링이다. MCTF 는 GOP(Group Of Picture)

단위로 이루어지며, 따라서 프레임율은 GOP 길이에 따라 가변할 수 있는 범위가 결정된다. 마지막으로 화질에 대한 스케일러빌리티는 CGS(Coarse Granular Scalability) 또는 FGS(Fine Granular Scalability)에 의해 구현될 수 있다. CGS 는 레이어 기반으로 부호화를 수행하는 측면에서 공간적 스케일러빌리티 부호화 방식과 동일하지만 공간적 부호화의 경우, 상위 레이어에서 공간 해상도가 올라가는 반면, CGS 에서는 공간 해상도의 변화없이 높은 화질로 부호화하는 차이가 있다. 반면 FGS 는 QP 값을 올리면서 부호화하여, 마치 비트플레인(bit plane)별로 부호화 하는 효과를 나타낸다.

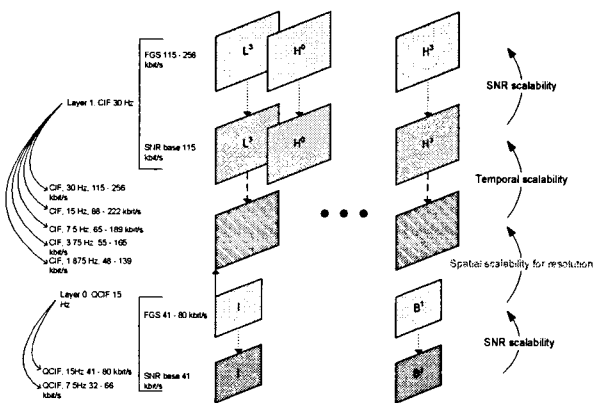


그림 1. SVC 에서의 스케일러빌리티

## 2.2 SVC 비트스트림 추출기

시간, 공간, 화질에 대한 스케일러빌리티를 가지는 SVC 비트스트림은 비트스트림 추출기를 통과하면서 스케일러빌리티가 변환된 새로운 비트스트림을 생성하게 된다. 현재 JSVM2.0 의 비트스트림 추출기는 특정 공간 해상도, 프레임율, 비트율의 세 가지 값을 지정하면 이에 맞추어 비트스트림을 추출해 낸다. 이때 특정 공간해상도, 프레임율, 비트율의 세 가지 값을 추출점(extraction point)라고 한다. 그림 2는 현재 JSVM2.0 의 비트스트림 추출기의 구조이다.

실제 비트스트림 추출기는 두 단계를 거쳐 비트스트림 추출을 수행한다. 첫 번째 단계는 비트스트림 분석하는 과정이며, 두 번째 단계는 분석결과를 근거로하여 추출기에 입력값으로 주어지는 추출점에 따라 추출할 NAL 유닛을 선택한 후, 비트율에 맞추어 FGS 레이어의 NAL 데이터를 처리하여 출력 비트스트림을 구성하는 단계이다.

첫 번째 단계에서의 분석은 scalability information

SEI 로부터 스케일러빌리티에 대한 정보를 얻고, 이로부터 추출점의 공간해상도와 프레임율을 나타내는 dependency id 와 temporal level 값을 알아낸다. 또한 입력 비트스트림에 포함된 각 공간적, 시간적, 화질적 스케일러빌리티에 따른 데이터 량을 계산한다. 다음 단계에서는 NAL 유닛 헤더에는 NAL 데이터의 공간 해상도와 프레임율을 나타내는 dependency id 와 temporal level 정보가 존재하는데, 추출기는 이 정보를 이용하여 dependency id 와 temporal level 의 값들 중 하나라도 추출점의 값들 보다 큰 NAL 유닛들을 제거한다. 마지막으로 제거되지 않은 NAL 들에 대해서는 앞 단계에서 분석한 데이터 량에 대한 결과를 토대로 추출점의 비트율에 맞추어 FGS 레이어의 NAL 유닛을 자르게 된다.

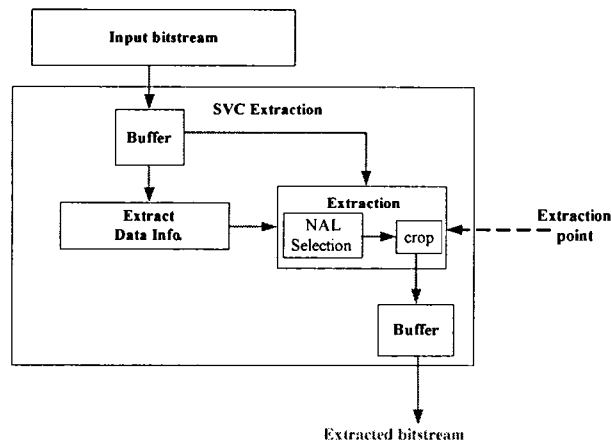


그림 2. 비트스트림 추출기의 구조

## III. SVC 비트스트림에서의 실시간 스케일러빌리티 변환

### 3.1 실시간 공간 해상도 스케일러빌리티 변환

SVC 에서는 움직임 예측(motion estimation)시에 참조 영상은 GOP 안의 영상을 이용하여 수행된다. 이것은 MCTF 에 의해 시간적 스케일러빌리티를 제공하기 위한 것이다. 그러므로 GOP 단위로 공간해상도를 가변을 하는 경우, 하나의 GOP 에서는 동일한 공간해상도를 가지므로 참조영상이 복호화되지 않는 문제는 발생하지 않는다.

키 프레임의 경우, 이전 GOP 의 키 프레임들을 참조하여 움직임 예측을 수행한다. 공간해상도를 축소시키는 경우, 축소된 프레임 이전의 프레임들이 움직임 보상에 의한 복호화 할 때 사용 가능하므로 문제를 발

생하지 않는다. 그러나 공간해상도를 향상시키는 경우, 향상된 프레임이 움직임 보상에 의해 복호화를 수행하면 문제가 발생할 수 있다. 즉 이전 프레임은 작은 해상도로만 복호화되어 있으므로 현재의 향상된 해상도에 맞는 참조영상은 복호화되어 있지 않아, 움직임 보상을 위해 참조할 영상이 존재하지 않는 것이다. 또 하나의 문제는 IDR 에 관련되어 있다. 고해상도를 가지는 복호레이어는 참조영상 리스트 버퍼를 초기화 할 필요가 있다. 따라서 공간해상도의 실시간 가변은 부호화시 고려되지 않는다면 비트스트림 추출기에서 지원이 불가능하다. 또한 현재의 SVC 복호화기는 출력영상의 해상도가 가변하며 복호화하는 기능을 지원하지 않기 때문에 이에 대한 고려가 필요하다.

### 3.2 비트스트림 추출

실시간 스케일러빌리티 가변 추출기는 실시간으로 추출점을 입력받아 적용할 필요가 있다. 입력 추출점에서 전송 비트율을 고려하기 위해서는 추출기에 입력될 비트스트림의 데이터량을 알 수 있어야 가능하다. 네트워크에서의 가용 비트율(available bandwidth)은 전송되는 데이터의 신호의 특정 시간 간격(time window)으로 확인할 수 있다[3]. 따라서 비디오 스트리밍 서버의 경우, 주기적인 시간간격에 따라 가용 비트율을 확인할 수 있으며, 이에 따라서 비트스트림 추출기에게 추출점을 알려 줄 수 있다.

SVC 비트스트림에서 프레임율은 MCTF 의 temporal level 에 따라 결정된다. GOP 내의 프레임들은 각각 특정 프레임율에 영향을 미치는 temporal level 값을 가지며, 그 값은 NAL 헤더에 기록된다. 따라서 프레임율을 결정하기 위해서는 GOP 단위의 비트스트림 추출이 필요하다. 따라서 비디오 스트리밍 서버에서는 GOP 의 정수단위로 가용 비트율을 예측하고, 이를 기반으로 비트스트림의 추출을 수행하는 것이 필요하다.

가용 비트율에 따른 추출점은 하나 이상이 있을 수 있다. 예를 들어 100 Kbps 의 비트율의 맞추기 위해 CIF, 30fps 또는 CIF, 15fps 로 추출할 수도 있다. 이러한 여러 가능한 추출점 가운데 하나를 선택하는 것은 각각의 추출점을 통해 얻어진 결과에 대한 화질을 비교함으로써 최종 선택점을 결정할 수 있다[4].

### 3.3 실시간 스케일러빌리티 가변 비트스트림 복호화

실시간 스케일러빌리티 가변 비트스트림의 기본적인 복호화는 스케일러블 비디오 복호화방법과 동일하다. 화질 가변의 경우, 복호화기는 단순히 FGS 레이어를 복호화하므로 FGS 레이어 관련 NAL 이 얼마나 절단되었는지는 복호화에 영향을 미치지 않는다. 프레임을 가변의 경우, MCTF 의 update step 과 화면 출력시 출력 타이밍과 관련된다. 현재 복호화기는 temporal level 의 가변에 따른 MCTF 의 update 가 가능하게 되어 있으므로, 가변 프레임율에 따른 문제는 복호화 후 출력 타이밍과 관련된다. 고정 프레임율의 경우, 출력 타이밍을 비디오 플레이어의 초기화과정에서 결정할 수 있지만, 가변 프레임율의 경우에는 각 프레임의 출력타이밍을 확인해야 한다.

렌더링은 사용자 단말에서 복호화된 영상을 시간에 맞추어 디스플레이하는 역할을 한다. 따라서 복호화기는 렌더러에게 복호화 영상 뿐만 아니라 복호화된 영상의 POC(Picture Order Count)정보와 IDR 프레임 정보를 알려주어야 한다. POC 는 직전 IDR 프레임에서부터의 프레임 거리를 나타내는데, 렌더링은 POC 와 IDR 프레임 정보로부터 시간상의 프레임 위치를 계산하여 렌더링 해야 한다.

## IV. 실험

본 논문에서 제안한 방법을 JSVM 2.0 에서 구현하였다. 실험에서는 SVC 테스트 영상인 “Crew” 영상을 사용하였고, QCIF, CIF 의 2 레이어, 3 FGS 레이어, GOP 길이는 16 Frames, Intra 로 부호화한 GOP 키 프레임간의 간격은 1 GOP, 총 144 프레임을 부호화하였다. QCIF 와 CIF 에서의 최대 프레임율은 각각 15fps, 30fps 이다. 그림 3은 실험에 사용된 “Crew” 영상이다.

실험에서는 총 재생 시간 약 4 초 내에서 비트스트림 추출기에 입력 추출점 값을 바꿔가면서 실시간 추출기를 통해 추출된 비트스트림을 얻었다. 이 때 입력 추출점은 입력 비트스트림의 2GOP(32 프레임) 간격마다 변경하였다. 실험에서 확인 하고자 하는 것은 추출된 결과 비트스트림의 실제 비트율과 입력 추출점의 비트율과의 비교와 실제로 추출된 비트스트림이 스케일러빌리티가 변하면서 복호화되는지를 확인하는 것이다. 표 1 은 실험에 사용된 입력 추출점의 변경 시점에 관한 정보를 나타낸다. 표 2 의 실험 결과에서와 같이, 실제 추출된 비트스트림의

비트율이 입력 추출점의 비트율을 거의 따라가는 것을 확인 할 수가 있다. 즉 제안한 방법에 의한 실시간 가변 스케일러빌리티가 가능한 비트스트림 추출에 의해 네트워크의 가용 비트율이 가변하는 환경에서 충분히 유용함을 보여준다. 또한 복화화 결과는 프레임율과 영상의 화질이 가변하면서 정상적인 복호화가 이루어짐을 알 수 있다.

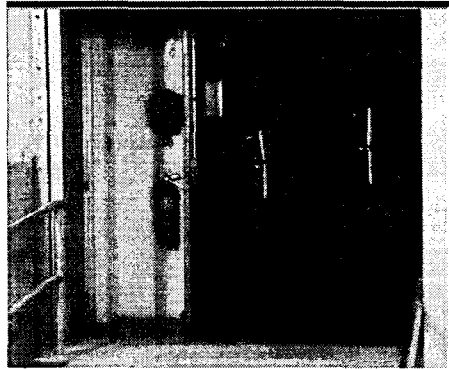


그림 3. "Crew" 테스트 비디오



(a)



(b)

그림 4. "Crew" 결과 비디오 : (a) 첫 번째 추출점 복호화 영상, (b) 두 번째 추출점 복호화 영상

표 1 시간에 따른 비트스트림 추출점변화

| 파라미터 변경구간   | 입력 추출점                    |
|-------------|---------------------------|
| 0 ~ 1th GOP | 352x288, 15프레임, 200 kbps  |
| 2 ~ 3th GOP | 352x288, 30프레임, 450 kbps  |
| 4 ~ 5th GOP | 352x288, 15프레임, 300 kbps  |
| 6 ~ 7th GOP | 352x288, 7.5프레임, 150 kbps |

표 2 각 테스트별 실험 결과

| 추출점 변경구간    | 입력비트율    | 추출 비트율   |
|-------------|----------|----------|
| 0 ~ 1th GOP | 200 Kbps | 198 Kbps |
| 2 ~ 3th GOP | 450 Kbps | 431 Kbps |
| 4 ~ 5th GOP | 300 Kbps | 313 Kbps |
| 6 ~ 7th GOP | 150 Kbps | 205 Kbps |

## V. 결론

본 논문에서는 SVC 비트스트림을 이용하여 시간적, 품질적 스케일러빌리티를 가변 시키면서 비트스트림을 실시간으로 추출할 수 있는 비트스트림 추출방법을 제안하였다. 실험을 통해 시간적으로 변하는 네트워크의 가용 비트율에 맞추어 비트스트림이 추출되는 것을 확인 할 수 있었으며, 이것은 비디오 스트리밍 서비스의 구현시 실제 가변적인 네트워크 환경에서 사용자에게 QoS 를 만족시는 중요한 기능을 할 수 있음을 보여준다. 그러나 현재 부호화 및 복호화기, 그리고 비트스트림 추출기는 실시간으로 공간 해상도 가변 비트스트림을 추출해 낼 수 없으므로 이에 대한 연구 및 표준모델의 수정이 요구된다.

## 참고문헌

- [1] ISO/IEC JTC 1/SC 29/WG 11, "Working Draft 2 of ISO/IEC 14496-10:2005/AMD1 Scalable Video Coding," N7086, Apr. 2005, Busan.
- [2] ISO/IEC JTC 1/SC 29/WG 11, "Joint Scalable Video Model (JSVM) 2.0 Reference Encoding Algorithm Description," N7084, Apr. 2005, Busan.
- [3] R. S. Prasad, M. Murray, C. Dovrolis, and K. Claff, "Bandwidth estimation: metrics, measurement techniques, and tools," *Trans. On IEEE Network*, Nov., Dec., 2003
- [4] 정용주, 김영석, Troung Cong Thang, 노용만, 김태희, 김재곤, "의미적 개념 기반 비디오 트랜스 코딩 방법 및 시스템," 한국방송공학회 학술대회논문집, pp. 59-63, Nov. 2004.