

HD급 H.264 기술의 발전 동향

The Research Trend of the H.264 Technology

임베디드 S/W 기술 동향 및
연구 개발 현황

| | |
|-----------------|-------------------|
| 석진욱 (J.W. Seok) | 실시간멀티미디어연구팀 선임연구원 |
| 김범호 (B.H. Kim) | 실시간멀티미디어연구팀 연구원 |
| 이정우 (J.W. Lee) | 실시간멀티미디어연구팀 선임연구원 |
| 조창식 (C.S. Cho) | 실시간멀티미디어연구팀 팀장 |

목 차

-
- I. 서론
 - II. H.264 기술
 - III. 고화질 H.264 기술
 - IV. ETRI 개발 HD급 H.264 시스템
 - V. 결론

H.264 비디오 표준은 2005년 현재, MPEG-2를 대체하는 새로운 비디오 표준으로서 세계 각국에서 신 개념 멀티미디어 서비스의 기본 기술로서 채택이 유력시 되고 있다. 한국에서는 DMB 서비스를 위한 기본 비디오 기술로서 알려져 있는 H.264는 미국 및 유럽 등에서는 차세대 HD 서비스를 위한 비디오 코덱으로 받아 들여지고 있으며 HD급 H.264 시스템 기술 경쟁이 날로 치열해지고 있는 상황이다. 본 논문에서는 HD급 고화질 H.264 기술에 대한 소개와 발전 동향을 설명하고 ETRI에서 독자 개발한 HD급 H.264 부호화기와 복호화기의 소개 및 성능에 대하여 소개하고자 한다.

I. 서론

2004년 한국에서는 세계 최초로 “손 안의 TV”를 구현한 위성 DMB 서비스가 시작되었다. 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.

하지만, 미국을 비롯한 유럽 각국에서 H.264 기술은 오히려 MPEG-2를 밀어내고 차세대 비디오 표준의 핵심 코덱으로 간주되고 있으며 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 본고장이라 할 수 있는 미국마저도 H.264 표준의 급격한 확장과 H.264 기술 자체의 발전에 제대로 대응을 하지 못하고 있다고 보아도 과언이 아니다.

2005년 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 어떠한 발전을 이루었기에 2005년~2006년 현재 HD급 비디오의 표준으로까지 각광 받게 되었는지를 알아보는 것은 H.264 표준이 어떠한 방향으로 발전할 것인가에 대한 전망과 향후 비디오 코덱이 어떠한 방향으로 발전하게 될 것인가를 가늠할 수 있는 길잡이가 될 것이다.

본고는 다음과 같이 구성된다. II장에서는 H.264 기술의 전반적인 내용을 간략히 설명하고 III장에서는 고화질 H.264 기술의 발전 동향과 내용을, IV장

에서는 한국전자통신연구원이 최근 개발한 HD급 소프트웨어 H.264 기술에 대한 내용을 서술하고 이어 결론을 서술한다.

II. H.264 기술

1. 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 압축기술의 집약체적 성격을 지니고 있었다면 2003년 7월 이후의 기술은 H.264 압축기술을 본격적으로 HD급 영상에 적용하기 위한 기술의 태동과 HD급 H.264 기술의 구현이라는 관점에서 발전되었다.

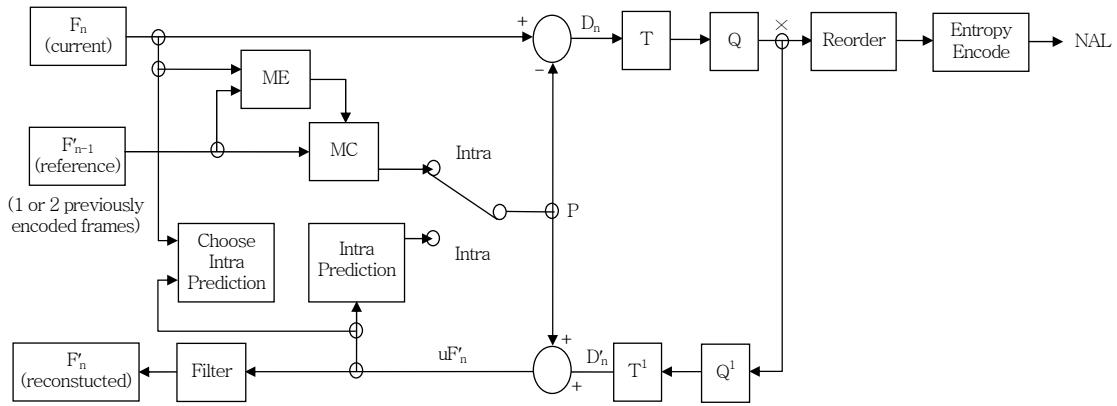
2003년 7월 이전의 H.264 기술은 우리나라에서 학계나 기업, 연구소 그리고 비디오 전문가들에게 알려진 바대로 여러 개의 블록 코딩 모드와 다중 참조 픽처 등의 특징을 사용하여 MPEG-2 보다 두 배의 압축률, MPEG-4 ASP 보다 40%의 압축률로 동일한 화질을 구현하는 비디오 압축 기술이다.

(그림 1)은 H.264 부호화기의 기본 구조도이다.

1998년 ITU-T를 중심으로한 비디오 압축 표준 그룹은 당시 제정된 ITU-T의 최신 비디오 코덱인 H.263의 절반 정도의 비트율에서 동등한 화질을 제공한다는 목표로 H.26L의 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 생각했고 동시에 H.263과의 호환성도 고려하지 않겠다는 것을 처음부터 내세우며 시작되었다.

이 과정에서 보다 더 부호화 효율을 높이기 위해 복잡도를 제한하지 않고 부호화하고자 하는 검토가 진행되었고 높은 복잡도 때문에 지금까지 목표규격에 채택되지 않았던 기술들도 적극적으로 도입되었다. 특히 하인리히 헤르쯔 연구소(HHI)의 토마스 위



(그림 1) H.264 부호화기 블록도

간드(Thomas Wiegand)에 의해 rate-distortion optimization 방식을 제안했는데 이로 인해 H.26L의 성능이 크게 향상되어 크게 주목을 받게 되었다.

Rate-distortion optimization의 원리는 DCT-양자화된 값에 대한 라그랑지안 계수의 곱과 대블록 혹은 블록간 차의 자승의 합(SSD)으로 이루어진 라그랑지안을 사용하여 이를 통해 블록 모드, 움직임 벡터, 참조 픽처 등을 결정하는 방식이다. 만일 움직임 벡터를 기술하는 데 10bit가 필요하고 DCT-양자화에 5bit가 필요하다면 차라리 DCT-양자화에 10bit를 할당하고 움직임 벡터 기술에 1~2bit를 사용한다면 전체적인 비트 사용량은 줄어들게 되어 원영상과 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 인코딩하는 경우와 차이가 나지 않는 것을 볼 수 있다.

H.26L의 성능이 알려지던 그 때 MPEG 진영에서는 MPEG-4 비디오 버전 1에서 4까지의 표준화 작업이 완료되었다. 그 이후 MPEG 진영의 참가인원 수가 줄게 되면서 MPEG 단독으로는 표준화 활동이 불가능해지자 MPEG은 ITU-T와 공동으로 표준화 작업을 하게 되었다. 이때, MPEG-4 ASP을 기준으로 P2, P3, P4 및 H.26L의 4가지 방식에 대

한 주관적 화질 테스트 결과 H.26L 방식의 우위가 확인되면서 H.26L 기반의 새로운 표준화가 시작되었고 이것이 H.264 혹은 MPEG-4 AVC의 시작이 되었다.

H.264의 표준 제정이 시작될 때 ITU-T 진영에서는 주로 영상회의 시스템과 관련된 영상압축 부호화를 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 부호화나 스케일러빌리티 등은 도입하지 않는다.
2. 높은 압축률을 표준화의 최대 목표로 한다.
3. 시간지연의 자유도를 높여 여러 서비스에 적용 가능하도록 한다.
4. 네트워크 친화성이 높아야 한다.
5. 에러 내성에 강해야 한다.
6. 부호화기와 복호화기의 규격을 스케일러블하게 변경할 수 있어야 한다.
7. 복호기 간에 복호 값이 같아야 한다.
8. 고화질 응용 시스템도 고려한다.
9. 저장 매체용 파일 기록 포맷을 정의한다.

이러한 원칙 하에 이루어진 표준화 작업 결과 현재 일반적으로 알려져 있는 H.264의 특징들로 구현

되었으며 가능한한 고압축률과 고화질을 위해 많은 기술들을 표준으로 채택하여 부호화 및 복호화에 있어 높은 복잡도를 지니게 되어 “매머드 코덱”이라는 별칭을 가지게 되었다.

그럼에도 불구하고 2003년 7월로 일단 주요 표준화 작업이 완료된 H.264 표준은 그때까지만 해도 주요 대상이 QCIF에서 SD급 영상까지였다.

2. 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 시도되었다. 특히 2004년 이전에만 하더라도 H.264와 같은 복잡도가 심한 코덱의 경우에는 ASIC 하드웨어 혹은 대규모 병렬 DSP 방식 외에는 부호화 및 복호화가 모두 어려울 것으로 예상되어 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 main profile을 확장하여 HD급 영상에 대응하기 위한 high profile의 제정과 H.264 코덱의 약점으로 알려진 color space의 보정 및 확장에 주안점이 두어진다.

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. 1pixel의 기본단위는 8bit이나 10, 12bit 역시 지원한다.
3. Loseless 코딩 및 stereo 코딩을 지원한다.
4. 기존 4×4 정수 DCT 연산에 부가하여 정수 8×8 DCT 연산을 추가한다.

5. Loseless 코딩을 위해 새로운 intra prediction 모드를 추가한다.(LPCM)
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, 무 손실 부호화를 위한 LPCM 방식이 FRExt의 주요 특징으로 채택되어 부호화기는 어떤 기준에 부합하다고 판단될 시 참조 픽처의 대블록을 그대로 복사하여 복원할 수 있는 모드가 추가되었다. 이 결과로 I-frame 부호화에서도 P-frame의 skip 모드, B-frame의 direct 모드와 같은 역할을 하는 모드가 추가되어 부호화 효율을 급격히 높일 수 있게 되었다.

한편 H.264 코덱의 약점으로 알려져 있던 색상 부호화 부분에 있어 새로운 색상공간을 도입하고 4:2:2, 4:4:4 부호화 및 복호화가 가능하도록 하여 색상 부호화의 효율을 높이는 동시에 HD급 영상장비와의 호환성을 높이도록 하였다.

마지막으로 새로운 정수 8×8 DCT 연산이 도입되었는데, 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 부동소수점 8×8 DCT에 익숙한 코덱 개발자들에게 보다 친숙한 DCT 연산을 도입한다는 취지에서 정수형 8×8 DCT가 도입되었다.

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

III. 고품질 H.264 기술

1. 고품질 H.264 기본 기술

H.264 비디오 코덱을 사용한 고품질 영상 부호화 기술은 주로 main profile을 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 부호화기에서는 5개의 서로 다른 라그랑지안을 사용하도록 규정되어 있다.

반면 rate-distortion optimization을 사용하지 않을 경우에는 SAD와 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 사용하는 경우가 사용하지 않는 경우보다 약 2~5% 정도 비트 수가 늘어난다. 반면 신호대 잡음비는 0.5dB 이상 높아짐에 따라 보다 고품질의 영상 부호화를 수행하려면 rate-distortion optimization은 필수적이라 하겠다.

그러나 rate-distortion optimization을 수행하기 위해서는 반드시 rate에 해당하는 DCT-양자화된 값의 부호화된 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 부호화까지 수행해야 한다는 의미이다.

이때 발생하는 연산량은 매우 많아 JM reference 부호화기의 경우 rate-distortion optimization을 수행하지 않는 경우에 대하여 약 10배 가까이 연산시간이 늘어난다.

그러므로 rate-distortion optimization을 고속화

하기 위한 알고리즘이 반드시 요청되며 이는 보통 DCT-양자화의 결과를 대 블록 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 부호화 연산량에 거의 영향을 미치지 않으면서 rate-distortion optimization을 비교적 훌륭히 수행할 수 있게 한다.

두번째로 필요한 부분은 B-frame의 사용이다. B-frame은 순방향 및 역방향 참조 픽처를 사용하여 순방향 예측, 역방향 예측, 양방향 예측을 할 수 있게 됨으로써 예측 부호화 효율을 크게 높일 수 있다. 보통 최적화된 H.264 부호화기의 경우 콘텐츠에 따라 P-frame의 10~20% 수준에서 부호화가 가능하다.

그러나, B-frame의 사용은 순방향, 역방향, 양방향 예측의 3개의 예측을 모두 수행해야 하기 때문에 P-frame 보다 일반적으로 많은 연산 부하가 발생한다. 그러므로 효율적인 B-픽처의 부호화를 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 드러난다는 점이다.

이는 부호화기의 실제 구현에 있어 고속화 알고리즘을 사용하면 할수록 뚜렷이 나타나는 것으로써 고속화 알고리즘의 거의 대부분이 이른바 early-skip 방식으로 이루어짐에 따라 대 블록간 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. 더군다나, 현재, H.264 표준에서 채택하고 있는 de-blocking 필터는 이웃 블록의 특성, 참조 픽처의 특성, I, P, B 부호화에 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.

2. HD급 H.264 부호화/복호화의 문제점

비록 H.264 비디오 코덱이 HD급 영상에 기존 코덱보다 강점을 가지고 있다 하더라도 H.264 비디오 코덱은 HD급 영상 부호화 및 복호화에 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 부호화 및 복호화에는 YUV 4:2:0 8bit 시스템을 사용한다고 가정했을 때 HD 720P 24fps의 경우 1픽처 당 약 500MB, HD 1080P 24fps의 경우 1픽처 당 2GB의 데이터가 요구된다.

일반적으로 serial-ATA 방식 하드 디스크의 초당 데이터 액세스 속도가 150~300MB이며 만일, HD급 영상을 PCI-X 2.0 266MHz에서 읽어 들인다 해도 266MHz에서의 데이터 전송속도는 초당 약 2GB이므로 HD 720P나 HD 1080P의 부호화 혹은 복호화는 사실상 하드웨어의 한계를 넘어선 것이라 할 수 있다.

이같은 이유 때문에 일부 H.264 부호화기나 복호화기는 실시간 인코딩을 위해서 부화소 연산을 아예 수행하지 않는 경우가 있는데 이 경우 비트 수 증가는 약 40% 이상이며 신호대 잡음비의 하락은 매우 심각한 정도까지에 이르러 사실상 MPEG-4 ASP나 MPEG-2와 아무런 차이가 없거나 혹은 더 못한 결과를 가져온다.(모드별 기술 데이터가 많기 때문)

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

반면, HD급 영상 부호화에 있어 정화소 추정 알고리즘은 큰 영향을 미치지 못한다. 일반적으로 16×16 규격의 탐색 범위를 가질 경우 그 이상의 탐색 범위와 비교하여 비트 수의 증가나 신호대 잡음비의 하락은 극히 미미하며 최악의 경우 4×4의 탐색 범위를 가지더라도 비트 수의 증가보다는 신호대 잡음비가 약간 떨어지는 정도이다.

최근 H.264 reference 부호화기에 채택된 이른바 hexagonal search 방법도 실제 구현 및 테스트 결과에 의하면 diamond 탐색 방법에 비해 약 3~5% 정도의 비트 수 하락과 미미할 정도의 신호대 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 부호화 등에서는 오히려 장애 요소일 뿐이다.

또한, 정화소 탐색 속도를 빠르게 하기 위한 각종의 예측 알고리즘의 경우에도 예측 알고리즘이 한 번이라도 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 부호화기와 복호화기 간의 상호 운용성 검증이 거의 이루어지지 않았다는 점이다. 대부분의 부호화기 생산자와 복호화기 생산자는 자사의 제품에 대해서는 최고의 성능을 발휘한다. 그러나, 다른 제품과 연동 시험에 들어갈 경우에는 심각한 화질 열화 현상(블로킹 현상, 화면 색상 변화 등)이 발생한다.

이러한 경우가 발생하는 원인은 각 부호화기 생산자나 복호화기 생산자의 경우 자신들만의 고속화 알고리즘을 채택하고 있으며 이것이 JM의 reference 부호화기나 복호화기와 서로 상이하기 때문이다. 가장 심각한 경우가 이른바 오픈 소스로 알려져 있는 VideoLAN사의 H.264 부호화기로서 이 부호화기로 부호화된 H.264 스트림은 아예 직접적으로 JM의 H.264 복호화기로 복호되지 않는다. 이는 매우 심각한 현상으로 만일 복호화기 업체가 하드웨어로 복호화기를 설계하고 판매한다고 가정했을 때 해당 부호화기로 만들어진 H.264 스트림은 아예 복호화되지 않을 수 있다는 위험성이 있다는 점이다.

따라서, 부호화기나 복호화기 업체들의 솔루션을 선택할 때 과연 이들 부호화기나 복호화기가 제대로 H.264 표준을 준수하고 있는가를 살펴보는 것은 매우 중요한 문제라 할 수 있다.

3. 기존 HD급 H.264 부호화/복호화기

이 같은 문제점에도 불구하고 2005년부터 HD급 H.264 부호화기 및 복호화기가 여러 연구소, 기업 등에서 최근 출시되고 있다. 거의 대부분의 HD급 H.264 부호화기의 경우 DSP 혹은 FPGA를 사용한 모델로서 DSP의 경우에는 Texas Instrument사의 DSP 혹은 LSI Logic의 DSP를 주로 사용한다. FPGA를 사용한 HD급 H.264 부호화기의 경우는 다수의 FPGA를 사용하여 HD급 인코딩을 수행하는 방식을 채택하고 있다. HD급 H.264 인코더는 독일의 Tandberg사에 의해 만들어진 EN5990이 있다 ((그림 2) 참조). 내부에 30여 개의 TI-DSP로 만들어진 것으로 알려져 있으며 Tandberg사 고유의 하드웨어 플랫폼인 ICE 위에 만들어져 있고 main profile을 지원하며 level은 4.0(HD 1080P)를 지원하는 것으로 되어 있다. 부호화 속도는 HD 720P 24fps, HD 1080i 30fields이다. HD급 복호화기 생산업체인 Conexant사의 HD급 연동 시험을 성공리에 마쳤으며 현재 세계에서 가장 널리 인지되어 있는 HD급 H.264 인코더 장비이며 JM reference 대비 약 15% 가량 성능이 떨어진다고 알려져 있다.



(그림 2) Tandberg사의 HD급 실시간 H.264 부호화기 EN5990

또 다른 HD급 실시간 인코더는 Modulus Video사의 ME6000 HD급 실시간 인코더이다((그림 3) 참조).

Modulus Video사의 제품은 6개의 Virtex-4 FPGA와 1개의 펜티엄-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 Xilinx사의 최신 FPGA를 사용하며 50MHz 이하에서 동작이 검증된 상태이다. 이들 업체는 120MHz 클럭에서 HD 720P 부호화가, 240MHz급에서는 HD 1080P 실시간 부호화가 가능하다고 주장하고 있으나 FPGA에서 이 정도의 클럭 주파수를 가지고 정상적인 동작을 하기는 대단히 어려운 일이므로 다소 신빙성이 떨어진다고 하겠다.

복호기의 경우 지금은 여러 업체들이 HD급 솔루션을 경쟁적으로 내놓고 있다. 가장 먼저 출시되어 많은 테스트가 이루어진 제품은 ST Micron사의



(그림 3) Modulus Video사의 HD급 실시간 H.264 부호화기 ME6000

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. 두번째 솔루션은 MPEG-2 솔루션부터 안정성에 있어 호평을 받아온 BroadCom의 7411 솔루션이다. BroadCom의 7411 솔루션은 그러나 메인 CPU로 자사의 7038 video processor를 사용하도록 되어 있는 2chip 솔루션이라는 점이 약점으로 지적된다. 세번째 솔루션은 Connexant의 Cx2418x 솔루션으로 현재 나와 있는 H.264의 모든 프로파일을 지원할 수 있는 라인업을 갖춘 것이 장점이며 동시에 Tandberg사 제품과 연동시험까지 거쳐 상호 운용성이 검증된 제품이라는 것이 강점이다. 네번째 솔루션은 SigmaDesign의 SMP8630이다. 한국에서는 Sigma Design의 인지도가 매우 높는데 그 이유는 Sigma Design의 MPEG-2 솔루션이 가격 대비 성능에서 매우 만족스러웠기 때문이다. 현재, 한국에서는 Sigma Design사의 제품이 정식으로 배포되기를 기다리고 있으나 칩 디버깅 작업 등에 의해서 Sigma Design이 애초에 약속했던 일정이 계속 뒤로 미루어지고 있는 형편이다. 이외에도 WIS, ATEME, Main Concept, Moonlight, VSoft, Apple 등의 하드웨어 및 소프트웨어 솔루션이 시장에 출시되어 있다.

IV. ETRI 개발 HD급 H.264 시스템

1. ETRI 개발 HD급 H.264 부호화기

ETRI에서는 지난 2005년 11월 HD 720P의 실시간 부호화가 가능한 소프트웨어 인코더를 개발했다. 소프트웨어 솔루션으로서는 현재, 세계에서 가장 빠른 부호화 속도를 가지고 있으며 최고 속도(25~35fps)에서 약 3~5% 정도의 비트 수 증가와 약간의 신호대 잡음비 하락이 나타난다. 한편, 최저

속도(15~18fps)에서는 JM reference 부호화기 보다 약 2% 정도 비트 수가 절감되고 있으며 신호대 잡음비는 같은 옵션의 JM reference와 비교해서 떨어지지 않는다.

JM reference 부호화기 9.6을 기반으로 구현되었으며 main profile을 지원하며 실시간으로는 level 3.1을 지원할 수 있을 것으로 예상되고 비 실시간으로는 level 4.0 이상도 가능하다.

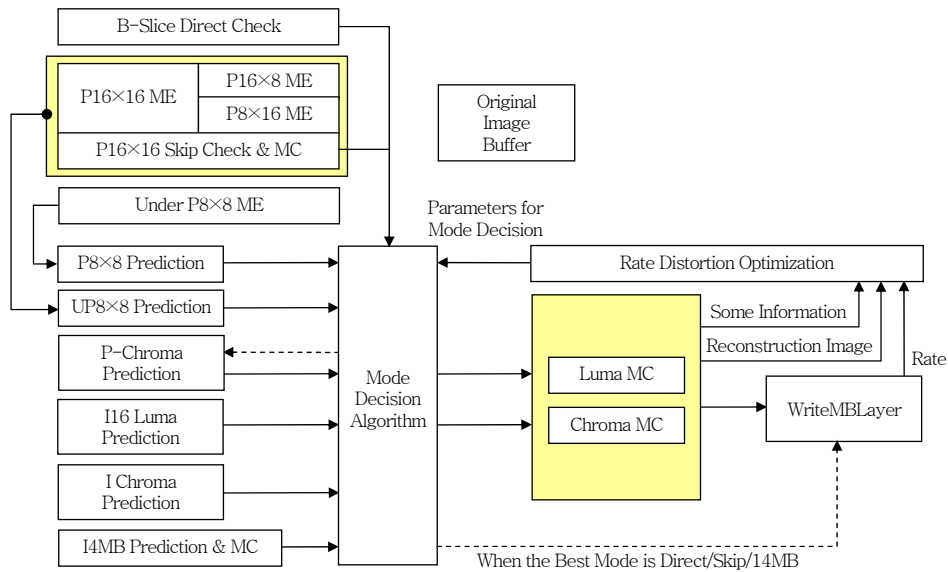
이러한 고속의 HD급 부호화기를 만들기 위해서는 여러 고속 알고리즘들이 구현되고 탑재되어야 하는데, 첫번째는 고속화 지원에 알맞은 H.264 부호화기의 구조화이다. JM reference 부호화기의 경우 모든 경우에 대한 부호화가 가능해야 하기 때문에 이러한 부호화기 구조화가 거의 이루어져 있지 않고 보아도 과언이 아니다. ETRI에서는 이러한 부호화기의 구조화를 위해 (그림 4)와 같이 부호화기를 구조화하고 있다.

또한 방대한 양의 HD급 데이터, 특히 부화소 데이터의 처리를 위해 부화소 움직임 벡터 추정시 부화소를 사용하지 않고 정화소의 SAD 값과 라그랑지안을 추정하여 부화소 예측을 수행한다. 이를 통해 B 픽처에서도 HD 720P의 움직임 추정 연산시간을 프레임 당 30msec 이내에서 수행할 수 있었다.

정화소의 HD급 데이터는 HD 720P 24fps의 경우 약 70MB, HD 1080P 24fps의 경우 150MB 수준이므로 다수의 HDD가 RAID-0으로 결합되어 있을 경우 실시간 부호화가 가능한 수준에 이른다.

또한, H.264의 특징을 살리기 위해 DCT 연산은 4×4만 사용했으며 4×4 DCT 연산의 속도를 증가시키기 위해 4개의 4×4 DCT 연산을 한 번의 8×8 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 대 블록 하나 당 4번의 SIMD 연산으로 수행이 가능하도록 구현 알고리즘을 최적화 및 새롭게 알고리즘을 개발하였다. 이로써, H.264의 표준을 그대로 따르면서도 HD 720P급 한 픽처



(그림 4) ETRI 개발 HD급 H.264 부호화기 기본 블록도

의 de-blocking 필터링은 I-픽처에서 15msec 수준이며 P 및 B에서는 15msec 이하에서 가능하다.

고속 부호화를 위해 중요한 요소인 블록 모드 결정 방법에는 I-블록, P-블록간 평균 rate-distortion 값 및 전체 블록의 평균 rate-distortion 값을 각각 Kalman 필터로 추정하고 16×16 대 블록의 움직임 보상 후 나타나는 CBP 값을 사용하여 P 블록의 모드를 결정하는 방법을 사용하였다. 이 방법을 각 단계별로 적용하여 고속 부호화의 단계를 나누었고 개발한 블록 모드 결정 방법을 사용하지 않았을 때보다 평균 30~50% 이상 부호화 속도가 향상되었으며 HD 영상에서의 비트율 증가는 3~ 5%, 신호대 잡음비 하락은 약 0.02~0.03dB 정도였다.

이로써 단일 프로세서에서 HD 720P 영상 부호화에 HyperThreading 지원 3.4GHz 펜티엄-4 프로세서에서는 6~10fps의 부호화 속도를 가지게 되었으며 dual core 3.0GHz 펜티엄-4 프로세서에서는 최대 14fps의 부호화 속도를 나타내었다.

실시간 HD 720P 부호화를 위해 ETRI에서는 인코더의 구조를 멀티 슬라이스 지원이 가능하도록 하였으며 슬라이스 단위로 멀티 스레드가 지원되도록 병렬 구조화시켜 Intel Xeon 3.66GHz 4-way 서버

에서 24~35fps의 부호화 속도를 확보하여 HD급 영상 캡처 보드에서 PCI-X 2.0 266MHz 혹은 PCI-Express 133MHz급을 지원할 경우 실시간 HD 720P 부호화가 가능하다.

2. ETRI 개발 HD급 H.264 복호화기

HD급 H.264 복호화기는 부호화기와는 달리 오픈 소스 진영의 복호화기를 ETRI 개발 부호화기에 맞도록 몇 가지를 수정하여 개발하였다. 멀티 슬라이스 구조를 가지지 않던 기존 오픈 소스 진영의 복호화기를 수정하여 멀티 슬라이스 구조의 H.264 스트림이 들어와도 이를 복호화 할 수 있도록 복호화기를 수정하였으며 몇 가지 불필요한 부분의 수정을 통해 HD급 main profile 멀티 슬라이스 구조의 스트림을 펜티엄-4 3.4GHz CPU에서 복호화 할 수 있도록 만들었다.

실험결과 H.264 비디오 복호화에서는 움직임이 크지 않은 영상에 대하여 인텔 모바일 펜티엄에서도 복호화는 가능하나 영상의 지연 현상이 발생하며 지연 현상 없이 오디오와 비디오간의 동기화까지 모두 보장하기 위해서는 적어도 HyperThreading 지원

이 가능한 펜티엄-4 3.6GHz 이상의 하드웨어 환경이 지원되어야 했다. 또한, 하드 디스크 역시 최소 초당 150MB 이상의 수행속도를 보장할 수 있는 고속의 하드 디스크가 필요했으며 HD급 영상 출력을 위해서 1920×1440 이상의 출력이 가능한 고성능 그래픽 카드가 요구된다. 따라서, 복호화 시 발열이 큰 문제가 되는데 이를 해결하기 위해서는 CPU 바로 위 상단에 CPU 크기의 1.2~1.5배 크기의 송풍구를 마련해야 하며 측면 혹은 후면에 열을 배출하기 위한 냉각팬을 별도로 설치해야 하며 케이스의 양 측면에도 반드시 송풍구를 만들어야 했다. X86 하드웨어 플랫폼에서 HD급 H.264 복호 시스템을 구성하기 위해서는 충분한 전력 공급과 시스템 냉각 시스템 설계가 반드시 필요하다.

V. 결론

2005년 12월을 기점으로 세계 각국, 특히 유럽을 중심으로 SD급에서 HD급 H.264 인코더 및 디코더 솔루션들이 대거 출시되고 있다. 이는 지금까지 H.264 표준화 관련 연구에서 H.264 실용화 연구 단계로 옮겨가고 있는 한국내 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 있음을 생각할 때 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 뒤처져 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 한다.

약어 정리

| | |
|-------|--|
| ASIC | Application Specific Integrated Circuit |
| ASP | Advanced Simple Profile |
| ATA | AT Attachment |
| ATSC | Advanced Television Systems Committee |
| AVC | Advanced Video Codec |
| CABAC | Context-based Adaptive Based Arithmetic Coding |
| CBP | Coded Block Pattern |
| CPU | Central Processing Unit |
| DSP | Digital Signal Processor |
| FIR | Finite Impulse Response |
| FPGA | Field Programmable Gate Array |
| FRExt | Fidelity Range Extension |
| HD | High Definition |
| HHI | Heinrich Hertz Institute |
| ITU-T | International Telecommunication Union-Telecommunication Standardization Sector |
| JM | Joint Model |
| JVT | Joint Video Team |
| MPEG | Moving Picture Expert Group |
| PCI | Peripheral Component Interconnect |
| PCM | Pulse Code Modulation |
| QCIF | Quarter Common Interchange Format |
| QVGA | Quarter VGA, Quarter Video Graphics Array |
| SAD | Sum of Difference |
| SD | Standard Definition |
| SSD | Sum of Squared Distortion |
| UVLC | Universal Variable Length Coding |
| WMV | Windows Media Video |

참고 문헌

- [1] Thomas Wiegand, Gary J Sullivan, Gisle Bjøntegaard, and 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, pp.560-576.
- [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, 2004.