

네트워크 혼잡 제어를 위한 H.264/SVC 스트림의 계층 선택 알고리즘

김남윤[†], 황기태^{††}

요 약

H.264/SVC는 기본 계층과 하나 이상의 향상 계층으로 구성된 확장 가능한 비디오 스트림을 제공한다. 따라서 일부 계층을 제거함으로써 네트워크 상황에 따라 효율적으로 대처할 수 있는 장점이 있다. 그러나 인터넷과 같은 동적인 환경에서 네트워크 혼잡으로 인한 임의의 패킷 손실은 SVC 스트림의 화질에 치명적인 영향을 줄 수 있다. 따라서 네트워크 혼잡을 피하기 위해서는 효율적으로 스트림 계층을 선택하여 전송율을 조정하여야 한다. 본 논문에서는 제한된 대역폭을 가진 네트워크 노드에서 스트림의 비트율-왜곡(rate-distortion) 특성을 이용하여 네트워크 혼잡을 피할 수 있는 세 가지 계층 선택 알고리즘을 제시한다. 그리고 시뮬레이션을 통해 본 논문에서 제안한 FS(Far-Sighted) 알고리즘이 스트림의 특성을 효율적으로 이용함으로써 전체 스트림의 PSNR 값을 최대화할 수 있음을 보인다.

Layer Selection Algorithms of H.264/SVC Streams for Network Congestion Control

Namyun Kim[†], Kitae Hwang^{††}

ABSTRACT

H.264/SVC provides scalable video streams which consist of a base layer and one or more enhancement layers. Thus, it can efficiently adapt encoded streams to individual network conditions by dropping some layers of bit streams. However, on a dynamic environment such as the Internet, random packet losses due to network congestion can cause drastic effect on SVC quality. To avoid network congestion, the rate of video streams should be adjusted by carefully selecting a layer of each stream. In this paper, we propose three layer selection algorithms which can avoid network congestion by using the rate-distortion characteristics of streams. Simulation results show that FS(Far-Sighted) algorithm can maximize the overall PSNR value of streams by efficiently using the characteristics of video streams.

Key words: H.264/SVC Streaming, Congestion Control(혼잡 제어), Layer Selection(계층 선택), Rate-Distortion Optimization(비트율-왜곡 최적화)

1. 서 론

멀티미디어 앱축과 네트워크 전송 기술의 발전으로 멀티미디어 스트리밍 서비스가 점차 확대되고 있

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

접수일 : 2010년 9월 28일, 수정일 : 2010년 11월 26일

다[1,2]. 비디오 컨퍼런스, 디지털 TV 방송, Video On Demand 서비스가 사용되고 있으며 앞으로 스마트 TV의 등장으로 양방향 디지털 방송이 활성화될 것으로 기대된다. 그런데, 스트리밍 서비스 환경은

완료일 : 2010년 12월 6일

[†] 정회원, 한성대학교 정보시스템공학과 부교수

^{††} 정회원, 한성대학교 컴퓨터공학과 교수

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

* 본 연구는 한성대학교 교내연구비 지원 과제임.

다양한 네트워크 인프라와 단말기로 인해 이질적인 속성을 띠고 있다. 즉, 3G 이동 통신망, WiFi, 초고속 인터넷과 같은 다양한 네트워크가 존재하고 스마트 폰, 태블릿 PC, 노트북과 같은 다양한 단말기가 존재하고 있다. 따라서 이질적인 환경에서 다양한 고객의 요구를 만족시킬 수 있는 스트리밍 기술은 필수적이라고 할 수 있다.

H.264의 확장으로서 JVT(Joint Video Team)에 의해 표준화된 SVC(Scalable Video Coding)는 압축 크기의 효율성과 더불어 확장 가능성을 제공하고 있다[3]. 비디오 스트림을 기본 계층(base layer)과 하나 이상의 향상 계층(enhancement layer)으로 인코딩할 수 있기 때문에, 동적인 상황에 적절한 계층을 선택할 수 있다. 특히 비디오 스트림들이 요구하는 대역폭이 가용 네트워크 대역폭보다 큰 혼잡 상황에서 단말기의 해상도나 스트림의 특성을 파악하여 일부 계층을 제거함으로써 점진적인 화질 저하를 유도할 수 있다. 그런데 SVC 스트림의 계층은 서로 의존성이 존재하고 계층별로 화질에 끼치는 영향이 상이 하므로 주의깊게 선택되어야 한다[4].

최근 계층간의 의존성이나 비트 스트림의 특성을 고려하여 비트 스트림을 추출하는 다양한 알고리즘이 제안되어 왔다[5-7]. 이러한 추출 알고리즘은 최적의 화질을 얻기 위해 하나의 비디오 스트림에 대해 추출 지점(extraction point)을 선택한다. 여기서 “추출 지점”이란 특정 공간/시간/SNR 해상도를 의미한다. 한편, 네트워크상의 중간 노드에서 다수의 비디오 스트림이 존재할 때, 혼잡으로 인한 스트림의 영향을 줄이는 다양한 기법이 제시되었다[8-10]. Dyahadray [8]는 공간/시간/SNR 계층의 특징을 이용한 계층간 스위칭 기법을 제시하였으며 Tappayuthpi-jarn [9]은 TFRC에 기반한 계층 제거 기법을 제시하여 패킷 손실과 지연 시간을 감소시켰다. 그러나 위의 두 가지 연구는 RD(rate-distortion) 관점에서 최적의 화질을 제공하기 위한 계층 선택 알고리즘을 제시하지는 못하였다. Mahalingam [10]은 가용 네트워크 대역폭하에서 높은 화질을 제공하기 위해 비디오 스트림의 추출 지점을 선택하는 알고리즘을 제시하였다. 이 기법은 현재의 추출 지점에서 다음 추출 지점을 선택할 때 증가하는 유털리티 값(PSNR 증가율/Rate 증가율)을 기반으로 정렬한 후 최대의 유털리티 값을 가지는 스트림 계층을 선택하기 때문에

효율적인 계층 선택 기법을 제공하고 있다. 그러나 다음에 이웃한 추출 지점만을 고려함으로써 최적의 화질을 제공하지 못하는 단점을 가지고 있다.

본 연구는 이전 연구[7]에서 제시한 단일 스트림의 계층 선택 알고리즘의 확장으로서 네트워크 혼잡 발생시 점진적인 화질 저하를 유도하기 위한 다중 스트림의 계층 선택 기법을 제시한다. 오프 라인 시 주어진 비트율에서 최대의 PSNR을 얻을 수 있는 추출 지점에 대한 정보를 획득한 후, 온라인 시 가용 네트워크 대역폭내에서 전체 스트림의 PSNR의 합을 최대화할 수 있도록 각 비디오 스트림의 계층을 선택한다. 본 알고리즘을 통해 선정된 계층에 의존하지 않는 NAL(Network Abstraction Layer) 패킷은 혼잡 발생시 제거된다.

본 논문의 구성은 다음과 같다. 2절에서는 비디오 스트리밍 시스템 모델과 설계 목표에 대해 설명하고 3절에서는 H.264/SVC에 대한 개요와 스트림의 비트율-왜곡(RD: rate-distortion) 특성에 대해 서술한다. 4절에서는 혼잡 제어를 위한 세 가지 계층 선택 알고리즘을 제안하고 5절에서는 알고리즘의 성능을 분석한다. 마지막으로 6절에서는 향후 연구과 함께 결론을 맺는다.

2. 혼잡 제어를 위한 비디오 스트리밍 시스템 구조

2.1 시스템 모델

그림 1은 인터넷을 통한 비디오 스트리밍 시스템의 구조를 보여주고 있다. 서버 노드의 데이터베이스에는 최고의 해상도로 인코딩된 SVC 스트림과 비디오의 추출 지점에 대한 정보를 가진 “RD 정보”가 저장되어 있다. RD(rate-distortion) 정보는 주어진 비트율에서 가장 높은 PSNR을 제공하는 추출 지점에 대한 정보를 담고 있는 헌트로서 계층 선택 알고리즘에서 사용된다. 비트 스트림 추출기(extractor)는 SVC 스트림을 바탕으로 서브스트림을 추출하며 스트리밍 서버는 서브 스트림과 RD 정보를 인터넷을 통해 전송한다.

중간 노드(intermediate node)에서는 비디오를 수신한 후, 가용한 네트워크 대역폭에 맞춰 비디오 스트림의 전송율을 재조정한다. 일반적으로 네트워크 혼잡시 일부 패킷이 손실되는데, 이러한 경우 수신자

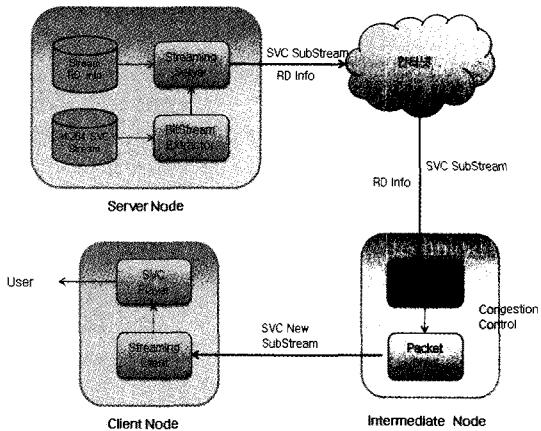


그림 1. 스트리밍 시스템의 전체 구조: 본 논문의 주요 대상은 중간 노드의 계층 선택(Layer Selection) 모듈이다.

측에서 에러 은폐(error concealment) 기법을 통해 비디오 화질을 보정할 수 있다. 그러나 SVC 스트림의 경우에는 계층간의 의존성으로 인해 일부 계층의 손실이 다른 계층의 디코딩을 불가능하게 할 수 있다. 따라서 중간 노드에서는 비디오 스트림의 계층을 재선택하여 전송율을 조정하여야 한다. 중간 노드의 계층 선택(layer selection) 모듈은 스트림의 RD 정보를 이용하여 전체 스트림의 PSNR 값을 최대화할 수 있도록 스트림의 계층을 선택한다. 그리고 패킷 제거(packet drop) 모듈은 선택된 계층에 의존적이지 않는 계층의 패킷, 즉, 상대적으로 덜 중요한 패킷을 제거한 후 클라이언트에게 전송한다.

클라이언트 노드는 중간 노드에서 변경된 새로운 서브 스트림을 수신한 후, SVC 플레이어를 이용하여 비디오를 재생한다.

2.2 스트림의 계층 선택 알고리즘의 목표

네트워크의 중간 노드에 유입되는 n 개의 비디오 스트림이 존재하고 노드의 네트워크 대역폭을 R_{NB} 이라고 하자(그림 2 참조). 비디오 스트림 k 의 새로운 계층이 요구하는 전송율을 R_k ($k=1,2,\dots,n$)이라고 하고 총 합을 R_{out} 이라고 정의한다. 이 때 $R_{out} = \sum_{k=1}^n R_k$ 이다. 그리고 비디오 k 의 PSNR은 $PSNR_k$ 이며 허용 가능한 품질을 보장하기 위한 최소 PSNR 요구 조건을 $PSNR_{min}$ 으로 정의한다.

결국, 시스템의 제약 조건은 $(R_{out} \leq R_{NB})$ 와 $(PSNR_k \geq PSNR_{min} \text{ for each video stream } k)$ 라고 표현할

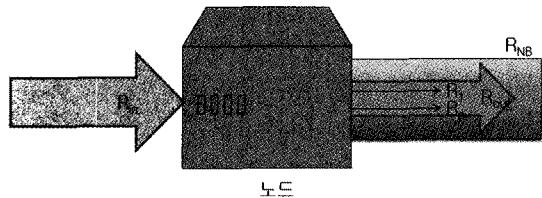


그림 2. n 개의 스트림을 가진 노드: R_{out} 은 R_{NB} 보다 크지 않아야 혼잡을 피할 수 있다

수 있다. 그리고 계층 선택 알고리즘의 목표는 전체적인 화질을 높이기 위해 모든 스트림의 PSNR 값의 합을 최대화하는 것으로 정의한다. 이를 수식으로 표현하면 다음과 같다.

$$\begin{aligned} & \text{maximize} \left(\sum_{k=1}^n PSNR_k \right) \\ & \text{subject to } (\sum_{k=1}^n R_k \leq R_{NB}) \text{ and } (PSNR_k \geq PSNR_{min} \text{ for each stream } k) \end{aligned} \quad (1)$$

3. H.264/SVC 스트림 분석

3.1 H.264/SVC 개요

H.264/SVC는 계층적인(layered) 구조에 기반하고 있으며 기본 계층과 하나 이상의 향상 계층으로 구성된다. 각 향상 계층은 비트 스트림의 화질을 개선하며 공간, 시간, SNR 확장 모드를 지원한다[11,12]. 그림 3에서 큐브는 특정 공간/시간/SNR 해상도를 가지는 계층을 의미하며, 최대의 공간/시간/SNR 해상도를 가지는 “global bitstream”에서 출발하여 추

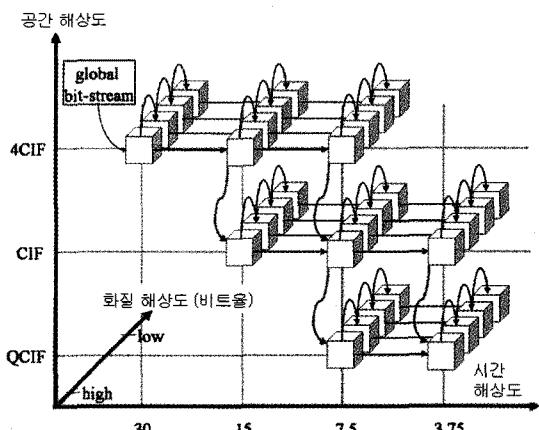


그림 3. SVC 확장 모드(공간/시간/SNR 해상도)

출 경로를 따라 다양한 서브 스트림이 생성된다.

SVC 스트림의 계층은 (D, T, Q)의 ID로 표현될 수 있는데, D는 공간적인 레벨, T는 시간적인 레벨, Q는 SNR 레벨을 의미한다. 그리고 SVC 스트림은 네트워크 전송을 위하여 NAL(Network Abstraction Layer)의 집합으로 표현된다. NAL은 헤더와 페이로드로 구성되어 있으며, 헤더는 (D, T, Q)에 대한 정보와 우선 순위에 대한 정보를 포함하고 페이로드에는 스트림 코드를 포함하고 있다. 한편, AU(Access Unit)은 NAL의 집합으로서 디코딩 후 정확히 하나의 팩쳐를 형성한다. 즉, AU는 한 시점에서 생성된 기본 계층과 향상 계층의 NAL을 포함한다.

SVC 표준안에서는 비트 스트림 추출 방법을 표준화하지 않지만, JVT에서는 JSVM(Joint Scalable Video Model)을 정의하고 비트 스트림 추출기 소프트웨어를 제공하고 있다[13]. JSVM 추출기(BitStreamExtractorStatic)는 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. 그림 4는 NAL의 우선 순위를 보여주고 있는데, 각 블록은 특정 계층에 속한 NAL을 의미한다. NAL 우선 순위는 다음과 같이 결정된다. 응용 혹은 사용자가 명시한 목표 공간/시간적인 해상도보다 작거나 같은 공간/시간적인 해상도를 가진 SNR 기본 계층 ($Q=0$)이 먼저 선택된다. 그리고 작은 공간 해상도를 가진 NAL에 대해 SNR 향상 계층($Q=1, 2, \dots$)이 시간적인 레벨(T) 순서로 선택된다. 마지막으로 목표 공간 해상도를 가진 NAL에 대해 SNR 향상 계층($Q=1, 2, \dots$)이 SNR 향상 레벨(Q) 순서로 선택된다. 그림 4의 스트림에 대한 예제를 살펴보면, 사용자가 지정한 목표 공간/시간 해상도가 D0/T1, D1/T1, D1/T2 일 때, 추출되는 계층의 순서는 다음과 같다.

- D0/T1: (0,0,0) → (0,1,0) → (0,0,1) → (0,1,1) → (0,0,2) → (0,1,2).
- D1/T1: (0,0,0) → (0,1,0) → (1,0,0) → (1,1,0) → (0,0,1) → (0,0,2) → (0,1,1) → (0,1,2) → (1,0,1) → (1,1,1) → (1,0,2) → (1,1,2).
- D1/T2: (0,0,0) → (0,1,0) → (0,2,0) → (1,0,0) → (1,1,0) → (1,2,0) → (0,0,1) → (0,0,2) → (0,1,1) → (0,1,2) → (0,2,1) → (0,2,2) → (1,0,1) → (1,1,1) → (1,2,1) → (1,0,2) → (1,1,2) → (1,2,2).

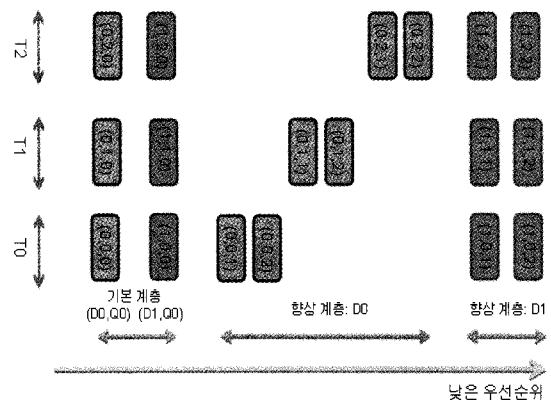


그림 4. JSVM에서 계층별 우선 순위

3.2 SVC RD(Rate-Distortion) 정보

JSVM는 목표 공간/시간 해상도와 사용 비트율을 입력으로 받은 후, 그림 4와 같은 우선 순위에 기반하여 서브 스트림을 추출한다. 실제로 추출기에서 사용되는 명령어와 옵션은 다음과 같다.

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

그림 5(a)는 “Soccer” 비디오 샘플에 대해, 주어진 공간/시간 해상도에서 비트율을 변화하면서 획득한 PSNR 값을 보여주고 있다. 주어진 공간/시간 해상도에서 비트율이 증가함에 따라 SNR 향상 계층이 추가되기 때문에 PSNR 값이 단조증가한다. 여기서 각 점들은 스트림의 추출 지점(extraction point)을 의미한다.

비디오 화질을 높이기 위해서는 주어진 비트율에서 최대의 PSNR을 갖는 추출 지점을 선택하여야 한다. 그림 5(b)에서 네트워크 사용 대역폭이 약 310 kbps일 때 선택할 수 있는 추출 지점으로는 ① 일부 화질 계층을 포함한 CIF 해상도/7.5 fps, ② CIF 해상도/15 fps, ③ 일부 화질 계층을 포함한 QCIF 해상도/30fps 이다. 이 때 최대의 PSNR 값을 얻기 위해서는 세 번째 추출 지점을 선택하여야 한다. 따라서 오프라인 시 컨텐츠를 분석하여 최대의 PSNR을 획득할 수 있는 추출 지점을 선택하는 것이 중요하다고 할 수 있다.

본 논문에서는 비트율이 주어졌을 때 최대의 PSNR을 제공하는 추출 지점을 “RD 포인트”라고 정의하며 이러한 “RD 포인트”的 집합을 “RD 정보”라고 한다. RD 정보는 그림 5 (b)에서 굵은 선으로

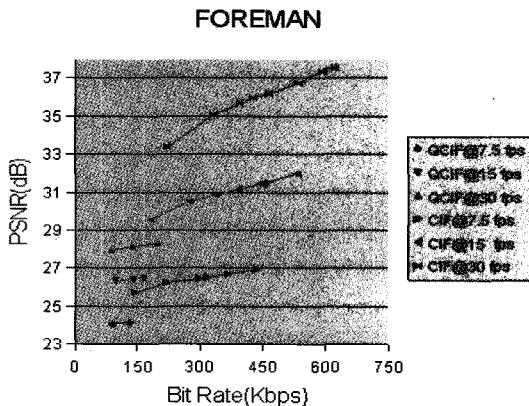
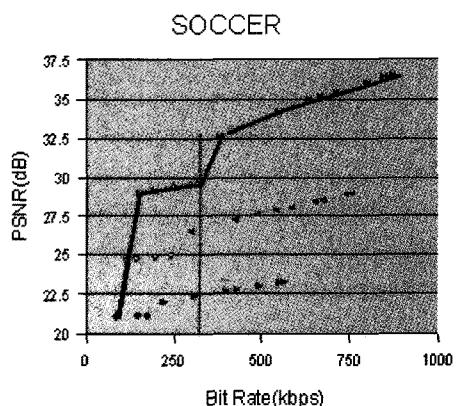


그림 5. (a) Soccer 서브스트림의 PSNR



(b) 주어진 비트율에서 최대의 PSNR 값을 가지는 추출지점에 대한 정보를 가지는 RD 정보

연결된 점들로 구성되며 외곽 곡선을 따라가면서 쉽게 획득될 수 있다. 이러한 특성을 가지는 RD 정보는 오프라인시에 획득된다. RD 포인트 p_i 가 주어졌을 때 RD 정보는 $[D_i, T_i, b_i, \text{PSNR}_i]$ 형식을 가진다. 예를 들어 SOCCER 비디오의 RD 정보의 내용은 다음과 같다.

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

4. H.264/SVC 계층 선택 알고리즘

네트워크 중간 노드에서 유입되는 비디오 스트림의 전송율의 합이 가용한 네트워크 대역폭보다 큰 경우 혼잡이 발생된다. 따라서 노드에서는 각 비디오 스트림의 새로운 계층을 선택함으로써 혼잡을 피해야 한다. 본 논문에서는 수식 1)에서 제시한 바와 같이 두 가지의 조건이 존재한다. 첫째, 재조정한 전송율의 합 R_{out} 이 노드의 네트워크 대역폭 R_{NB} 보다 작거나 같아야 한다. 둘째, 모든 비디오 스트림의 허용 가능한 화질을 보장하기 위해 PSNR_{min} 이상이어야 한다.

본 절에서는 이러한 제약 조건하에서 전체 비디오 스트림의 PSNR 값을 최대화할 수 있는 세 가지 알고리즘을 제시한다. 알고리즘에서 사용되는 주요 기호를 정리하면 다음과 같다(그림 6 참조).

네트워크의 중간 노드에 유입되는 비디오 스트림의 집합을 S라고 하고 스트림 수를 n 이라고 하자 ($|S|=n$). 그리고 스트림 k 의 RD 포인트에 대한 두 개의 인덱스를 p_k, p'_k 라고 하자. 만약 포인트 p_k 에서 p'_k 을 선택할 경우, 전송율 및 PSNR의 변화량은 다음과 같이 표현할 수 있다.

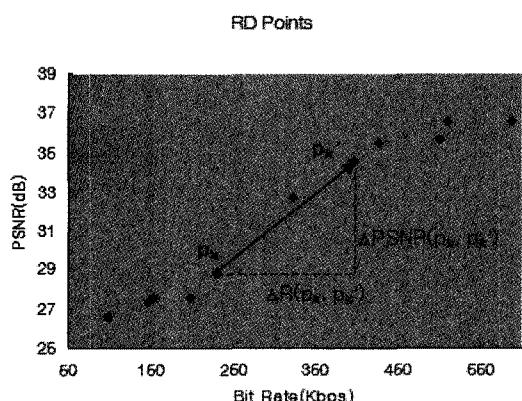
$$\text{전송율 변화량: } \Delta R(p_k, p'_k) = R(p'_k) - R(p_k)$$

$$\text{PSNR 변화량: } \Delta \text{PSNR}(p_k, p'_k) = \text{PSNR}(p'_k) - \text{PSNR}(p_k)$$

한편, 포인트 p_k 에서 p'_k 으로 움직였을 때, 시스템에서 얻을 수 있는 효용성(utility)은 다음과 같이 정의될 수 있다.

$$\text{효용성: } U(p_k, p'_k) = \frac{\Delta \text{PSNR}(p_k, p'_k)}{\Delta R(p_k, p'_k)} \quad (2)$$

그림 6은 스트림 k 의 포인트 p_k 에서 p'_k 를 선택하

그림 6. p_k 에서 p'_k 선택시 효용성

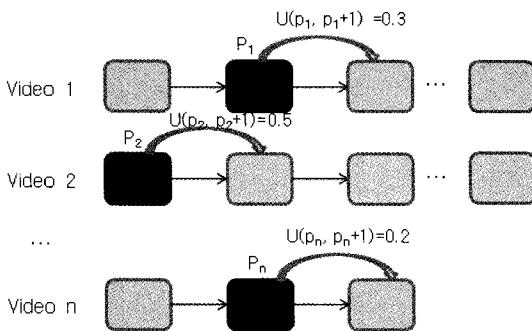


그림 7. NS 알고리즘의 예제: $U(p_2, p_2+1)$ 이 가장 크기 때 문에 p_2 는 1 증가한다.

였을 때, 전송율 및 PSNR 값의 변화량을 보여주고 있다. 효용성은 직선의 기울기와 같다.

각 비디오 스트림에 대해 최소한의 화질을 제공하기 위해 $PSNR_{min}$ 을 정의하였다. 스트림 k 에 대해 $PSNR(p_k) \geq PSNR_{min}$ 을 만족하는 최소 인덱스 p_k 가 존재할 때, 초기에 할당되는 네트워크 대역폭은 다음과 같이 계산된다.

$$\sum_{k=1}^n PSNR(p_k)$$

4.1 NS(Near-Sighted) 알고리즘

비디오 스트림 집합 S 에 속한 모든 스트림 k 에 대해, p_k 에서 다음 포인트인 p_k+1 까지의 효용성 $U(p_k, p_k+1)$ 을 계산한다. 그리고 n 개의 스트림 중에

서 가장 높은 효용성을 가진 스트림 j 를 선택한다. 이 스트림은 단위 비트율당 가장 높은 PSNR을 제공한다. 마지막으로, 다음 포인트 p_j+1 을 선택함으로써 증가하는 전송율 $\Delta R(p_j, p_j+1)$ 이 가용 대역폭을 넘지 않으면 p_j 를 1 증가하고 가용 대역폭을 감소시킨다. 이 작업을 가용 네트워크 대역폭이 없을 때까지 반복한다. 이 기법은 현재 포인트에서 다음 포인트까지의 효용성만을 고려함으로써 근시안(NS: Near-Sighted) 알고리즘이라고 명명하며 간단한 의사 코드는 표 1과 같다. 알고리즘 수행 후 결정되는 최종 포인트 p_k 의 (D_k , T_k , b_k)를 통해 비디오 스트림의 계층이 결정된다.

4.2 Fair 알고리즘

비디오 스트림 k 에 대해 $PSNR(p_k) \geq PSNR_{min}$ 을 만족하는 RD 포인트 p_k 가 존재할 때, 초기에 가용한 네트워크 대역폭은 $R_{AB} = R_{NB} - \sum_{j=1}^n PSNR(p_j)$ 으로 계산된다.

Fair 알고리즘은 각 스트림에 $PSNR_{min}$ 을 보장하기 위해 할당한 전송율을 제외하고 남는 대역폭을 각 비디오에게 공평하게 할당하는 알고리즈다. 따라서 비디오 k 에게 할당하는 대역폭 R_k 는 다음과 같이 계산된다.

$$R_k = R(p_k) + [R_{NB} - \sum_{j=1}^n PSNR(p_j)]/n$$

표 1. NS 알고리즘의 의사 코드

Algorithm 1: Near-Sighted

Input: RD points for each video stream,
Output: index of selected RD points for each stream

```

 $p_k$  is a variable indicating an index of RD points for stream  $k$ ;
for each video stream  $k \in S$ 
    initialize  $p_k$  such that  $PSNR(p_k) \geq PSNR_{min}$  for stream  $k$ ;
     $R_{AB} = R_{NB} - \sum_{k=1}^n R(p_k)$ ; //  $R_{AB}$ : available bandwidth,  $R_{NB}$ : bandwidth of a node
    while ( $S$  is not empty and  $R_{AB} > 0$ )
        Find stream  $j$  such that  $j = \arg \max_{k \in S} U(p_k, p_k+1)$ ;
        if ( $\Delta R(p_j, p_j+1) < R_{AB}$ )
             $R_{AB} = R_{AB} - \Delta R(p_j, p_j+1)$ ;
             $p_j = p_j + 1$ ;
            else
                remove stream  $j$  from video stream set  $S$ ;
        return a set of index of RD points  $\{p_1, p_2, \dots, p_n\}$ ;
    
```

Fair 알고리즘은 비디오 스트림 k 의 할당 가능한 대역폭 R_k 내에서 가장 높은 PSNR 값을 가지는 RD 포인트를 선택하여 계층을 선택한다.

4.3 FS(Far-Sighted) 알고리즘

NS 알고리즘은 비디오 스트림 k 의 현재 인덱스 p_k 에서 다음 인덱스인 p_k+1 까지의 효용성 $U(p_k, p_k+1)$ 을 고려하였다. 그러나 FS 알고리즘은 p_k+1 뿐만 아니라 p_k 보다 큰 모든 RD 포인트를 대상으로 한다. 그럼 8에서 p_1 에서 3개의 효용성, p_2 에서의 2개의 효

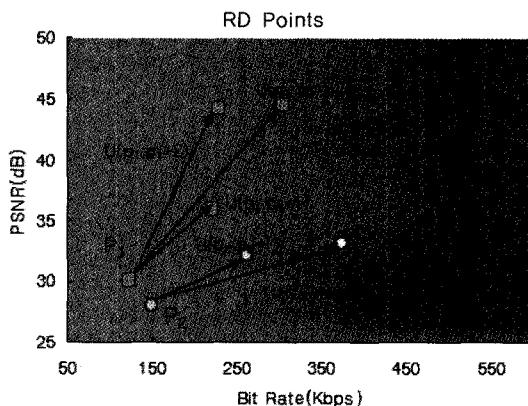


그림 8. FS 알고리즘에서 효용성 검사: $U(p_1, p_1+2)$ 이 가장 크기 때문에 p_1 은 2 증가한다.

표 2. FS 알고리즘의 의사 코드

Algorithm 2: Far-Sighted

Input: RD points for each video stream,
Output: index of selected RD points for each stream

```

 $p_k$  is a variable indicating an index of RD points for video stream  $k$ ;
for each video stream  $k \in S$ 
    initialize  $p_k$  such that  $PSNR(p_k) \geq PSNR_{min}$  for stream  $k$ ;
     $R_{AB} = R_{NB} - \sum_{k=1}^n R(p_k)$ ; //  $R_{AB}$ : available bandwidth,  $R_{NB}$ : bandwidth of a node
    while ( $S$  is not empty and  $R_{AB} > 0$ )
        for each stream  $k \in S$ 
             $t_k = \arg \max_{j>0} U(p_k, p_k + j)$ 
            Find stream  $i$  such that  $i = \arg \max_{k \in S} U(p_k, p_k + t_k)$ ;
            if ( $\Delta R(p_i, p_i + t_i) < R_{AB}$ )
                 $R_{AB} = R_{AB} - \Delta R(p_i, p_i + t_i)$ ;
                 $p_i = p_i + t_i$ ;
            else
                remove stream  $i$  from video stream set  $S$ ;
return a set of index of RD points  $\{p_1, p_2, \dots, p_n\}$ ;

```

용성을 검사하여 최대값을 선정한다. 여기서는 5개의 효용성 중에서 $U(p_1, p_1+2)$ 이 가장 크기 때문에 p_1 은 2가 증가하게 된다. 이 알고리즘은 비디오 스트림의 전체 특성을 고려할 수 있기 때문에 좀 더 좋은 화질을 제공할 수 있는 장점이 존재한다. FS 알고리즘의 의사 코드는 표 2와 같다.

5. 성능 분석

J SVM 소프트웨어를 통해서 추출된 SVC 비트 스트림을 바탕으로 RD 정보를 결정한 후, 세 가지 계층 선택 알고리즘에 대한 실험 결과를 제시한다. 이를 바탕으로 알고리즘의 특징을 분석한다.

5.1 실험 환경

실험 대상 비디오 파일을 움직임의 정도에 따라 분류한 후, 대표적인 샘플 세 가지를 (CITY.yuv, FOREMAN.yuv, SOCCER.yuv) 선택하였다[14]. 각 비디오 파일은 모두 YUV 형식으로 300 프레임으로 구성되어 있으며 해상도는 CIF 포맷이다(표 3 참조).

YUV 포맷의 비디오 파일은 인코딩 설정 파일을 통하여 SVC 스트림으로 변환된다. 먼저 공간 확장성을 위해 QCIF, CIF 해상도 2개의 계층을 제공하고, SNR 확장성을 위해 기본 계층과 3 개의 MGS 계층

표 3. 비디오의 종류 및 특징

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

을 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.

고정된 공간/시간 해상도 하에서 대역폭의 값을 100에서 900까지 50 단위로 변경하면서 비트 스트림을 추출한 후, 생성된 스트림의 비트율과 PSNR 값을 얻었다. 그림 5 (a)는 “Soccer” 비디오 스트림에 대한 결과를 보여주고 있으며 이를 바탕으로 RD 포인트는 그림 5 (b)와 같이 얻어진다.

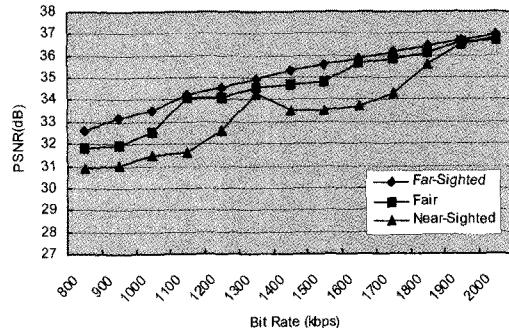
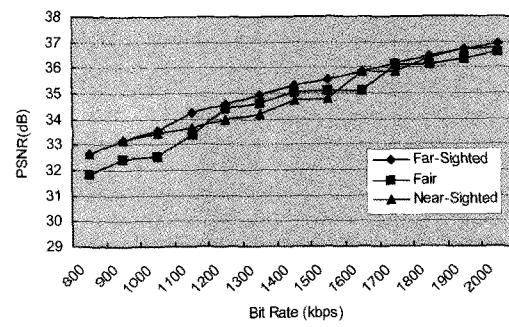
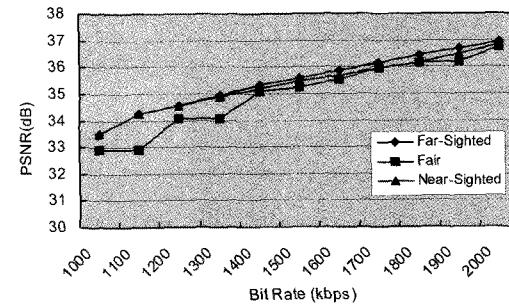
계층 선택 알고리즘은 비디오 스트림의 RD 정보를 입력으로 노드의 네트워크 대역폭 R_{NB} 값을 변화시키면서 RD 포인트를 결정한다. RD 포인트는 $[D_i, T_i, b_i, PSNR_i]$ 로 구성되며 (D_i, T_i, b_i) 정보는 그림 1의 “Packet Drop” 모듈에 전달되어 혼잡을 제거한다.

5.2 실험 결과 및 분석

그림 9~11은 R_{NB} 값을 변화시키면서 생성된 서브 스트림의 평균 PSNR 값을 보여주고 있다. 그림 9는 $PSNR_{min}=24$ 일 때, 그림 10은 $PSNR_{min}=28$ 일 때, 그림 11은 $PSNR_{min}$ 일 때 결과를 나타낸다.

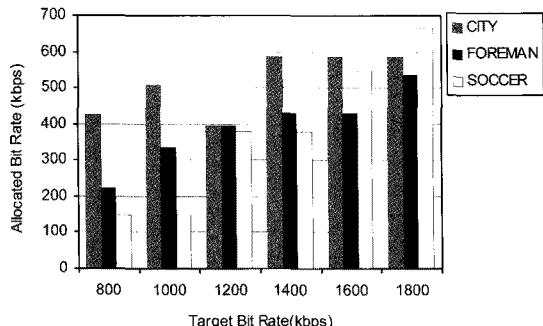
실험 결과를 바탕으로 세 가지 알고리즘의 특징을 요약하면 다음과 같다.

- FS(Far-Sighted) 알고리즘이 가장 좋은 성능을 나타내고 있다. 그 이유는 비디오 스트림의 RD 포인트 전체를 고려하여 효용성을 검토하기 때문에 스트림의 화질이 가장 좋은 것으로 나타나고 있다. 비디오 스트림의 계층 수가 40 개일 경우 RD 포인트가 대략 15 개 정도로 구성되기 때문에 복잡도가 높지 않다고 할 수 있다.
- NS(Near-Sighted) 알고리즘은 Mahalingam[10]이 제시한 방식과 유사하며 이웃한 RD 포인트의 효용성을 검토한다. 대체로 Fair 알고리즘보다 우수한 성능을 나타내고 있으나 $PSNR_{min}$ 이 24인 경우와 같이 단조증가하지 못하는 경우가 있다. 그

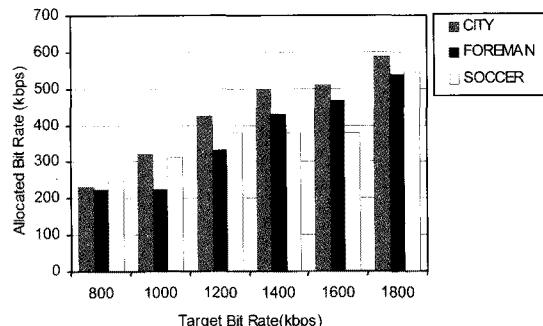
그림 9. 알고리즘의 성능 비교($PSNR_{min}=24$)그림 10. 알고리즘의 성능 비교($PSNR_{min}=28$)그림 11. 알고리즘의 성능 비교($PSNR_{min}=32$)

이유는 다음 가능한 RD 포인트의 효용성이 매우 작을 경우에는 지속적으로 선택되지 못하기 때문이다. 따라서 비디오 스트림의 특성에 많은 영향을 받는 것으로 나타났다.

- Fair 알고리즘은 전체적인 성능은 떨어지지만 스트림 간 공평한 전송율을 할당하는 특성을 가지고 있다. 그림 12는 FS 알고리즘과 Fair 알고리즘에서 세 가지 비디오 스트림이 실제로 할당받은 전송율을 보여주고 있다. 따라서 공평성이 중요한



(a) FS 알고리즘



(b) Fair 알고리즘

그림 12. 비디오 스트림간의 공평성 측정 ($\text{PSNR}_{\min} = 28$)

응용에서는 FS 알고리즘과 Fair 알고리즘을 혼합함으로써 좋은 성능을 기대할 수 있다. 예를 들면, 가용한 대역폭의 일부분을 공평하게 분배하고 나머지 부분은 FS 알고리즘에 따라 분배할 수 있다.

6. 결 론

네트워크의 다양성과 가변성으로 인해 혼잡이 발생하는 경우 H.264/SVC 스트림에 치명적인 영향을 줄 수 있다. 그 이유는 H.264/SVC는 기본 계층과 향상 계층으로 구성되어 있고 계층간 의존성이 존재하기 때문에 일부 계층의 손실은 다른 계층에 영향을 줄 수 있기 때문이다. 본 논문에서는 혼잡 발생 시 계층간의 의존성을 만족하면서 최대의 화질을 얻기 위한 세 가지 계층 선택 알고리즘을 제안하였다. 실험 결과 FS(Far-Sighted) 알고리즘이 가장 좋은 성능을 나타내는 것으로 나타났다. 한편 NS(Near-Sighted) 알고리즘은 복잡도는 낮으나 스트림의 특성에 영향을 받으며 FS보다 낮은 특성을 제공하는 것으로 나타났다.

향후에는 실제 네트워크 상황을 시뮬레이션하기 위해 가상 네트워크 환경을 구성한 후, 각 알고리즘을 분석할 예정이다. 또한 TCP와 UDP가 혼재되어 있는 경우 공평하면서 좋은 성능을 제공할 수 있는 알고리즘을 개발할 예정이다.

참 고 문 헌

- [1] Dapeng Wu, Yiwei Thomas Hou, Wenwu Zhu, Ya-Qin Zhang, and Jon M. Peña, "Streaming

Video Over the Internet: Approaches and Directions," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.11, No.3, March 2001.

- [2] Huifang Sun, Anthony Vetro, and Jun Xin, "An Overview of Scalable Video Streaming," *Wireless Communications and Mobile Computing*, Vol.7, Issue 2, Jan 2007.
- [3] T. Wiegand, G. Sullivan, 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] J. Monteiro, C. Calafate, and M. Nunes, "Evaluation of the H.264 Scalable Video Coding in Error Prone IP Networks," *IEEE Transaction on Broadcasting*, 54(3), September 2008.
- [5] 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*, 17(9):1186 - 1193, 2007.
- [6] Y Wang, L. Chau and K. Yap, "Spatial Resolution Decision in Scalable Bitstream Extraction for Network and Receiver Aware Adaptation," *IEEE International Conference on Multimedia and Expo*, 2008.
- [7] 김남윤, 황호영, "H.264 SVC에서 비트 스트림 추출을 위한 공간과 시간 해상도 선택 기법,"

멀티미디어 학회 논문지, 13(1), January 2010.

- [8] A. Dyahadray, H. Shringarpure, N. Mulay, and G. Saraph, "Congestion Control using Scalable Video Streaming," National Conference on Communications, 2008.
- [9] K. Tappayuthpijam, G. Liebl, T. Stockhammer, and E. Steinbach, "Adaptive Video Streaming over a Mobile Network with TCP-Friendly Rate Control," *International Conference on Wireless Communications and Mobile Computing*, 2009.
- [10] R. Mahalingam, W. Tu, and E. Steinbach, "RD-Optimized Rate Shaping for Multiple Scalable Video Streams," *IEEE International Conference on Multimedia and Expo*, 2007.
- [11] H. Schwarz, D. Marpe, and T. Wiegand. "Hierarchical B pictures," Joint Video Team, Doc. JVT-P014, July 2005.
- [12] H. Kirchhoffer, H. Schwarz, and T. Wiegand. "CE1: Simplified FGS," Joint Video Team, Doc. JVT-W090, Apr. 2007.
- [13] HHI, SVC reference software, http://ip.hhi.de/imagecom_G1/savce/downloads/SVC-Reference-Software.htm.
- [14] <ftp://tnt.uni-hannover.de/pub/svc/testsequences>.



김 남 윤

1992년 2월 서울대학교 컴퓨터공
학과 졸업
1994년 2월 서울대학교 컴퓨터공
학과 석사
2000년 2월 서울대학교 컴퓨터공
학과 박사

1999년 9월~2002년 2월 삼성전자 무선사업부 책임연구원
2002년~현재 한성대학교 정보시스템공학과 부교수
관심분야: 멀티미디어 통신, 웹 검색, 모바일 통신 및 응용



황 기 태

1986년 2월 서울대학교 컴퓨터공
학과 졸업
1988년 2월 서울대학교 컴퓨터공
학과 석사
1994년 2월 서울대학교 컴퓨터공
학과 박사

2000년~2001년 University of California, Irvine 방문 교수
1994년~현재 한성대학교 컴퓨터공학과 교수
관심분야: 분산 시스템, 모바일 시스템