목차

VR 관련 기술

- 스트리밍 기술

- 인코딩 기술

최신 VR기술

- 타일 분할 스트리밍 기술

- 분산 컴퓨팅 기반 스트리밍

VR 서비스 관련

- VR 개요

- VR 서비스 요구사항

- VR 서비스의 구조

- 해결 과제

VR 관련 기술

스트리밍

# VR 영상의 스트리밍 기술

## MPEG-DASH

MPEG-DASH란 무엇일까? 먼저 MPEG란 Moving Picture Expert Group의 약자로 동화상 전문가 그룹이라는 뜻이다. [[1]](#endnote-1) DASH란 ‘Dynamic Adaptive Streaming over HTTP’의 약자로 adaptive 라는 말에서 알 수 있듯이, 클라이언트가 자신의 단말기 쪽 네트워크 상황에 맞추어 적절한 비디오 화질을 선택하여 서비스를 받을 수 있는 기술을 말한다.

즉 MPEG에서 DASH라는 프로토콜을 만들었다는 것이다. 2008년부터 HTTP 기반의 Adaptive Streaming에 대해 MPEG에서 뜨거웠는데, 이에 따라 MPEG에서 MPEG-DASH라는 그룹을 만들고 관련된 Streaming Package 포맷에 대해 표준화를 시작한 것이다.

MPEG-DASH에는 두가지 요소가 있다.

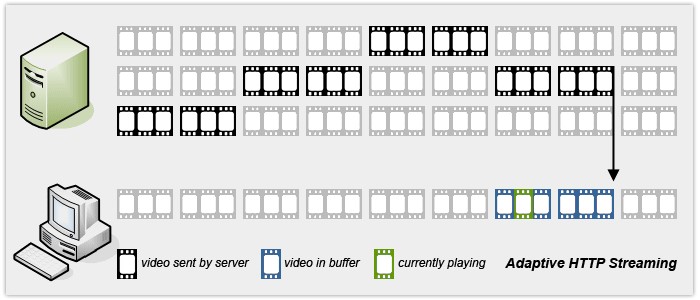
* Media Presentation (MP)
* Media Presentation Description (MPD)

MP는 오디오/비디오를 하나의 파일로 묶거나 각각의 파일로 분리 할 수 있다. 그리고 MPD는 스트림의 정보를 나타내는 XML 이다.

DASH의 기본 동작에 대해 알아보자.

DASH 서버에서는 비디오 콘텐츠를 각기 다른 비트 rate를 갖는 여러 버전으로 인코딩 한 다음, 인코딩된 비디오 데이터를 작은 세그먼트(또는 chunk)단위로 분할한다. 이 세그먼트들이 DASH 클라이언트의 요청에 따라 전달이 되는데, DASH 클라이언트에서는 자신의 상황에 따라서 bit rate 적용 알고리즘을 수행하여 네트워크의 가용 대역폭을 계산하고 그 결과를 토대로 다음 비디오 세그먼트의 bit rate를 결정한다. 즉, DASH 클라이언트의 비디오 플레이어가 현재 단말 네트워크의 상태 및 QoE를 고려하여, 적절한 비디오의 화질을 선택하여 서비스 하도록 한다는 것이다. [[2]](#endnote-2) [[3]](#endnote-3)

## Adaptive HTTP streaming

 [[4]](#endnote-4)

Adaptive HTTP Streaming 이라는 단어에서 알 수 있듯이, HTTP를 이용한 스트리밍 방식이며 Adaptive란 유저의 bandwidth 환경에 따라 그에 맞는 동영상을 보내주는 기술이다.

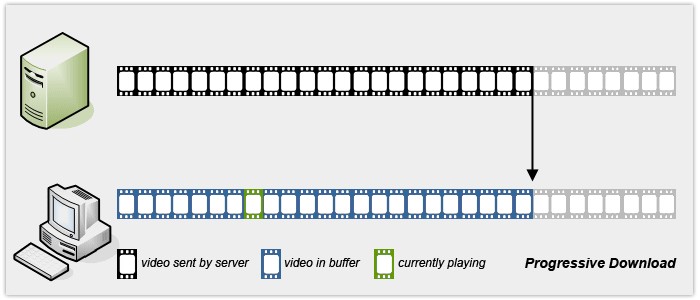
RTMP/RTSP 스트리밍과 프로그레시브 다운로드의 장점을 결합한 방식으로 서버는 작은 조각 단위(Chunk : 몇초 단위) 의 동영상을 가지고 스트리밍하게 되며, 유저는 이 비디오 조각을 연속된 스트림으로 연결시킨다.

장점으로는

1. 다양한 환경에 있는 다양한 유저들에게 적절한 동영상을 보내줄 수 있다.
2. Bandwidth / CPU 상황 등을 체크하여 그에 맞는 화질과 음질을 내려주는 방식으로 품질개선 및 안전성을 확보할 수 있다.
3. (Progressive 장점) 미디어 서버와 같은 비싼 장비와 RTMP 프로토콜 같은 복잡한 기술을 사용하지 않는다.
4. (RTMP 장점) Bandwidth 효율성이 뛰어나고, 비디오 화질 변경이 용이하다.

단점으로는 유저에 따라 적절한 동영상을 보내주어야 하기에 서비스 공급자는 여러 bit-rate 로 인코딩하여 저장해야 하는 불편이 있다. [[5]](#endnote-5)

* Progressive Streaming

[[6]](#endnote-6)

Progressive Streaming이란 Progressive 에서도 알 수 있듯이, 점진적으로 파일을 로컬에 다운로드를 한 후 재생하는 방식이다. 이미 다운로드가 된 부분은 재생이 가능하다.

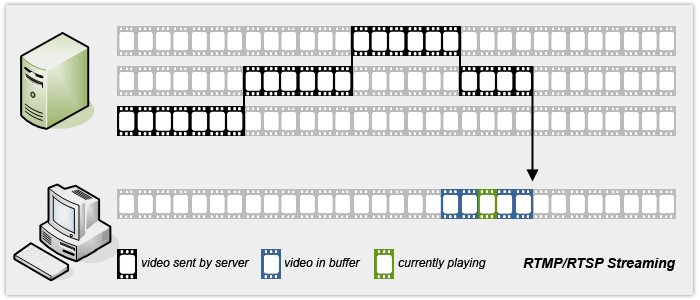
Adaptive HTTP Streaming 이 착안한 Progressive Streaming의 장점은

1. 구현이 쉽고 비용이 저렴하다. 미디어 서버가 아니라 웹 서버에 동영상 파일을 저장하기 때문이다. 따라서 유저는 웹 서버를 통해서 파일을 다운받게 된다. 또한 당연히, HTTP 프로토콜을 사용한다.

그러나 단점 또한 있다.

1. 다운로드를 한 후에 재생하는 방식이기 때문에, 영상을 끊기지 않고 보기 위해서는 유저의 네트워크의 bandwidth가 웹서버에서 제공하는 bit-rate보다 높아야 한다.
2. 파일 전체를 다운로드 하기 때문에, 일부만 보더라도 파일 전체에 대한 데이터가 소모된다. 예를 들어, 1GB 사이즈의 동영상을 본다고 할 때, 사용자가 300MB 만 시청하고 나갔을 경우, 나머지 700MB 의 동영상도 이미 다운로드 완료되는데, 이것은 실제 시청한 데이터에 비해 전송 비용이 들었다는 것이며, 큰 규모에서는 고비용을 의미한다.
3. 다운로드 후에는 파일을 화질이 고정된다. 따라서 동영상을 보는 중간에 화질을 변경할 수 없다.

* RTMP/RTSP

[[7]](#endnote-7)

RTMP는 Real Time Messaging Protocol의 약자로, 오디오, 비디오 및 기타 데이터를 인터넷을 통해 스트리밍 할 때 사용한다.

Adaptive HTTP Streaming 이 착안한 RTMP의 장점은

1. 각기 다른 네트워크 환경에 있는 유저들 에게 적절한 영상을 보내 줄 수 있다. 이유는 Progressive Streaming과 달리 서버에 서로 다른 bit-rate로 인코딩된 파일을 올려놓았기 때문이다. 따라서 중간에 화질 변경 또한 가능하다.
2. Seeking이 가능하다. Progressive Streaming과 달리 시청자가 요청하려는 일부분의 frame 단위로 전송하기 때문이다.

그러나 단점 또한 있다.

1. Flash 기반이라, 모바일 기기에서는 사용이 어렵다.
2. 장점 1번을 구현하기 위해서는 스트리밍 서버(미디어 서버)의 구축이 필요하며, 이를 관리할 관리자가 필요하다.
3. FMS (Flash Media Server) 와 Wowza Media Server 라는 유료 소프트웨어가 필요하다.

다음으로 RTSP는 Real Time Streaming Protocol의 약자로 오디오, 비디오 등의 멀티미디어 데이터를 포함하는 미디어 서버를 원격 조작하기 위한 프로토콜이다. 명령어는 “PLAY”, “PAUSE” 같이 VCR 동작하고 비슷하며 시간 정보를 바탕으로 서버에 접근한다.

따라서 RTMP/RTSP 란 RTMP로 구축해 놓은 환경에서 RTSP프로토콜을 이용하여 클라이언트에게 동영상을(프레임을) 보내주는 것 이라 할 수 있다. [[8]](#endnote-8)

인코딩

# 인코딩, 디코딩 그리고 코덱

코덱을 간단히 한마디로 말하자면 영상을 재생하기 위해 필요한 프로그램이라고 할 수 있다. 코덱은 영어로 Codec 이라 하는데, 이는 Coding과 Decoding의 합성어 이다. 그렇다면 Coding은 무엇이고 Decoding은 무엇일까?

Coding은 coder 라는 기능을 하는 것, Decoding은 decoder 라는 기능을 하는 것을 말한다. 먼저 coder란 아날로그 신호를 디지털 신호로 바꾸는 일을 하는 것이고 decoder는 coder가 바꾼 디지털 신호를 다시 아날로그 신호로 바꾸는 것이다.

coder가 아날로그 신호를 디지털 신호를 바꾸는 것을 영상 원본을 압축한다고 볼 수 있는데 (즉 영상에 변화가 일어난다), 압축해야 하는 이유는 영상 자체의 파일은 크기가 너무 커 바로 전송할 수 없어, 파일을 작게 만들기 위함이다. 다음으로 decoder가 디지털 신호를 다시 아날로그 신호로 바꾼다는 것은, coder로 변환한 영상을 재생하기 위함이다. [[9]](#endnote-9) [[10]](#endnote-10)

전자기기이(가) 표시된 사진

높은 신뢰도로 생성된 설명

# 다양한 코덱들

코덱은 하나의 프로그램이다. 따라서 서로 다른 코덱은 서로 다른 알고리즘을 쓰기 때문에 호환되지 않는다.

코덱은 다양한 종류가 존재하는데, 이는 기술의 발전과 다양한 기기 그리고 서로 다른 필요에 의해 개발되었다. Divx, Xvid, MPEG, INDEO, QuickTime, FLV, DV, MP4, H.263, H.264, Cinepak, M-JPEG 등이 있다. [[11]](#endnote-11)

# MPEG의 코덱 통합

* MPEG-1
* 디지털 저장매체용 (최대 1.5 Mbit/s)
* 1992년말 국제표준 [IS 11172-1,2,3,4,5]

1992년 국제표준으로 확정

최대 약 1.5 Mbit / s의 디지털 저장 매체 용 동영상 및 관련 오디오 코딩

* MPEG-2
* 디지털방송, 통신, 저장매체용 (2~30 Mbit/s)
* 1994년말 국제표준 [IS 13818-1,2,3,4,5,6,7,9]

동영상 및 관련 오디오 정보의 일반적인 코딩

1994년 국제표준으로 확정

* MPEG-4
* 초저속 전송, 고압축 저장용 (64 kbit/s 이하)
* 1995년시작, 1998년 국제표준 [IS 14496-1,2,3,4,5,6]

1993년부터 논의가 시작 CDROM을 응용 목표로 한 1.5Mbits/sec의 MPEG-1 표 준과 Digital TV를 응용 목표로 한, 5Mbits/sec의 MPEG-2가 다루지 않는 초저비트율 오디오/비디오 압축을 목표로

ITU-T의 H.263 표준보다 더 좋은 성능의 표준을 만드는 것이 어렵다고 판단

새로운 목표를 통신, 컴퓨터, 방송의 융합 환경의 분석 을 통해, 내용기반의 오디오/비디오 데이터 부호 표준 으로 정립하였다

* + MPEG-4 버전 1 비디오

전체 영상 중에서 특정 객체만을 분리 하여 압축할 수 있는 객체 기반 동영상 압축을 기반으로 하며, 객체의 모양정보의 압축 기술을 특징으로 한다.

이동통신 등의 에러 환경에 대응하기 위한 에러 내구성 기술과 같은 내용의 동영상을 다양한 단말과 통신망의 환경에 따라 효율적으로 적당한 비트율로 전송/재생할 수 있는 다양한 신축 부호(scalable coding) 기술을 표준화

시청각 객체의 코딩

1998년 국제표준으로 확정

* + 멀티미디어압축표준(A/V 객체부호화)
* MPEG-7
* 멀티미디어컨텐츠 정보의 표현 방법
* 2001년 국제표준 [IS 15938-1,2,3,4,5,6,7]

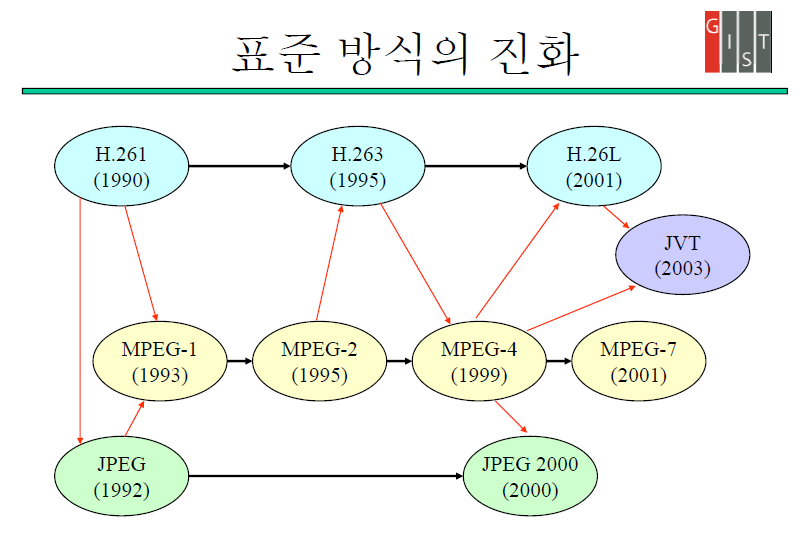
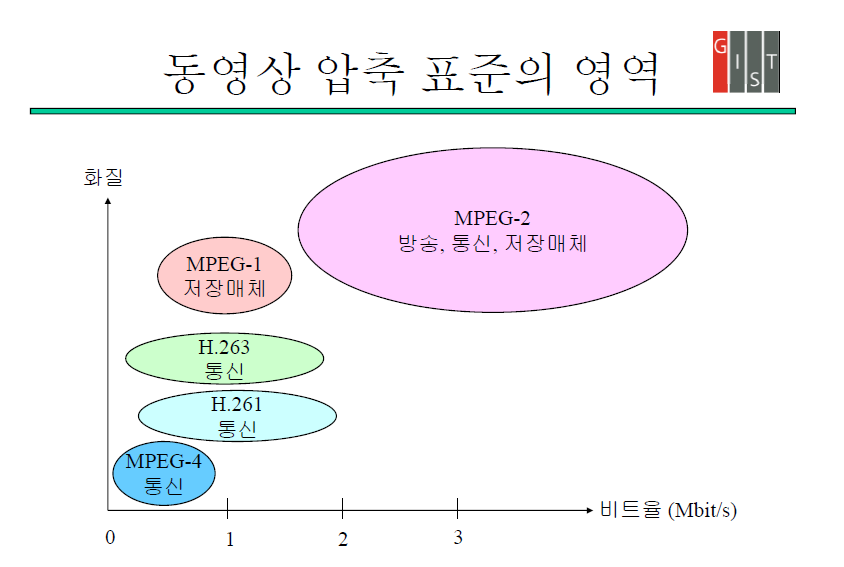
멀티미디어 콘텐츠 설명 인터페이스

2001년 국제표준으로 확정

* MPEG-21

멀티미디어 프레임워크

2003년 국제표준으로 확정

|  |  |  |  |
| --- | --- | --- | --- |
| 국제 표준 | 전송률 | 처리 신호 | 응용 분야 |
| JPEG | 20:1 | 정지영상 | 사진 저장, 전송 |
| H.261 | P x 64 kbps | 동영상 | 영상전화, 영상회의 |
| H.263 | P x 64 kbps | 동영상 | 영상전화, 영상회의 |
| MPEG-1 | 1 – 1.5 Mbps | 동영상/오디오 | 디지털 저장매체 |
| MPEG-2 | 1-30 Mbps | 동영상/오디오 | 방송, 통신, 저장 |
| MPEG-4 | <= 4 Mbps | 동영상/오디오 | 인터넷, 이동통신 |

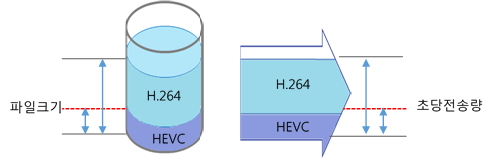
# HEVC의 사용

## HEVC란?

2013년 4월에 발표된 MPEG-H Part.2 규격으로 기존보다 압축 효율이 높아서 HEVC(High Efficiency Video Coding)라고도 한다. HEVC를 표준화시킨 그룹은 JCT-VC(Joint Collaborative Team on Video Coding)이며 여기엔 ITU-T/ISO/IEC가 참여했다.

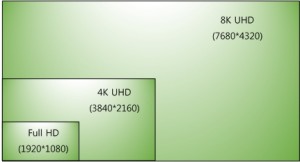
## HEVC의 등장

최근 보편화된 H.264와 근래 이슈가 되고 있는 HEVC(High Efficiency Video Coding/H.265)의 성능 차이를 그림으로 비교해보면,



이와 같은 성능의 차이가 난다.

다음 예를 보자.



H.264로 제작된 Full HD 영상을 4Mbps로 서비스하고 있다고 가정했을 때, 이를 해상도를 4K 해상도로 높이면 단순 산술 계산시 필요한 네트웍 용량은 16Mbps가 되며, 8K 해상도로 가면 32Mbps가 된다. 이러한 상황에서 압출률이 H.264대비 50% 개선되었다 가정한 HEVC(H.265) 를 활용한다고 하면 4K 서비스에 필요한 용량은 8Mbps가 되며, 8K 서비스에는 64Mbps가 되는 것이다. [[12]](#endnote-12)

물론, 현재 보편화 된 H.264도 4K UHD에 대응하지 못하는 것은 아니지만 4K UHD에 이어 다가올 8K UHD까지 대응하려면 H.264 보다 높은 압축율이 필요하고 이를 위해 만들어 낸 것이 HEVC, 즉 H.265다.

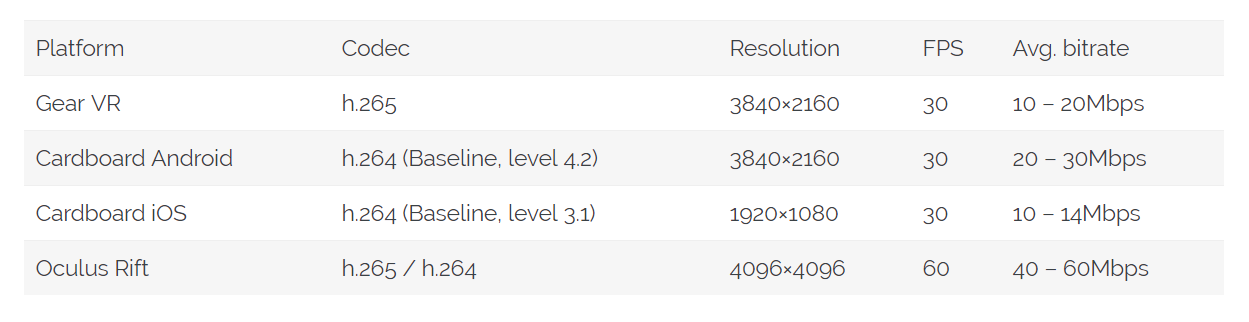
## HEVC의 사용

2014년 11월 발매된 삼성의 미러리스 카메라 NX1

2015년 8월에 표준 규격으로 발표된 UHD 블루레이

# VR 영상의 코덱

VR 영상에서 사용되는 코덱으로는 H.265, H.264, H.264 (Baseline, level 4.2), H.264 (Baseline, level 3.1) 가 있다. [[13]](#endnote-13)



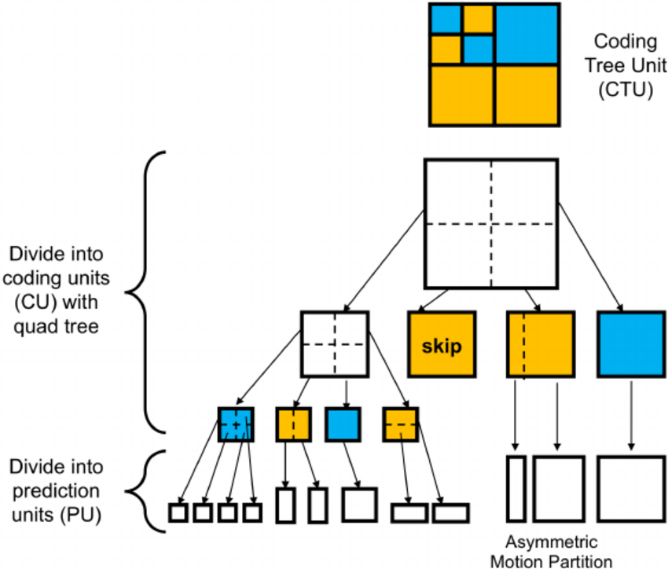
최신 VR 기술

타일 분할 스트리밍 기술

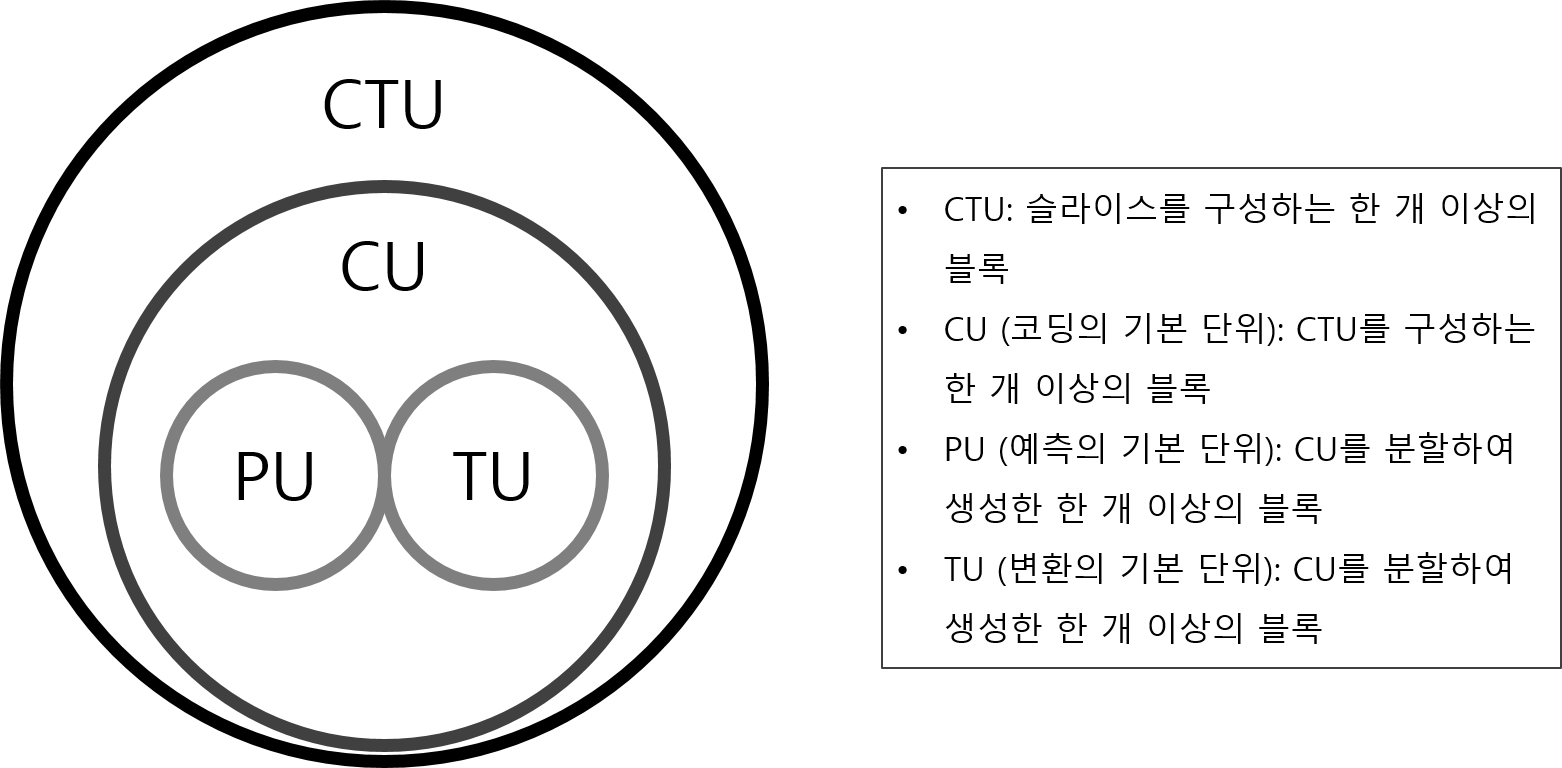
## HEVC (타일 위주의 설명)

HEVC는 High Efficiency Video Coding 의 약자로, ISO/IEC MPEG과 ITU-T의 영상 부호화 전문가 그룹(Video Coding Experts Group)이 Joint Collaborative Team on Video Coding (JCT-VC) 라는 팀에서 2013년에 새롭게 제정한 국제 비디오 표준이다.

기존 H.264/AVC 대비 약 50% 압축 효율이 향상되었고 4K/8K UHD 비디오를 지원한다.

 [[14]](#endnote-14)

HEVC는 Block-based hybrid video coding 구조로 설계되었다. 이에 따라 16x16이라는 고정적인 크기의 Macroblock를 기반으로 encoding 및 decoding을 하는 H.264/AVC와 달리 64x64, 32x32, 16x16의 가변적인 크기를 갖는 CTU를 사용한다.



### New Parallelization Tools ‘Tiles’

* Tile Parallelization

HEVC에서 쓰일 새로운 병렬처리 기법이다. 이것은 Picture를 가변적인 크기의 tile로 나누는데, 이 tile들은 서로 의존성을 갖지 않으며 헤더를 추가로 사용하지 않는다.

* Tile-Based Encoding

병렬처리를 지원하는 Tile기법을 적용하여 encoding하는 것이다. Viewport 기반 타일 선택을 위해 MCTS(motion-constrained tile set)를 적용한다. 이에 따라 stream시 원하는 tile만 선택하기 때문에 bandwidth의 효율이 증가할 수 있다.

* Tile-Based Decoding

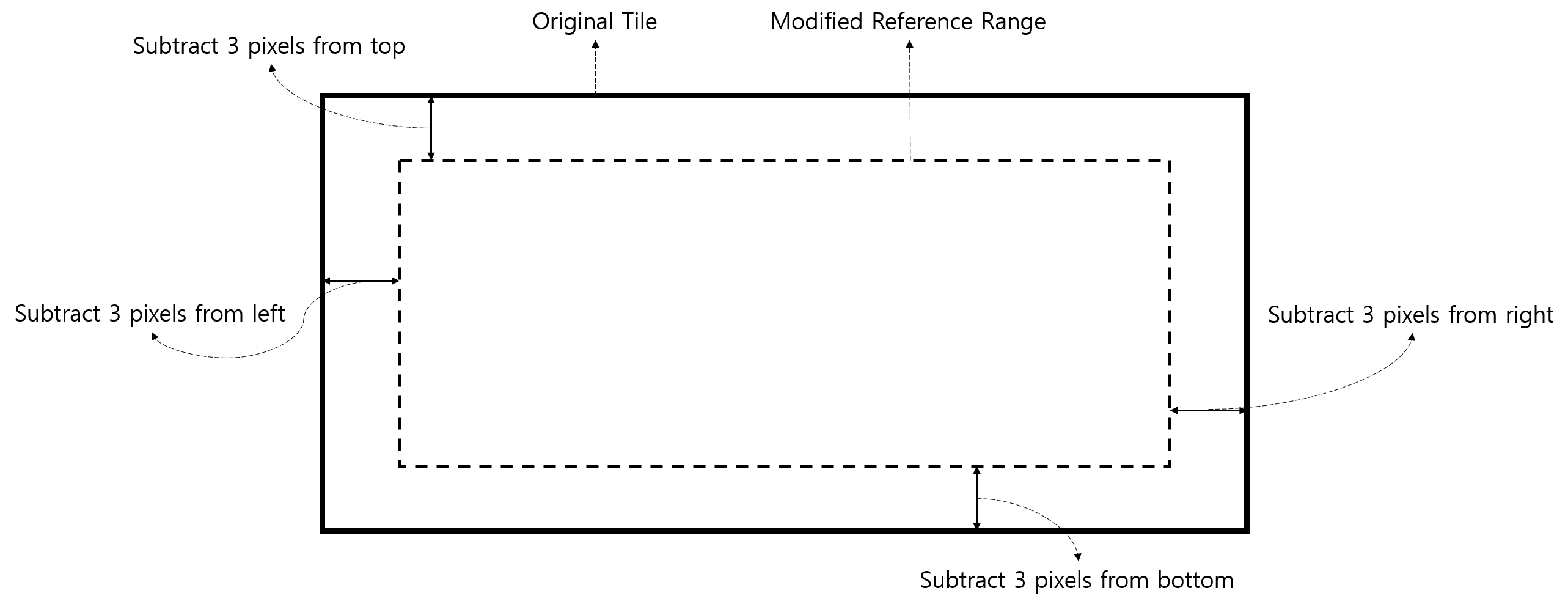
Viewport를 기반으로 하여 부분 타일을 decoding하는 것이다. 이에 따라 storage의 효율이 증가할 수 있다. 또한 Viewport에 해당하는 tile은 고화질로, 아닌 tile은 저화질로 decoding한다.

* Motion-Constrained Tile Set (MCTS)

(뭔지 모르겠다…)

* + Tile Boundary Consideration of Current MCTS

보간을 고려하여 참조 범위를 수정한다.



* + MCTS Performance Comparison

|  |  |  |  |
| --- | --- | --- | --- |
| Test Sequence  (8192x4096) | Original bitrate  (9 Tiles) | Proposed bitrate  (4 Tiles) | Proposed bitrate  (1 Tiles) |
| KiteFlite | 60 mpbs | 28 mpbs(-53%) | 7 mpbs(-88%) |
| Harbor | 22 mpbs | 10 mpbs(-53%) | 2 mpbs(-88%) |
| Trolley | 32 mbps | 16 mpbs(-50%) | 4 mpbs(-87%) |
| GasLamp | 10 mpbs | 5 mpbs(-47%) | 1 mpbs(-86%) |

* + Tile Extractor for MCTS

병렬처리와 NAL단위 추출을 위해 타일과 슬라이스를 병합하여 인코딩 된다. 추출된 슬라이스 단위의 NAL을 분석하여 원하는 타일(Viewport 기반)을 decoding한다.

## HEVC의 문제점

H.265는 높은 압축율로 기존 H.264 대비 비트레이트를 절반으로 낮출 수 있는 장점이 있지만 이로 인해 기존 보다 더 높은 컴퓨팅 파워를 요구하는 단점이 있다. 언제나 새로운 압축 기술이 개발되면 그에 맞게 요구하는 컴퓨팅 파워도 높아져 왔으니 그리 놀랄 일은 아니지만 이번엔 좀 정도가 심하다.

H.265는 2배의 압축율을 달성하기 위해 5배 이상 복잡한 계산을 처리해야 한다. 실질적인 인코딩 작업 시간 기준으로 기존 보다 대략 200배 가까이 느리다는 주장도 있어 현재의 PC로는 실시간 인코딩은 꿈도 꾸지 못할 정도라고 한다. [[15]](#endnote-15)

# References

1. https://ko.wikipedia.org/wiki/MPEG [↑](#endnote-ref-1)
2. http://unipro.tistory.com/109 [↑](#endnote-ref-2)
3. http://donghoson.tistory.com/48 [↑](#endnote-ref-3)
4. http://www.jwplayer.com/blog/what-is-video-streaming/ [↑](#endnote-ref-4)
5. http://blog.kollus.com/?p=144 [↑](#endnote-ref-5)
6. http://www.jwplayer.com/blog/what-is-video-streaming/ [↑](#endnote-ref-6)
7. http://www.jwplayer.com/blog/what-is-video-streaming/ [↑](#endnote-ref-7)
8. http://blog.kollus.com/?p=144 [↑](#endnote-ref-8)
9. <http://allpcpc.com/encording/2001>, [↑](#endnote-ref-9)
10. https://m.blog.naver.com/PostView.nhn?blogId=angelacha3&logNo=20010030103&proxyReferer=https%3A%2F%2Fwww.google.co.kr%2F [↑](#endnote-ref-10)
11. http://bwessay.tistory.com/199 [↑](#endnote-ref-11)
12. http://blog.kollus.com/?p=430 [↑](#endnote-ref-12)
13. https://purplepill.io/blog/best-encoding-settings-resolution-for-4k-360-3d-vr-videos/ [↑](#endnote-ref-13)
14. https://www.google.co.kr/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&ved=0ahUKEwjW6N\_Gz67WAhXHXLwKHWNcCcgQjhwIBQ&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2F290180658\_fig5\_Fig-6-A-coding-tree-unit-is-subdivided-into-CUs-along-the-associated-coding-quadtree&psig=AFQjCNGrOJJVwWZgEswBqBKhbr8YLDh-lg&ust=1505820193509220 [↑](#endnote-ref-14)
15. http://www.kbench.com/?q=node/123783 [↑](#endnote-ref-15)