

임베디드 SW 시스템 개발을 위한 리눅스 기반 솔루션 소개 ⑥

임베디드 리눅스 기반의 H.264 코덱 개발을 통한 IP 셋톱박스

H.264 비디오 표준은 MPEG-2를 대체하는 새로운 비디오 표준이며 세계 각국에서 신 개념 멀티미디어 서비스의 기본 기술로 주목받고 있다. 한국에서는 DMB 서비스를 위한 기본 비디오 기술로서 알려져 있는 H.264는 미국 및 유럽 등에서는 차세대 HD 서비스를 위한 비디오 코덱으로 받아들여지고 있다. 따라서 HD급 H.264 시스템 기술 경쟁은 날로 치열해지고 있는 상황이다. 이번 회에서는 HD급 고화질 H.264 기술에 대한 소개와 발전 동향을 설명하고 동시에 이를 지원하기 위한 H.264 인코더 및 디코더, 그리고 IP 셋톱박스 기술의 발전 동향 및 주요기술을 살펴본다.

글 : 석진욱 선임연구원 / ETRI 실시간 멀티미디어 연구팀
 jnwseok@etri.re.kr / www.etri.re.kr

소 개

한국에서는 지난 2004년에 세계 최초로 “손 안의 TV”를 구현한 위성 DMB 서비스를 시작했다. 위성 DMB 서비스에서 가장 두드러진 특징 중 하나는 비디오 코덱으로 H.264(MPEG-4 AVC) 표준을 사용했다는 점이며, 이 때문에 한국에서 H.264는 소형 비디오 서비스 등에 사용하는 기술로 일반인들에게 알려지게 되었다.

미국을 비롯한 유럽 각국에서 H.264 기술은 오히려 MPEG-2를 밀어내고 차세대 비디오 표준의 핵심 코덱으로 간주되고 있다. 특히, HD 방송에서 오히려 후발 주자라고 할 수 있는 유럽 각국에서는 HD 방송을 위한 비디오 표준으로 H.264 기술을 사실상 지정해 놓고 있는 실정이다.

MPEG-2 비디오 표준과 ATSC 시스템으로 HD 방송을 일찍 시작했던 우리나라로서는 이러한 유럽 각국은 물론 ATSC의 본고장이자 할 수 있는 미국마저도 H.264 표준의 급격한 확장과 H.264 기술 자체의 발전에 제대로 대응을 하지 못하고 있다고

보아도 과언이 아니다.

현재 세계 각국은 각 나라별로 나름의 독자적인 SD급(720×480) 이상의 H.264 솔루션을 확보했거나 확보하기 위한 치열한 기술전쟁을 벌이고 있다. 하지만, 한국의 경우 저렴한 비용의 SD급 이상의 고화질 H.264 기술을 독자적으로 완전히 확보하지 못하고 있는 실정이다.

따라서 2003년 이전만 하더라도 QVGA(320×240) 혹은 QCIF(352×288)급에서 논의되고 있던 H.264 기술이 얼마만큼 발전을 이루었기에 최근에는 HD급 비디오의 표준으로까지 각광받게 되었는지를 알아보도록 하겠다. 이는 H.264 표준이 어느 방향으로 발전할 것인가에 대한 전망은 물론, 향후 비디오 코덱의 발전 방향까지 가늠할 수 있는 길잡이가 될 것이다.

이번 회의는 다음과 같이 구성된다. 글의 전반부에서는 H.264 기술의 전반적인 내용과 고화질 H.264 기술의 발전 동향과 내용을 설명하고, 후반부에서는 한국전자통신연구원이 최근 개발한 HD급 S/W H.264 기술의 내용 및 H.264 디코더 솔루션을 사용한 IP 셋톱박스 기술을 소개하도록 한다.



새로운 유비쿼터스 컴퓨팅 시대가 우리 주변에서 소리없이 진행되고 있습니다. 임베디드 SW 연구단에서는 이전부터 이러한 유비쿼터스 컴퓨팅 시대의 핵심 소프트웨어인 임베디드 소프트웨어 기술을 확보하고 있으며, 이 기술들을 토대로 우리나라 산업이 세계적인 경쟁력을 갖도록, 표준, 마이크로, 나노형 임베디드 SW 플랫폼 개발에서 미들웨어 플랫폼 및 응용 소프트웨어 개발, 나아가 다양한 응용 시제품 개발까지 그 연구 영역을 확대해 나가고 있습니다.

* 연재순서

임베디드 SW 시스템 개발을 위한 리눅스 기반 솔루션 소개

1회 임베디드 SW 표준 플랫폼의 기반이 되는 임베디드 운영체제 Qplus

2회 임베디드 SW 개발도구 Esto

3회 센서 네트워크용 나노 운영체제 Nano Qplus

4회 임베디드 리눅스 기반의 로봇, 텔레매틱스 SW 플랫폼

5회 임베디드 리눅스 기반의 엔터테인먼트 센터

6회 임베디드 리눅스 기반의 H. 264 코덱 개발을 통한 IP Set top Box

H.264 비디오 기술

2003년 이전의 H.264 기술 동향

H.264 기술은 2003년과 2004년을 기점으로 내용이 크게 달라지기 시작한다. 일반적으로 2003년 이전의 H.264 기술이 MPEG-2, MPEG-4 ASP의 연장선상에서 그 동안 발전해온 비디오 압축 기술의 집약체적 성격을 지니고 있었다면 2003년 7월 이후의 기술은 H.264 압축 기술을 본격적으로 HD급 영상에 적용하기 위한 기술의 태동과 HD급 H.264 기술의 구현이라는 관점에서 발전되었다. 2003년 7월 이전의 H.264 기술은 우리나라에서 학계나 기업, 연구소 그리고 비디오 전문가들에게 알려진바 대로 여러 개의 블록 코딩 모드와 다중 참조 픽처 등의 특징을 사용하여 MPEG-2 보다 두 배의 압축률, MPEG-4 ASP 보다 40%의 압축률로 동일한 화질을 구현하는 비디오 압축 기술이다.

1998년 ITU-T를 중심으로 한 비디오 압축 표준 그룹은 당시 제정된 ITU-T의 최신 비디오 코덱인 H.263의 절반 정도의 비트율에서 동등한 화질을 제공한다는 목표로 H.264의 제정 작업을 시작한다. 당시만 하더라도 H.263의 압축 효율은 MPEG-2 보다 높았기 때문에 H.263 보다 우수한 압축효율을 가진 비디오 코덱을 만든다는 것은 보다 장기적인 연구를 필요로 한다고 생각했다. 동시에 H.263과의 호환성도 고려하지 않겠다는 것을 처음부터 내세우며 시작됐다.

이 과정에서 보다 인코딩 효율을 높이기 위해 복잡도를 제한

하지 않고 인코딩하고자 하는 검토가 진행되었고 높은 복잡도 때문에 지금까지 목표규격에 채택되지 않았던 기술들도 적극적으로 도입되었다. 특히 하인리히 헤르츠 연구소(Heinrich Hertz Institute: HHI) 의 토마스 위간드(Thomas Wiegand)에 의해 Rate-Distortion Optimization 방식을 제안했는데 이로 인해 H.264의 성능이 크게 향상되어 크게 주목을 받게 됐다.

Rate-Distortion Optimization의 원리는 DCT-양자화된 값에 대한 라그랑지안 계수의 곱과 대 블록 혹은 블록 간 차의 제곱의 합(SSD : Sum of Squared Distortion)으로 이루어진 라그랑지안을 사용하여 이를 통해 블록 모드, 움직임 벡터, 참조 픽처 등을 결정하는 방식이다. 만일 움직임 벡터를 기술하는데 10bit가 필요하고 DCT-양자화에 5bit가 필요하다면 차라리 DCT-양자화에 10bit를 할당하고 움직임 벡터 기술에 1~2bit를 사용한다면 전체적인 비트 사용량은 줄어들게 되어 원 영상과 복원 영상간의 차이가 큰 부분에 보다 많은 비트를 할당할 수 있게 되어 화질이 향상될 수 있다.

실제로 H.264에 적용되고 있는 움직임 벡터 부분의 Rate-Distortion Optimization 부분을 삭제하거나 약화시키면 비트 수의 증가가 거의 100% 가깝게 늘어나 사실상 MPEG-2로 인코딩 하는 경우와 차이가 나지 않는 것을 볼 수 있다.

H.264의 성능이 알려지던 그때 MPEG 진영에서는 MPEG-4 비디오 버전 1에서 4까지의 표준화 작업이 완료됐다. 그 이후 MPEG 진영의 참가인원 수가 줄게 되면서 MPEG 단독으로는 표준화 활동이 불가능해지자 MPEG은 ITU-T와 공동으로 표준화

작업을 하게 됐다. 이때, MPEG-4 ASP을 기준으로 P2, P3, P4 및 H.26L의 4가지 방식에 대한 주관적 화질 테스트 결과 H.26L 방식의 우위가 확인되면서 H.26L 기반의 새로운 표준화가 시작되었고 이것이 H.264 혹은 MPEG-4 AVC의 시작이었다.

H.264의 표준 제정이 시작될 때 ITU-T 진영에서는 주로 영상의 시스템과 관련된 영상압축 인코딩을 위한 반면 MPEG 진영은 범용의 영상압축 규격을 원했다. 따라서 ITU-T 진영과 MPEG 진영은 다음과 같은 규칙에 합의하여 H.264 표준 제정에 들어가게 된다. H.264 코덱에서의 개발 원칙은 다음과 같다[1].

1. 비디오 코덱의 기본으로 돌아간 기술개발로서 객체 인코딩이나 스케일러빌리티 등은 도입하지 않는다
2. 높은 압축률을 표준화의 최대 목표로 한다
3. 시간지연의 자유도를 높여 여러 서비스에 적용가능 하도록 한다
4. 네트워크 친화성이 높아야 한다
5. 에러 내성에 강해야 한다
6. 인코더와 디코더의 규격을 스케일러블 하게 변경할 수 있어야 한다
7. 복호기 간에 복호 값이 같아야 한다
8. 고화질 응용 시스템도 고려한다
9. 저장 매체용 파일 기록 포맷을 정의한다

이러한 원칙하에 이루어진 표준화 작업 결과 현재 일반적으로 알려져 있는 H.264의 특징들로 구현되었으며 가능한 한 고압축률과 고화질을 위해 많은 기술들을 표준으로 채택했다. 따라서 인코딩 및 디코딩에 있어 높은 복잡도를 지니게 되어 “메머드 코덱”이라는 별칭을 얻게 됐다. 그럼에도 불구하고 2003년 7월로 일단 주요 표준화 작업이 완료된 H.264 표준은 당시만 해도 주요 타깃이 QCIF에서 SD급의 영상이 한계였다.

2003년 이후의 H.264 기술 동향

2003년 7월까지의 1차 표준화 작업이 완료되기 이전부터 H.264 코덱의 성능이 널리 알려지면서 H.264 코덱의 실제 구현이 광범위하게 시도됐다. 특히 2004년 이전에만 하더라도

H.264와 같은 복잡도가 심한 코덱의 경우에는 ASIC 하드웨어 혹은 대규모 병렬 DSP 방식 외에는 인코딩 및 디코딩이 모두 어려울 것으로 예상됐다. 따라서 소프트웨어적인 접근보다는 주로 ASIC에 의한 하드웨어 접근과 DSP를 사용한 접근이 주류를 이루었다. 이 과정에서 SD급 테스트 영상뿐 아니라, HD급 영상에 대한 테스트가 이루어지면서 H.264 코덱의 성능이 유감없이 발휘되자 JVT에서는 HD급 영상에 H.264 코덱 적용을 위한 표준화 작업에 들어간다.

FRExt로 알려진 2003년 7월 이후의 표준화 작업에서는 주로 메인 프로파일(Main Profile)을 확장하여 HD급 영상에 대응하기 위한 하이 프로파일(High Profile)의 제정과 H.264 코덱의 약점으로 알려진 Color Space의 보정 및 확장에 주안점을 뒀다.

2004년 7월 Draft-3이 발표되면서 윤곽이 드러난 H.264 FRExt의 주요 사항은 다음과 같다[2].

1. YUV 4:2:0을 기본으로 하되 4:2:2 및 4:4:4 포맷 역시 지원하며 보다 선명한 색상복원을 위해 Alpha Blending을 지원할 수 있다
2. 1 Pixel의 기본단위는 8bit이나 10, 12Bit 역시 지원한다
3. Loseless코딩 및 Stereo코딩을 지원한다
4. 기존 4×4 정수 DCT연산에 부가하여 정수 8×8 DCT연산을 추가한다
5. Losless 코딩을 위해 새로운 Intra Pre-dition 모드를 추가한다(I_PCM)
6. 새로운 정수 색상공간(Color Space)을 지원한다(Y, Cg, Co)
7. 보다 고품질의 영상복원을 위한 양자화 매트릭스를 선택적으로 입력할 수 있다

이러한 H.264 FRExt의 특징들은 기존 HD급 영상 장비와의 호환성을 다분히 염두에 둔 것으로서 많은 HD급 영상 장비들이 주로 10bit 데이터 처리를 하며 Alpha Blending을 수행하는 것에 대응하기 위해서이다.

또한, 무손실 인코딩을 위한 I-PCM 방식이 FRExt의 주요 특징으로 채택되어 인코더는 어떤 기준에 부합하다고 판단될 시 참조 픽처의 대블록을 그대로 복사하여 복원할 수 있는 모드

가 추가됐다. 이 결과로 I-Frame 인코딩에서도 P-Frame의 SKIP 모드, B-Frame의 Direct모드와 같은 역할을 하는 모드가 추가되어 인코딩 효율을 급격히 높일 수 있게 됐다.

한편 H.264 코덱의 약점으로 알려져 있던 색상 인코딩 부분에 있어 새로운 색상공간을 도입하고 4:2:2 및 4:4:4 인코딩 및 디코딩이 가능하도록 하여 색상 인코딩의 효율을 높이는 동시에 HD급 영상장비와의 호환성을 높였다.

마지막으로 새로운 정수 8×8 DCT 연산이 도입됐다. 이는 기존의 4×4 DCT 연산이 개발자들의 생각 이상으로 많은 연산 부하를 필요로 하고, 하드웨어의 발전에 따라 정수 연산을 도입하였다고 해서 특별히 연산속도가 빨라지지 않으며, 기존 부동 소수점 8×8 DCT에 익숙한 코덱 개발자들에게 보다 친숙한 DCT 연산을 도입한다는 취지에서 정수형 8×8 DCT가 도입됐다.

그러나 정수형 8×8 DCT는 H.264의 장점이자 HD 영상 처리를 위해 매우 중요한 세밀한 영상 인코딩 및 디코딩에 큰 약점을 가지고 있어 이를 주도적으로 사용하는 마이크로소프트사의 WMV-9(혹은 VC-1코덱)이 H.264에 비해 주관적 화질 비교에서 HD급 영상 복원에서 떨어진다는 의견이 많다.

고화질 H.264 기술

고화질 H.264 기본 기술

H.264 비디오 코덱을 사용한 고화질 영상 인코딩 기술은 주로 메인 프로필을 사용하여 I, P, B 프레임에 적절한 비트 수를 할당하면서 압축률을 그대로 유지하도록 하는 기술이다.

이를 위해서는 반드시 Rate-Distortion Optimization 기술이 필수적인데, 그 이유는 H.264 코덱이 가지는 여러 가지 모드를 적은 비트 수를 가지면서도 원 영상에 가장 근접하도록 선택하여야 하기 때문이다. 이 때문에 JM의 Reference 인코더에서는 5개의 서로 다른 라그랑지안을 사용하도록 규정되어 있다.

반면 Rate-Distortion Optimization을 사용하지 않을 경우에는 SAD(Sum of Difference)와 움직임 벡터와 연계된 라그랑지안과 참조 픽처 선택을 위한 라그랑지안의 두 라그랑지안을 사용하도록 되어 있다. 그 결과, HD급 영상에서는 콘텐츠에 따라 차이는 있으나 Rate-Distortion Optimization을 사용하

는 경우가 사용하지 않는 경우보다 약 2~5% 정도 비트 수가 늘어난다. 반면 신호 대 잡음비는 0.5dB 이상 높아짐에 따라 보다 고화질의 영상 인코딩을 수행하려면 Rate-Distortion Optimization은 필수적이라 하겠다.

그러나 Rate-Distortion Optimization을 수행하기 위해서는 반드시 Rate에 해당하는 DCT-양자화된 값의 인코딩된 비트 수 및 선택된 모드를 기술하기 위한 비트 수가 반드시 요구된다. 이는 Rate-Distortion Optimization을 위해서는 각 모드마다 DCT-양자화를 수행해야 한다는 의미이며 또한 임시로 CABAC 혹은 UVLC 인코딩까지 수행해야 한다는 의미다.

이때 발생하는 연산량은 매우 많아 JM Reference 인코더의 경우 Rate-Distortion Optimization을 수행하지 않는 경우에 대하여 약 10배 가까이 연산 시간이 늘어난다. 그러므로 Rate-Distortion Optimization을 고속화하기 위한 알고리즘이 반드시 요청되며 이는 보통 DCT-양자화의 결과를 대블록 Distortion 값의 1차 및 2차 확률 모멘트와 선형 혹은 비선형 모델을 사용한 비트 수 예측 모델의 파라미터를 추정하는 방식이 사용되어야 한다. 이 경우, 예측 알고리즘의 연산량이 인코딩 연산량에 거의 영향을 미치지 않으면서 Rate-Distortion Optimization을 비교적 훌륭히 수행할 수 있게 한다.

두 번째로 필요한 부분은 B-Frame의 사용이다. B-Frame은 순방향 및 역방향 참조 픽처를 사용하여 순방향 예측, 역방향 예측, 양방향 예측을 할 수 있게 됨으로써 예측 인코딩 효율을 크게 높일 수 있다. 보통 최적화된 H.264 인코더의 경우 콘텐츠에 따라 P-Frame의 10~20% 수준에서 인코딩이 가능하다.

그러나 B-Frame의 사용은 순방향, 역방향, 양방향 예측의 3가지 예측을 모두 수행해야 하기 때문에 P-Frame 보다 일반적으로 많은 연산 부하가 발생한다. 그러므로 효율적인 B-픽처의 인코딩을 위한 최적화 기술이 반드시 요청된다.

마지막으로 요청되는 부분은 H.264의 De-Blocking Filter이다. 주관적인 화질 시험을 수행했을 때 일반적인 의견은 확실히 De-Blocking Filter를 사용했을 때와 사용하지 않았을 때 화질의 차이가 뚜렷이 드러난다. 이는 인코더의 실제 구현에 있어 고속화 알고리즘을 사용하면 할수록 뚜렷이 나타나는 것으로서 고속화 알고리즘의 거의 대부분이 이른바 Early-Skip 방

식으로 이루어짐에 따라 대 블록 간 영상의 차이가 확연히 드러나기 때문이다. 그러나 HD 영상의 경우 HD 720P의 영상 크기는 1280×720, HD 1080P의 경우 1920×1080이므로 De-Blocking 필터 부분이 수행하는 연산의 양은 매우 많다. 더구나 현재 H.264 표준에서 채택하고 있는 De-Blocking 필터는 이웃 블록의 특성, 참조 픽처의 특성, I, P, B 인코딩에 따라 De-Blocking 필터 계수의 선택이 매우 다양하게 되어 있다. 이는 필터 계수 선택에 “If” 명령어를 많이 사용하도록 하여 실제 De-Blocking 필터 구현에 있어 Pipelining에 의한 연산 고속화를 거의 기대하지 못하게 만든다.

따라서 많은 상용 H.264 인코더들은 실시간 인코딩을 위해서 De-Blocking 필터를 일부러 사용하지 않는 경향이 있는데 이는 시간이 지나면 지날수록 매우 치명적인 화질 열화를 낳는다.

HD급 H.264 인코딩/디코딩의 문제점

비록 H.264 비디오 코덱이 HD급 영상에 기존 코덱보다 강점을 가지고 있더라도 H.264 비디오 코덱은 HD급 영상 인코딩 및 디코딩에 매우 치명적인 약점을 가지고 있다. 그것은 H.264 코덱이 채택하고 있는 부화소 모드 때문이다. H.264 코덱은 6개의 정화소를 FIR 필터를 통해 얻어지는 반화소와 이렇게 얻어진 반화소와 인접 정/반화소간의 산술 평균을 통해서 만들어지는 1/4 화소를 사용한다.

이 때문에 HD급 H.264 인코딩 및 디코딩에는 YUV 4:2:0 8 Bit 시스템을 사용한다고 가정했을 때 HD 720P 24fps의 경우 1픽처 당 약 500Mbytes, HD 1080P 24fps의 경우 1픽처 당 2GBytes의 데이터가 요구된다. 일반적으로 Serial-ATA 방식 하드 디스크의 초당 데이터 접근 속도가 150Mbytes~300Mbytes이다. 만일, HD급 영상을 PCI-X 2.0 266MHz에서 읽어 들인다 해도 266MHz에서의 데이터 전송 속도는 초당 약 2GBytes이므로 HD 720P나 HD 1080P의 인코딩 혹은 디코딩은 사실상 하드웨어의 한계를 넘어선 것이라 할 수 있다.

이 같은 이유 때문에 일부 H.264 인코더나 디코더는 실시간 인코딩을 위해서 부화소 연산을 아예 수행하지 않는 경우가 있다. 이 경우 비트 수 증가는 약 40% 이상이며 신호 대 잡음비의 하락은 매우 심각한 정도까지에 이르러 사실상 MPEG-4 ASP

나 MPEG-2와 아무런 차이가 없거나 혹은 더 못한 결과를 가져온다(모드 별 기술 데이터가 많기 때문).

실제 H.264의 경우 움직임 벡터의 거의 90%이상에서 1/4 화소를 포함하고 있으며 알고리즘으로 부화소의 점유율을 떨어뜨리면 떨어뜨릴수록 거의 신호 대 잡음비는 크게 하락하며 비트 수 증가는 부화소 점유율의 하락에 정확히 반비례한다.

반면, HD급 영상 인코딩에 있어 정 화소 추정 알고리즘은 큰 영향을 미치지 못한다. 일반적으로 16×16 규격의 탐색 범위를 가질 경우 그 이상의 탐색 범위와 비교하여 비트 수의 증가신 신호 대 잡음비의 하락은 극히 미미하다. 최악의 경우 4×4의 탐색 범위를 가지더라도 비트 수의 증가보다는 신호 대 잡음비가 약간 떨어지는 정도이다.

최근 H.264 Reference 인코더에 채택된 이른바 Hexagonal Search 방법도 실제 구현 및 테스트 결과에 의하면 Diamond 탐색 방법에 비해 약 3~5% 정도의 비트 수 하락과 미미할 정도의 신호 대 잡음비의 향상을 보여준다. 비록 Hexagonal 탐색 방법이 전역 탐색에 근접한 성능을 보여준다 하더라도 기존 Diamond 탐색 방법에 대하여 뚜렷한 성능 향상을 보여주지 못하는데다가 Diamond 탐색에 비하여 약 4배 이상의 연산 시간을 필요로 하기 때문에 실시간 인코딩 등에서는 오히려 장애 요소일 뿐이다.

또한, 정화소 탐색 속도를 빠르게 하기 위한 각종의 예측 알고리즘의 경우에 예측 알고리즘이 한번이라도 참조 픽처의 데이터를 필요로 하게 될 경우에는 역으로 움직임 예측을 위한 연산 시간을 대폭 증가시키게 된다. 이는 Cache Memory에 연관성이 떨어지는 데이터가 입력됨에 따라 Cache Data에 대한 적중률을 떨어뜨려 연산 시간증가의 주요 요소가 된다.

마지막으로 심각한 문제는 인코더와 디코더간의 상호운용성 검증이 거의 이루어지지 않았다는 점이다. 대부분의 인코더 생산자와 디코더 생산자는 자사의 제품에 대해서는 최고의 성능을 발휘한다. 그러나 다른 제품과 연동 시험에 들어갈 경우에는 심각한 화질 열화 현상(블록킹 현상, 화면 색상 변화 등)이 발생한다.

이러한 경우가 발생하는 원인은 각 인코더 생산자나 디코더 생산자의 경우 자신들만의 고속화 알고리즘을 채택하고 있으며 이것이 JM의 Reference 인코더나 디코더와 서로 상이하기

때문이다. 가장 심각한 경우가 이른바 오픈 소스로 알려져 있는 VideoLAN 사의 H.264 인코더로서 이 인코더로 인코딩된 H.264 스트림은 아예 직접적으로 JM의 H.264 디코더로 복호되지 않는다. 이는 매우 심각한 현상으로 만일 디코더 업체가 H/W로 디코더를 설계하고 판매한다고 가정했을 때 해당 인코더로 만들어진 H.264 스트림은 아예 디코딩되지 않을 수 있다는 위험성이 있다는 점이다.

따라서 인코더나 디코더 업체들의 솔루션을 선택할 때 과연 이들 인코더나 디코더가 제대로 H.264 표준을 준수하고 있는가를 살펴보는 것은 매우 중요한 문제라 할 수 있다.

기존 HD급 H.264 인코딩/디코더

이 같은 문제점에도 불구하고 2005년부터 HD급 H.264 인코더 및 디코더가 여러 연구소, 기업 등에서 최근 출시되고 있다. 거의 대부분의 HD급 H.264 인코더의 경우 DSP 혹은 FPGA를 사용한 모델로서 DSP의 경우에는 Texas Instrument 사의 DSP 혹은 LSI Logic의 DSP를 주로 사용한다. FPGA를 사용한 HD급 H.264 인코더의 경우는 다수의 FPGA를 사용하여 HD급 인코딩을 수행하는 방식을 채택하고 있다. HD급 H.264 인코더는 독일의 Tandberg 사에 의해 만들어진 EN-5990이 있다. 내부에 30여개의 TI-DSP로 만들어진 것으로 알려져 있으며 Tandberg 사 고유의 하드웨어 플랫폼인 ICE 위에 만들어져 있고 메인 프로필을 지원하며 Level은 4.0(HD 1080P)을 지원하는 것으로 되어 있다. 인코딩 속도는 HD 720P 24fps, HD 1080i 30fields이다. HD급 디코더 생산업체인 Connexant 사의 HD급 연동 시험을 성공리에 마쳤고 현재 세계에서 가장 널리 인지되어 있는 HD급 H.264 인코더 장비이며 JM Reference 대비 약 15% 가량 성능이 떨어진다고 알려져 있다. 또 다른 HD급 실시간 인코더는 Modulus Video 사의 ME6000 HD급 실시간 인코더이다.

Modulus Video 사의 제품은 6개의 Virtex-4 FPGA와 1개의 Pentium-4 호스트 CPU로 이루어진 제품으로서 최근 High Profile, Level 4.0을 지원한다고 알려져 있다.

최근에는 많은 FPGA 디자인 업체들이 HD급 H.264 인코더 IP를 출시하였는데 대부분의 경우 Xilinx 사의 최신 FPGA를 사용하며 50MHz 이하에서 동작이 검증된 상태이다. 이들 업체는

120MHz 클럭에서 HD 720P 인코딩이, 240MHz 급에서는 HD 1080P 실시간 인코딩이 가능하다고 주장하고 있으나 FPGA에서 이 정도의 클럭 주파수를 가지고 정상적인 동작을 하기는 대단히 어려운 일이므로 다소 신빙성이 떨어진다 하겠다.

복호기의 경우 지금은 여러 업체들이 HD급 솔루션을 경쟁적으로 내놓고 있다. 가장 먼저 출시되어 많은 테스트가 이루어진 제품은 ST Micron사의 STB7100 제품으로 한국에서도 많은 업체들이 STB7100을 베이스로 IP_STB를 만들고 있다. 그러나 STB7100의 경우 호스트 CPU로 SH4-202 RISC CPU를 사용하고 있는데 현재 리눅스 진영에서의 지원이 미비하다는 약점을 지니고 있다.

두 번째 솔루션은 MPEG-2 솔루션부터 안정성에 있어 호평을 받은 BroadCom의 7411 솔루션이다. BroadCom의 7411 솔루션은 그러나 메인 CPU로 자사의 7038 Video Processor를 사용하도록 되어 있는 2칩 솔루션이라는 점이 약점으로 지적된다.

세 번째 솔루션은 Connexant의 Cx2418x 솔루션으로 현재 나와 있는 H.264의 모든 프로파일을 지원할 수 있는 라인업을 갖춘 것이 장점이며 동시에 Tandberg 사 제품과 연동시험까지 거쳐 상호 운용성이 검증된 제품이라는 것이 장점이다. 네 번째 솔루션은 SigmaDesign의 SMP8630이다. 한국에서는 Sigma Design의 인지도가 매우 높는데 그 이유는 Sigma Design의 MPEG-2 솔루션이 가격 대비 성능에서 매우 만족스러웠기 때문이다. 이외에도 WIS, ATEME, Main Concept, Moonlight, VSoft, Apple 등의 하드웨어 및 소프트웨어 솔루션이 시장에 출시되어 있다.

HD급 H.264 인코더를 위한 주요 기술

ETRI개발 HD급 H.264 인코더

ETRI에서는 지난해 11월에 HD 720P의 실시간 인코딩이 가능한 소프트웨어 인코더를 개발했다. 소프트웨어 솔루션으로서 현재, 세계에서 가장 빠른 인코딩 속도를 가지고 있으며 최고 속도(25fps~35fps)에서 약 3~5% 정도의 Bit 수 증가와 약간의 신호 대 잡음비 하락이 나타난다. 한편, 최저 속도(15fps~18fps)에서는 JM Reference 인코더 보다 약 2% 정도 비트수가

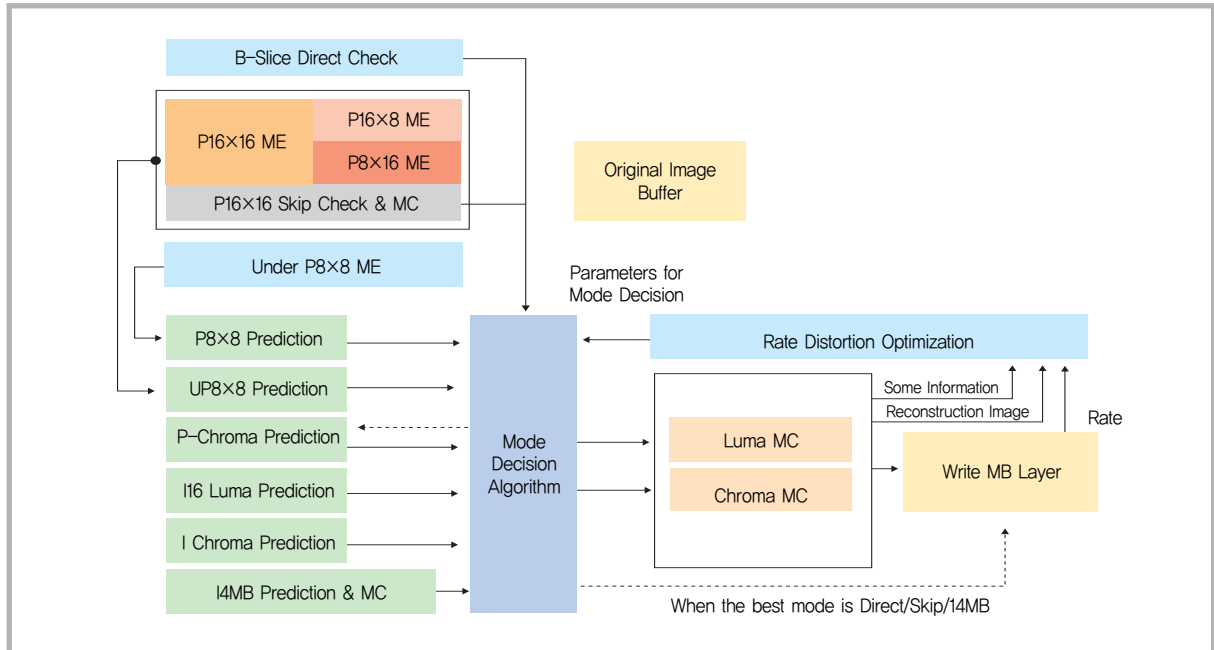


그림 1. ETRI 개발 HD급 H.264 인코더의 구조도

절감되고 있으며 신호 대 잡음비는 같은 옵션의 JM Reference와 비교해서 떨어지지 않는다.

JM Reference 인코더 9.6을 기반으로 구현되었으며 메인 프로파일을 지원하며 실시간으로는 Level 3.1을 지원할 수 있을 것으로 예상되고 비 실시간으로는 Level 4.0 이상도 가능하다.

이러한 고속의 HD급 인코더를 만들기 위해서는 여러 고속 알고리즘들이 구현되고 탑재 되어야 하는데, 첫 번째는 고속화 지원에 알맞은 H.264 인코더의 구조화이다. JM Reference 인코더의 경우 모든 경우에 대한 인코딩이 가능해야 하기 때문에 이러한 인코더 구조화가 거의 이루어져 있지 않다고 보아도 과언이 아니다. ETRI에서는 이러한 인코더의 구조화를 위해 그림 1과 같이 인코더를 구조화 하고 있다.

또한 방대한 양의 HD급 데이터, 특히 부화소 데이터의 처리를 위해 부화소 움직임 벡터 추정 시 부화소를 사용하지 않고 정화소의 SAD값과 라그랑지안을 추정하여 부화소 예측을 수행한다. 이를 통해 B 픽처에서도 HD 720P의 움직임 추정 연산 시간을 프레임당 30ms 이내에서 수행할 수 있었다.

정화소의 HD급 데이터는 HD 720P 24fps의 경우 약

70Mbytes, HD 1080P 24fps의 경우 150Mbytes 수준이므로 다수의 HDD가 RAID-0으로 결합되어 있을 경우 실시간 인코딩이 가능한 수준에 이른다. 또한, H.264의 특징을 살리기 위해 DCT 연산은 4x4만 사용했으며 4x4 DCT 연산의 속도를 증가시키기 위해 4개의 4x4 DCT 연산을 한번의 8x8 DCT 연산으로 수행할 수 있는 고속화 병렬 알고리즘을 개발 구현하였다.

최적화가 진행될수록 많은 연산량을 가지게 되는 또 하나의 요소인 De-Blocking Filter의 경우 8개의 16bit 연산을 통해 대블록 하나당 4번의 SIMD 연산으로 수행이 가능하도록 구현 알고리즘을 최적화, 및 새롭게 알고리즘을 개발하였다. 이로써, H.264의 표준을 그대로 따르면서도 HD720P급 한 픽처의 De-Blocking 필터링은 I-픽처에서 15msec 수준이며 P 및 B에서는 15msec 이하에서 가능하다.

고속 인코딩을 위해 중요한 요소인 블록 모드 결정 방법에는 I 블록 및 P 블록 간 평균 Rate-Distortion 값 및 전체 블록의 평균 Rate-Distortion값을 각각 Kalman 필터로 추정하고 16x16 대 블록의 움직임 보상 후 나타나는 CBP 값을 사용하여 P 블록의 모드를 결정하는 방법을 사용하였다. 이 방법을 각 단계

별로 적용하여 고속 인코딩의 단계를 나누었으며 개발한 블록 모드 결정 방법을 사용하지 않았을 때보다 평균 30~50% 이상 인코딩 속도가 향상 되었으며 HD 영상에서의 비트율 및 신호 대 잡음비는 그림과 같다.

이로서 단일 프로세서에서 HD 720P 영상 인코딩에 Hyper Threading 지원 3.4GHz 펜티엄 4 프로세서에서는 6~10fps 의 인코딩 속도를 가지게 되었으며 Dual Core 3.0 GHz 펜티엄 4 프로세서에서는 최대 14fps의 인코딩 속도를 나타내었다. 특히 최근 인텔에서 출시된 Intel Micro Architecture 기반 듀얼코어 CPU 에서는 불과 1.8GHz의 CPU 클럭임에도 불구하고 HD720P 인코딩에서 최대 20fps까지 나오는 놀라운 인코딩 속도를 보였다.

실시간 HD720P 인코딩을 위해 ETRI에서는 인코더의 구조를 멀티 슬라이스 지원이 가능하도록 하였으며 슬라이스 단위로 멀티 쓰레드가 지원되도록 병렬 구조화 시켜 Intel Xeon 3.66GHz 4-way Server에서 24~35fps의 인코딩 속도를 확보하여 HD급 영상 캡처 보드에서 PCI-X 2.0 266MHz 혹은 PCI-Express 133MHz 급을 지원할 경우 실시간 HD 720P 인코딩이 가능하다.

HD급 인코더를 위한 주요 기술

HD급 H.264에 관한 주요 표준화가 완료되어 있는 현 시점에도 H.264 비디오를 통해 실제로 방송이나 HD-DVD, 블루레이 DVD와 같은 차세대 영상 저장매체를 위한 인코딩에는 H.264 표준 적용에 있어 세심한 주의가 요청된다.

첫 번째 문제는 FRExt 혹은 High-Profile과 같은 HD급 H.264 비디오 표준의 세심한 적용이다. High-Profile의 경우 색차 정보를 기존 MPEG-2 HD 영상과 같은 4:2:2 방식으로 제정할 경우 이에 의해 색차 정보가 두 배정도 늘어나 이를 시간으로 처리하기 위한 추가적인 기능이 요청된다. 하지만 이 부분은 최근 진행된 기술 개발에 의해 어느 정도 해소된 상황이므로 큰 문제가 되지 않는다.

두 번째 부분은 현행 8bit 기반의 영상신호를 HD 신호 복원을 위한 10bit, 12bit 처리를 할 때 발생하는 비트 미스매치 부분이다. 현재로서는 8bit 기반 하드웨어에 의한 HD급 H.264 처리가 보편화 되어 있는 실정이며 대부분의 인코더 기기들이

10bit 이상 지원을 하지 못하고 있음에 따라 앞으로도 10bit 이상의 HD 신호 처리는 특수한 분야에서 제한적으로 사용될 것으로 생각된다.

가장 치명적인 부분은 FRExt에서 사용되는 I-PCM의 적용 여부이다. I-PCM 모드는 I-프레임임에도 디코더 혹은 인코더의 이전 프레임에 저장된 데이터를 사용하여 I-프레임상의 대블록 일부를 복원하는 모드이다. 이는 H.264에서 I-프레임이 가장 큰 비트 점유율을 차지하고 있기 때문에, 만일, 복원된 영상과 현재 영상이 아무런 차이가 없다면 구매여 이를 I-프레임을 위한 예측 모드에 적용하여 압축 정보를 보낼 필요 없이 복원된 영상으로 현재 영상을 대체하여 비트 수를 절약하기 위한 모드이다.

I-PCM 모드는 이 자체로 인코딩, 디코딩에 들어가는 연산수를 줄일 수 있으며 비트수를 절약할 수 있다는 장점이 있다. 그러나 I-PCM 모드가 일반 방송 스트림에 적용될 경우 발생할 수 있는 패킷 손실에 대하여 매우 취약하다는 단점이 있기 때문에 적용하기가 어렵다.

예를 들어, I-프레임이 매 15프레임 마다 기존 방식대로 들어가 있다고 가정하면 패킷 손실로 화면이 깨지는 현상이 발생하더라도 0.5초 내로 다시 깨진 화면이 I-프레임에 의해 복구될 수 있으나 I-PCM 모드가 적용되면 I-프레임이 적용된다 하더라도 깨어진 영상이 I-PCM이 적용된 구간에서는 복구되지 않기 때문에 계속 깨진 영상이 디코딩 될 수밖에 없는 현상이 발생한다. 따라서 방송 혹은 VoD와 같은 응용에서는 I-PCM의 사용은 부적절하다.

반면, 저장 매체의 경우에는 방송이나 VoD와 같은 분야 보다 이러한 패킷 손실이 일어날 확률이 매우 작으므로 I-PCM의 사용이 충분히 가능하다. 또한 이렇게 절약된 비트수를 사용하여 보다 작은 양자화 계수를 사용하게 되면 보다 선명한 화질을 얻을 수 있으므로 I-PCM 모드의 사용은 블루레이 DVD나 HD-DVD등 저장 매체 분야에 제한적으로 사용되어야 할 것이다.

HD급 H.264 디코더의 주요 기술

HD급 H.264 디코더의 경우 2006년 8월 현재 충분히 상용화가 가능할 수준으로 세계 각국에서 많은 솔루션들이 개발되었

다. 그러나 이들 업체들의 광고에도 불구하고 몇 가지 주의할 점이 있다.

첫 번째의 경우는 성능이다. 많은 업체들이 HD급 H.264 디코더의 성능을 보증하고 있으며 거의 대부분 성능에서 큰 문제가 없다. 그러나 어떤 업체의 솔루션의 경우 충분히 나쁜 콘텐츠에 대한 성능 테스트를 거치지 않았는지, 제대로 HD급 콘텐츠를 디코딩 하지 못하는 경우도 있다.

이러한 문제점은 대부분 업체에서 제공하는 표준 OS나 SDK 만을 사용하면 해결될 수 있으나 뒤에 설명할 각종 응용들을 모두 돌린다고 가정할 경우 충분한 성능을 보장 못하는 경우가 발생할 수 있다. 대부분의 H.264 디코더의 경우 처음 테스트나 데모 과정에서 매우 정적인 화면을 디코딩 하는 것을 볼 수 있다. 또한 동적인 화면이라 해도 급격하게 움직이는 부분은 얼마 되지 않는 부분도 있으며 영화에서 나오는 격렬한 추격 장면 등을 통해 빠르게 움직이는 영상도 충분히 디코딩 할 수 있다고 여기는 경우가 많다. 그러나 이러한 경우 모두는 디코더 성능 테스트에 적합하지 않은 경우이다. 정적인 영상의 경우 아무리 화질이 좋다 하더라도 기본적으로 정적인 영상이기 때문에 데이터의 양이 얼마 되지 않는다. 따라서 이러한 경우 H.264 디코더의 디코딩 능력을 충분히 살펴볼 수 없다. 에어쇼 혹은 자동차 추격신은 움직임이 많으므로 좋은 디코더 테스트 영상으로 생각하기 쉬우나 에어쇼의 경우 실제로 움직이는 대상은 화면 전체에서 얼마 되지 않으며 대부분 하늘이 나타나기 때문에 실제로 디코딩이 일어나는 영상은 얼마 되지 않는다. 자동차 추격 장면의 경우 움직임이 매우 격렬하여 많은 디코딩 연산이 일어나는 것으로 생각하기 쉬우나 매우 격렬한 영상의 경우 실제로 화면을 분석해보면 80% 이상이 I-블록에 해당된다. 디코더 입장에서 I-block의 디코딩은 연산 부하가 크지 않으므로 이 역시 충분한 디코더 성능 테스트라고 볼 수 없다. 디코더의 연산 부하가 가장 큰 경우는 군중신의 경우가 해당되는데 HD급 H.264의 경우 군중 속의 사람 한 명, 한 명이 사실상 움직임 벡터를 가질 수 있으므로 이러한 영상을 디코딩 하는 데는 많은 연산 부하가 걸린다. 예를 들면 많은 군중들에 의한 군무나 단체 마라톤과 같은 영상에 대한 테스트가 디코더의 성능을 평가하는데 좋은 영상이다.

두 번째는 OS 문제이다. 대부분의 디코더 솔루션들은 일단

마이크로소프트 사의 모바일용 OS를 지원하도록 설계되어 있으므로 MS 사의 OS를 사용할 경우 큰 문제는 발생하지 않는다. 그러나 MS 사의 OS를 사용할 경우 셋톱박스의 원가 상승 요인이 되므로 많은 디코더 솔루션 업체들이 임베디드 리눅스를 OS로 지원한다. 그런데 이러한 OS들은 하드웨어 성능을 최적하게 만들기 위해 내부적으로 많은 수정이 된 상태이므로 이 상태에서 비주얼한 프로그램 가이드 등을 만들어 넣기 위해서는 많은 어려움이 발생한다. 즉, 기존의 그래픽 라이브러리들이 잘 동작하지 않을 수 있으며 이 경우 디코더 솔루션 업체가 제공하는 라이브러리를 사용할 수밖에 없어 셋톱박스 업체 고유의 차별화된 솔루션을 만들기 어렵다는 난점이 있다.

세 번째는 미들웨어 부분이다. 이 중에서도 특히 논란이 되는 것은 OCAP이나 ACAP 같은 DRM 연동 미들웨어 부분인데 케이블 TV, IP-TV, 위성TV 등의 과금 문제 등 사용자 인증을 위한 모듈 수행을 위해 필요한 부분이다. 그런데 이러한 많은 경우 이러한 미들웨어들은 하드웨어 플랫폼의 범용성 문제 때문에 자바 언어로 구성된 경우가 많다. 이 때문에 디코더 솔루션은 내부에 최소 233MHz 이상의 클럭을 가지는 CPU를 내장해야 하는데 이것이 칩 당 단가 상승 요인이 된다. 자바 가상 머신을 충분히 돌리기 위해서는 적어도 300MHz 정도의 클럭이 필요하며 200MHz의 경우는 자바 가상 머신 및 미들웨어를 최적화함으로써 해결을 볼 수 있다. 그러나 100MHz대의 경우, 130nm 이상 공정을 사용하므로 칩 당 단가는 가장 저렴할 수 있으나 미들웨어 탑재에서 문제를 일으킬 수 있다. 이 문제를 해결하기 위해서는 자바로 짠 미들웨어 프로그램을 C나 기계어로 바꾸어 느린 CPU에서도 빠르게 돌아갈 수 있도록 최적화 하는 작업이 필요하다. 그러나 많은 개발 기간과 우수한 프로그램 인력이 필요하다는 점에서 여전히 어려운 과제이다.

네 번째는 인코더와의 연동 문제이다. 대부분의 디코더들은 기본적으로 1080i에서 최적화 되어 동작할 수 있도록 사전에 튜닝이 되어 있는 상태이다. 그런데, 720P 혹은 1080P를 지원해야 할 경우도 발생한다(화질은 720P가 1080i 보다 더 뛰어나다. 최고급의 화질은 1080P가 제공한다). 이 경우에는 H.264 비디오에 대한 충분한 이해가 선행되어야 한다. 하드웨어 솔루션의 특성상 비록 디코딩은 가능할 수 있다 하더라도 기본적으로 인

터레이스 방식으로 출력이 맞추어져 있다면 프레임/필드 간 불일치에 의해 화면에서 심한 떨림 현상이 발생할 수 있다. 또한 디코더가 자체적으로 해상도를 낮춰 내보낼 수도 있어 좋은 화질을 내보내기 위해 만든 프로그레시브 방식이 거꾸로 인터레이스 방식보다 안 좋은 화질을 보일 수도 있다. 더구나, 대부분의 디코더 솔루션은 초당 30프레임 디코딩에 초기 상태가 세팅되어 있기 때문에 720P 60프레임 혹은 1080i 60필드가 입력될 시 화면 떨림 현상이 크게 일어날 수 있다. 따라서 디코더 솔루션 엔지니어들은 720P 24~60프레임, 1080i 30~60필드에 대한 모든 경우의 수를 사전에 인지하고 이에 대응할 수 있도록 디코더 플랫폼을 조정해 줘야 한다.

다섯 번째는 방송용 혹은 VoD용 H.264 시퀀스를 위한 MPEG-2 디멀티플렉서이다. MPEG-2 디멀티플렉서의 경우는 대부분 하드웨어 업체에서 SDK 형태로 제공된다. 그런데 2006년 8월 현재까지 한국을 비롯하여 거의 대부분의 국가에서는 HD급 H.264의 방송 혹은 VoD 서비스를 위한 MPEG-2 TS의 세부 표준을 정하지 못하고 있다. MPEG 내부에서는 H.264를 MPEG-2 TS에 실어 보내기 위해서 2005년도에 ISO/IEC 13818-1 amendment-3이라는 표준으로 H.264를 위한 MPEG-2 TS의 표준을 지정했다. 그러나 ATSC와 같이 서비스에 따르는 세부 표준은 아직 지정되지 않아 시범 방송을 하려는 유럽 각국 및 방송사, 서비스 업체에 따라 관련 표준이 난립해 있는 상황이다. 한국의 경우도 사정은 이와 같은바, 만일 한국이 특정 회사의 MPEG-2 TS 표준을 따르게 되면 기술 종속이 심하게 우려되기 때문에 최소한 이 부분만은 한국의 독자적인 시스템 표준을 갖추어야 한다. 그러므로 H.264 디코더 관련 엔지니어들은 항상 국내의 H.264 및 IP-TV나 그 외 서비스 동향에 대한 표준화 진척상황에 큰 관심을 기울이며 이에 대응하여 자체적으로 MPEG-2 디멀티플렉서를 수정할 수 있는 기술력을 갖추어야 한다.

결론

지난해 12월을 기점으로 세계 각국, 특히 유럽을 중심으로 SD급에서 HD급 H.264 인코더 및 디코더 솔루션들이 대거 출

시되고 있다. 이는 지금까지 H.264 표준화 관련 연구에서 H.264 실용화 연구 단계로 경우 옮겨가고 있는 국내 비디오 코덱 관련 연구 동향에 비하여 세계 각국은 훨씬 앞서서 H.264 실용화 단계에 접어들고 있음을 의미한다. 이미 비디오 코덱 자체 연구 분야는 H.264의 다음 버전인 H.265의 연구 그룹이 결성되어 활동에 들어가고 있다. 이는 한국의 현재 실정은 비디오 관련 원천기술 및 구현기술 모두 세계적 연구동향에 대하여 크게 뒤떨어져 있음을 보여준다.

ETRI는 HD급 소프트웨어 HD급 H.264 솔루션을 개발하여 한국이 상대적으로 뒤쳐져 있는 고화질 H.264 헤드엔드 기술에 대하여 선진 각국의 기술 수준을 넘어서기 위해 노력하고 있으며 이를 통해 얻어진 시스템 구현 기술과 많은 실측 데이터 및 고화질 구현 기술을 발전시켜 향후 고화질 비디오 압축 기술 개발 및 표준화를 적극 주도하고자 노력 중이다. ^{RF}time

약어 정리

HD	High Definition
ATSC	Advanced Television Systems Committee
SD	Standard Definition
QVGA	Quarter VGA, Quarter Video Graphics Array
QCIF	Quarter Common Interchange Format
ASP	Advanced Simple Profile
ITU-T	International Telecommunication Union-Telecommunication Standardization Sector
AVC	Advanced Video Codec
JVT	Joint Video Team
FRExt	Fidelity Range Extension
DSP	Digital Signal Processor
FPGA	Field Programmable Gate Array
JM	Joint Model
PCI	Peripheral Component Interconnect

참고자료

- [1] Thomas Wiegand, Gary J Sullivan, Gisle Bjøn-tegaard, Ajay Luthra "Overview of the H.264/AVC Video Coding Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, July 2003, pp560-576.
- [2] ISO/IEC 14496-10:2004 FDAM 1, Information technology - Coding of audio-visual objects-part 10: Advanced Video Coding, Amendment 1 : AVC professional extensions