

H.264 SVC에서 비트 스트림 추출을 위한 공간과 시간 해상도 선택 기법

김남윤[†], 황호영[‡]

요 약

H.264 SVC(Scalable Video Coding)는 디스크 저장 공간 효율성과 높은 확장성을 제공할 수 있는 장점이 있다. 그러나 스트리밍 서버나 단말기는 비트 스트림을 효율적으로 추출해야 한다. 본 논문에서는 네트워크 가용 대역폭을 넘지 않으면서 최대의 PSNR을 얻기 위한 SVC 비트 스트림 추출 기법을 제시한다. 이를 위하여 오프라인시에 최대의 PSNR을 얻기 위한 추출 지점에 대한 정보를 획득한 후, 온라인시에 네트워크 가용 대역폭을 만족하는 비트 스트림의 공간/시간 해상도를 결정한다. 이러한 공간/시간 해상도 정보는 네트워크 가용 대역폭과 함께 비트 스트림 추출기의 입력 파라미터로 사용된다. JSVM 참조 소프트웨어를 활용한 실험을 통하여 본 논문에서 제시한 추출 기법이 높은 PSNR을 제공함을 증명하였다.

Spatial and Temporal Resolution Selection for Bit Stream Extraction in H.264 Scalable Video Coding

Namyun Kim[†], Hoyoung Hwang[‡]

ABSTRACT

H.264 SVC(Scalable Video Coding) provides the advantages of low disk storage requirement and high scalability. However, a streaming server or a user terminal has to extract a bit stream from SVC file. This paper proposes a bit stream extraction method which can get the maximum PSNR value while date bit rate does not exceed the available network bandwidth. To do this, this paper obtains the information about extraction points which can get the maximum PSNR value offline and decides the spatial/temporal resolution of a bit stream at run-time. This resolution information along with available network bandwidth is used as the parameters to a bit stream extractor. Through experiment with JSVM reference software, we proved that proposed bit stream extraction method can get a higher PSNR value.

Key words: H.264 SVC(H.264 확장 가능한 비디오 코딩), Bit Stream Extraction(비트 스트림 추출), Spatial/Temporal Resolution Selection(공간/시간 해상도 선택), Rate-Distortion Optimization(비트-왜곡율 최적화)

1. 서 론

네트워크 전송과 데이터 압축 기술의 발전으로 인해 인터넷을 통한 멀티미디어 전송이 대중화되고 있

* 교신저자(Corresponding Author) : 김남윤, 주소 : 서울 성북구 삼선동 2가 389(136-792), 전화 : 02)760-4348, FAX : 02)760-4347, E-mail : nykim@hansung.ac.kr

접수일 : 2009년 8월 10일, 수정일 : 2009년 11월 14일

완료일 : 2009년 12월 11일

다[1,2]. VoD, IPTV, VoIP 기술은 모두 멀티미디어 응용의 대표적인 예라고 할 수 있다. 이러한 멀티미디어 응용에서 등장하는 문제 중 하나는 네트워크이나 사용자 단말기의 다양성으로 인한 멀티미디어의

[†] 정회원, 한성대학교 정보시스템공학과 부교수

[‡] 정회원, 한성대학교 멀티미디어공학과 조교수

(E-mail : hyhwang@hansung.ac.kr)

* 본 연구는 2008년도 한성대학교 교내연구비 지원 과제임

적용 가능성이다. 단말기에 따라 CPU 처리 능력, 배터리, 디스플레이 해상도가 상이하고 네트워크 인프라는 초고속에서 모바일까지 서로 상이하며 동적으로 사용 대역폭이 변하는 특성이 있다. 따라서 이질적이고(heterogeneous) 동적인 환경에 적용할 수 있는 멀티미디어 압축 및 전송 기술은 점차 중요해지고 있다.

이질적인 환경에 적합한 멀티미디어 전송 기술의 해결책으로서 다양한 시나리오에 맞게 콘텐츠를 멀티 버전으로 준비하는 방법이 있다. 이러한 방법은 기술적인 측면에서 단순하지만 저장 장치의 낭비와 확장 가능성이 제한되는 단점이 있다. 한편, 다양한 계층을 포함하는 확장 가능한(scalable) 콘텐츠를 구성하여 네트워크이나 단말기의 상태에 따라 일부 계층을 제거하는 기법이 있다. 이 기법은 하나의 멀티 미디어 콘텐츠를 사용하여 다양한 요구 조건(고화질에서 저화질 콘텐츠)을 만족시킬 수 있는 것으로 MPEG-2 이래로 활발히 연구되어 왔다.

MPEG-2에서 제시된 확장 가능한 압축 기법은 기본(base) 계층과 향상(enhancement) 계층을 콘텐츠에 포함시켜 네트워크 대역폭이 작을 경우에는 기본 계층만을 전송하고 대역폭이 충분할 경우에는 기본 계층과 향상 계층을 모두 전송하는 방법이다. 코딩 기법으로는 공간 확장성(spatial scalability), 시간 확장성(temporal scalability), 화질 확장성(SNR scalability)을 제공한다. 그러나 MPEG-2에서의 확장 가능한 압축 기법은 코딩 효율성이 많이 떨어져 그다지 주목을 받지 못하였다. 그러나 최근 H.264/AVC의 확장으로서 SVC(Scalable Video Coding) 기술이 표준화되었는데[3-5] 이 기술은 단일 계층을 가진 H.264/AVC에 비해서 코딩 효율성을 저하시키지 않으면서 보다 높은 시간/공간/화질 확장성을 제공하는 장점이 있다.

한편, H.264 SVC를 통해 압축된 멀티미디어 스트림은 네트워크 대역폭을 고려하여 적절히 추출되어야 한다. 즉, 스트리밍 서버나 중계기에서는 네트워크 상태를 파악한 후, SVC 파일에서 서브스트림을 추출하여 전송하여야 한다. 또한 단말기에서는 디코딩하기 전에 해상도에 적합한 서브 스트림으로 추출된 후 디스플레이된다. 따라서 최근에 계층 간의 의존성이나 비트 스트림의 특성을 고려하여 비트 스트림을 추출하는 다양한 알고리즘이 제안되어 왔다

[6-8]. 관련 연구 [6]은 패킷의 중요도와 손실 가능 여부를 고려하여 SVC 계층의 대역폭을 계산한 후, 전송 채널과 가까운 계층을 선택하는 방법이지만 PSNR에 대한 고려가 부족하였다. 관련 연구 [7,8]에서는 주어진 목표 비트율과 공간/시간 해상도에서 최대의 PSNR을 얻기 위한 기법이 제시되었다. 그러나 SVC 비트 스트림은 다양한 추출 지점(extraction point)이 존재한다. 여기서 추출 지점은 공간 해상도/시간 해상도/화질 수준으로 표현된다. 예를 들어 낮은 네트워크 대역폭에서는 저화질로 구성하거나 낮은 프레임을 혹은 낮은 공간 해상도를 이용하여 서브 스트림을 생성할 수 있다. 따라서 SVC 스트림에서 PSNR을 최대화할 수 있는 최적의 추출 지점을 찾는 작업은 무엇보다 중요하다고 할 수 있다. 본 논문에서는 콘텐츠를 오프라인시에 분석하여 주어진 비트율 하에서 최대의 PSNR을 가지는 추출 지점에 대한 정보 즉, RD(rate-distortion) 정보를 유지한다. 그리고 온라인시 네트워크 사용 대역폭이 주어졌을 때, RD 정보를 활용하여 사용 대역폭보다 낮은 비트율을 가지는 추출 지점을 선택함으로써 최대의 PSNR을 얻을 수 있는 공간 해상도/시간 해상도를 결정한다. 이러한 정보는 비트 스트림 추출기의 파라미터로 입력되어 최적의 비트 스트림을 추출할 수 있다. 본 논문에서 제안한 알고리즘은 JSVM에서 제공하는 비트 스트림 추출기에 활용될 수 있으며 실험을 통해 성능이 향상됨을 보인다.

본 논문의 구성은 다음과 같다. 2절에서는 H.264 SVC의 개요와 JSVM 참조 소프트웨어에 대해 설명한다. 3절에서는 JSVM의 비트 스트림 추출기를 이용하여 추출된 서브 스트림의 특징을 분석한다. 4절에서는 RD 정보를 바탕으로 공간/시간 해상도를 결정하여 비트 스트림을 추출하는 과정을 설명한 후, 실험을 통해 성능을 분석한다. 마지막으로 5절에서는 향후 연구 및 결론에 대해 서술한다.

2. H.264 SVC 개요

2.1 확장 모드(Scalability Mode)

JVT(ITU-T의 VCEG과 ISO/IEC의 MPEG)는 확장 가능한 비트 스트림 생성을 위해 H.264/AVC의 확장으로 SVC 표준을 만들었다. SVC는 계층적인 구조에 기반하고 있으며 비트 스트림은 기본 계층

(base layer)과 하나 이상의 향상 계층(enhancement layer)으로 구성된다. 각 향상 계층은 비트 스트림의 품질을 개선하며 다음 세 가지 확장 모드에 의해 만들어진다: 시간적인 확장 모드(temporal scalability mode), 공간적인 확장 모드(spatial scalability mode), 화질 확장 모드(SNR scalability mode).

2.1.1 시간적인 확장 모드

B 프레임을 수직적인 구조로 구성한 후 특정 단계의 B 프레임을 추가/제거함으로써 비트 스트림의 재생율(frame rate)을 조정한다[9]. 그림 1의 X축에서 3.75 fps는 기본 계층을 형성하고 7.5/15/30 fps는 B 프레임의 추가를 통하여 향상 계층의 역할을 수행한다.

2.1.2 공간적인 확장 모드

기본 계층은 프레임의 낮은 공간 해상도를 제공하며 향상 계층을 통해 고해상도를 제공한다. 각 계층은 “motion-compensated prediction”과 “intraprediction”과 같은 “intralayer prediction”을 사용하고 계층간에는 코딩 효율을 높이기 위해 “interlayer prediction”을 사용한다[4]. 그림 1의 Y축은 기본 계층인 QCIF(176×144)와 향상 계층인 CIF(352×288), 4CIF(704×576) 해상도를 보여주고 있다.

2.1.3 화질 확장 모드

향상 계층을 통해 점진적으로 화질을 개선할 수 있으며 “coarse grain scalability(CGS)”와 “medium grain scalability(MGS)”가 존재한다[10]. MGS는 향상 계층을 좀 더 세밀하게 분류할 수 있는데, 하나의 공간/시간 해상도에 대해 최대 16개의 향상 계층을 제공한다. 그림 1의 Z축은 기본 계층과 3개의 향상 계층을 보여주고 있다.

SVC 비트 스트림은 연속적인 NAL(Network Abstraction Layer)로 구성된다. NAL은 헤더와 페이로드를 포함하고 있다. 한편, AU(Access Unit)은 NAL의 집합으로서 디코딩 후 정확히 하나의 픽쳐를 형성한다. 즉, AU는 한 시점에서 생성된 기본 계층과 향상 계층의 NAL을 포함한다. 각 NAL은 특정 공간/시간/화질 계층에 속한다. 이러한 정보는 NAL 헤더 중 dependency_id, temporal_id, quality_id 항목에 저장된다.

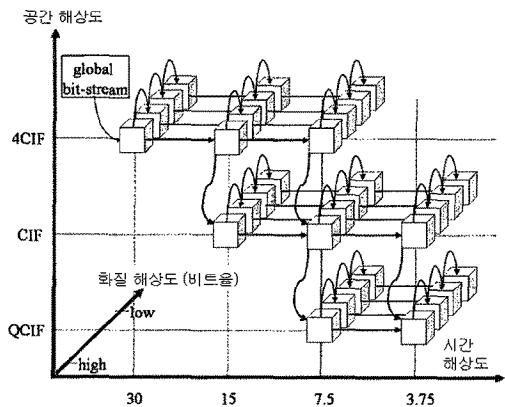


그림 1. SVC 확장 모드(시간/공간/화질)

2.2 JSVM 참조 소프트웨어

JSVM(Joint Scalable Video Model)에서는 SVC를 위한 참조(reference) 소프트웨어를 제작하였다[5]. 그림 2는 YUV 포맷의 비디오 스트림을 설정 파일(configuration file)을 이용하여 인코딩한 후, 목표 비트율에 맞게 비트 스트림을 추출하는 과정을 보여주고 있다. 그리고 추출된 비트 스트림의 품질을 측정하기 위해 디코딩한 후, 원 영상과의 PSNR 값을 계산한다.

SVC 표준안에서는 비트 스트림 추출 방법을 표준으로 정하고 있지 않아서 다양한 방법이 가능하다. 본 절에서는 JSVM의 비트 스트림 추출기(BitStreamExtractorStatic.exe) 알고리즘에 대해 간략히 서술한다.

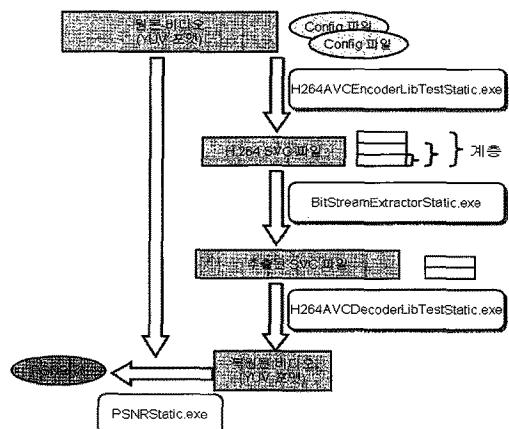


그림 2. JSVM 소프트웨어 사용 시나리오

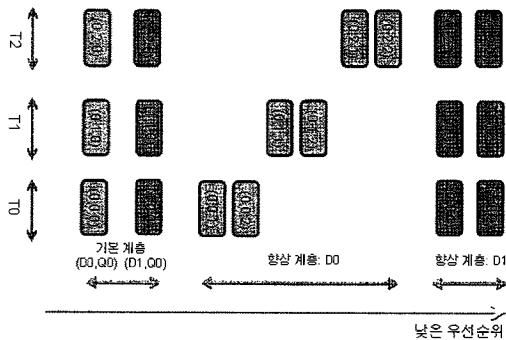


그림 3. JSVM에서 계층별 우선 순위

JSVM 추출기는 우선 순위에 기반하여 계층을 추출한다. 즉, 특정 계층 (D, T, Q)에 속한 NAL은 미리 정의된 순서로 정렬되어 있으며 목표 비트율에 도달될 때까지 순서대로 추출된다. 여기서 D는 공간적인 레벨, T는 시간적인 레벨, Q는 화질적인 레벨을 의미한다. 그림 3은 NAL의 우선 순위를 보여주고 있는데, 각 블록은 특정 계층에 속한 NAL을 의미한다. NAL 우선 순위는 다음과 같이 결정된다. 용광 혹은 사용자가 명시한 목표 공간/시간적인 해상도 보다 작거나 같은 공간/시간적인 해상도를 가진 화질 기본 계층 ($Q=0$)이 먼저 선택된다. 그리고 작은 공간 해상도를 가진 NAL에 대해 화질 향상 계층 ($Q=1, 2, \dots$)이 시간적인 레벨(T) 순서로 선택된다. 마지막으로 목표 공간 해상도를 가진 NAL에 대해 화질 향상 계층($Q=1, 2, \dots$)이 화질 향상 레벨(Q) 순서로 선택된다.

3. JSVM을 통해 추출된 SVC 비트 스트림 분석

본 절에서는 JSVM 소프트웨어(BitStreamExtractorStatic.exe)를 통해서 추출된 SVC 비트 스트림의 평균 비트율과 PSNR 값을 실험을 통해 분석한다. 사용된 소프트웨어의 버전은 JSVM 9.15 이다.

3.1 비디오 파일

실험 대상 비디오를 움직임의 정도에 따라 분류한 후, 대표적인 샘플 세 가지를 (CITY.yuv, FOREMAN.yuv, SOCCER.yuv) 선택하였다[11]. 각 비디오 파일은 모두 YUV 형식으로 300 프레임으로 구성되어 있

표 1. 비디오의 종류 및 특징

	CITY	FOREMAN	SOCCER
움직임	Low	Medium	High
프레임수	300	300	300
공간 해상도	CIF	CIF	CIF
크기(Byte)	45,619,200	45,619,200	45,619,200

으며 해상도는 CIF 포맷이다(표 1 참조).

3.2 인코딩 설정 파일

YUV 포맷의 비디오 파일은 인코딩 설정 파일을 통하여 SVC 스트림으로 변환된다. 사용된 인코딩 설정 파일의 구성은 다음과 같다. 먼저 공간 확장성을 위해 QCIF, CIF 해상도 2개의 계층을 제공하고, 화질 확장성을 위해 기본 계층과 3 개의 MGS 계층을 제공한다. 그리고 시간 확장성을 위해 1.875/3.75/7.5/15/30 fps 5개의 계층을 제공한다. 결국, 총 40개 ($=2*4*5$)의 계층을 가진 비트 스트림을 생성한다. 세 개의 비디오에 대한 인코딩 설정 파라미터는 모두 동일하며 그림 4는 CITY SVC 파일을 생성하기 위한 메인 설정 파일, QCIF/화질 기본 계층(Layer0.cfg), QCIF/화질 MGS 계층(Layer1.cfg), CIF/화질 기본 계층(Layer2.cfg), CIF/화질 MGS(Layer3.cfg) 설정 파일을 각각 보여주고 있다. 메인 설정 파일에서 인코딩 출력 파일의 이름은 CITY.264이며 GOP 크기는 16으로 설정되어 있다. Layer1.cfg와 Layer3.cfg에서 MGS 파라미터 벡터는 16개의 transform 계수의 수를 4, 4, 8개로 분류하고 있다. 계층 설정 파일에서 양자화 파라미터(QP) 값은 각각 32, 26, 34, 28로 설정되어 있다. 이러한 인코딩 설정 파일을 바탕으로 JSVM의 인코더(H.264AVCEncoderLbTestStatic.exe)를 이용하여 H.264 SVC 파일을 생성한다.

3.3 비트 스트림 추출기 파라미터 설정

본 논문에서는 JSVM 비트 스트림 추출기의 옵션으로서 “-e”를 사용하였다. 이 옵션은 사용자가 지정한 공간/시간 해상도와 네트워크 대역폭을 바탕으로 2.2절에서 설명한 알고리즘을 이용하여 비트 스트림을 추출한다. 아래는 -e 옵션의 형식을 보여주고 있다.

```
BitStreamExtractorStatic.exe input.svc output.svc -e D@T:B
```

```

# JSVM Main Configuration File
OutputFile CITY.264
FrameRate 30.0
FramesToBeEncoded 300
GOPSize 16
CgsSnrRefinement 1
...
NumLayers 4
LayerCfg layer0.cfg
LayerCfg layer1.cfg
LayerCfg layer2.cfg
LayerCfg layer3.cfg

# Layer0 Configuration File
InputFile CITY_QCIF30.yuv
QP 32

# Layer1 Configuration File
InputFile CITY_QCIF30.yuv
...
MGSVectorMode 1
MGSVector0 4
MGSVector1 4
MGSVector2 8
QP 26

# Layer2 Configuration File
InputFile CITY_CIF30.yuv
QP 34

# Layer3 Configuration File
InputFile CITY_CIF30.yuv
...
MGSVectorMode 1
MGSVector0 4
MGSVector1 4
MGSVector2 8
QP 28

```

그림 4. 실험에서 사용된 SVC 설정 파일 예제

실험에서는 공간 해상도 파라미터인 D는 $176 \times 144 / 352 \times 288$, 시간 해상도 파라미터인 T는 $7.5 / 15 / 30$, 그리고 네트워크 대역폭 파라미터인 B는 $100 / 150 / \dots / 850 / 900$ (kbps)값을 이용하였다.

3.4 PSNR 값 계산

원 영상에 비해 추출된 비트 스트림의 품질을 평가하기 위해 PSNRStatic 도구를 사용한다. 이 때 추출된 영상은 공간/시간 해상도가 원 영상보다 작을 수 있으므로 해상도를 확장할 필요가 있다. 먼저 공간 해상도의 업 샘플링은 JSVM DownConvertStatic을 이용하면 업 샘플링을 수행할 수 있다. 한편, 시간 해상도의 업 샘플링은 JSVM 도구에서 제공되지 않기 때문에 자체적으로 제작하였다. 이 프로그램은 이전 팩터를 복사하여 삽입함으로써 시간적인 업 샘플링을 수행한다. 예를 들어 3.75 fps를 30 fps로 확장하기 위해서는 하나의 팩터를 7번 복사하여 추가한다. 공간/시간 해상도를 업 샘플링한 후 PSNRStatic 도구를 이용하여 원 영상과 업샘플된 영상을 비교하여 PSNR 값을 계산한다.

3.5 실험 결과 및 분석

고정된 공간/시간 해상도 하에서 네트워크 대역폭 B의 값을 100에서 900까지 50 단위로 변경하면서 비트 스트림을 추출한 후, 생성된 스트림의 비트율과 PSNR 값을 얻었다. 그림 5는 세 개의 비디오에 대한 결과이며 각 점은 [D, T, B]에 해당되는 추출 지점을 의미하며 x축과 y축은 추출 지점에서 생성된 스트림의 비트율과 PSNR 값을 보여주고 있다.

예를 들어 SOCCER 영상에서 고정된 D(= CIF: 352×288)와 T(= 15 fps)에서 B의 값을 100~900

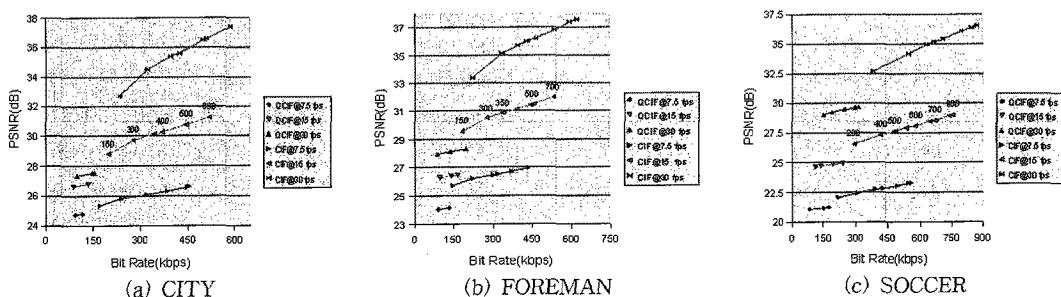


그림 5. JSVM 비트 스트림 추출기를 이용하여 추출한 스트림의 비트율과 PSNR

표 2. SOCCER SVC의 계층 우선순위: D=CIF, T=15 fps

화질 기본 계층: (높은 우선순위) QCIF, CIF	D=0, T=0~3, Q=0 D=1, T=0~3, Q=0	(0,0,0), (0,1,0), (0,2,0), (0,3,0) (1,0,0), (1,1,0), (1,2,0), (1,3,0)
화질 향상 계층: (중간 우선순위) QCIF	D=0, T=0, Q=1~3 D=0, T=1, Q=1~3 D=0, T=2, Q=1~3 D=0, T=3, Q=1~3	(0,0,1), (0,0,2), (0,0,3) (0,1,1), (0,1,2), (0,1,3) (0,2,1), (0,2,2), (0,2,3) (0,3,1), (0,3,2), (0,3,3)
화질 향상 계층: (낮은 우선순위) CIF	D=1, T=0~3, Q=1 D=1, T=0~3, Q=2 D=1, T=0~3, Q=3	(1,0,1), (1,1,1), (1,2,1), (1,3,1) (1,0,2), (1,1,2), (1,2,2), (1,3,2) (1,0,3), (1,1,3), (1,2,3), (1,3,3)

(kbps)까지 변화했을 때 JSVM 추출기에 의해서 추출된 스트림의 비트율과 PSNR 값은 그림 5의 (C)에서 오른쪽 가운데 실선으로 표시되어 있다. 실선위의 숫자는 네트워크 대역폭 B를 의미한다. 비록 D, T가 고정되어 있지만 B의 값에 따라 화질 계층을 추가되어 어 비트율과 PSNR 값이 변하게 된다. B의 값에 따른 계층의 추가는 그림 3의 계층 간의 우선순위에 의해 좌우되는데, CIF/15 fps의 경우에 JSVM의 추출 알고리즘에 의해 결정되는 계층의 우선 순위는 표 2와 같다. 여기서 계층은 (D,T,Q)로 표현된다. D는 0 (=QCIF), 1(=CIF1), T는 0(=1.875 fps), 1(=3.75 fps), 2(=7.5 fps), 3(=15 fps), 4(=30 fps), Q는 0(=기본 계층), 1(=MGS0 계층), 2(=MGS1 계층), 3(=MGS2 계층)의 값을 가진다.

추출기는 표 2와 같은 우선 순위를 바탕으로 네트워크 대역폭 B를 고려하여 비트 스트림을 추출한다. 표 3은 B의 값에 따라 추출되는 스트림의 계층 수, 비트율, 그리고 PSNR 값을 보여주고 있다. B=100, 200일 경우, Q=0인 모든 계층인 (0,0,0)에서 (1,3,0) 까지의 8개의 계층이 포함된다. 그런데 추출된 스트림의 비트율은 295.5 kbps로서 B 보다 크다. 그 이유는 그림 3과 같이 명시된 D, T의 해상도 즉, CIF/15fps와 작거나 같은 공간/시간 해상도를 가지는 화질 기본 계층을 모두 포함하기 때문이다. 한편, B=300, 400 일 경우 (0,0,0)에서 (0,3,3)까지 총 20개

의 계층이 포함되는데, 그 이유는 낮은 해상도 (QCIF)를 가지는 화질 향상 계층 (0,0,1)에서 (0,3,3)을 모두 포함하기 때문이다. 이 때 생성된 스트림의 비트율은 420.8 kbps이며 PSNR 값은 27.32로 증가하게 된다.

JSVM 추출기에 의해 추출된 스트림의 특징은 다음과 같이 요약할 수 있다.

- 추출된 스트림의 비트율은 사용자가 지정한 비트율 B를 초과할 수 있다. 예를 들어 표 3에서 B=100 인 경우 생성된 스트림은 295.5 kbps이다. 즉 CIF/15 fps와 작거나 같은 공간/시간 해상도를 가지는 화질 기본 계층을 모두 포함하므로 비트율 B를 초과할 수 있다. 따라서 목표 비트율 B를 초과하지 않도록 공간/시간 해상도를 주의깊게 선택하여야 한다.

- 비디오의 특성에 따라 공간/시간 해상도를 적절히 선택해야 한다. 그림 5에서 CITY 비디오는 D 와 T를 QCIF/30fps 로 설정했을 경우 항상 CIF/15fps보다 PSNR 값이 작지만 SOCCER의 경우에는 QCIF/30 fps보다 CIF/15 fps가 PSNR이 높다. 그 이유는 SOCCER의 경우 움직임이 많아서 공간 해상도보다 시간 해상도가 더욱 중요하기 때문이다. 따라서 각 비디오 별로 특정 비트율에서 최대 PSNR을 가지는 공간/시간 해상도를 유지할 필요가 있다.

표 3. SOCCER SVC에서 스트림 추출 결과: D= CIF, T= 15 fps

B(kbps)	100	200	300	400	500	600	700	800
Layer 개수	8	8	20	20	21	24	28	32
bit rate(kbps)	295.5	295.5	420.8	420.8	485.4	581.4	672.1	755.3
PSNR(dB)	26.54	26.54	27.32	27.32	27.64	28.06	28.52	28.98

4. 비트-왜곡율(Rate-Distortion) 최적화를 위한 비트 스트림 추출 알고리즘

비디오 스트림의 데이터 비트율이 네트워크 가용 대역폭을 초과할 때 일부 비디오 패킷을 제거하여야 한다. 이를 위해 본 논문에서는 오프라인시에 그림 5와 같이 추출 지점에서의 비트율과 PSNR 값을 계산한다. 그리고 주어진 비트율에서 최대의 PSNR 값을 가지는 추출 지점만을 선택하여 파일에 저장한다. 온라인시에는 네트워크 대역폭을 초과하지 않는 추출 지점을 선택함으로써 최대의 PSNR을 가지는 서브스트림을 생성한다. 그림 6은 본 논문에서 제시하는 비트 스트림 추출 과정을 보여주고 있다. 현재 가용한 네트워크 대역폭 B 가 주어졌을 때 공간/시간 해상도 선택 알고리즘은 컨텐츠별 RD(rate-distortion) 정보를 활용하여 D , T 파라미터를 결정한다. 컨텐츠별 RD 정보는 주어진 비트율에서 최대의 PSNR 값을 얻을 수 있는 공간/시간 해상도에 대한 정보로서 오프라인시 분석을 통하여 생성되며 구성 요소는 [공간 해상도, 시간 해상도, 스트림의 데이터 비트율, PSNR 값]이다. 그리고 JSVM 추출기는 이러한 파라미터를 활용하여 SVC 비트 스트림을 생성한다.

4.1 공간/시간 해상도 선택 알고리즘

본 절에서는 컨텐츠별 RD 정보의 생성 과정을 설명하고 이를 바탕으로 공간/시간 해상도 선택 알고리즘에 대해 설명한다.

컨텐츠별 RD 정보는 각 비트율에서 최대의 PSNR을 가지는 추출 지점의 집합으로서 다음과 같이 생성된다. 그림 5에서 각 점들은 RD(rate-distortion) 포인트로서 추출 지점(extraction point)을 나타내고 있

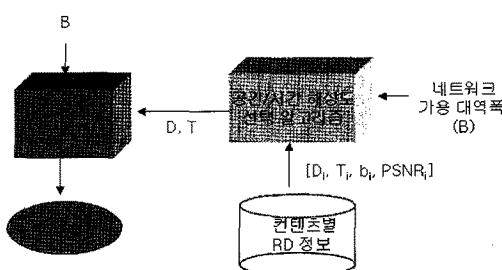


그림 6. SVC 비트 스트림 추출 과정

으며 이러한 추출 지점은 파라미터 $[D, T, B]$ 를 입력으로 추출기를 통해 결정되는 계층들의 집합으로서 생성된 서브스트림은 비트율과 PSNR 속성을 가진다. 이 때 고려할 점은, 비트율은 네트워크 가용 대역폭을 초과하지 않아야 하며, PSNR 값은 최대화하여야 한다. 따라서 비디오 품질을 개선하기 위해서는 추출 지점을 하기 위해서하여야 한다. 예를 들어, 그림 7에서 네트워크 가용 대역폭이 약 310 kbps일 때 선택할 수 있는 추출 지점은 다음과 같다: 1) 일부 화질 계층을 포함한 CIF 해상도/7.5 fps, 2) CIF 해상도/15 fps, 3) 일부 화질 계층을 포함한 QCIF 해상도/30fps. 이 때 최대의 PSNR 값을 얻기 위해서는 세 번째 추출지점을 선택하여야 한다.

컨텐츠별 추출 지점을 p_i 로 정의하고 추출 지점의 집합을 S 라고 정의한다. 그리고 p_i 에서 요구되는 데이터 비트율을 b_i 라고 하고 B 를 네트워크 가용 대역폭이라고 하자. 네트워크 가용 대역폭 B 가 주어졌을 때 최대의 PSNR을 가지는 추출 지점 p^* 은 다음과 같이 정의된다.

$$p^* = \arg_{p_i \in S} \max PSNR_i \\ \text{subject to } b_i \leq B$$

p^* 은 그림 7에서 긁은 선으로 연결된 것으로 추출 지점의 외곽 곡선을 따라 가면서 쉽게 획득될 수 있다. 본 논문에서는 이러한 특성을 가지는 추출 지점을 오프라인시에 찾아서 RD 정보를 구성한다. 추출 지점 p_i 가 주어졌을 때 RD 정보는 $[D_i, T_i, b_i, PSNR_i]$ 형식을 가진다. 예를 들어 SOCCER 비디오의 RD 정보의 내용은 다음과 같다.

[1 4 870.38 36.53],	[1 4 845.84 36.40],
[1 4 796.78 36.06],	[1 4 707.70 35.40],
[1 4 664.06 35.15],	[1 4 633.45 34.94],
[1 4 547.12 34.18],	[1 4 380.28 32.63],
[0 4 314.90 29.68],	[0 4 296.33 29.63],
[0 4 246.62 29.45],	[0 4 186.41 29.22],
[0 4 148.06 29.00],	[0 3 141.68 24.75],
[0 3 114.90 24.67],	[0 2 84.08 21.10]

스트리밍 서버는 각 멀티미디어 컨텐츠를 분석하여 RD 정보는 유지하며 전송시에 RD 정보를 활용하여 SVC 스트림을 추출한다. 만약 네트워크 중계

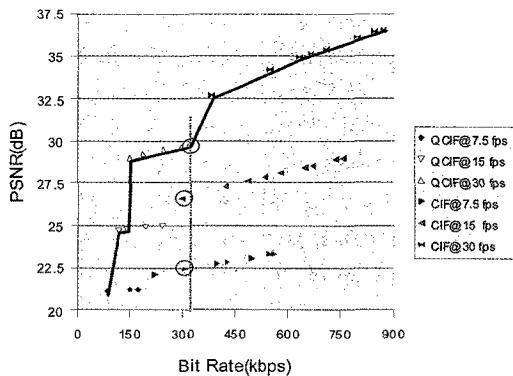


그림 7. 주어진 비트율하에서 최대의 PSNR 값을 가지는 RD 포인트(SOCCER)

기나 단말기에서 스트림을 추출할 필요가 있을 시에는 SVC 파일과 함께 RD 정보가 전송된다. 일반적으로 RD 정보는 작은 크기로서 전송 오버헤드는 매우 작다. RD 정보를 이용한 공간/시간 해상도 알고리즘의 동작 과정은 다음과 같다.

- Step 1: RD 정보 $[D_i, T_i, b_i, \text{PSNR}_i]$ 을 b_i (혹은 PSNR_i 값)에 따라 정렬한다.
- Step 2: 네트워크 가용 대역폭 B가 주어졌을 때 이 값보다 작은 b_i 를 갖는 RD 정보를 찾는다.
- Step 3: RD 정보에서 D_i, T_i 를 결정한 후, $[D_i, T_i, B]$ 를 BitStreamExtractorStatic의 파라미터로 설정한다.

4.2 성능 분석

공간/시간 해상도 선택 알고리즘을 이용한 비트 스트림 추출기의 성능을 분석하기 위해서 아래의 두

개의 옵션을 사용한다. PropAlgo에서는 공간/시간 해상도 선택 알고리즘을 통해서 나온 파라미터를 바탕으로 “-e” 옵션을 사용하며 BasicAlgo는 네트워크 가용 대역폭을 입력으로 “-b” 옵션을 사용하였다. “-b” 옵션은 주어진 가용 대역폭 B보다 작거나 같은 서브스트림을 추출하는데 사용된다.

PropAlgo: BitStreamExtractorStatic input.svc
output.svc -e D@T:B

BasicAlgo: BitStreamExtractorStatic input.svc
output.svc -b B

그림 8은 네트워크 대역폭을 100 kbps에서 900 kbps 까지 변화시켰을 때 추출된 스트림의 비트율과 PSNR 값을 보여주고 있다. BasicAlgo에서는 비디오의 특성에 따른 PSNR 값을 고려하지 않고 단지 비트율만 고려하여 추출하기 때문에 PSNR 값이 급격하게 변하는 것을 알 수 있다. 그러나 PropAlgo에서는 비트율이 증가함에 따라 PSNR 값이 단조 증가하며 항상 BasicAlgo보다 높은 PSNR 값을 제공하는 것을 알 수 있다.

5. 결 론

본 논문에서는 주어진 비트율하에서 최대의 PSNR 값을 가지는 RD(rate-distortion) 포인트에 대한 정보를 바탕으로 공간/시간 해상도 선택 기법을 제안하였다. 본 논문에서 제안한 알고리즘은 JSVM에서 제공하는 비트 스트림 추출기의 파라미터로 활용 가능하다. 그리고 실험을 통하여 본 논문에서 제안한 기법이 높은 PSNR을 제공하면서 안정적인 성능 향상을 나타내는 것을 보였다. 향후 연구 과제로

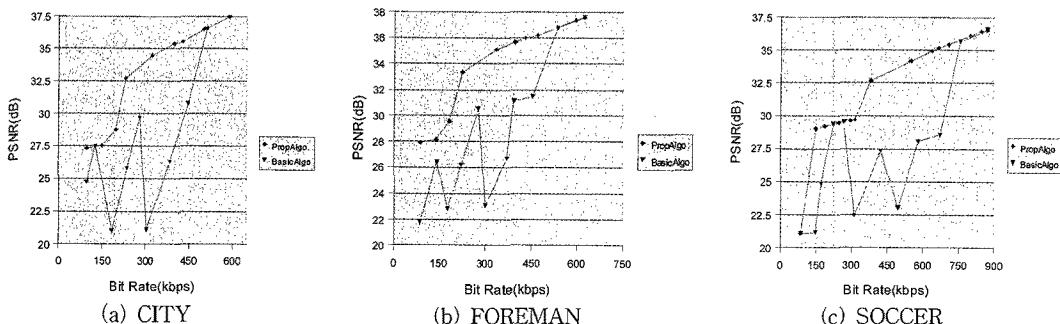


그림 8. RD 성능 비교: 제안한 방식 vs JSVM 기본 추출기

서 네트워크 오류로 인한 패킷 손실이 PSNR에 끼치는 영향을 분석할 예정이다.

참 고 문 헌

- [1] Dapeng Wu, Y. T. Hou, Wenwu Zhu, Ya-Qin Zhang and Peha, J.M., "Streaming Video Over the Internet: Approaches and Directions," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.11, No.3, Mar. 2001.
- [2] Hufang Sun, Anthony Vetro and Jun Xin, "An Overview of Scalable Video Streaming," *Wireless Communications and Mobile Computing*, Vol.7, No.2, Jan. 2007.
- [3] T. Wiegand, G. Sllivan, J. Reichel, H. Schwarz and M. Wien, "Joint Draft ITU-T Rec. H.264 - ISO/IEC 14496-10/ Amd.3 Scalable Video Coding," 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 Transactions on Circuits and Systems for Video Technology*, Vol.17, No.9, 2007.
- [5] HHI, SVC reference software http://ip.hhi.de/imagecom_G1/savce/downloads/SVC-Reference-Software.htm.
- [6] 류은석, 이정환, 유혁, "다중 경로 네트워크에서 H.264 SVC에 기반한 비디오 스트리밍 추출 및 전송 기법," 정보과학회 논문지, 정보통신 제35권 6호, 2008.
- [7] I. Amonou, N. Cammas, S. Kervadec, and S. Pateux, "Optimized rate-distortion extraction with quality layers in the scalable extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technology*, Vol.17, No.9, 2007.
- [8] E. Maani and A. K. Katsaggelos, "Optimized Bit Extraction Using Distortion Estimation in the Scalable Extension of H.264/AVC," *IEEE International Symposium on Multimedia (ISM)*, Berkeley CA, Dec. 2008.
- [9] H. Schwarz, D. Marpe, and T. Wiegand, "Hierarchical B pictures," Joint Video Team, Doc. JVT-P014, July 2005.
- [10] H. Kirchhoffer, H. Schwarz, and T. Wiegand, "CE1: Simplified FGS," Joint Video Team, Doc. JVT-W090, Apr. 2007.
- [11] <ftp.tnt.uni-hannover.de/pub/svc/testsequences>.



김 남 윤

1992년 2월 서울대학교 컴퓨터 공학과 졸업
 1994년 2월 서울대학교 컴퓨터 공학과 석사
 2000년 2월 서울대학교 컴퓨터 공학과 박사
 1999년 9월 ~ 2002년 2월 삼성전자 무선사업부 책임연구원
 2002년 ~ 현재 한성대학교 정보시스템공학과 부교수
 관심분야 : 멀티미디어 통신, 웹 검색, 모바일 통신 및 응용



황 호 영

1989년 ~ 1993년 서울대학교 컴퓨터공학과 공학사
 1993년 ~ 1995년 서울대학교 컴퓨터공학과 석사
 1995년 ~ 2003년 서울대학교 전기 컴퓨터공학부 공학박사
 2003년 ~ 2007년 안양대학교 디지털미디어학부 조교수
 2007년 ~ 현재 한성대학교 멀티미디어공학과 조교수
 관심분야 : 정보통신, 무선 및 이동통신망, 센서네트워크, 멀티미디어시스템 등