

논문 2006-43CI-6-8

# 스케일러블 비디오 코딩에서의 실시간 스케일러빌리티 변환

## (Dynamic Full-Scalability Conversion in SVC)

이동수(배태면\*, 노용만\*\*)

(Dong Su Lee, Tae Meon Bae, and Yong Man Ro)

### 요약

유비쿼터스 환경에서의 비디오 적응 변환을 위하여 현재 MPEG-4 AVC/SE SVC에서는 공간적, 시간적, 품질적 스케일러빌리티를 지원하지 않는다. 이러한 스케일러빌리티의 변환은 가변적인 대역폭을 가지는 네트워크 특성에 따라 실시간으로 이루어져야 한다. 그러나 현재의 SVC에서는 품질적 스케일러빌리티에 대해서는 실시간으로 스케일러빌리티를 변환할 수 있지만 공간적, 시간적 스케일러빌리티에 대해서는 이를 제대로 지원하지 못한다. 이에 본 논문에서는 이와 관련된 문제점들을 자세히 알아보고 이를 해결하기 위한 방법을 제안하였다. 부호시에는 주기적으로 IDR NAL을 삽입하였다. 추출시에는 실시간 추출에 필요한 정보들을 미리 분석한 후 실제 추출시에 활용하여 실시간 추출을 가능하도록 하였다. 마지막으로 복호시에는 실시간 스케일러빌리티를 지원할 수 있는 프로세스를 정의하였다. 이를 JSM4.0에 구현하였고 주어진 네트워크 대역폭에 스케일러빌리티를 통하여 적응된 비트스트림의 비트레이트, PSNR, 스케일러빌리티를 측정하여 제안된 방법의 유효성을 검증하였다.

### Abstract

Currently, Scalable Video Coding (SVC) is being standardized. By using scalability of SVC, QoS managed video streaming service is enabled in heterogeneous networks even with only one original bitstream. But current SVC is insufficient to dynamic video conversion for the scalability, thereby the adaptation of bitrate to meet a fluctuating network condition is limited. In this paper, we propose dynamic full-scalability conversion method for QoS adaptive video streaming in H.264/AVC SVC. To accomplish full scalability dynamic conversion, we propose corresponding bitstream extraction, encoding and decoding schemes. On the encoder, we newly insert the IDR NAL to solve the problems of spatial scalability conversion. On the extractor, we analyze the SVC bitstream to get the information which enable dynamic extraction. By using this information, real time extraction is achieved. Finally, we develop the decoder so that it can manage changing bitrate to support real time full-scalability. The experimental results showed that dynamic full-scalability conversion was verified and it was necessary for time varying network condition.

Keywords : SVC, QoS, 스케일러빌리티

## I. 서론

현재 비디오 스트리밍 서비스 환경은 유비쿼터스 시대에 맞추어 다양한 네트워크 환경 및 장치 간의 전송에 대한 필요성이 증가되고 있는 실정이다<sup>[1]</sup>. 특히 DMB, DVB-H, 무선 LAN과 같은 환경하에서의 비디오 전송에 대한 요구성이 증대되었다<sup>[2][3]</sup>. 이러한 요구의 증가에 따라 복잡한 변환 없이 적응할 수 있는 비

디오 코딩 기술의 필요성이 증대되었다. 이를 위한 비디오 코딩의 특성으로 스케일러빌리티가 요구되었으나 기존의 표준인 MPEG-1은 스케일러빌리티 특성이 존재하지 않으며, MPEG-2, MPEG-4에서 일부 스케일러빌리티를 지원하나 제한점이 많다<sup>[4]</sup>. 이에 맞추어 MPEG에서는 MPEG-4 AVC를 기반으로 MPEG-4 AVC/Scalable Extension(SE) Scalable Video Coding(SVC)<sup>[5]</sup>가 현재 표준화되고 있다<sup>[6]</sup>.

SVC는 비디오 적응을 위하여 공간적, 시간적, 품질적 스케일러빌리티를 지원할 수 있는 구조로 부호화된다. 이 부호화된 비디오 스트림은 추출기에서 네트워크의 특성에 맞추어 적절히 추출하여 사용할 수 있으므로

\* 학생회원, \*\* 정회원, 한국정보통신대학교 공학부 (School of Engineering, Information and Communications University)

접수일자: 2006년7월12일, 수정완료일: 2006년10월30일

비디오 적응을 위하여 복잡한 변환이 필요 없게 된다. 즉, 한 번의 부호화 후 추출을 통하여 각 네트워크 환경 및 기기 특성에 맞는 비트스트림을 제공할 수 있는 것이다.

네트워크를 통한 비디오 스트리밍 서비스는 시간에 따라 가변하는 네트워크 대역폭의 특성 때문에 일정 QoS를 보장하기 어렵다<sup>[7]</sup>. 무선 네트워크 환경과 같이 대역폭의 가변폭이 큰 경우의 QoS적응적인 비디오 스트리밍을 위하여 SVC의 공간적, 시간적, 품질적 스케일러빌리티는 실시간으로 지원될 수 있어야 한다. 그러나 현재의 SVC에서는 실시간 품질적 스케일러빌리티는 원활히 지원할 수 있으나 공간적, 시간적 스케일러빌리티의 변환과 관련하여 여러 문제점이 존재한다. 이는 3장에서 자세히 서술하겠다.

본 논문에서는 원활한 비디오 스트리밍 서비스의 지원을 위하여 부호기, 추출기, 복호기에 대하여 각각 방안을 제시하였다. 이를 위해 부호기에서는 주기적으로 IDR NAL을 삽입하도록 하였다. 추출기에서는 미리 실시간 스케일러빌리티 지원에 필요한 정보를 비트스트림 분석을 통하여 저장 후 실제 전송 시에 사용하도록 하였다. 복호시에는 스케일러빌리티 가변에 대응할 수 있는 과정을 정의하여 스케일러빌리티 변화를 처리하도록 하였다. 이러한 방안을 통해 SVC의 실시간 스케일러빌리티 변환을 지원할 수 있게 된다. 제안된 방법을 JSVM4.0에 구현하였고 실험을 통하여 유효성을 검증하였다.

본 논문은 다음과 같이 구성되었다. II장에서는 SVC의 스케일러빌리티 지원에 대해 설명하고 III장에서는 이러한 스케일러빌리티를 실시간으로 지원하려고 할 때의 문제점에 대해서 자세히 서술하겠다. IV장에서는 이러한 문제점을 해결하기 위한 방법을 제안하고 V장에서는 실험을 통하여 제안된 방법의 유효성을 검증하였다. 끝으로 VI장에서는 전체 결론을 내리겠다.

## II. SVC의 스케일러빌리티 지원

### 1. SVC의 스케일러빌리티

SVC는 다양한 네트워크 환경에서 쉽게 적용할 수 있도록 공간적, 시간적, 품질적 스케일러빌리티를 제공하고 있다<sup>[8][9]</sup>. 공간적 스케일러빌리티는 계층적 부호화 방법을 통해 이루어지며 시간적 스케일러빌리티는 계층적 B 픽처 구조를 통해 이루어진다. 마지막 품질적 스케일러빌리티는 Fine Granular SNR(FGS)<sup>[10]</sup> 또는

Coarse Granular SNR(CGS)를 통해 제공된다.

#### 가. 공간적 스케일러빌리티

SVC는 공간적 스케일러빌리티를 지원하기 위해 계층적 부호화 방법을 이용하고 있다. 그림 1은 2계층의 공간적 스케일러빌리티를 지원하기 위한 SVC 코덱의 구조를 보여주고 있다. SVC의 계층적 코덱에는 하위 계층으로서 H.264/AVC 호환가능 코덱과 상위 계층으로서 H.264/AVC Scalable Extension(SE) 코덱이 있다. H.264/AVC SE는 H.264/AVC의 구조에 계층 간 중복되는 정보를 제거하는 부분이 추가되었다. 즉, 하위 계층의 움직임 정보, 텍스처 정보, 잔여 신호 정보를 상위 계층 정보의 예측에 사용하고 있다. 상위 계층과 하위 계층의 해상도가 다르다면 이러한 하위 계층의 정보들은 업샘플링되어 이용된다<sup>[11]</sup>.

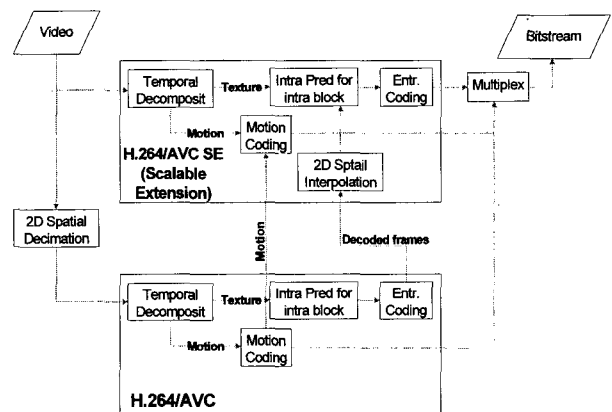


그림 1. SVC 부호기 구조

Fig. 1. Structure of SVC Coder.

#### 나. 시간적 스케일러빌리티

시간적 스케일러빌리티는 계층적 B 픽처 코딩<sup>[12]</sup>으로서 GOP 단위로 지원되어진다. GOP에서 key 픽처는 시간적으로 마지막에 위치한다. 계층적 B 픽처 코딩에서

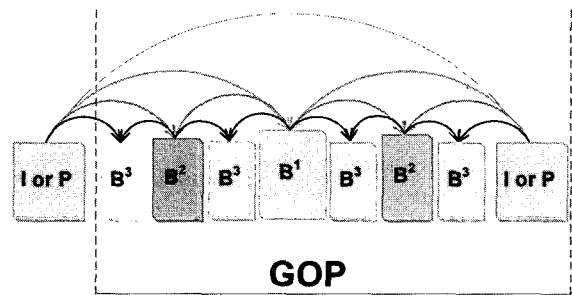


그림 2. 계층적 B픽처 코딩 구조

Fig. 2. Structure of Hierarchical B Picture Coding.

이전 GOP의 key 픽처와 현재 GOP의 key 픽처간의 예측이 가장 먼저 일어난다. 나머지 계층적 예측은 그림 2와 같이 수행된다. 그림 2에서는 실시간적으로 3계층까지 지원될 수 있다.

다. 품질적 스케일러빌리티

품질적 스케일러빌리티는 FGS와 CGS로 지원될 수 있다. FGS는 각 공간적 계층에서 모든 프레임의 계수를 정제하여 얻어진다. 이것은 quantization step size를 줄여가면서 부호화하여 원본 신호와의 차이를 얻어서 이를 sub-bitplane 방식과 유사하게 부호화하여 이루어진다. 그림 3은 이렇게 얻어진 신호를 부호화한 FGS 슬라이스 구조를 개념적으로 나타내고 있다.

CGS는 같은 공간적 해상도를 사용한다는 것을 제외하고는 공간적 스케일러빌리티를 지원하는 공간적 부호화 구조와 같다. 상위 계층에서 quantization step size를 하위 계층보다 작게 하여 수행함으로써 상위 계층을 사용할 시 품질적 측면에서 보다 좋은 화질을 보이게 되는 것이다.

이러한 공간적, 시간적, 품질적 스케일러빌리티는 폭넓은 범위의 공간적, 시간적, 품질적 스케일러빌리티를 제공하기 위하여 쉽게 결합될 수 있다<sup>[13]</sup>. 그림 4와 같

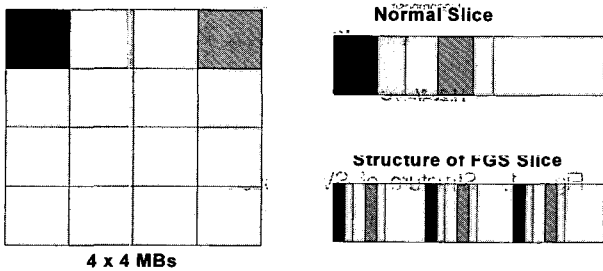


그림 3. FGS NAL의 구조  
Fig. 3. Structure of FGS NAL.

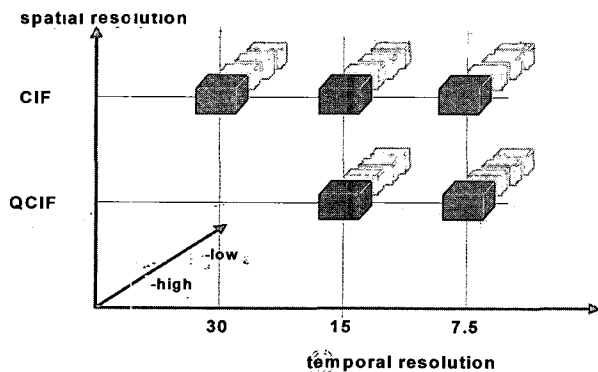


그림 4. 결합된 스케일러빌리티의 예  
Fig. 4. Example of Combined Scalability

이 결합된 스케일러빌리티 공간을 살펴보면 공간적 해상도로 CIF, QCIF가 지원가능하고 시간적 해상도로써 30fps, 15fps, 7.5fps가 지원가능하다. "high", "low"로 표시된 부분은 품질적 스케일러빌리티 지원을 나타내며 그림 4에서는 3개의 FGS를 나타내고 있다.

라. 스케일러빌리티 지원을 위한 SVC의 비트스트림 구조

SVC 비트스트림은 Network Abstract Layer(NAL)로 되어져 있다. NAL은 크게 Video Coding Layer(VCL) NAL과 non-VCL NAL로 나눌 수 있다. non-VCL NAL은 프로파일, 레벨, 프레임의 높이, 넓이 등과 같은 복호화에 필요한 파라미터나 정보 데이터를 포함한다. VCL NAL은 비디오 코딩 데이터를 포함하는 NAL으로써 한 슬라이스 단위를 포함하고 있다. VCL NAL의 헤더에는 SVC의 스케일러빌리티 정보를 그림 5와 같이 가지고 있다. dependency\_id는 3비트로 공간적 계층을 나타내고, temporal\_level은 역시 3비트로 시간적 계층을 가리킨다. 또한 quality\_level은 2비트로써 FGS로 지원되는 품질적 계층을 나타낸다. 이러한 정보를 이용하여 추출시에 추출지점에 맞게 데이터를 추출하는 것이 가능해진다.

그림 6과 같이 SVC의 비트스트림은 먼저 복호화에 필요한 정보를 담고 있는 non-VCL NAL들이 먼저 나타내고 이후 VCL NAL들이 나타나게 된다. Supplemental Enhancement Information(SEI) NAL을 통해 각 계층의 지원될 수 있는 픽처와 해상도, 프레임 등을 나타내고 Sequence Parameter Set(SPS)으로 전체 비디오 시퀀스의 정보를 나타내며 Picture Parameter

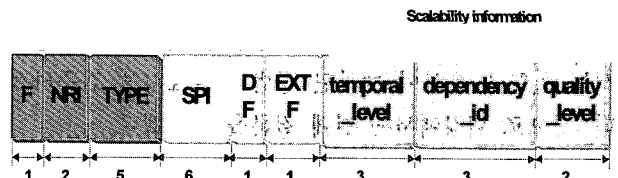


그림 5. NAL 헤더의 구조  
Fig. 5. Structure of NAL Header.

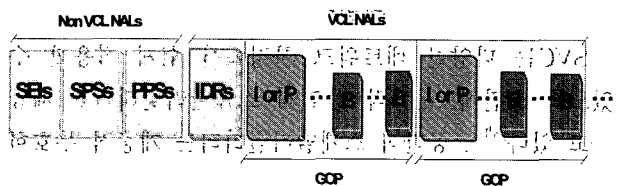


그림 6. SVC 비트스트림 순서  
Fig. 6. Order of SVC Bitstream.

Set(PPS)을 통해 픽처의 정보를 나타낸다.

### 2. 비디오 스트리밍과 SVC

그림 7은 유비쿼터스 네트워크 환경에서의 SVC를 이용한 비디오 스트리밍 구조를 나타내고 있다. PC, Notebook, PDA등과 같은 다양한 장치의 해상도를 고려해야 하며 LAN, 무선 등의 다양한 네트워크 환경을 고려하여 비디오는 적응되어야 하고 이러한 적응은 SVC를 이용하여 부호화 된 경우에는 추출기에서 스케일러블하게 데이터를 추출함으로써 손쉽게 이루어질 수 있는 것을 볼 수 있다. 이러한 스케일러빌리티는 네트워크의 가변적인 대역폭을 고려할 때 실시간으로 이루어져야하나 현재의 SVC에서는 이와 관련된 문제점이 존재한다. 다음 장에서부터 이를 자세히 서술하겠다.

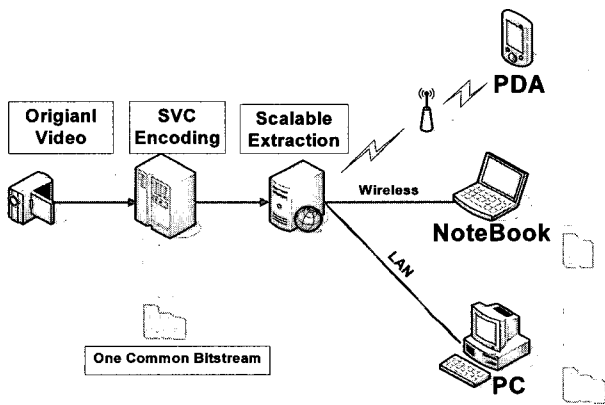


그림 7. 비디오 스트리밍과 SVC  
Fig. 7. Video Streaming and SVC.

### III. 실시간 스케일러빌리티 지원의 문제점

앞에서 살펴본 바와 같이 SVC는 공간적, 시간적, 품질적 스케일러빌리티를 제공하고 있다. 비디오 스트리밍시 네트워크의 가변적인 대역폭 특성을 고려할 때 이러한 스케일러빌리티는 실시간으로 지원되어야 한다. 그러나 이를 지원하기에는 현재의 SVC에서는 문제점이 존재한다. 이를 공간적, 시간적, 품질적 스케일러빌리티 지원 그리고 네트워크 대역폭에서의 적응과 관련된 문제점들을 각각 살펴보겠다.

#### 1. 공간적 스케일러빌리티 지원의 문제점

부호화기에서는 SVC의 비트스트림의 공간적 계층의 추출에 상관없이 복호화할 수 있어야 한다. 공간적 계층 변환의 지원시에 나타날 수 있는 문제점은 다음과

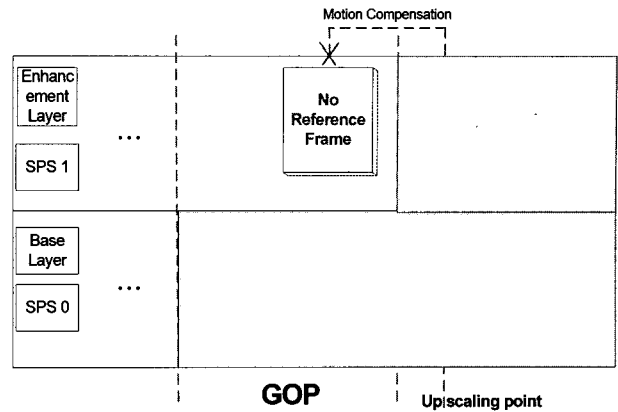


그림 8. 공간적 해상도를 높일 때의 문제점  
Fig. 8. Problem of Spatial Layer Changing to Higher Layer.

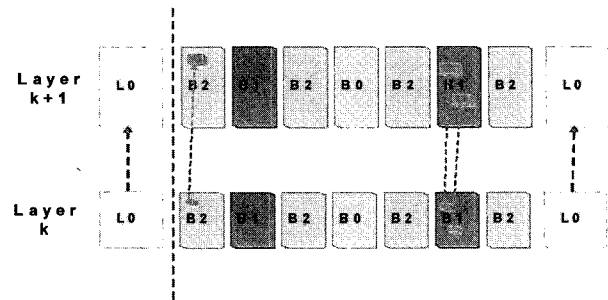


그림 9. Single Loop Decoding Mode.  
Fig. 9. Single Loop Decoding Mode.

같다. 우선 공간적 계층이 변화되었을 때 이를 인지할 수 있는 과정이 있어야 한다. 만약 공간적 계층이 변화되었다면 영상의 복호화 최상위 계층을 나타내는 Reconstruction Layer를 이에 맞추어 설정해주어야 한다. 또한 이에 맞는 영상의 넓이, 높이 등의 픽처 복원과 관련된 정보를 관리하는 SPS를 활성화시켜 주어야 하는데 이와 관련된 절차가 존재하지 않는다는 것이다. 둘째로 영상의 공간적 해상도를 높이는 경우를 생각해 보면 그림 8과 같이 새로운 상위 계층의 프레임의 경우 이전의 같은 계층의 프레임이 존재하지 않아 움직임 보상이 참조 프레임의 부재로 인하여 제대로 복호가 이루어지지 않게 된다.

공간적 해상도를 낮추는 경우에는 그림 9와 같은 single loop decoding mode<sup>[14]</sup>를 제외하고는 문제가 생기지 않는다. single loop decoding mode은 복호화시의 연산량을 줄이기 위하여 최상위 계층에서만 움직임 보상을 수행하게 된다.

single loop decoding mode를 사용하는 경우에는 움직임 보상이 최상위 계층에서만 일어난다. 이에 따라 그림 10와 같이 계층이 낮아졌을 때 이전 GOP의 동일

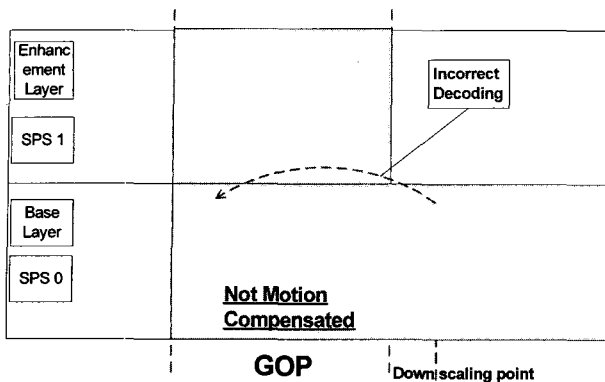


그림 10. Single Loop Decoding Mode에서 공간적 해상도를 낮출 때의 문제점 발생

Fig. 10. Problem of Spatial Layer Changing to Lower Layer with Single Loop Decoding Mode.

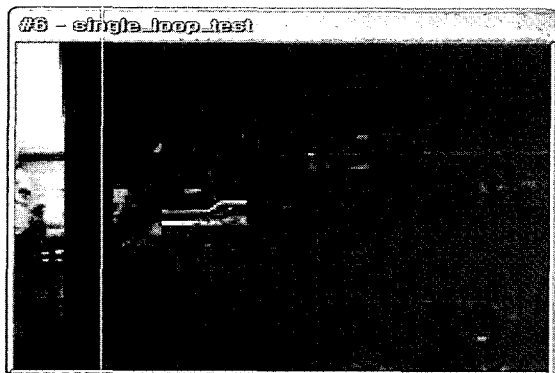


그림 11. Single Loop Decoding Mode에서 공간적 해상도를 낮추었을 때의 결과

Fig. 11. Result of Spatial Layer Changing to Lower Layer with Single Loop Decoding Mode.

계층에서 픽처를 참조해야하는데 이 이전 GOP의 동일 계층 픽처는 single loop decoding mode에서는 제대로 복호화가 이루어지지 않은 상태이다.

그림 11은 single loop decoding mode에서 공간적 해상도를 낮추었을 때의 결과를 나타내고 있다. 설명한 바와 같이 참조 프레임이 제대로 복호화 되어 있지 않아 그림 11과 같은 결과를 보이게 된다.

공간적 스케일러빌리티와 관련된 문제점들은 참조 프레임의 문제이므로 이와 관련되어 참조를 수행하지 않고 복원할 수 있도록 하여야 한다.

## 2. 시간적 스케일러빌리티 지원의 문제점

시간적 스케일러빌리티는 GOP단위의 계층적 B 픽처 구조를 통해 이루어지므로 이를 실시간으로 지원하기 위해서는 GOP단위의 NAL의 수를 알 수 있어야 한다. 이를 위해 비트스트림을 분석하여야 하는데 이는 많은 시간을 소요하게 된다.

## 3. 품질적 스케일러빌리티 지원의 문제점

품질적 스케일러빌리티의 경우 FGS와 CGS를 통해 지원될 수 있는데 FGS의 경우에는 프레임 단위로 지원되므로 문제점이 발생하지 않는다. CGS의 경우에는 공간적 계층의 문제점과 같은 문제점이 발생하게 된다.

## 4. 네트워크 대역폭에서의 적응 문제

네트워크 대역폭은 가변적이므로 이에 맞추어 추출 스케일러빌리티 또한 실시간으로 가변될 수 있어야 한다. 이를 위해서는 현재의 네트워크 대역폭을 알 수 있다고 하였을 때 이에 맞는 스케일러빌리티의 비트스트림 비트레이트를 계산할 수 있어야 하며 이 단위 또한 정할 수 있어야 한다. 이를 위해서 역시 비트스트림을 분석하여야 한다.

## IV. 실시간 스케일러빌리티 지원 해결 방안

### 1. 부호기에서의 제안

공간적 스케일러빌리티 지원의 문제점인 참조 프레임의 부재 문제는 부호화시에 주기적으로 IDR<sup>[15]</sup> NAL을 삽입하고 이 IDR NAL을 기점으로 공간적 스케일러빌리티의 변환을 허용함으로써 해결할 수 있다. IDR NAL은 intra 슬라이스로 부호화되며 이 IDR NAL이 복호화되었을 때에는 모든 참조 관계를 초기화하게 되어 독립적인 부호화, 복호화를 수행하여야 한다. 그러나 IDR은 I슬라이스로 이루어져 있기 때문에 비트스트림의 비트레이트가 증가하게 되는 단점이 있다. single loop decoding mode에서의 문제점 역시 IDR을 주기적으로 삽입함으로써 해결할 수 있게 된다.

### 2. 추출기에서의 제안

네트워크 상황에 맞추어 실시간 추출을 하기 위해서는 비트스트림 정보를 먼저 분석할 필요가 있다. 네트워크 대역폭에 맞는 스케일러빌리티를 결정하기 위해서는 각 스케일러빌리티 계층별 데이터를 알아야 하는데 이는 시간이 많이 걸리는 일이므로 비디오 스트리밍 전송 전에 미리 이루어져야 한다. 이를 위하여 추출기는 그림 12와 같이 비트스트림 분석기와 NAL 추출기로 구성할 수 있다. 비트스트림 분석기는 비트스트림을 분석 후 실시간 추출에 필요한 정보를 저장하게 된다.

실시간 추출에 필요한 정보는 표1과 같다. 이러한 정보들은 크게 두 갈래로 나뉠 수 있다. Scalability SEI, SPS, PPS는 추출 가능한 공간적, 시간적, 품질적 계층

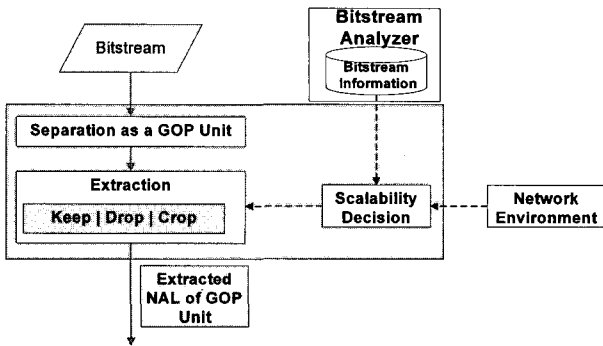


그림 12. 제안하는 추출기의 구조  
Fig. 12. Structure of Proposed Extractor.

표 1. 실시간 추출에 필요한 비트스트림 정보  
Table 1. Necessary Information for Real-time Extraction.

| Contents                                    |
|---|
| Scalability Information SEI로부터 얻어진 정보       |
| SEIs, SPS, PPS NAL들의 수와 IDR NAL들의 개수        |
| 각 GOP의 VCL NAL의 개수                          |
| 전체 NAL의 개수                                  |
| 각 공간적, 시간적, 품질적 스케일러빌리티에 해당하는 NAL의 데이터 량의 합 |
| 각 공간적, 품질적 스케일러빌리티에 해당하는 프레임 개수             |
| GOP에 속하는 프레임의 개수                            |

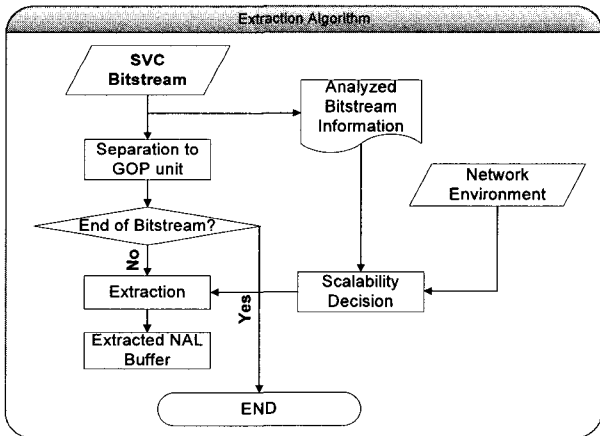


그림 13. 비트스트림 추출 알고리즘  
Fig. 13. Bitstream Extraction Algorithm.

에 대한 정보를 제공하고 GOP단위의 NAL의 수, 픽처 수, 계층별 데이터 량 등의 정보들은 네트워크에 적용할 수 있는 GOP단위의 계층별 비트레이트를 계산할 수 있게 해준다.

NAL 추출기에서는 이러한 정보를 바탕으로 실제 추출을 담당하게 된다. 추출 알고리즘은 그림 13과 같다.

제안된 알고리즘에서는 GOP를 기본 연산단위로 한다. 이는 SVC에서 시간적 스케일러빌리티를 위해 사용하는 계층적 B 픽처가 GOP단위로 일어나기 때문이며

공간적 계층 변화를 GOP의 키 픽처에서만 일어나도록 하였기 때문이다. 먼저 입력된 비트스트림은 bitstream separation module에서 GOP단위로 나뉘어진다. 추출 모듈은 decision 모듈에서 결정된 공간적, 시간적, 품질적 스케일러빌리티를 나타내는 추출 지점에 따라 나뉘어진 비트스트림을 다시 추출하게 된다. 이러한 작업은 비트스트림이 끝이 날 때까지 반복되게 된다. 스케일러빌리티 결정은 네트워크 대역폭과 장치의 지원 가능한 해상도를 입력 파라미터로 결정하게 된다. 주어진 네트워크 대역폭을 NB(Kbytes/sec)라고 하고, 해당 픽처의 수를 NOP, 해당 프레임율을 fr이라고 할 때 NB에 대응되는 비트스트림의 데이터량 DNB는 식 (1)과 같이 나타낼 수 있다.

$$D(NB) = NB / fr * NOP \quad (1)$$

이 때 각 스케일러빌리티 계층에 대하여 얻을 수 있는 데이터 량 D(SL)은 식 (2)와 같이 나타낼 수 있다.

$$D(SL) = \sum_{d=0}^D \sum_{t=0}^T \sum_{q=0}^Q SB(d, t, q) \quad (2)$$

여기서 SB는 부호화된 비트스트림을 의미한다. d, t, q는 각기 공간적, 시간적, 품질적 스케일러빌리티를 의미한다. 주어진 네트워크 대역폭 NB와 장치의 지원 가능한 해상도 DR에 대하여 스케일러빌리티 d, t, q를 결정하는 것은 식 (3)과 같이 나타낼 수 있고 이는 전체 데이터 량에서 계층별 데이터량을 뺀 것이 주어진 대역폭 NB에 대응하는 데이터량보다 작거나 같아야 하는 제한을 나타내는 식 (4)를 만족하는 d, t, q를 찾는 일이다.

$$(d, t, q) = SD(NB, DR) \quad (3)$$

$$TD - D(SL) \leq D(NB) \quad (4)$$

### 3. 복호기에서의 제안

복호기는 이러한 공간적, 시간적, 품질적 변화에 대응할 수 있어야 한다. 다른 품질적 변화에서는 문제가 없지만 공간적 스케일러빌리티 변화를 지원하기 위해서는 복호기에서 이러한 변화를 감지하고 이에 대응하는 절차를 지원하여야 한다.

먼저 복호기에서 공간적 계층의 변화를 감지하기 위하여 그림 14와 같은 알고리즘을 사용한다. 제안된 방법에서는 공간적 계층의 변화는 IDR에서만 일어날 수

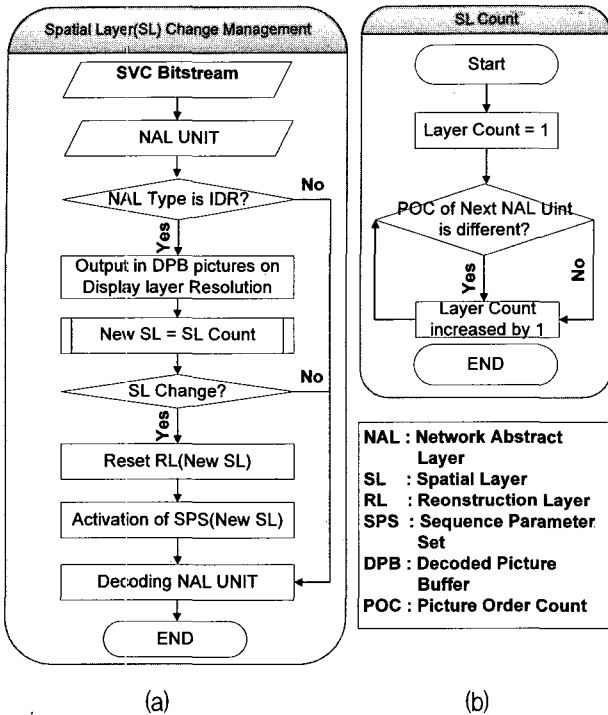


그림 14. 계층 변화 대응 알고리즘 : (a) 계층 변화 대응도, (b) 계층 조사도

Fig. 14. Spatial Layer Change Managing Algorithm : (a) Layer Change Managing Diagram, (b) Layer Check Diagram.

있으므로 이를 알기 위하여 Access Unit(AU)의 공간적 계층을 조사한다. 이 AU의 공간적 계층은 GOP단위의 공간적 계층과 동일하다. 현재의 공간적 계층과 이전까지의 공간적 계층이 다르다면 공간적 계층의 변화가 일어난 것이다. 이에 따라 복호화되는 계층의 최상위 계층을 나타내는 Reconstruction Layer를 설정하고 이에 맞추어 복원되는 영상의 해상도와 관련된 정보를 관리하는 SPS를 활성화시켜준다. 마지막으로 전송되는 비트스트림의 공간적 계층이 달라지더라도 실제 보여지는 영상의 해상도는 동일해야 하므로 전송 영상의 공간적 해상도에 상관없이 일정 해상도로 맞추어 주어야 한다.

4. 시스템

지금까지 실시간 스케일러빌리티를 지원하기 위한 부호기, 추출기, 복호기에서의 방법을 제안하였다. 제안된 방법으로 실시간 스케일러빌리티를 지원할 수 있어 이에 따라 폭 넓은 네트워크 대역폭에서의 적응 변환이 가능해졌다. 그림 15는 제안된 방법으로 구성된 전체 시스템을 나타낸다. 이와 같은 시스템을 통하여 네트워크 대역폭에 적응할 수 있도록 실시간 스케일러빌리티 지원이 가능하게 되었다.

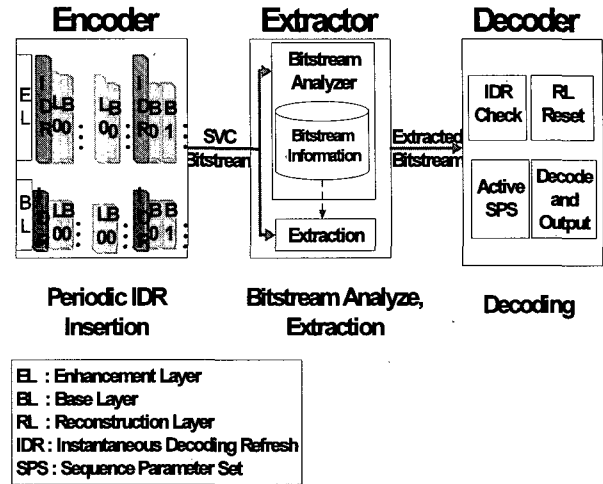


그림 15. 실시간 스케일러빌리티 지원을 위한 전체 시스템 구조

Fig. 15. System Structure for Real-time Scalability Supporting.

V. 실험

제안된 방법을 검증하기 위해 실험을 수행하였다. 첫 번째 실험은 기존의 SVC에서 구성된 방법과 본 논문에서 제안한 방법을 비교하여 제안된 방법의 타당성을 검증하였다. 두 번째 실험으로는 제안된 방법의 유효성을 보이기 위한 실험을 수행하였다. 전체 실험은 펜티엄 PC (3G dual, Ram 3G CPU)을 이용하였다.

1. 객관성 검증 실험

제안된 방법의 객관성 입증을 위하여 SVC의 참조 프로그램인 JSVM 4.0에서의 single loop decoding mode에서 공간적 계층을 낮추었을 때의 복원 결과와 제안한 방법을 추가로 JSVM 4.0에 구현하였을 때의 결과를 살펴보았다. 실험 환경은 표2와 같으며 그림 16과 17은 공간적 계층을 CIF에서 QCIF로 낮추었을 때의 결과이다.

그림 16은 본래의 JSVM4.0에서는 공간적 계층을 낮추었을 때 III장 1절에서 설명한 공간적 스케일러빌리티

표 2. 실험 환경

Table 2. Experiment Environment.

| 영상         | City      |
|------------|-----------|
| 각 계층별 해상도  | QCIF, CIF |
| 각 계층의 프레임율 | 30fps     |
| 각 계층의 FGS  | 3개의 FGS   |
| IDR 삽입 주기  | 3GOP      |
| 프레임 수      | 225       |

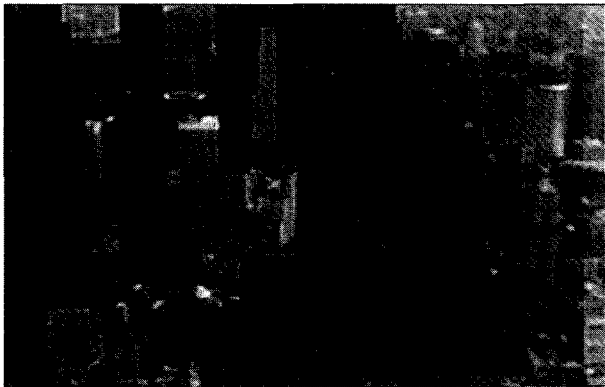


그림 16. 본래의 JSVM4.0에서 공간적 계층을 낮추었을 때의 복원 영상

Fig. 16. Spatial Layer Lowered Sequence with Original JSVM4.0.

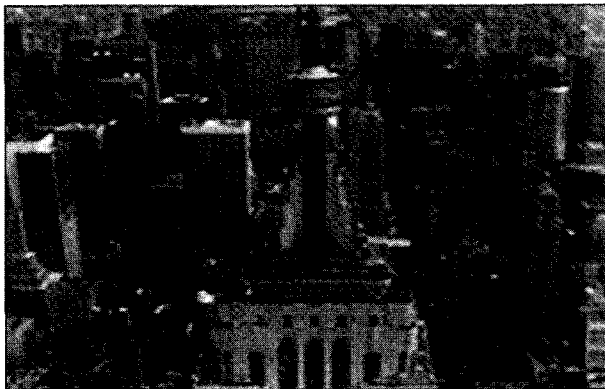


그림 17. 제안된 방법을 JSVM4.0에 구현하여 공간적 계층을 낮추었을 때의 복원 영상

Fig. 17. Spatial Layer Lowered Sequence with Proposed Methods Implemented JSVM4.0.

의 지원 문제에 따라 16과 같이 제대로 영상이 복원되지 않는 것을 볼 수 있다. 이에 반해 제안된 방법을 사용한 결과인 그림 17은 같은 상황에서 제대로 복원이 되어 제안한 방법이 제대로 동작함을 확인할 수 있다.

## 2. 유효성 검증 실험

제안된 방법의 효과를 보이기 위하여 표3의 실험환경에서 그림 18에 나타난 대역폭에서의 적응 실험을 수행하였다.

다음 3가지 경우에 대하여 추출된 데이터의 비트레이트, 스케일러빌리티, PSNR을 관찰하였다. 본래의 SVC에서도 품질적 스케일러빌리티는 FGS를 통해 실시간으로 변경할 수 있으므로 이를 CASE1으로 하여 품질적 스케일러빌리티만으로 적응 변환하는 경우와 제안된 방법을 통해 시간적, 품질적 스케일러빌리티로 적응 변환하는 경우를 CASE2, 공간적, 시간적, 품질적 스

표 3. 실험 환경

Table 3. Experiment Environment.

| 영상         | City      |
|------------|-----------|
| 각 계층별 해상도  | QCIF, CIF |
| 각 계층의 프레임율 | 30fps     |
| 각 계층의 FGS  | 3개의 FGS   |
| GOP의 크기    | 32 frames |
| IDR 삽입 주기  | 3GOP      |
| 프레임 수      | 225       |

케일러빌리티 모두 이용하여 적응 변환하는 경우를 CASE3로 하였다. 표4는 각 경우의 추출 지점 (d, t, f)로써 나타내고 있다. d는 dependency\_id로써 공간적 계층을 나타내고 t는 temporal\_level로써 프레임율을 나타내고 f는 quality\_layer로써 FGS계층을 나타낸다. d값 0은 QCIF, 1은 CIF를 나타내며 t 5, 4, 3은 각기 30, 15, 7.5fps를 나타낸다. f는 FGS의 양으로써 3개의 FGS를 사용하였기 때문에 최대 값은 3이다.

네트워크 대역폭 NB와 기기의 지원 가능한 해상도 DR에 대하여 만족될 수 있는 추출지점 d, t, q의 여러 조합이 나타날 수 있다. 이 때 최적 d, t, q의 지점을 찾아내는 것은 본 논문의 범위를 벗어나는 것이다. 따라서 여기서는 다음과 같은 간단한 비트레이트 추출 알고리즘을 사용한다. 우선 품질적 스케일러빌리티의 감소를 통해서 비트레이트에 적응할 수 있는지를 살펴 보기 위해 FGS NAL의 cropping 또는 drop을 통해 NB에 만족할 수 있는지를 살펴본다. 만약 품질적 적응만으로 주어진 NB에 적응할 수 없는 경우에는 시간적 스케일러빌리티를 낮추어 비트레이트에 적응시킨다. 이로서도 만족할 수 없는 경우에 공간적 계층을 감소시킨다. 이러한 계산은 시간적 계층을 지원하는 계층적 B 픽처가 GOP단위로 이루어지므로 이를 지원하기 위고, 네트워크 대역폭을 쉽게 추정하기 위해 GOP단위로 이루어진다. CASE3의 추출시 15fps를 최소 프레임율로 하였고 그 이하로 낮추어야 할 경우에는 공간적 해상도를 낮추도록 하였다.

요구된 대역폭에 따른 각 경우의 추출 비트레이트는 그림 18과 같다. 대역폭은 GOP 3~5구간을 제외하면 요구되는 비트레이트에 맞게 추출되고 있다. 그러나 GOP 3~5구간을 살펴보면 CASE1의 경우는 FGS를 모두 낮추었음에도 불구하고 추출된 데이터의 비트레이트가 주어진 대역폭을 넘어서는 것을 볼 수 있다. 이는 비디오 스트리밍 서비스 자체가 불가능한 경우이다. CASE2에서는 GOP 3~5구간에서 주어진 대역폭에 적응 가능한 추출 데이터량을 가지는 것을 볼 수 있다. 그러



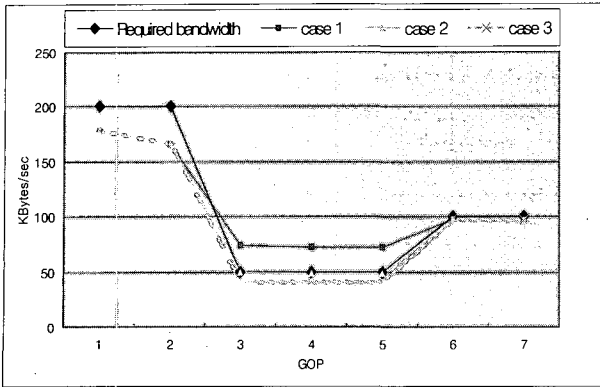


그림 18. 주어진 대역폭과 적응 데이터량  
Fig. 18. Given Bandwidth and Adapted Data Amount.

표 4. 각 경우의 추출 지점  
Table 4. Extraction Points for Each Case.

| Interval<br>(GOP Number) | Extraction point |           |           |
|--------------------------|------------------|-----------|-----------|
|                          | CASE 1           | CASE 2    | CASE 3    |
| 0 - 2                    | 1, 5, 3.0        | 1, 5, 3.0 | 1, 5, 3.0 |
| 3 - 5                    | 1, 5, 0.0        | 1, 3, 0.0 | 0, 5, 3.0 |
| 6 - 7                    | 1, 5, 1.1        | 1, 5, 1.1 | 1, 5, 1.1 |

나 대역폭을 맞추기 위해서 FGS를 모두 낮추었고 표 4에서 보는 바와 같이 프레임율을 7.5fps까지 떨어져야 하는 것을 볼 수 있다.

GOP 3~5 구간에서 FGS 계층 차이 때문에 CASE2와 CASE3의 PSNR은 다른 것을 볼 수 있다. 그러나 CASE2와 CASE3의 공간적 해상도가 다르기 때문에 직접적인 PSNR값 비교는 의미가 없다. 따라서 CASE3의 경우 업샘플링을 통하여 영상의 해상도를 같게 하여 PSNR 28.3dB를 얻을 수 있었다. 결과 화면인 그림 19와 20을 비교해 보았을 때 CASE2의 화질이 더 좋은 것을 볼 수 있다. 그러나 CASE3은 공간적 해상도를 낮추는 대신 프레임율을 유지할 수 있는 것을 볼 수 있다.

영상의 PSNR과 프레임율에 대한 주관적인 화질 연구<sup>[16]</sup>에 따르면 빠른 영상일 경우, 어느 정도의 화질 열화까지는 프레임율을 높게 유지하는 쪽을 더욱 선호한다고 조사된 바 있다. CITY영상은 전체적으로 빠른 모션을 가지므로 빠른 영상으로 볼 수 있으며 7.5fps의 영상을 관찰해볼 경우 끊어짐 현상이 일어나 깜빡임 현상이 심한 것을 볼 수 있다. 이에 따라 CASE2의 결과보다 CASE3의 경우에서 주관적 화질 만족도가 높을 것을 알 수 있다.

제안된 방법을 통하여 서술한 바와 수 있는 것을 볼 수 있었다. 일반적으로 임의 접근같은 이득을 얻을



그림 19. GOP 3에서의 CASE 2 영상  
Fig. 19. CASE 2 Sequence in GOP 3.



그림 20. GOP 3에서의 CASE 3 영상  
Fig. 20. CASE 3 Sequence in GOP 3.

표 5. 각 경우의 복원된 PSNR  
Table 5. PSNR for Each Case.

| Interval<br>(GOP Number) | CASE 1 | CASE 2 | CASE 3 |
|--------------------------|--------|--------|--------|
| 0~2                      | 40.1   | 40.1   | 40.1   |
| 3~5                      | 31.3   | 31.8   | 38.1   |
| 6~7                      | 33.8   | 33.8   | 33.8   |

(random access)과 전송중의 에러에 대비하기 위해 보통 0.5~2초 간격으로 IDR을 삽입하므로 이런 IDR 삽입시 제안된 방법으로 구현한다면 0.5~2초 간격에서 허용되는 실시간 변환에 대해 추가적으로 늘어나는 데이터양이 존재하지 않는다. 실험은 3GOP마다 IDR을 삽입하여 약 3초마다의 IDR 삽입이므로 추가적인 데이터량이 없다. 또한 표3의 환경으로 부호화된 비트스트림에 대하여 원본 JSVM4.0과 추가로 구현한 JSVM4.0에서의 복호 시간을 측정하였다. 이의 결과로 원본 JSVM4.0에서는 37.941697 sec, 제안 방법을 추가로 구현한 JSVM4.0에서는 37.942476 sec가 측정되었다. 이의 차이는 0.000779 sec으로 나타났다.

## VI. 결 론

본 논문에서는 SVC의 실시간 스케일러빌리티를 지원하기 위한 방안을 제안하였다. 이를 위하여 부호화시에는 주기적으로 IDR NAL을 삽입하도록 하였고 추출기에서는 미리 비트스트림을 분석하여 실시간 스케일러빌리티 지원을 위해 필요한 정보를 저장하여 사용하도록 하였다. 복호화시에는 실시간 스케일러빌리티 지원을 위해 Reconstruction 계층의 설정, SPS의 활성화와 같은 일련의 과정을 정의하였다. 실험 결과를 분석해보았을 때 제안된 방법의 유효성을 검증할 수 있었다. 추가적으로 추출 지점의 결정에 대한 연구가 필요하다.

## 참 고 문 헌

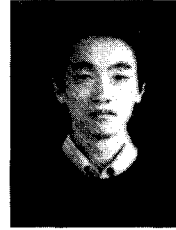
- [1] Chang, S.F., and Vetro, A., "Video adaptation: concepts, technologies, and open issues," *Proceedings of the IEEE*, Vol. 93, Issue 1, pp. 148-158, Jan. 2005.
- [2] T. Stockhammer and M. M. Hannuksela, "H.264/AVC video for wireless transmission," *IEEE wireless communication*, vol. 12, pp. 6-13, Aug. 2005.
- [3] A. K. Katsaggelos et al., "Advances in efficient resource allocation for packet-based real-time video transmission," *Proceedings of the IEEE, Special Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, pp. 135-147, Jan. 2005.
- [4] Ohm, J.R., "Advances in scalable video coding," *Proceedings of the IEEE*, Vol. 93, Issue 1, pp. 42-56, Jan. 2005
- [5] H. Schwarz, D. Marpe, and T. Wiegand, "MCTF and Scalability Extension of H.264/AVC," *Proc. of PCS*, SanFrancisco, CA, USA, Dec. 2004.
- [6] E. Francois, J. Vieron, and G. Boisson, "MPEG SVC: Why a New Video Coding Standard?," *Broadband Europe 2004 Conference*, Session 10, pp. 1-5, Brugge, Belgium, Dec. 2004.
- [7] M. Mrak, M. Grgic, and S. Grgic, "Scalable Video Coding in Network Applications," *Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp. 205-211, June 2002.
- [8] ISO/IEC JTC 1/SC 29/WG 11, "Working Draft 4 of ISO/IEC 14496-10:2005/AMD3 Scalable Video Coding," N7555, Nice, France, Oct. 2005.
- [9] W. H. Peng, C. Y. Tsai, T. Chiang and H. M. Hang, "Advances of MPEG Scalable Video Coding Standard," *KES Int'l Workshop on Intelligent Information Hiding and Multimedia Signal Processing*, Melbourne, pp. 889-895, Sep. 2005.
- [10] W.Li, "Overview of Fine Granular Scalability in MPEG-4 Video standard," *IEEE Trans. Circuits Syst. Video Techno.*, 11(3), Vol. 11, no. 3, Mar. 2001.
- [11] ISO/IEC JTC 1/SC 29/WG 11, "Joint Scalable Video Model (JSVM) 4.0 Reference Encoding Algorithm Description," N7556, Nice, France, Oct. 2005.
- [12] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical B pictures and MCTF," *ICME*, Toronto, Canada, Jul. 2006.
- [13] Schwarz, H, Marpe, D., Schierl, T., and Wiegand, T., "Combined scalability support for the scalable extension of H.264/AVC," *Multimedia and Expo, ICME 2005. IEEE International Conference*, pp. 6-8 July, 2005.
- [14] Schwarz, H, Hinz, T., Marpe, D., and Wiegand, T., "Constrained inter-layer prediction for single-loop decoding in spatial scalability," *IEEE International Conference on Image Processing*, Vol. 2, 11-14 pp. 870-873, Sept. 2005.
- [15] ISO/IEC JTC 1/SC 29/WG 11, "Text of ISO/IEC FDIS 14496-10: Advanced Video Coding 3rd Edition," N6540, Redmond, July 2004.
- [16] Yong Ju Jung, and et. El, "Analysis of Human Perception for Semantic Concept-based Video Transcoding," *IWATE 2005*, pp. 251-256, 2005.

저 자 소 개



이 동 수(학생회원)  
 2005년 아주대학교 정보 및 컴퓨터공학부 학사.  
 2006년 한국정보통신대학교 공학 석사.  
 2006년~현재 동 대학원 박사과정

<주관심분야 : SVC, Video Codec, AI>



배 태 면(학생회원)  
 1996년 경북대학교 전자공학과 학사.  
 1998년 경북대학교 전자공학과 석사.  
 2000년 8월 경북대학교 박사과정  
 2000년 9월~2001년 3월 Togabi 기술 연구원

2001년 3월~2006년 3월 서울대학교 컴퓨터기술 연구소 연구원

2002년 9월~현재 한국정보통신대학교 박사과정  
 <주관심분야 : semantic video analysis, SVC, Video Adaptation, Watermarking>



노 용 만(정회원)  
 1985년 연세대학교 전자공학과 학사.  
 1987년 KAIST 전기공학과 석사.  
 1992년 KAIST 전기공학과 박사.  
 1992년~1995년 Dept. of Radiological Science University of California Irvine 초빙연구원

1996년 Dept. of Electronical Eng. and Computer Science, University of California, Berkeley 연구원

1997년~현재 한국정보통신 대학교 정교수  
 <주관심분야 : 이미지/비디오 처리 및 분석, MPEG-7, 특징인식, 이미지/비디오 인덱싱>