H.264/SVC를 기반으로 한 ROI 화질 개선 방법

이정환O 유혁
고려대학교 컴퓨터학과
{jhlee, hxy} @os.korea.ac.kr

H.264/SVC Based ROI Quality Improvement Method

Jung-Hwan LeeO Hyuck Yoo
Department of Computer Science and Engineering, Korea University

요 약

본 논문에서는 H.264/SVC(Scalability Video Coding)를 기반으로 한 ROI(Region Of Interest) 화질 개선 방법을 제안 한다. H.264/SVC는 공간적, 시간적, 화질적 확장성을 지원하고 있다. 이로 인해 네트워크 환경과 모바일 기기에 더욱 적합하게 되었다. 하지만 이러한 확장성은 객관적인 화질 개선을 고려하여 설계 되었다. 본 연구는 주관적인 화질 개선을 위해 관심 영역을 설정 하였으며, 이를 위하여 관심 영역을 설정 하기 위한 방법과 이를 효율적으로 전송 하기 위한 효과적인 전송 방법을 제안 한다. 또한 기존 SVC 확장성까지도 활용하는 화질 개선 방법을 제시 한다.

1. 서 론

H.264/SVC(scalable video coding)은 ITU-T Video Coding Experts Group(VCEG)의 Joint Video Team(JVT)와 ISO/IEC moving Picture Experts Group(MPEG)이 합작으로 표준화를 시작하였다. 확장성(scalability)은 MPEG-2, H.263, MPEG-4(part2) Visual[1]에서도 시도가 있었지만, 기존 방법에 비해 부호화 효율과 환경적 요소 때문에 성공하지 못하였다. 하지만 H.264/SVC는 기존 H.264/AVC의 높은 부호화 효율과 확장성(scalability)의 꾸준한 요구로 인해 표준안이 본격적으로 논의되기 시작했다.

H.264/SVC의 장점은 단일 압축 부호화 방식에서 벗어나 다중 압축 부호화를 한다는 것과 세 가지 (공간적, 시간적, 화질적)형태의 확장성(scalability)를 지원하고 있다. 또한 각기 다른 사용자의 확장성(scalability)을 지원하기 위한 재인코딩 또는 코드의 수정이 필요 없다는 장점이 있다. 이러한 점은 기존의 비디오 코딩 방법 보다 사용자의 기기와 네트워크 환경에 적합하다.

다음 세대의 네트워크 환경은 여러 개의 망을 이용하며, 사용자의 요구에 따라 망을 선택하게 될 것이다. 실예로 수도권을 중심으로 구축중인 Wi-bro, WCDMA를 확장한 HSDPA, 도시 곳곳에 설치 되어 있는 AP를 통해 연결 되는 WLAN등 현재에도 여러 가지 망이 존재한다[2]. 이에 따라 최근에 출시된 사용자의 기기들에도 여러 종류의 네트워크 인터페이스들이 지원하고 있다. 그 예로 I-Phone은 Wi-Fi(802.11g), GSM, Bluetooth 2.0 세 개의 무선 네트워크 인터페이스를 지원하고 있다.

이렇게 각기 다른 네트워크 망과 사용자 기기의 다양성을 지원하며 비디오 스트리밍을 하기 위해서 H.264/SVC의 활용 범위가 더욱 확대 될 것이다.

더욱이 여러 개의 망을 통해 H.264/SVC를 보다 효율적으로 이용하기 위해서는 네트워크 환경과 인코딩된 비디오 정보와의 상호 보안성이 요구된다. 이에 대한 연구로는 SVC로 인코딩된 비디오 정보의에 우선순위를 두어 채널의 신뢰도에 따라 구분하여 전송하는 다중 전송 방법[2], 네트워크 환경에 따라 GOP(group of picture)의 크기를 변경하는 방법[3], 콘텐트의 내용과 네트워크 환경에 따라 FMO(flexible macroblock ordering)방식을 달리 하여 전송 하는 방법이 있다[4].

하지만 이러한 방법들은 화면의 객관적인 화질 개선에 관해서만 고려 하고 있다. 사실 사용자들은 화면 전체를 보면서도 관심 있게 보는 부분과 그렇지 않은 부분으로 나뉜다. 예를 들어 골프 선수가 공을 칠 때 사용자들은 주변의 갤러리들 보다는 선수와 공에 집중할 것이다. 이런 부분을 ROI(Region of Interest)라고 하며 ROI에 관한 기술은 H.263과 MPEG-4 FGS 에서도 이미 존재하였다. 하지만 본 논문에서는 SVC에서 네트워크 환경에 따른 능동적인 ROI구성과 효과적인 전송 방법에 그 의의가 있다.

2. 관련 연구

2-1. ROI 선택 방법

ROI(Region Of Interest)코딩에 관한 기술은 다양하지만 그 목적은 높은 화질의 ROI영역과 낮은 화질의 ROI가 아닌 부분을 배경 부분으로 지정 하는 데 있다. ROI를 지정하는 데에는 두 가지 방법으로 분류

할 수 있다. 첫 번째는 능동적인 ROI 코딩 방법, 두 번째는 수동적인 ROI 코딩 방법이다.
수동적인 방법은 인코딩 시에 미리 정의 되어 있는 영역을 환경에 변화에 상관없이 ROI를 설정 하는 것이다. 이 방법의 예로는 화면 전체가 고정 되어 있는 CCTV의 경우 사용자가 특정 화면을 지정할 때 쓰인다. 이 방법은 max-shift에 속하여 인코딩 시에 정의 된다.[5] 능동적인 방법은 ROI가 지정 되어 있는 것이 아니라 환경이나 콘텐트에 따라 변할 수 있다. 이런 방법으로는 그림 1의 좌측과 같이 인코딩 시에 Motion Vector를 활용하여 움직임이 가장 많은 영역을 ROI로 지정하는 방법, 그림 1의 우측과 같이 Macroblock의 범위로 지정하는 방법, 필터를 이용하여 고역/저역 지역으로 나눠 고주파 영역을 ROI로 지정 하는 방법 등이 있다. 본 논문에서는 네트워크 환경과 사용자 기기의 네트워크 인터페이스 별로 능동적인 ROI 지정 방법을 이용하였으며, ROI 지정 방법은 Motion Vector값을 이용하였다.

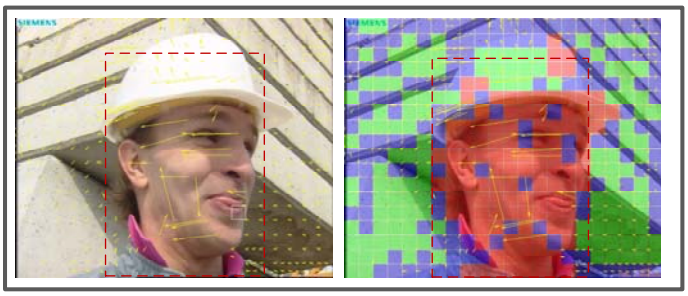

그림 1 ROI assign method

2-2. 다중 채널 전송 기법

앞 장에서 정의된 ROI 코딩 방법은 H.264/SVC 인코딩 시 계층화된 비디오 데이터로 구성된다. 이렇게 인코딩 정보를 다양한 네트워크 인터페이스를 지원하는 모바일 기기를 지원하려면 네트워크에서 다중 채널 데이터 전달 기법인 Path diversity 연구를 활용할 수 있다.[6] 즉, 계층화된 비디오 전송에 다중채널 기법 및 멀티호밍(Multihoming)기법을 도입함으로써 네트워크 및 하드웨어 자원의 사용을 최적화 할 수 있다. 구체적인 방법으로는 H.264/SVC로 인코딩 된 정보들은 확장성(scalability)에 따라 하나의 비디오 콘텐트가 다양한 레이어들로 구성 되어 있다. 하지만 이러한 레이어들은 그림 2과 같이 상위 레이어(enhancement layer)가 하위 레이어(base layer)를 참조 하여 복호화(decoding) 되는 방식으로 설계되어 있다. 따라서 하위 레이어 없는 상위 레이어는 의미가 없다. 때문에 이를 이용하여 레이어 별로 우선순위를 두고 이종 망 간의 네트워크 상황 정보를 수집한 후 분석하여 선택적으로 전송 한다[7]. 더욱이 모바일 네트워크는 네트워크 별로 Delay와 Throughput, 채널 별 전송 에러가 물리적 환경에 따라 다르다. 또한 에러 컨트롤 기술이 다르고, 환경적 요인에 의해 시시각각 변화 하기 때문에 주기적인 채널 상태 정보를 받아 스트림 서버에 전송 하는 과정이 필요 하다[7]. 이렇게 받은 정보를 미디어 스트림 서버에서 그림 3와 같이 하나의 비디오 데이터를 사용자 기기와 네트워크 인터페이스 별로 적응적으로 전송한다. 이렇게 다중 채널로 계층화된 레이어를 전송 한다면 네트워크 자원과 모바일 기기의 자원을 최대로 활용 할 수 있을 것이다.
본 논문에서의 다중 채널 전송 기법은 전송 하려는 비디오 데이터의 특성(ROI scalability)을 위해 채널 환경이 다른 이종 망의 네트워크 특성과 환경적 요소의 고려가 필요하다.

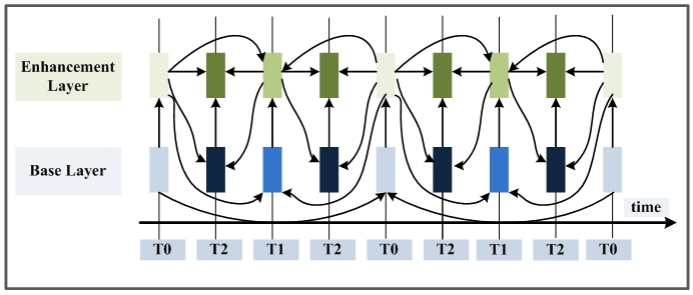

그림 2. SVC temporal prediction structure

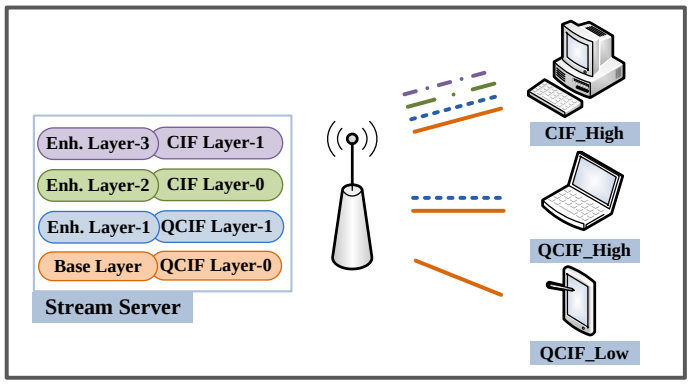

그림 3. Multi channel transmission to SVC

3. ROI를 이용한 Scalability

ROI 확장성(scalability)를 위한 다른 방법으로는 원본 이미지에서 ROI만을 하위 레이어(base layer)로 만들고 그 나머지 부분을 상위 레이어(enhancement layer)로 나누는 방법이 있다. 하지만 이를 위해서는 기존의 비디오 코드를 수정해야 한다는 점과 ROI가 아닌 영역에 대해서는 전혀 정보를 받지 못할 수도 있다는 단점이 있다[8].
하지만 본 논문에서는 채널 환경에 따른 ROI의 영역의 화질적 확장성(scalability)을 이용하여 기존의 코드를 그대로 사용 할 수도 있으며, 주변 정보에 관한 전송도 가능 하다.

구체적인 방법으로는, 앞 장에서 설명한 다중 채널 전송 기법은 네트워크 환경과 종류에 따라 채널별로 차별을 두어 전송 하였다. 이렇게 채널별로 적응적 스트림 모델을 기반으로 ROI(Region Of interest)기법을 적용하였다. 우선 2-1절에서 설명한 ROI 선택 방법에 따라 영역을 설정한다. 다음 단계로 ROI와 그렇지 않은 영역을 2-2절에서 설명한 다중 채널 전송 기법을 이용하여 ROI에 대해 우선순위를 두어 전송한다. ROI에 관해서는 CGS(Coarse Grain Scalability)를 이용하여 화질 별 확장성(scalability)를 두어 더욱 능동적인 스트림 모델을 만들었다. 이때, ROI 영역의 확장성(scalability)와 전체 영역에 관한 확장성(scalability)에 관한 옵션을 동일하게 적용 시킨다. 이 두 가지 영역에 대한 인코딩 방법 역시 H.264/SVC로 동일 하게 한다. 이러한 방법은 사용자 기기와 비디오 콘텐츠에 따라 ROI영역과 전체 영역에 대해 채널의 개수와 환경에 따라 큰 화질 차이를 나타낸다. 알고리듬은 ROI 선택 방법에 의해서 구해진 ROI영역에 대해서 roi_size_left, roi_size_right, roi_size_top, roi_size_bottom 네 개의 ROI 크기에 대한 변수를 갖는다. 또한 네트워크 인터페이스의 개수를 나타내는 n_count, 네트워크 대역폭을 나타내는 n_bandwidth 정보를 기록 한다. 이렇게 구해진 정보들은 화면 내에 동일한 크기의 ROI영역의 위치를 변화 하여 스트림 서버에 저장된다. 또한 네트워크 정보들은 스트림 매니저를 통해 각 기기별로 전송된다. 결국 하나의 비디오 콘텐츠에 그림 4와 같이 두 개의 영역에 대한 인코딩을 한다.

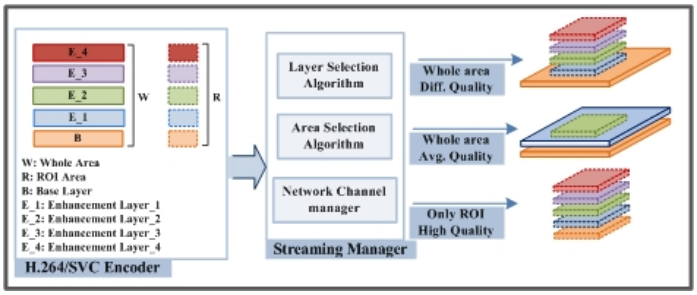

그림 4 ROI Scalability Model

이렇게 구분된 레이어들은 그림 4의 스트림 매니저를 통해 분석된 후 채널에 할당 되어 전송하게 된다. 이때 네트워크 상황에 따라 할당 되는 레이어 별 채널에 변화가 생긴다. 그 한가지 예로 신뢰성 높은 채널(Reliability Channel)의 대역폭이 전체 영역의 하위 레이어(Base layer)와 상위 레이어(Enhancement layer)를 합쳤을 때의 비트 레이트에 못 미칠 경우, 상위 레이어보다 낮은 비트 레이트를 갖는 ROI의 상위 레이어로 대체되게 된다. 나머지 전체 상위 레이어들은 신뢰성 낮은 채널(Error-Prone Channel)로 전송 되는데 이때 에러발생에 대체하기 위해 여러 정류의 FEC(Forward Error Correction)기법이 채널 특성에 따라 적용 될 수 있다. 이렇게 전송 된 두 가지 영역의 상위 레이어는 각각의 채널 버퍼에 저장된 후 시간의 동시성을 맞춘 다음 복호 된다.

그 후 채널 상황에 따라 전송 하기 위해서는 채널 상황 분석기의 정보에 의해 채널이 정의된 후 하위 레이어(Base layer)는 신뢰성 높은 채널(Reliability Channel)로 전송되고 나머지 상위 레이어(Enhancement layer)들은 신뢰성 낮은 채널(Error-Prone Channel)로 전송된다. 하지만 두 가지 영역 레이어의 차이 때문에 경계부분에 화면 왜곡이 나타날 수 있다. 이 문제를 해결 하기 위해 SVC 1/4 샘플 움직임 추정을 수행한다. 이를 위해 1/2샘플은 정수위치 샘플로부터 보간법을 사용하여 생성한다. 신뢰성 낮은 채널(Error-prone Channel)에 대해서는 FEC(Forward Error Correction)를 적용 하여 에러에 대처 하지만, 원본 데이터에 추가적인 데이터가 포함되면 전체 비디오 데이터가 커지는 단점이 있다. 신뢰성 높은 채널(Reliability Channel)은 비디오 데이터의 전체 화면에 대한 하위 레이어(Base Layer) 정보를 포함하여 복호에 대한 기본 설정 정보를 전송한다.

4. 실험

본 실험은 CIF사이즈, 초당 30프레임 비디오 콘텐트를 이용하여 같은 대역폭으로 전송 하였다. 부호화는 H.264/SVC의 Reference Software인 JSVM 8.9를 이용하였다[9]. 실험은 JSVM으로 인코딩된 비디오 정보에 ROI기술을 적용한 영상과 그렇지 않은 영상을 비교 하였다. 또한 ROI부분은 영상의 1/4를 넘지 않게 설정 하였다. 실험 환경은 표 1과 같다.

표1. 실험 부호화 옵션

Attribute	Value
GOPSize	8
FrmaesToBeEncoded	100
Quantization Parameter (1)	18
Quantization Parameter (2)	32
NumberReferenceFrames	1

표 1과 같이 인코딩된 정보들은 두 단계의 화질적 확장성(scalability)과 8단계의 시간적 확장성(scalability)만을 적용하여 그림 5과 같은 레이어로 구성 되었다.

Layer	Resolution	Framerate	Bitrate	MinBitrate	DTQ
0	352x288	3.7500	143.00	143.00	<0,0,0>
1	352x288	7.5000	187.72	187.72	<0,1,0>
2	352x288	15.0000	231.72	231.72	<0,2,0>
3	352x288	30.0000	282.72	282.72	<0,3,0>
4	352x288	3.7500	1163.00		<0,0,1>
5	352x288	7.5000	1568.52		<0,1,1>
6	352x288	15.0000	2007.52		<0,2,1>
7	352x288	30.0000	2547.52		<0,3,1>

그림 5 Layer formation(Soccer content)

그림 5 에서와 같이 같은 시간적 확장성(scalability)을 갖고 있더라도 화면 전체에 관한 화질적 확장성(scalability)때문에 레이어 0~3 과 4~7의 비트 레이트가 많이 차이 나게 된다. 더욱이 표 1의 부호화 옵션의 Quantization Parameter의 (1)과 (2)의 차이가 비트 레이트에 큰 영향을 미치고 있다.

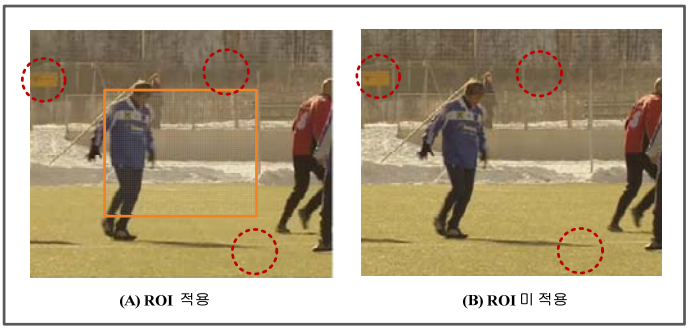

그림 6 ROI scalability quality comparison

그림 6는 Soccer 콘텐트에 ROI를 적용한 것이다. ROI은 (A)가운데 네모난 상자 이며, 점선은 두 화면간의 화질적 차이를 보여준다. 왼쪽에 있는 (A)화면의 PSNR (Peek Signal to Noise Ratio)은 32.56dB이고 (B)화면의 PSNR은 43.56dB이다. 하지만 ROI를 적용하면 사용자들이 주요하고 보는 영역의 화질적 확장성(scalability)를 높게 줌으로써 더 낮은 비트 레이트와 PSNR로도 화질을 개선 할 수 있다.

5. 결론

ITU-T와 ISO는 다양한 네트워크 망과 각기 사용자 기기에 적합한 미디어 스트림을 하기 위한 목적으로 H.264/SVC를 표준화 하였다. 본 논문에서는 H.264/SVC에 ROI 기술을 적용하기 위해 해결해야 몇 가지 문제를 살펴 보았다. 또한 H.264/SVC에 다중 채널 기법을 적용 하는 방법과 그에 따른 장점도 알아보았다. 기존의 SVC기술은 객관적 화질 개선에 관한 확장성(scalability)을 지원하였다. 하지만 ROI를 기존의 공간적, 시간적, 화질적 확장성(scalability)와 같이 사용한다면 객관적 화질 개선뿐만 아니라 주관적인 화질 개선 효과도 기대 할 수 있을 것이다. 또한 기존 부호화 방식의 변화나, 재 부호화 과정이 필요 없이 성능을 향상 시킬 수 있을 것이다. 앞으로 다중 ROI설정과 능동적인 ROI영역 설정에 대한 연구가 필요 할 것으로 기대 된다.

6. 참고 문헌

[1] Huifang Sun, Antiony Veto, Jun Xin ”An overview of scalable video steaming” *WIRELES COMMUNICATIONS AND MOBILE COMPUTING*, 159-172, 2007

[2] Thomas Schierl, Thomas Stockhammer, Thomas Wiegand “Mobile Video Transmission Using Scalable Video Coding” *IEEE TRANSACTIONS*, vol. 17, No. 9, September 2007

[3] Jun-Ren Ding, Jar-Ferr Yang, Ji-Kun Lin “Motion-based Adaptive GOP Algorithms for Efficient H.264/AVC Compression” *JCIS 2006*, October 2006

[4] P. Lambert, W. De Neve, Y. Dhondt, and R. Van de Walle, “Flexible Macroblock ordering in H.264/AVC,” *Journal of Visual Communication & Image Representation*, vol. 17, pp. 358 – 375, January 2006.

[5] Said Benierbah, Mohamed Khamadja “A New Technique for Quality Scalable Coding with H.264”, *IEEE Transactions on On page*, pp. 1332- 1340, Volume: 15, Issue: 11, Nov. 2005

[6] John G. Aposolopoulos, Mitchell D. Trott, “Path diversity for enhanced media streaming”, *IEEE Communications Magazine*, Vol. 42, No. 8,pp. 80-87, Aug. 2004

[7]Amir Naghdinezhad, M. R. Hashemi “A Novel Adaptive Unequal Error Protection Method for Scalable Video over Wireless Networks” *ISCE 2007*,pp. 1-6 , June 2007

[8] P. Lambert, W. De Neve, D. De Schrijver, Y. Dhondt, and R. Van de Walle, “Using placeholder slices and MPEG-21 BSDL for ROI extraction in H.264/AVC FMO-encoded bitstreams,” in *Proceedings of SIGMAP 2006*, pp. 9–16 , August 2006

[9] SVC Reference Software, [Online]. Available: CVS server(host: garcon.ient.rwth-aachen.de)