

SVC 다중 ROI 를 이용한 인터랙티브 비디오 시스템 구현

최정화, 배태면, 노용만
한국정보통신대학교 멀티미디어 그룹
영상 및 비디오 시스템 연구실
mail : ahssen@icu.ac.kr, yro@icu.ac.kr

Implementation of User Interactive Video System using Multiple ROI in SVC

Jeong Hwa Choi, Tae Meon Bae, Yong man Ro
Dept. of Multimedia, Image and Video sYstem Lab.
Information and Communications University

요 약

본 논문은 사용자의 인터랙션(user interaction)을 고려하여, 사용자가 관심을 가지는 영역에 대해서 최적화된 비디오 콘텐츠를 제공받을 수 있는 비디오 서비스 시스템을 제안한다. 본 논문에서 제안하는 사용자 인터랙티브를 고려한 서비스 시스템은 SVC(Scalable Video Coding)의 다중 ROI 를 사용한다. 또한 제안된 시스템에서는 다양한 방식의 비디오 품질과, 다중 ROI 에 대한 선택 과정에서 콘텐츠 소비 주체와 시스템간의 쌍방향 제어 수단을 개발하고 콘텐츠의 관심 영역 지정 과정에서 사용자의 인터랙션 정보를 이용하는 방법을 제안한다. 제안하는 시스템의 유효성을 테스트하기 위해 사용자의 인터랙션에 의해 변환하는 SVC 콘텐츠를 만들었고, 사용자 단말에 인터랙션 기능으로 SVC 콘텐츠를 디코딩하는 시스템을 구현하여 유용성을 검증하였다.

1. 서론

유비쿼터스 컴퓨팅 환경이 될수록 다양한 비디오 품질과 디스플레이 장치가 사용된다. 이처럼 다양화된 콘텐츠 소비 환경에서는 사용자의 편의성 증대를 위한 소비자 중심의 비디오 시스템의 중요성 및 필요성이 점차 증가되고 있다.

네트워크 성능, 디스플레이 장치에 따라 최적화된 서비스를 제공하기 위해서는 다양한 타입의 콘텐츠 소비 환경에 대한 고려와 사용자 개개인의 선호도에 대한 이해가 요구된다. 따라서 각기 다른 타입의 환경에 대해 최적화된 비디오가 필요하다. 이런 필요성에 의해 대두된 것이 스케일러블 비디오 코딩(SVC: Scalable Video Coding)이다. SVC에서는 하나의 콘텐츠에 대해서 한번만 부호화를 하면 공간적, 시간적, 품질적으로 다양한 콘텐츠를 별도의 재부호화 작업을 거치지 않고서도 제공할 수 있다.

SVC에서는 화면 해상도, 컴퓨팅 자원, 네트워크

대역폭 등이 제한된 상황일 경우, 사용자의 소비 환경에 특화된 비디오를 제공한다. 그러나 기존의 SVC에서는 스케일러빌리티 변환 과정에서 사용자의 콘텐츠 소비 환경에 대한 정보만 반영되고, 사용자의 비디오 영상에 대한 품질 변화에 대한 요구, 관심 영역의 이동 등은 반영되지 않았다. 따라서 비디오를 소비하는 과정에서 사용자의 즉각적인 품질 요구의 변화, 관심 영역의 변화 등이 실시간 반영의 필요성이 제기된다. 이를 위해 사용자와 시스템 간의 쌍방향 제어 수단의 개발을 통해 비디오의 스케일러빌리티에 사용자의 의견이 반영이 요구된다.

본 논문에서는 현재 표준화 진행중인 스케일러블 비디오 코딩(SVC)의 다중 ROI 를 사용하여, 콘텐츠 소비 주체와 시스템간의 쌍방향 제어 수단을 시스템에 반영하기 위해 사용자 인터랙션 정보를 사용하여 사용자 단말에 SVC 콘텐츠를 디코딩하는 시스템을 제안한다. 그리고 제안된 방법에 대해서 시스템에 적

용하여 유용성을 증명하였다.

본 논문의 구성은 다음과 같다. 2 절에서는 관련된 연구에 대해 설명하고, 3 절에서는 제안된 스키마에 대해 설명한다. 4 절에서는 이를 이용한 시스템의 구현에 대해 설명하고, 5 절에서는 결론을 기술한다.

2. 관련 연구

2.1 스케일러블 비디오 코딩

ISO/IEC MPEG 와 ITU-U VCEG 에서 공동으로 표준화가 진행중인 스케일러블 비디오 코딩(SVC : Scalable Video Coding)은 공간적, 시간적, 품질적 확장성을 가진 SVC 콘텐츠 생성을 목적으로 한다.[1].

SVC 관련 표준화 내용들은 JSVM(Joint Scalable Video Model)으로 구현된다. 한편, 휴대폰이나 PDA, PMP 같이 낮은 해상도와 제한된 네트워크 등의 한정된 자원을 통해 사용자가 콘텐츠를 소비할 때, 사용자는 비디오에서 관심을 가지고 자세히 보기를 원하는 부분에 대해서 최상의 해상도로 보기를 원한다. 이를 해결하기 위해서 공간적, 시간적, 품질적 확장성을 지닌 SVC 비트스트림을 사용하여 여러 개의 관심영역 (ROI : Region Of Interest)을 인코딩 과정에서 정의한다. 그리고 사용자의 인터랙션을 통한 제어에 따라 특정 ROI를 포함한 SVC 비트스트림을 추출한 후, 사용자 단말에서 디코딩한 결과를 제공하는 것이 해결 방법이다[2]. 현재 표준화가 진행중인 SVC에서는 SVC 비트스트림에 ROI를 정의하여 이를 구현하는 방법이 제안되었다.

이처럼 ROI를 사용하는 SVC에서 다중 ROI가 사용될 때는 어떤 방법으로 ROI를 선택하여 그 ROI의 공간적, 시간적, 품질적인 스케일러빌리티를 결정할지 문제가 된다.

2.2 관심 영역

관심영역(ROI : Region Of Interest)은 FMO(Flexible Macroblock Ordering)를 이용하여 슬라이스 그룹 형태로 정의된다. 이 때 사용되는 슬라이스 그룹은 H.264에 정의된 Slice Group Map Type 2로 정의된다. 16x16 사이즈의 매크로 블록으로 이루어진 직사각형 형태로 구성된 슬라이스 그룹으로 정의하고 처리함으로써 중복 영역(overlapped Region) 등에 대한 독립적인 복호화가 가능해진다. [3, 4]

SVC는 FMO를 이용하여 한 ROI에 대해 하나의 슬라이스 그룹으로 정의한다. 그리고 ROI가 속한 슬라이스 그룹의 ID를 배정하는 방법으로 ROI를 제어한다. JSVM에서 특정 ROI를 부호화될 때는 해당 슬라이스 ID와 동일한 ID를 가지는 매크로 블록만이 블록화된다. 따라서 특정 ROI를 추출하여 스케일러빌리티 변환이 가능하다.

서로 다른 ROI 사이에 중복 영역이 존재할 경우에는 중복된 영역을 별도의 슬라이스 그룹으로 정의한다. 따라서 한 영상 내에서 다수의 ROI를 지정할 수 있으며 인코딩 과정에서 미리 결정된 다수의 ROI 중에 특정 ROI에 대해서 선택적인 스케일러빌리티가 가능하다. 축구와 같은 스포츠 영상의 경우에는 축구

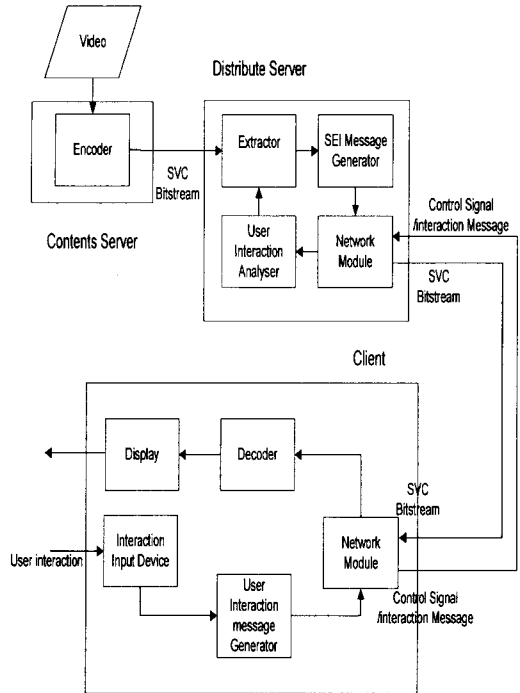
공에 대한 ROI 지정을 통해서 축구공에 대해 최적화된 SVC 콘텐츠를 제공될 수 있다[5]

그러나 surveillance 비디오와 같이 콘텐츠에 따라 ROI가 다수가 존재하거나 사용자의 관심 영역이 빠르게 변화 경우가 있다. 이러한 경우 영상 자체에 최적화된 스케일러빌리티는 사용자의 만족도가 저하된다. 따라서 스케일러빌리티 과정에 사용자의 관심 영역의 변화가 반영될 필요성이 제기된다.

3. 사용자 인터랙션을 사용한 콘텐츠 서비스 시스템

3.1 시스템 구조

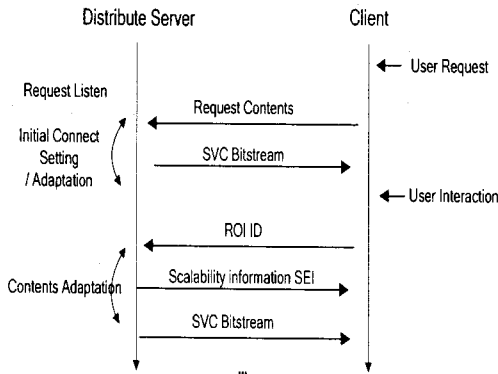
앞 절에서 언급된 내용을 근거로 SVC 콘텐츠에서 ROI를 포함하는 영상 추출 과정에서 사용자의 인터랙션 정보를 반영한다면 좀 더 소비자의 만족도가 높은 개인화된 콘텐츠 서비스 제공이 가능하다. 본 논문에서 제안하는 시스템은 실시간으로 변화하는 사용자의 인터랙션에 의한 사용자 관심 정보가 ROI의 선택과 ROI에 대한 스케일러빌리티 과정에 반영한다.



(그림 1) 제안하는 시스템의 구조

그림 1은 제안하는 시스템의 구조이다. 전체 시스템은 클라이언트, 분배 서버, 콘텐츠 서버의 세 부분으로 이루어진다. 각 부분의 기능은 다음과 같다. 클라이언트 부분에서는 사용자 인터랙션 입력 모듈과 인터랙션 분석 모듈, SVC 비트스트림에 대한 복호화 등을 담당하게 된다. 분배 서버(distribution server)에서는 클라이언트에서 보낸 제어 신호(Control Signal)와 ROI 정보 등에 대한 추출(Extraction) 작업을 담당한다. 콘텐츠 서버(contents Server)에서는 비디오 영상

에 대한 부호화를 담당하여 분배 서버의 추출기 (extractor)로 보내주는 역할을 한다.



(그림 2) 시스템의 프로토콜 구조

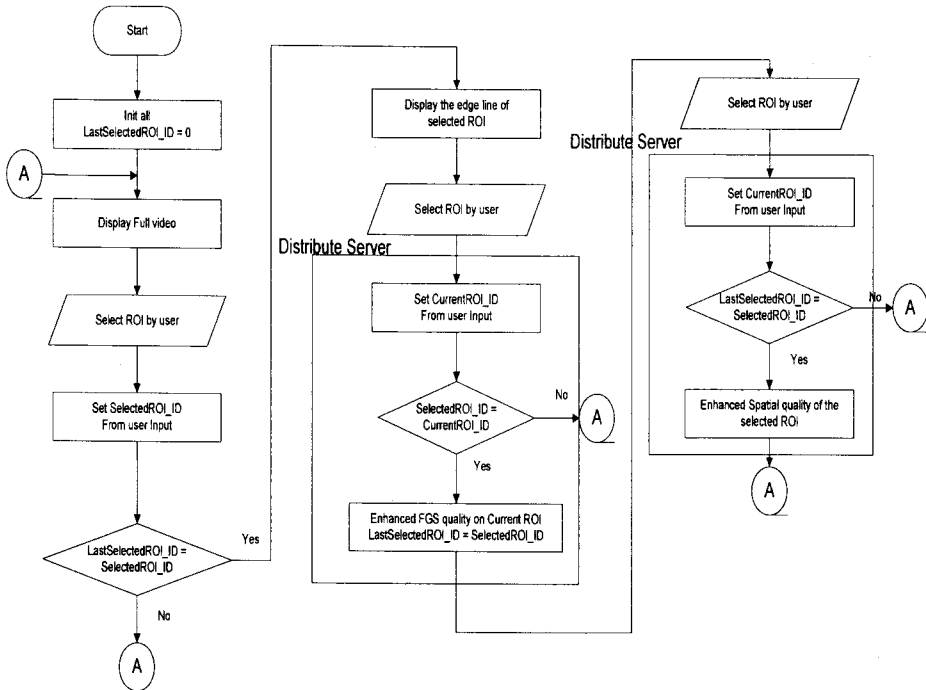
사용자 단말에서 실시간으로 인터랙션 정보가 입력 되면, 클라이언트는 인터랙션에 의해 선택된 ROI 에 대한 정보를 포함한 ROI ID를 분배 서버로 전달한다. 추출기에 의해 변경된 스케일러빌리티에 대한 정보는 Scalability Information SEI 메시지를 통해서 변환

된 SVC 비트스트림과 함께 클라이언트로 보내진다. 변환된 SVC 비트스트림은 클라이언트에서 디코딩되어 제공된다. 따라서 사용자는 선택한 ROI 에 대해 변환된 비디오를 실시간으로 제공받는다. 그림 2 는 분배 서버와 클라이언트 사이의 프로토콜을 나타낸다.

3.2 시스템 기능의 흐름도

그림 3 은 사용자의 인터랙션에 의해서 SVC 의 ROI 스케일러빌리티가 변환되는 과정을 나타내는 흐름도이다. 시스템에서 비디오가 사용자에게 최초로 제공된 때는 스케일러빌리티 변환 이전의 전체영상이 제공된다. 그러나 콘텐츠 소비 도중에 사용자가 단말에 인터랙션을 입력하면 입력된 단말 상의 x, y 좌표 정보를 통해서 사용자의 관심영역에 해당하는 ROI 의 ID가 결정된다.

사용자가 관심을 가진 부분에 유효한 ROI 가 존재하면 후보 ROI 로 선택된다. 또한 후보 ROI 가 속한 슬라이스 그룹의 위치와 넓이를 단말에 표시된다. 이때 인터랙션 기능에 의해 선택된 후보 ROI 가 처음으로 선택된 경우, 선택된 ROI 의 FGS 정보를 추가로 전송하여 선택된 ROI 의 화질을 향상시킨다.



(그림 3) 사용자 인터랙티브티에 의해 ROI 가 결정되는 시스템 동작에 대한 알고리즘

화질이 향상된 ROI 를 또 다시 선택하게 되면, 현재 해상도보다 높은 해상도를 표현할 수 있는 레이어를 추가로 전송하여 ROI 의 해상도를 향상시킨다. 이미 선택된 ROI 가 있으면 사용자가 다른 ROI 를 선택하면 ROI 의 선택은 해지되어 전체영상을 나타내는 SVC 비트스트림을 클라이언트에게 제공한다.

제한된 ROI 를 추출하는 과정에서 디스플레이 장치의 해상도에 맞춰 보간법(interpolation)이 아닌 SVC 시스템에 의한 상위 레이어 정보를 제공을 통해 품질의 확장이 이루어짐으로써 효율적인 시스템이 된다.

4. 시스템 구현 및 고찰

본 논문에서 제안한 SVC 의 다중 ROI 를 이용한 인터랙티브 비디오 시스템은 JSVM 1.0 기반에서 구현하였다. 유용성 검증에 사용된 입력 테스트 영상은 QCIF, CIF, 4CIF 로 구성되었으며 각 레이어는 3FGS 레이어로 구성되었다. GOP 사이즈는 16 이고 최대 프레임율 30 으로 설정하여 부호화하였다. 입력 테스트 영상에서 다중 ROI 설정은 그림 4 에서와 같이 3 개로 설정했다.



그림 5 는 ROI1 을 후보 ROI 로 선택하여서 사용자에게 ROI1 의 위치 정보를 표시하고 있는 결과 화면을 보여주고 있다. 후보 ROI 에 대해 사용자가 인터랙션 기능에 의해 스케일러빌리티 확장을 요구하는 경우 선택된 ROI 에 대해 그림 5 에서 나타난 해상도보다 높은 해상도를 제공해 주게 된다.

그림 6 은 ROI 에 대한 품질 비교 결과를 보여준다. 공간해상도를 향상시킨 그림 6 (a) 와 영상 보간에 의해 공간 해상도를 높인 그림 6 (b) 에 대한 비교 결과이다. 단순히 보간법을 사용하여 QCIF 영상을 4CIF 로 확대한 그림 6 (b) 영상보다 공간 해상도 향상을 통한 그림 6 (a) 영상이 영상의 화질이 블록화와 화면의 뭉개짐이 없이 양호함을 알 수 있다.



(a) (b)
(그림 6) 해상도 확장 결과 비교, (a) 공간 해상도를 향상시킨 영상 (b) 영상 보간에 의해 공간 해상도를 높인 영상

5. 결론

본 논문에서는 스케일러블 비디오 코딩(SVC : Scalable Video Coding)의 다중 ROI 를 정의하여, 사용자의 인터랙션 정보를 통한 스케일러빌리티 과정의 제어 방법을 제안하였다. 그리고 JSVM 1.0 기반에서 제안한 시스템을 구현하여, 실험을 통해 시스템에 대한 유효성을 검증하였다.

다중 ROI 를 정의하여 인코딩 한 후 영상에 대해서 스케일러빌리티 과정에 사용자의 인터랙션 기능을 사용하였다. 사용자 인터랙션 의한 ROI 선택 과정의 제어를 통하여 사용자의 단말에서 SVC 콘텐츠를 디코딩하는 시스템을 구현하였으며 그 결과 비디오에 추가적인 스케일러빌리티를 제공하는 것과 같은 유용할 수 있었다.

감사의 말

본 연구는 한국정보통신대학교 디지털미디어연구소의 정보통신연구개발사업의 연구비 지원에 의하여 수행되었음

참고문헌

[1] ISD/IEC JTC 1/SC 29/WG 11, "Joint Scalable Video Model(JSVM) 1.0 Reference Encoding Algorithm Description," N6899, Hong Kong, Jan. 2005.
 [2] Y. J. Jung, Y. S. Kim, D. Y. Kim, Y. M. Ro, J.-G. Kim, J. Hong, "Analysis of Human Perception for Semantic Concept-based Video Transcoding," in Proc. International Workshop on Advanced Image Technology (IWAIT 2005), Jeju Island, Korea, Jan. 2005.
 [3] Iain E.G Richardson, "H.264 and MPEG-4 Video Compression," John Wiley, 2003
 [4] Tae Meon Bae, Truong Cong Thang, Duck Yeon Kim, Yong Man Ro, Jung Won Kang, and Jae Gon Kim, "Multiple Region-of-Interest Support in Scalable Video Coding," ETRI Journal, vol.28, no.2, pp.239-242, Apr. 2006.
 [5] 김덕연, 배태면, 노용만, 강정원, 김재곤, "스케일러블 비디오 코딩에서 ROI 추출에 관한 연구," 한국멀티미디어학회 추계학술발표대회, pp 49-52, 2005.