

# 디지털 시네마 영상의 최적 압축을 위한 H.264/AVC와 JPEG-2000 비교 연구

\*김철현, \*\*강남오, \*\*\*김용환, \*\*\*\*정수연, \*\*\*\*\*임상희, \*\*\*\*\*백준기

중앙대학교 첨단영상대학원 시각 및 지능시스템연구소

\*windows286@gmail.com

## Optimum Compression of Digital Cinema by Comparing H.264/AVC and JPEG-2000

\*Kim, Chul-Hyun \*\*Knag, Nam-Oh \*\*\*Kim, Yong-Hwan \*\*\*\*Jung, Soo-Yeun \*\*\*\*\*Lim, Sang-Hee \*\*\*\*\*Paik, Joon-Ki

Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University

### 요약

현재 디지털 시네마의 압축 표준으로는 DCI(Digital Cinema Initiatives)에 의해 정지 영상 압축의 국제 표준인 JPEG-2000이 선정되어 있다. 하지만 영상 압축에 있어 또 하나의 국제 표준인 H.264/AVC는 현재까지 동영상 압축에 있어 과거 어떤 동영상 압축 코덱보다 큰 효율을 달성해왔다. 본 논문에서는 이 두 가지 압축 방법을 사용하여 디지털 시네마 압축에서 요구하는 시각적 무손실 압축을 연구하였다. 즉 실험 영상에 다양한 파라미터 값으로 JPEG-2000과 H.264/AVC 압축을 수행 하였고 산출된 결과를 시각적 무손실 압축의 환경 조건하에서 제시하였다. JPEG-2000의 경우 RATE 0.02, 0.03 이전의 PSNR 변화 값이 낮은 편이었으며, DCI 규격에도 만족시키는 결과를 보여주었다. H.264의 경우 압축 성능에 있어서 JPEG-2000보다 다소 우월한 결과를 보여주었다. 계층적 코딩을 통해 매체간 높은 호환성을 보장하는 H.264/AVC의 장점은 디지털 시네마의 배포에 적용 가능성을 함께 고려할 때 JPEG-2000과 함께 디지털 시네마를 위한 압축 표준으로 고려해 볼 만 하다.

### 1. 서론

HD영상 제작의 본격화와 더불어 촬영 및 디스플레이 장비의 비약적인 발전으로 말미암아 디지털 시네마의 현실화가 뜨겁게 논의 중이다. 기술적인 면에서 볼 때, 현재 카메라의 경우 1080p 촬영을 지원하는 다양한 제품이 발표되었으며, 해상도로는 4K까지 촬영 가능한 디지털 시네마 카메라가 상용화되었다. 그리고 디스플레이 장치에 있어서도 2K를 넘어 4K까지 지원하는 제품이 판매되고 있다. 하지만 이에 비해 콘텐츠를 제작하는데 필요한 압축, 배급, 패키징 기술은 상대적으로 빈약하며, 특히 고해상도 데이터의 효율적인 압축은 디지털 시네마의 현실화를 위해서 속히 풀어야할 문제로 대두되고 있다. 본 논문에서는 현재 제시된 압축 기술들을 간의 비교 실험을 통하여 고용량 고해상도의 디지털 시네마 압축에 적합한 기술 및 요소를 제시한다.

먼저 디지털 시네마라 함은 촬영-편집-배급의 전 과정이 디지털화되는 것을 의미한다. 이중 압축은 편집 후 배급에서 반드시 필요한 과정이며 기존의 압축 연구와의 차이점은 2K급 이상의 고해상도 영상을 압축해야 한다는 특징이 있다.<sup>1)</sup> 대부분의 영화 제작자는 HD로 촬영하고 후반 작업을 하는 동안 무압축을 선호하며 실제 작업도 무압축으로 이루어지고 있다.<sup>[4]</sup> 하지만 스크리닝에 있어서는 문제가 다르다. 현재 DCI(Digital Cinema Initiatives)는 4K해상도를 가진 디지털 시네마를 적극 권장하고 있다. 이 기준에 따라서 1.896:1의 스크린 비

를 가지는 디지털 시네마의 해상도는 4,096 × 2,160에 이르며 한 픽셀당 12-bit depth를 가질 경우 한 프레임은 318,504,960bit에 이르게 된다. 이는 한 프레임당 약 38메가에 달하는 용량이다. 2시간 영화가 172,800프레임 (24프레임 × 60초 × 120분)을 가진다고 했을 경우 약 6.4TB의 용량을 가지게 된다. 오늘날의 컴퓨터 환경에서 6.4TB는 그렇게 큰 용량이 아니며 전문적인 후반 작업실에서는 충분히 보유할 만한 용량이다. 하지만 모든 극장이 6.4TB의 용량을 가지기는 어려운 것이 현실이다. 또한 멀티 상영관이 주류를 이루는 오늘날 스크린 환경을 고려해 볼 때, 10개의 상영관을 보유한 극장의 서버는 최소 64TB 이상을 확보해야 한다. 이는 현실적으로 확보하기 힘든 서버 용량이다. 따라서 디지털 시네마 배급을 위한 최종 DCDM(Digital Cinema Distribution Master) 단계에서 배급을 효율적으로 하기 위해서는 압축 기술이 반드시 필요하다.

현재 디지털 시네마의 압축 표준으로는 정지영상의 국제 표준인 JPEG-2000이 DCI에 의해서 선정되었다.<sup>[5]</sup> JPEG-2000은 이산 웨이블릿 변환을 기반으로 하며 8×8블록 이산 코사인 변환을 기반으로 한 기존의 JPEG보다 훨씬 더 뛰어난 성능을 보이고 있다. 동일한 화질을 고려했을 때 JPEG보다 더 작은 용량의 압축된 파일을 제공한다. 손실 압축과 무손실 압축이 동일 코딩 아키텍처에서 제공 가능하다는 것, 압축된 형태의 비트스트림에서 영상의 특정 영역에 임의로 접근 가능하다는 것 등은 JPEG-2000이 가진 두드러진 특징들이다. 그러나 영상 압축의 또 하나의 국제 표준인 H.264/AVC 또한 고압축 고효율을 달성해 왔다. 추가된 High 프로파일은 HD급 해상도에서 화질에 대한 높은 신뢰도를 유지한다. 압축된 형태의 비트스트림을 사용자의 특성에 맞게 화질을 조절할 수 있는 계층적 코딩 (Scalable

1) 본 논문에서 말하는 2K, 4K영상은 Digital Cinema Initiatives(DCI)에서 정한 규격을 의미한다. 2K는 2,048 × 1,080 픽셀이며, 4K는 4,096 × 2,160이다.

coding)을 통해 HD급에서 mobile 전송을 위한 영상까지 한 번의 압축으로 사용 가능한 점은 컨버전스 환경에서 높은 호환성을 보장한다.

본 논문은 고해상도 영상압축을 필요로 하는 디지털 시네마에서 H.264/AVC와 JPEG-2000 성능을 비교 분석하고 적절한 압축의 범위를 제시하는 것이 목적이다. 본 논문의 구성은 다음과 같다. 2장에서 압축의 필요성, 시각적 무손실 압축의 개념을 설명하고 압축 표준으로 JPEG-2000과 H.264/AVC를 각각 소개한다. 3장에서는 실험을 통해 두 압축 방법의 압축 성능을 비교하고, 4장에서 본 논문의 결론을 맺는다.

## 2. 디지털 시네마 압축 관련 기술

### 가. 시각적 무손실 압축

압축에 있어서 먼저 고려되어야 할 부분은 손실 압축(lossy coding)을 할 것이냐 무손실 압축(lossless coding)을 할 것이냐 라는 문제이다. 무손실 압축은 정보를 압축했다 재생해도 원래 정보에는 전혀 손실을 주지 않는 압축 방식이다. 하지만 영상에 있어서 무손실 압축은 복잡한 영상일수록 압축률이 낮아진다는 약점이 있다. 물론 압축을 하지 않을 때와 비교해 보면 무손실 압축을 통해 많은 용량이 줄어들지만 디지털 시네마에서 실제 배급 가능할 정도의 용량으로는 줄어들지 않는다.



그림 1. 원본 영상 중 일부



그림 2. JPEG-2000을 이용해 무손실 압축한 영상

그림 1은 1920 × 1080, 8-bit depth를 가진다. 영상 원본의 크기는 6,220,856 Byte(약 5.9MB)이며, JPEG-2000을 이용해 무손실 압축을 할 경우 파일 사이즈는 1,975,847 Byte(약 1.9MB)로 줄어들어, 약 310%정도의 압축률을 보인다. 앞서 계산한 4K 시네마 한편에 적용할 경우 을 1:1 무손실 압축할 경우 약 2.1TB의 용량으로 압축된다. 2.1TB는 현재 컴퓨터 시스템에서 배급용으로 사용하기에는 현실적으로 어려운 용량이다. 특히 배급에 있어서 광대역 인터넷망과 위성 등을 사용한다고 했을 경우 불가능하다고 판단된다. 따라서 디지

털 시네마 배급에 있어서는 영상 원본의 손상을 감수하는 손실압축이 불가피하다. 손실압축이 영상의 질을 손상시키지만, 인간의 눈이 인식하는 부분을 감안하여 압축을 할 경우 큰 문제가 되지 않는다. 그리고 초당 24프레임에서 48프레임까지 투자는 디지털 시네마에서 실제 손실의 차이를 눈으로 인지하는 것은 거의 불가능하다.[4]



그림 3. JPEG-2000을 이용해 20:1 손실 압축한 영상

그림 3은 20:1로 영상을 압축한 예이다. 원래 6,220,856Byte였던 영상이 310,869Byte로 줄어들었다. 대략 2,001%에 달하는 압축률이다. 하지만 사람의 육안으로 영상의 차이를 구별하는 것은 거의 불가능하다. 이 정도의 압축을 거칠 경우 최초 6.4 TB에 달하는 디지털 시네마 한편이 약 350MB 정도로 줄어들게 되며, 위성, 광대역 인터넷 등을 이용하는 디지털 시네마 보급에 적절한 수준이 될 것이다.



그림 4. JPEG-2000을 이용해 1000:1 손실 압축한 영상

하지만 손실압축을 심하게 할 경우 영상의 화질이 위의 그림 4처럼 심하게 저하될 수 있다. 위의 그림은 원래 영상을 1000:1압축을 실시한 결과이다. 크기는 6,183 Byte로 줄어들었지만 영상이 심하게 훼손되어 사용이 불가능하다. 본 논문에서는 이 같은 시각적 무손실 압축의 범위를 찾는 것이 목표이며, 효과적인 실험을 통해 JPEG-2000과 H.264/AVC에서 적절한 압축률을 제시하고 비교하고자 한다.

### 나. JPEG-2000

DCI의 기술권고안인 DCI Specification에서 압축 표준으로 정한 JPEG-2000은 International Organization for Standardization (ISO)와 International Electrotechnical Commission (IEC)의 공동의 노력으로 만들어진 영상 압축을 위한 국제표준이다. MPEG-2나 H.264에서 사용하는 블록단위의 DCT변환이 아닌 웨이블릿 변환을 사용하는 JPEG-2000 표준은 총 12개의 파트들로 구성되어 있다. 12개의 파트들 중 Part 1은 JPEG-2000 표준 중에서 가장 핵심적인 코딩 시스

템이 기술 되어 있고, Part 3은 모션 JPEG-2000으로써 Part 1에 기술된 알고리즘을 사용한 동영상 압축에 관한 내용이 기술되어 있으며 이는 고화질의 프레임 기반 압축을 사용하는 디지털 시네마와 같은 응용을 목표로 하고 있다.

JPEG-2000 Part 1에는 JPEG-2000 표준 중에서 가장 핵심적인 코딩 시스템이 디코더 관점에서 기술되어 있다. 본 절에서는 JPEG-2000 Part 1에 기술 되어 있는 핵심 코딩 시스템을 인코더 관점에서 간략하게 살펴보고자 한다. 핵심 코딩 시스템의 전체적인 코딩은 전처리(pre-preprocessing), 압축(compression), 압축된 비트 스트림 구성(compressed bitstream formation) 등 크게 세 부분으로 나뉜다.[1,2]

### 1) 전처리

전처리는 선택적인 사항으로써 JPEG-2000을 사용하여 압축을 하는 응용에 따라 실행될 수도 있고 그렇지 않을 수도 있다. 전처리는 타일링(tiling), 디씨 레벨 쉬프팅(DC level shifting), 멀티컴포넌트 변환(multicomponent transformation)으로 구성되어 있다.

타일링은 코덱에서 사용하는 메모리의 크기가 압축할 영상의 크기보다 더 작을 때 사용하는 방법으로써 영상을 더 작은 사각형의 영상으로 나누는 작업이다. 이 때 각각의 사각형 영상을 타일(tile)이라 부르며, 각각의 타일은 독립적으로 코딩된다.

보통 영상의 픽셀들은 부호가 없는 정수로 표현되어 있다. 그러나 수학적인 계산을 하기 위해선 각각의 픽셀들은 부호가 있는 정수로 표현되어야 하며, 이를 위해 디씨 레벨 쉬프팅 처리를 하게 된다. 디씨 레벨 쉬프팅은 각각의 픽셀 값에  $2^{P-1}$ 를 차분함으로써 구현된다. 여기서 P는 하나의 픽셀 값을 표현하기 위해 사용된 비트수를 의미한다.

컬러 영상의 경우 보통 3가지 컴포넌트(예를 들면 R, G, B)로 구성되어 있다. 이런 컬러 영상은 압축하기 전에 컴포넌트간의 상관성을 줄이기 위해 멀티컴포넌트 변환을 거치게 된다. JPEG-2000 Part1에서는 멀티컴포넌트 변환을 위해 2가지 변환을 제공하고 있다. 그 중 하나는 가역 컬러 변환(Reversible Color Transform)이고 나머지 하나는 비가역 컬러 변환(Irreversible Color Transform)이다. 가역 컬러 변환은 손실 압축(lossy compression)과 무손실 압축(lossless compression) 모두에 사용될 수 있는 반면, 비가역 컬러 변환은 손실 압축에서만 사용할 수 있다. 가역 컬러 변환과 비가역 컬러 변환의 순방향, 역방향 변환 공식은 다음과 같이 주어진다.[6]

<가역 컬러 변환>

\*순방향 변환 공식 식(1)

$$Y_r = [R+2G+B]/4$$

$$U_r = B-G$$

$$V_r = R-G$$

\*역방향 변환 공식 식(2)

$$G = Y_r - [U_r + V_r]/4$$

$$R = V_r + G$$

$$B = U_r + G$$

<가역 컬러변환>

\*순방향 변환 공식 식(3)

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299000 & 0.587000 & 0.114000 \\ -0.168736 & -0.331264 & 0.500000 \\ 0.500000 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

\*역방향 변환 공식 식(4)

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299000 & 0.587000 & 0.114000 \\ -0.168736 & -0.331264 & 0.500000 \\ 0.500000 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

## 2) 압축

전처리 과정을 거친 후 각각의 타일은 독립적으로 압축된다. 만약 원래의 영상이 컬러 영상일 경우, 즉 영상이 3가지 컴포넌트(R, G, B)를 지니고 있을 경우 각각의 타일 또한 3가지 컴포넌트를 가지고 있으며 3가지 컴포넌트는 멀티컴포넌트 변환을 거치게 되고, 변환된 3가지 컴포넌트는(Y, Cb, Cr) 각각 독립적으로 압축된다. 즉, 하나의 컴포넌트만 지닌 영상(예를 들면, 흑백영상)일 경우 타일링 과정을 거쳐 발생된 타일들은 각각 독립적으로 압축과정을 거치고, 여러 컴포넌트를 지닌 영상(예를 들면, 컬러영상)일 경우 타일링 과정을 거쳐 발생된 타일들은 다시 멀티컴포넌트 변환 과정을 거친 후, 발생되는 컴포넌트들은 각각 독립적으로 압축과정을 거치게 된다.

압축과정은 이산 웨이블릿 변환(Discrete Wavelet Transform), 양자화(Quantization), 엔트로피 코딩(Entropy Encoding)으로 크게 세 부분으로 나눌 수 있다.[6]

### 이산 웨이블릿 변환

JPEG에서는 8x8 블록 단위의 이산 코사인 변환(Discrete Cosine Transform)을 사용하는 반면, JPEG-2000에서는 이산 웨이블릿 변환을 사용한다. 영상 압축에 사용되는 2차원 이산 웨이블릿 변환은 1차원 웨이블릿 변환을 영상의 열 방향으로 적용한 데이터에 다시 행 방향으로 1차원 웨이블릿 변환을 적용함으로써 얻어진다. 영상에 웨이블릿 변환을 1번 적용시키면 영상은 4개의 부밴드(subband)로 분해된다. 4개의 부밴드는 LL1, LH1, HL1, HH1이고 LL1은 원 영상을 가로와 세로 방향으로 2:1로 샘플링 한 영상의 근사 영상을 나타낸다. 이 LL1 부밴드에 다시 2차원 웨이블릿 변환을 적용할 수 있으며 그 결과 LL1 부 밴드는 다시 4개의 부밴드로(LL2, LH2, HL2, HH2) 분해된다. LL2 부밴드에 한 번 더 2차원 웨이블릿 변환을 적용하면 LL2 부밴드는 다시 4개의 부밴드로 분해되며, 이러한 방식으로 계속해서 2차원 웨이블릿 변환을 적용할 수 있다. JPEG-2000 Part1에서는 하나의 타일 컴포넌트에 2차원 웨이블릿 변환을 최대 32회 적용할 수 있도록 하였다.

### 양자화

무손실 압축의 경우 양자화 과정을 거치지 않지만, 손실압축일 경우 이산 웨이블릿 변환 과정을 거쳐 발생된 모든 부밴드들은 양자화 과정을 거친다. JPEG-2000 Part 1에서는 원점 근처에 데드-존(dead-zone)이 있는 균일 스칼라 양자화(uniform scalar

quantization)를 사용한다. 스텝 사이즈가  $\Delta b$ 일 경우 데드-존의 길이는  $2\Delta b$ 가 되며, 각각의 부 밴드에 대해 각기 다른 양자화 스텝 사이즈를 사용할 수 있다.

### 엔트로피 코딩

무손실 압축의 경우 양자화 과정을 거치지 않은 데이터가 엔트로피 코딩되고, 손실 압축일 경우 양자화 과정을 거친 데이터가 엔트로피 코딩된다. 무손실 압축이든 손실 압축이든 엔트로피 코딩과정에 들어가기 전의 데이터는 부호가 있는 정수이다. 이 부호가 있는 정수는 엔트로피 코딩과정을 거치기전 부호-크기 표현으로 (sign-magnitude representation) 바뀐 후 엔트로피 과정을 거치게 된다. 엔트로피 코딩과정을 거치기 위해 각각의 부밴드들은 더 작은 블록(code-block이라고 불림)으로 나누어지게 되며, 코드 블록 별로 엔트로피 코딩과정을 거치게 된다. JPEG-2000에서 엔트로피 코딩과 압축된 비트스트림의 형성은 Tier-1 코딩과 Tier-2 코딩으로 나누어져 있다.

Tier-1 코딩에서는 각각의 코드 블록을 독립적으로 인코딩(encoding) 한다. 코드 블록은 비트 플레인 코딩(bit-plane coding)과 변종 이진산술 코딩(binary arithmetic coding)을 거쳐 압축된다. JPEG-2000은 비트 플레인 코딩을 위해 EBCOT(Embedded block coding with optimized truncation) 알고리즘을 채택하였다. 코드 블록은 비트 플레인(bit-plane)들로 분해되고 각각의 비트 플레인은 EBCOT를 거쳐 컨텍스트(context)와 이진 결정 값(binary decision value)을 얻는다. 컨텍스트들과 이진 결정 값들은 변종 이진 산술 코딩의 입력이 되어 압축된 코드들을 얻는다.

### 3) 압축된 비트스트림 형식

각각의 코드 블록들은 Tier-1 코딩을 거쳐 압축된 비트들로 만들어 지고, 각각의 코드 블록의 블록 요약 정보(block summary information)와 레이어(layer)는 Tier-2 코딩을 거쳐 효율적으로 표현된다. 레이어는 각각의 코드 블록에 대한 연속적인 비트 플레인 코딩 절차들로 구성되어 있으며, 블록 요약 정보는 코드 블록의 압축된 코드워드들의 길이, 비트 스트림 레이어들 사이에서의 절단 포인트 등으로 구성되어 있다.

### 다. H.264/AVC

MPEG-4/AVC 또는 H.264(이하 H.264/AVC)는 시각 정보의 부호화된 표현을 위한 표준안이다. 각 표준안은 기본적으로 두 가지를 정의하는 문서로서, 압축된 형태의 시각 데이터를 나타내는 부호화된 표현(또는 syntax)과 시각 정보를 복원하기 위해 디코딩하는 방법을 정의하고 있다.

H.264의 표준화 노력은 일련의 비디오 통신 표준안 시리즈를 개발했던 국제 통신 위원회(International Telecommunication Union; ITU)의 연구 그룹인 Video Coding Experts Group (VCEG)에 의해 시작되었다. H.264 표준안 개발의 최종 단계는 Joint Video Team(JVT)에 의해 수행되고 있으며, VCEG와 MPEG의 협동적인 노력으로 ISO/IEC (MPEG-4 Part 10)와 ITU-T (Recommendation H.264)의 공동 후원 하에 2003년 최종 표준안을 발표하는 것이 가능해졌다.

H.264 표준안은 압축 효율(이전의 어떤 표준안보다 훨씬 뛰어난 압축 성능을 제공한다), 전송 효율(내정되어 있는 여러 가지 기능을 사용하여 다양한 채널과 네트워크를 통한 신뢰성 있고 강력한 전송을 지원한다.), 그리고 인기 있는 비디오 압축 응용제품 생산에 초점을 맞추고 있다. 현재 4개의 프로파일이 지원되며, Baseline 프로파일은 화상회의와 같은 "대화를 위한" 응용분야에 특히 유용하고, Extended 프로파일은 네트워크를 통한 비디오 스트리밍에 유용한 추가적인 도구들을 포함하고 있으며, Main 프로파일은 비디오 방송 및 저장과 같은 소비자 응용제품에 적합한 도구들을 포함한다. 추가된 High 프로파일은 고해상도에서 높은 신뢰도를 얻기 위한 기능들을 포함하며 방송용 HD 시퀀스나 디지털 시네마에 적합하다.

H.264/AVC의 진보된 인트라 예측 알고리즘은 I-프레임의 압축 효율을 높이기 위해 제안되었다. 또한 High Profile에 소개된 8x8 DCT는 공간상에서의 높은 압축 효율뿐만 아니라 필름 입자를 보존하여 고해상도 영상과 필름 입자의 보다 디테일 한 표현을 가능하게 하였다. 그럼에도 불구하고 디지털 시네마의 압축 표준으로 제대로 된 논의가 이루어 지지 않은 것이 현실이다. 본 절에서는 디지털 시네마 압축 표준으로 적합한 H.264/AVC의 기술적 요소와 성능을 소개한다.[3]

### 1) H.264/AVC 주요기술

H.264를 설명하기 전에, 비디오 프레임의 인코딩 하는 주요 과정을 설명하고자 한다.

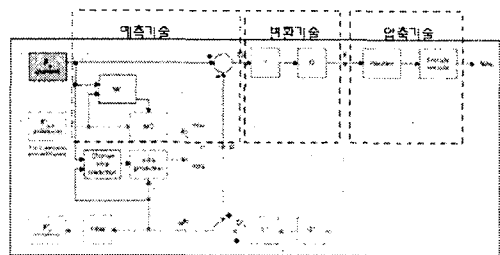


그림 5. H.264/AVC 인코더

입력 프레임 또는 필드  $F_n$ 은 매크로 블록 단위로 처리된다. 각 매크로 블록은 인트라 또는 인터 모드로 인코딩되며, 매크로 블록 내의 각 블록에 대한 예측 블록 PRED(그림 5에서 'P'로 표시)는 복원된 영상 샘플에 의해 생성된다. 인트라 모드에서 PRED는 현재 슬라이스에 존재하는 이전에 인코딩되고 디코딩되어 복원된 샘플로부터 생성된다. 예측 블록 PRED는 현재 블록과의 차이로 생성된 오차 블록  $D_n$ 을 생성하고, 오차(차이) 블록은 변환되고 양자화 되어  $X$ 로 표현되는 양자화 된 변환 계수가 되며, 변환 계수들은 스캔에 의해 재배치되고 엔트로피 인코딩된다. 엔트로피 인코딩된 계수들은 매크로 블록 내의 각 블록을 디코딩하는데 필요한 추가적인 정보들(예측 모드, 양자화 파라미터, 움직임 벡터 정보 등)과 함께 압축된 비트스트림을 형성하여 Network Abstraction Layer(NAL)를 통해 전송되거나 저장된다.

### 2) 예측, 변환, 양자화, 엔트로피 부호화

2) 매크로블록은 DCT 블록보다 큰 블록으로 정의된다. 보통 16x16 크기의 블록이 사용된다.

## 예측 코딩

예측 코딩이란 이전(또는 미래)프레임을 기반으로 현재 프레임의 영역에 대한 예측 데이터를 생성하고, 현재 영역으로부터 예측 영역을 빼서 오차 데이터를 만들어 내는 것을 의미한다. 예측이 성공적으로 수행되면 오차 데이터는 원본 프레임의 에너지보다 적게 되고, 오차 데이터는 원본 프레임보다 적은 비트로 표현될 수 있다. 예측은 시간적, 공간적 예측으로 나뉘는데 시간적 예측은 디지털 시네마를 위한 JPEG-2000과의 비교에 적합하지 않으므로 생략하기로 한다.

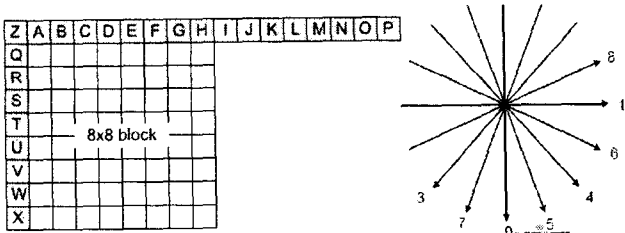


그림 6. H.264/AVC의 인트라 예측 모드

공간 예측(인트라 모드)에서 예측 블록 P는 이전에 인코딩 되고 복원된 블록에 의해 생성되고 인코딩하기 전에 현재의 블록으로부터 빼진다. 휘도 샘플의 경우 P는 각각의 4x4 블록 또는 16x16 매크로 블록에 대해 생성된다. 각각의 4x4 휘도 블록에 대해 9개의 선택 가능한 예측 모드가 존재하고, 16x16 휘도 블록과 색차 성분에 대해 각각 4개의 모드가 존재한다. 인코더는 일반적으로 각 블록에 대해 P와 인코딩 되는 블록 사이의 차이를 최소화하는 예측 모드를 선택한다. 그림 6의 오른쪽 화살표는 각 모드의 예측 방향을 나타낸다. 예측된 샘플은 예측 샘플 A-M의 가중치 평균으로부터 생성된다.

## 변환

영상 또는 비디오 코덱에서 변환 과정의 목적은 영상 또는 움직임 보상된 오차 데이터를 다른 영역 즉, 변환 영역(transform domain)으로 변환하기 위한 것이다. H.264는 코딩 될 오차 데이터의 종류에 따라 3가지 변환을 사용하는데 16x16 모드로 예측된 인트라 매크로 블록 내의 휘도 DC계수들의 4x4배열을 위한 Hadamard변환, (모든 매크로 블록 내의) 색차 DCT 계수들의 2x2배열을 위한 Hadamard변환 그리고 다른 모든 4x4 블록의 오차 데이터를 위한 DCT 기반 변환이 사용된다. H.264/AVC의 DCT 변환은 정수 변환으로 근사화된 버전이어서 과거 국제표준 코덱의 문제점 중의 하나였던 encoder/decoder mismatch 문제를 해결함과 동시에 연산을 +, -, shift로만 표현함으로써 변환의 연산량도 줄이는 효과를 가지고 있다.

High Profile 추가된 8x8 DCT는 HD 영상에서 필름의 입자를 보다 세밀하게 표현하는데 도움을 준다. 그림 7에서 보이는 것처럼 4x4 transform만 사용했을 때 보다 8x8을 함께 사용한(c)그림이 원본에 더 가까울 뿐만 아니라 압축율도 높은 것을 알 수 있다.[7]

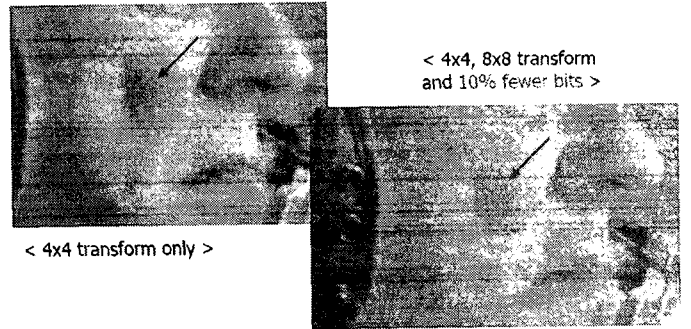
## 양자화

양자화는 어떤 범위의 값을 갖는 X를 보다 적은 범위의 값을 갖는 양자화 된 신호 Y로 대체한다. 양자화 된 신호는 원래의 신호보다 적은 범위의 값을 가지므로 보다 적은 비트로 표현이 가능해야 한다.

스칼라 양자화는 하나의 입력 샘플을 하나의 양자화 된 출력 값으로 대체하고, 벡터 양자화는 하나의 그룹의 입력 샘플을 양자화 된 값들의 그룹으로 대체한다.



(a) 원본 영상



(b) 4x4 transform 사용

(c) 4x4, 8x8 transform 사용

그림 7. HD 영상의 변환 비교 영상

H.264는 스칼라 양자화를 사용한다. 순방향 및 역방향 양자화의 방법은 나눗셈 그리고 부동 소수점 연산을 사용하지 않고 post-스케일링과 pre-스케일링 행렬에 통합시키기 위한 조건에 의해 복잡해졌다.

## 엔트로피 부호화

엔트로피 인코더는 비디오 영상의 요소를 나타내는 연속적인 부호를 전송하거나 저장하기 용이한 압축된 비트스트림으로 변환한다. 입력 부호는 양자화 된 변환 계수, 움직임 벡터, 마커, 헤더 그리고 부가적인 정보를 포함할 수 있다. H.264에서 사용하는 가변 길이 인코더는 입력되는 심볼들을 연속적인 코드워드(variable length codes 또는 VLC)로 변환한다. 각 심볼은 코드워드로 변환되는데, 코드워드의 길이는 가변적이지만 각각 정수개의 비트를 포함해야 한다. 자주 발생하는 심볼들은 짧은 VLC로 표현되고, 자주 발생하지 않는 심볼들은 긴 VLC로 표현된다. 매우 많은 수의 심볼들에 대해 이러한 인코딩 과정을 수행하면 데이터 압축이 발생한다. H.264/AVC에는 CAVLC (Context-based Adaptive Variable Length Coding)와 CABAC (Context-based Adaptive Binary Arithmetic Coding)라는 두 가지 엔트로피 방법을 제공한다. CABAC은 CAVLC에 비해서 30~40%의 연산량 증가에도 불구하고 10%정도의 압축률 개선이라는 효과를 가지고 있기 때문에, 고화질, 고압축을 요구하는 HDTV 등의 애플리케이션에서는 많이 이용되고 있다.

### 3) 계층적 코딩

인터넷 및 상이한 단말기를 통한 비디오의 전송이 늘어나면서 네트워크의 상황에 따른 적응적 비트율 제어 및 단말기의 기능에 따른 화질제어를 위한 스케일러빌리티(scalability) 기능의 중요성이 증대되고 있다. 비디오 데이터의 스케일러블 인코딩은 디코더가 인코딩된 비트스트림의 일부분만을 선택적으로 디코딩하는 것을 가능하게 해준다. 코딩된 스트림은 여러 개의 레이어로 배열되어 있는데, 'Base' 레이어와 하나 또는 그 이상의 'Enhancement' 레이어를 포함한다.



그림 8. 원본 영상



그림 9. Base 레이어로 인코딩되는 서브 샘플링 된 프레임



그림 10. Base 레이어 프레임(디코딩되고 업샘플링된 프레임)

그림에서처럼 디코더는 Base 레이어만 전송받아 '기본' 화질의 비디오 영상만을 디코딩할 수 있다. 또한 다른 디코더는 모든 레이어들을 전송받아 고화질의 영상을 디코딩한다. Scalable 비디오 코딩은 다양한 응용이 가능한데, 예를 들어 낮은 복잡도의 디코더는 Base 레이어만 디코딩할 수 있고 낮은 전송율의 비트스트림은 한정된 용량을 가지는 네트워크를 통한 전송을 위해 추출될 수 있으며 에러에 민감한 Base 레이어는 Enhancement 레이어보다 높은 우선순위를 가지고 전송될 수 있다.

### 3. JPEG-2000과 H.264/AVC를 이용한 시각적 무손실 압축 실험

#### 가. 실험환경

첫 번째 실험은 해상도 측정을 위한 ISO-12233차트를 사용하였다. 원래 디지털 카메라의 해상도를 측정하기 위해서 제작된 것으로 본 실험에서는 압축률에 따른 해상도의 변화를 측정하기 위해 사용하였다. 두 번째 실험은 DSClabs에서 제작된 CamAlign-BF-SW-v4.1 차트와 CamAlign-ChromaDuMonde 차트를 사용하였다. CamAlign-BF-SW-v4.1 차트는 카메라의 포커스 테스트용으로 만들어진 것이며, CamAlign-ChromaDuMonde 차트는 스킨톤 테스트를 위해서 만들어진 차트이다. 각각의 차트를 무압축 16bit TIFF형식으로 Canon D100카메라를 이용해 촬영하였으며 이후 JPEG-2000과 H.264/AVC로 압축하며 얼마나 원본 영상에 얼마나 손실이 일어나는지 측정하였다. 해상도 손실을 정확히 측정하기 위해서 Olympus HVRes V3.1을 이용해 측정하였으며, PSNR 측정을 같이 수행하였다.

JPEG-2000은 jasper-1.701.0 소프트웨어를 사용했다. jasper에서 TIFF를 지원하지 않는 관계로 RAS 포맷으로 변환하여 사용했다. RAS는 비트맵 영상의 무압축 포맷으로 칼라 정보 RGB가 BGR로 변환된다. H.264/AVC FRExt의 코딩을 위해 참고 소프트웨어는 JM 10.2 버전을 사용했다.

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \quad (\text{식 7})$$

PSNR(Peak Signal to Noise Ratio)(식 7.)은 로그 단위로 측정되며  $(2^n - 1)^2$  (영상에 존재할 수 있는 최대 샘플 크기의 제곱, 여기서 n은 영상 샘플의 비트수를 나타낸다.)과 원본 영상과 손상된 영상 또는 비디오 프레임 사이의 평균제곱오차(MSE: mean squared error)의 비율에 의해 결정된다.

#### 나. JPEG-2000실험

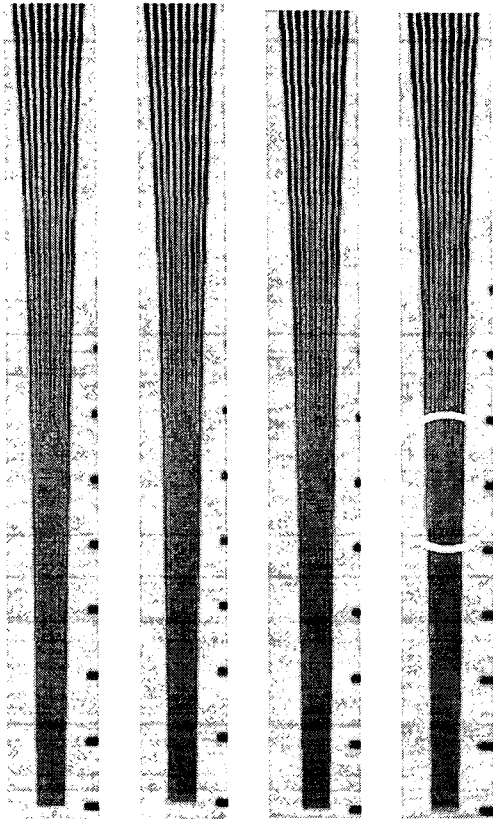
획득한 ISO-12233의 slant 해상도는 992로 나왔다. 각 rate별 slant 해상도는 표1과 같다.

표 1: ISO-12233 차트의 해상도 변화

input rate	Resolution of slant
(무손실)	992
0.200	992
0.100	992
0.090	992
0.080	992
0.070	992
0.060	992
0.050	992
0.040	992
0.030	988
0.020	991
0.010	920
0.009	921
0.008	921
0.007	917
0.006	831
0.005	low
0.004	low
0.003	low
0.002	-
0.001	-

- : 미실험, low : 측정불가

해상도 변화는 rate 0.04까지는 손실 없이 나타났다. 이후 0.02까지는 큰 변화가 없다. 실제 육안으로 확인 시에도 0.02까지는 큰 변화를 느끼기 어렵다.



(a) 무손실 (b) rate 0.04 (c) rate 0.02 (d) rate 0.01  
그림 11: ISO-12233의 세로 해상도 라인의 변화

그림 11에서 보듯이 시각적인 판단으로 볼 경우 rate 0.02까지 큰 훼손은 없어 보인다. 그러나 0.01부터는 뭉침 현상이 일어나면서 영상에 왜곡이 뚜렷이 일어나기 시작하였다. 그림 11 (d)의 원내부에 보면 뭉침 현상을 육안으로 확인할 수 있다. PSNR측정은 표2와 같다.

PSNR값은 압축이 일어나면서 계속 낮아지기 시작한다. 그러나 각 단계마다 변화폭은 균일하지 않았다. 특정한 rate값에서 손실이 많이 일어났다.

표 3을 보면 0.02에서 0.01로 변할 때 가장 큰 PSNR의 변화가 보이는 것을 알 수 있다. 그러므로 0.02이하의 압축은 데이터양의 변화에 비해 많은 손실을 가져오는 것을 확인할 수 있다. 그러나 이러한 PSNR의 변화는 고주파성분이 많은 ISO-12233차트와 CamAlign-BF-SW-v4.1 차트에서는 크게 나타났지만 CamAlign-ChromaDuMonde 차트에서 심하게 나타나지 않았다.

표 2: PSNR 측정값

input rate (무손실)	R	G	B
	inf	inf	inf
0.200	53.40	56.05	53.65
0.100	47.08	48.15	47.13
0.090	46.59	47.79	46.59
0.080	46.03	47.38	45.91
0.070	44.98	46.15	44.92
0.060	44.31	45.30	44.17
0.050	43.52	44.74	43.35
0.040	42.17	43.48	42.02
0.030	40.57	42.23	40.60
0.020	38.08	39.99	38.37
0.010	33.12	35.19	33.86
0.009	32.44	34.36	33.17
0.008	31.64	33.31	32.24
0.007	31.03	32.12	31.29
0.006	30.23	30.90	30.29
0.005	29.13	29.72	29.33
0.004	28.05	28.24	28.00
0.003	26.47	26.10	25.99
0.002	24.56	23.65	23.67
0.001	21.70	20.58	20.66

표 3: ISO-12233 차트의 PSNR 변화량

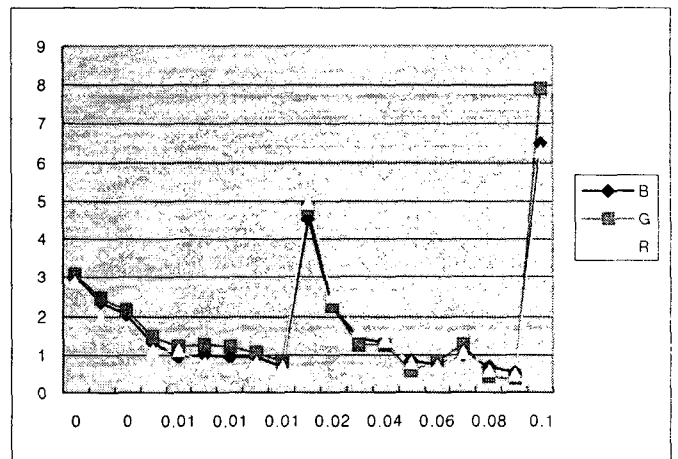
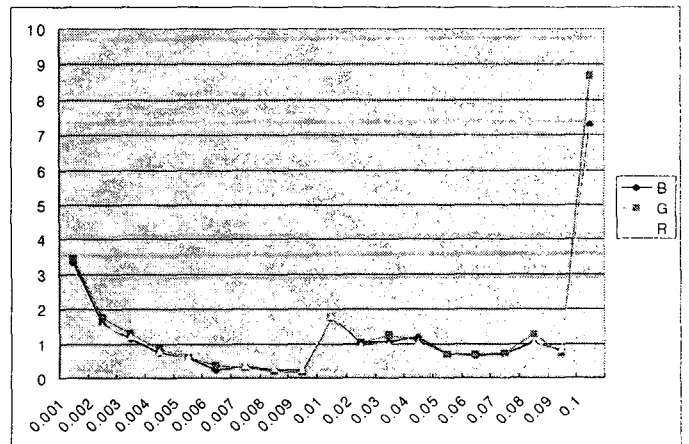


표 4: CamAlign-ChromaDuMonde 차트의 PSNR 변화량



### 3-3. H.264실험

H.264/AVC 실험은 표 5와 같은 환경에서 수행되었다.

표 5: H.264/AVC 실험환경

profile	High 4:4:4 Intra
Input file format	yuv
SATD(Hadamard)	On
RDO	On (fast algorithm)
BitDepthLuma	8bit
BitDepthChroma	8bit
QP	0-36(interval 2)
Coding mode	Common mode
Encoder	JFVM 11.1(JM9.5기반)

JPEG-2000은 압축률(rate)값에 따라서 진행되지만 H.264의 경우 QP값에 따라서 압축률이 달라진다. 해상도의 경우 JPEG-2000과 달리 우수한 값을 보였다.

표 5: ISO-12233 차트의 해상도 변화

QP	Resolution of slant
0	999
2	-
4	-
6	-
8	-
10	-
12	-
14	-
16	-
18	-
20	999
22	994
24	986
26	990
28	993
30	970
32	950
34	928
36	low

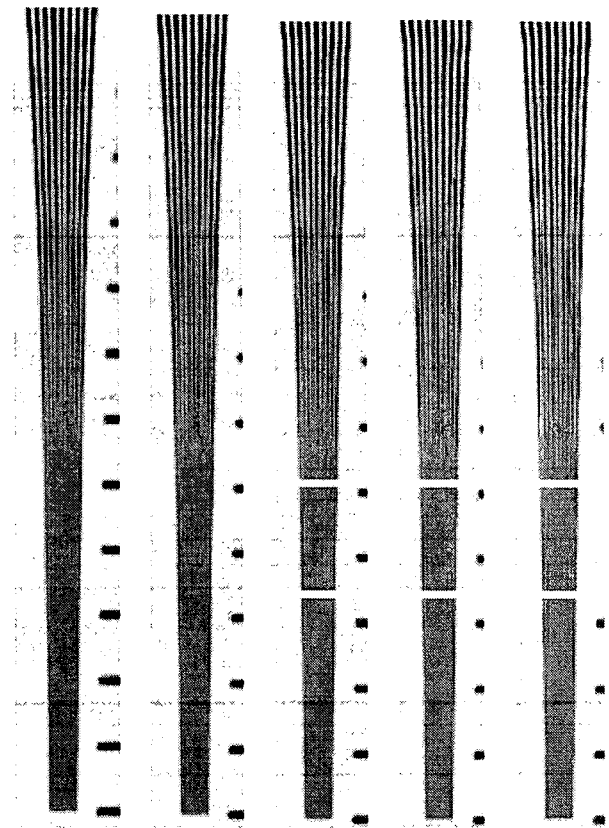
- : 미실험 / low : 측정불가

JPEG-2000에서 rate 0.02일때 해상도 991이며, H.264에서 QP 26이 990으로 서로 유사하다. 파일 사이즈는 JPEG-2000이 284,442Byte이며, H.264가 214,251Byte로서 다소 H.264의 효율이 더 뛰어나 보인다.

JPEG-2000처럼 어느 순간부터 급격한 훼손을 찾기는 어렵지만 QP 28이상에서는 뭉침 현상이 나타나기 시작했다. 그림 12의 (c)부터 (e)사의 노란 사각형 내부를 보면 뭉침 현상이 조금씩 심해지는 것을 육안으로 확인할 수 있다.

PSNR의 변화는 표6, 표7과 같다.

PSNR의 변화는 JPEG-2000에 비해 균일하게 나타났다. JPEG-2000이 특정한 rate에서 PSNR의 변화 눈에 띄게 나타나는 특징이 있었으나 H.264의 경우 QP의 변화에 따라 균등하게 변하는 특징을 가지고 있었다.



(a)QP 0 (b)QP 26 (c)QP 28 (d)QP 30 (e)QP 34

그림 12: ISO-12233의 세로 해상도 라인의 변화

표 6: PSNR 측정값

QP	R	G	B
0	42.43	50.74	40.54
2	42.38	50.45	40.48
4	42.28	49.87	40.40
6	42.08	49.01	40.21
8	42.01	48.30	40.18
10	41.95	47.51	40.20
12	41.54	46.82	39.71
14	41.36	45.89	39.55
16	41.40	45.01	39.76
18	40.58	44.12	38.78
20	40.15	43.25	38.37
22	39.89	42.42	38.17
24	39.14	41.42	37.45
26	38.28	40.42	36.81
28	37.54	39.44	36.17
30	36.78	38.29	35.49
32	35.90	37.05	34.79
34	34.89	35.89	34.02



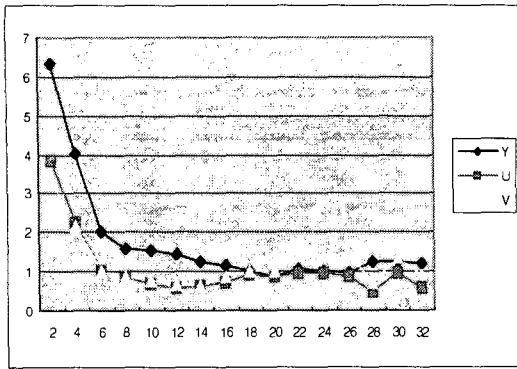


표 7: ISO-12233 차트의 PSNR 변화량

#### 4. 결론

이상과 같은 실험 결과를 볼 때 JPEG-2000 압축은 rate 0.02와 0.01 사이에서 가장 큰 변화 값을 보였다. 그러므로 디지털 시네마 압축에서 rate 0.01이하의 사용하지 않는 것이 적합하다고 판단되며, 0.02 혹은 0.03을 효율적인 압축범위로 제시할 수 있다. 4K 영상의 경우가 DCI가 규정한 전송률(250Mbps/sec)을 고려하면 0.02가 효율적인 압축률로 제안된다. 그러나 2K 24 FPS의 경우는 더 0.03, 0.04값도 충분히 고려할 수 있으며, 특히 컴퓨터 애니메이션과 같은 고주파 성분이 많은 영상에서는 해상도의 변화가 없는 0.04 정도의 값이 적절하다고 판단된다.

JPEG-2000이 특정한 부분에서 해상도와 PSNR의 변화가 큰 반면에 H.264/AVC는 균등한 변화를 보여주었다. QP 26 정도가 적절한 값으로 판단된다. 그러나 아직까지 H.264를 이용한 디지털 시네마에 대한 규정이 없기 때문에 정확한 값을 제안하기에는 어려움이 따른다.

이번 실험을 통해 현재 디지털 시네마의 압축 표준으로 제안된 JPEG-2000의 압축 성능이 뛰어난 것은 사실이지만 H.264/AVC의 압축 성능이 다소 우월한 것으로 나타났다. 또한 H.264가 현재도 활발히 성능을 개선하고 있으며 최근에 H.264/AVC High 4:4:4 Intra/Inter profile의 표준화가 완료된 만큼 인코더 최적화를 통하여 더 나은 압축 성능을 기대할 만하다.

본 연구는 영화진흥위원회의 2006년도 문화콘텐츠기술개발사업과 서울시산학연협력사업 지원으로 수행되었습니다.

#### 5. 참고문헌

- [1] Tinku Acharya, Ping-Sing Tsia, "JPEG-2000 Standard for Image Compression : Concepts, Algorithms and VLSI Architectures.", wiley-Interscience, pp. 137-162, 2005
- [2] Michael D. Adams, "The JPEG-2000 Still Image Compression Standard", ISO/IEC JTC 1/SC 29/WG 1 N2412, December 2002
- [3] Richardson, H.264 & MPEG-4 차세대 영상압축기술, 홍릉출판사, 2004
- [4] 찰스 S 스와츠, 디지털 시네마의 이해, 책과길, 2005
- [5] Digital Cinema System Specification V.1.0 , DCI , July 20, 2005

[6] 정수연, 임상희, 김철현, 유윤중, 백준기, "디지털 시네마에서 H.264/AVC와 JPEG-2000의 압축 성능 비교 연구." 문화콘텐츠기술연구소 논문집, 2권, 1호, pp. 143-151, 2006년 6월.

[7] Gordon, "ABT for Film Grain Reproduction in High Definition Sequences", JVT-H029, May 2003