

일반논문-06-11-1-05

스케일러블 비디오 코딩에서의 다중 ROI의 구현

배태면^{a)}, 김덕연^{a)}, T. C.Thang^{a)}, 노용만^{a)†}, 강정원^{b)}, 김재곤^{b)}

Multiple ROI Support in the Scalable Video Coding

Tae Meon Bae^{a)}, Duck Yeon Kim^{a)}, Truong Cong Thang^{a)}, Yong Man Ro^{a)†},
Jung Won Kang^{b)}, and Jae-Gon Kim^{b)}

요약

본 논문에서는, 스케일러블 비디오 코딩(SVC)에 다중 ROI(Region Of Interest)를 부호화 및 복호화 할 수 있는 새로운 기능을 지원하는 방법을 제안한다. 현재의 스케일러블 부호화는 공간, 시간, 화질에 대한 확장성을 가지는 비트스트림을 생성하는 것을 목적으로 하고 있다. ROI는 영상 내에서 중요한 의미를 가지는 영역을 나타내는데, 네트워크나 단말의 영상크기의 제약이 발생하는 경우, ROI만을 전송하여 QoS가 가능토록 할 수 있다. 본 논문에서는 H.264의 FMO를 이용하여 SVC에서 다중 ROI를 정의할 수 있는 syntax와 방법을 제시하고, 실제로 부호화 및 복호화하는 방법을 제시한다. 제안한 방법은 실제 JSVM1.0에서 구현하였으며 실험을 통해 본 논문에서 제안한 방법에 대한 검증은 수행하였다.

Abstract

In this paper, we propose a new functionality to Scalable Video Coding (SVC), which is the support of multiple ROIs for heterogeneous display resolution. Scalable video coding is targeted at giving temporal, spatial, and quality scalability for the encoded bit stream. Region of interest (ROI) is an area that is semantically important to a particular user, especially users with heterogeneous display resolutions. The bitstream containing the ROIs could to be extracted without any transcoding operations, which may be one of way to satisfy QoS. To define multiple ROI in SVC, we adapted FMO, a tool defined in H.264, and based on it, we propose a way to encode and decode ROIs. The proposed method is implemented on the JSVM1.0 and the functionality is verified using it.

Keyword: MPEG-4, SVC, JSVM, ROI, scalability

I. 서론

ISO/IEC MPEG 와 ITU-T VCEG의 JVT에서는 MCTF

- a) 한국정보통신대학교 멀티미디어 그룹
Multimedia Group., Information and Communications University
b) 한국전자통신연구원 방송미디어연구그룹
Broadcasting Media Research Group, Electronics and
Telecommunications Research Institute

※ 본 과제는 ETRI의 "비디오 정보의 다양한 스케일러빌리티 지원과 이를 위한 적응적 전송 및 시스템 기술"에 관한 연구 과제에 의해 지원 되었음.

† 교신저자 : 노용만(yro@icu.ac.kr)

(Motion Compensated Temporal Filtering) 와 H.264를 확장한 SVC(Scalable Video Coding)을 표준화하고 있다^[1]. 현재 표준화 되고 있는 SVC는 JSVM(Joint Scalable Video Model)으로 구현되어 있으며, 공간, 시간, 품질상의 스케일러빌리티를 가진 비트스트림을 제공하고 있다^[2]. SVC의 목적은 공간적, 시간적, 품질적 특성의 확장성을 제공하는 비트스트림을 생성하여, 거시적으로는 다양한 이종의 네트워크 환경 및 단말에 대해 QoS(Quality of Service)가 보장되

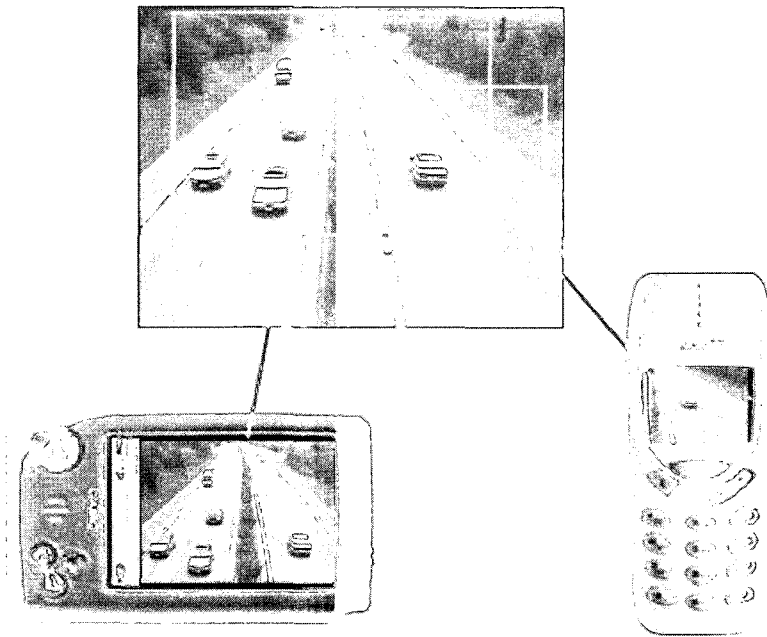


그림 1. 다중 ROI 사용 예
Fig. 1 example of multiple ROI

는 비디오 스트리밍 서비스를 사용자에게 제공하는 것이고, 미시적으로는 가변적인 네트워크 특성에서 적응성이 높은 비디오 서비스를 제공하는 것이다.

휴대폰과 PDA와 같이 해상도와 디스플레이 크기가 제한된 상황에서, 사용자는 비디오에서 자신이 보기를 원하는 부분에 대해서는 자신이 가진 장치가 제공 가능한 최대한의 해상도로 보기를 원한다. 이 같은 상황에서 ROI (Region Of Interest)를 정의하고 사용자에게 제공하는 것은 전체 영상의 해상도를 변경하여 단말의 디스플레이 크기에 맞추는 것보다 나은 환경을 제공할 수 있다. 그림 1은 교통 감시화면을 PDA와 휴대폰으로 확인하는 예이다. PDA와 휴대폰의 경우, 단말의 낮은 해상도로 인해 원 영상의 해상도를 낮추어서 제공받아야 한다. 그러나 ROI가 정의되어 있는 경우, ROI만을 수신하여 출력함으로써 각각의 단말에 맞게 관심영역만을 확인하는 것이 가능하다. 이러한 이유로 SVC표준화 그룹에서는 SVC의 ROI의 지원을 고려하고 있다^[3]. 또한 감시 비디오의 경우, 경우에 따라서는 사용자가 동일 비디오 내에서 여러 관심영역을 가질 수가 있는데 이것은 영상 내에 여러 개의 ROI를 정의하여 제공함으로써 해결이 가능하다.

본 논문에서는 SVC에 다중 ROI를 정의하고, 각 ROI에 대해 SVC가 제공하는 스케일러빌리티를 가지면서 독립적으로 복호화가 가능한 방법을 제안한다. 먼저 H.264의 FMO (Flexible Macroblock Ordering)를 적용하여 다중 ROI를 정의하고 syntax상에서 표현하는 방법을 제안한다. 그리고 이를 기반으로 ROI만을 포함하는 비트스트림의 추출을 위한 signaling 방법과 추출방법을 제안한다.

또한, 제안한 FMO기반의 ROI가 독립적으로 복호화되기 위해 필요로 하는 요구사항과 문제점을 분석하여, 움직임 추정시 움직임 벡터의 검색 영역에 대한 제약, ROI 경계의 깨짐 현상 방지를 위한 ROI 경계 부분에서의 처리 방법을 제안한다. 제안한 방법은 JSVM1.0에 구현되었으며, 실험을 통해 제안한 기능을 검증하였다.

II. 기존 동영상 부호화기에서의 ROI의 구현

다중 ROI를 정의하고, 각 ROI가 독립적으로 부호화 및 복호화 되도록 하게 하는 동영상 부호화에는 H.263과 MPEG-4

object based coding(객체기반 부호화)이 있다^{4,5)}. H.263의 경우, 기본 부호화 방법과 추가적인 부호화할 수 있는 방법을 옵션을 통해 제공하고 있다. 그 가운데 'Independent Segment Decoding'(ISD) 모드는 영상내의 특정 영역을 독립적으로 복호화 할 수 있는 방법을 제공하고 있다. MPEG-4 객체기반 부호화와 H.263의 ISD 모드는 독립적으로 부호화할 영역을 하나의 완전한 영상으로 보고 부호화를 수행한다는 점에서 공통된 기본 개념을 가지고 있다. 그러나 부호화 방법의 차이로 실제 고려되는 요구사항은 차이가 있다.

1. H.263의 Independent Segment Decoding 모드

H.263에서는 영상내의 특정 영역을 독립적으로 복호화할 수 있는 방법을 옵션(Annex R)으로 제공하고 있다. 특히 H.263에서는 확장된 계층적 부호화를 이용하여 공간적, 시간적, 품질적 특성의 확장성을 제공할 수 있는데, SVC의 CGS에 해당하는 방법을 통해 품질적 확장성 및 가변 프레임임을 지원하는 형태의 시간적 확장성을 제공하고 있다. 공간적인 확장성의 경우, SVC과 유사하게 각 해상도별로 부호화하면서 하위 해상도의 영상을 이용한 계층간 부호화(Inter-layer coding)을 수행한다. ISD모드는 이러한 두 모드 중 하나와 같이 사용되어, 특정영역이 복호화될 때 영역의 경계에서 처리해야 하는 문제들을 고려하도록 하고 있다. 먼저, 움직임 벡터가 부호화 영역 밖을 가르키는 것을 금지하도록 하는 제약이 있다. 이것은 시간적으로 다른 위치에 있는 영역간 참조시 영역밖의 정보를 사용하는 것을 금지하는 것이다. H.264에서는 많은 부호화 옵션을 가지고 있으며, 이러한 옵션과 동시에 적용해야 할 때 영역의 경계를 영상의 경계와 동일하게 처리하도록 한다.

2. MPEG-4의 객체기반 부호화

MPEG-4의 객체기반 부호화의 경우, 임의의 형태를 가지는 객체영역을 VOP로 정의하는 사각형영역으로 부호화를 수행한다. 이 같은 방식에서 영상내의 특정영역은 ROI와 같이 독립적으로 복호화가 가능한 형태로 부호화가 수행된다. 독립적인 복호화를 위해 특별히 고려되는 것은

VOP경계영역의 처리로써, 그림 2와 같은 방법을 통해 경계영역을 확장하는 방법을 사용한다.

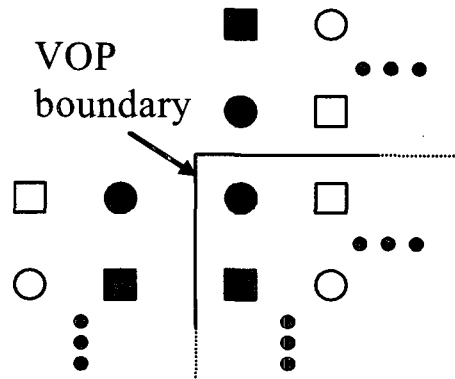


그림 2. MPEG-4 VOP의 경계영역에 대한 패딩방법
Fig. 2 MPEG-4 VOP padding method

III. SVC에서 다중 ROI의 구현

1. 슬라이스와 FMO

현재 ISO/IEC MPEG 와 ITU-T VCEG에서 표준화가 진행중인 SVC는 기본적인 비트스트림의 구조를 H.264에 기반하고 있다. H.264에서는 연속되는 매크로블럭들의 집합을 하나의 슬라이스로 표기하고, 하나 이상의 슬라이스를 슬라이스 그룹으로 하여 부호화와 복호화의 단위로 처리할 수 있다. 이 때, 하나의 프레임은 하나 이상의 슬라이스 그룹으로 이루어진다. 한편 H.264에서는 데이터의 에러 및 손실에 대한 내성을 위해 FMO(Flexible Macroblock Ordering 또는 다중 슬라이스 그룹)가 사용되고 있다. FMO를 이용하여 프레임 내에 있는 매크로블럭을 slice group map type에 정의된 다양한 형식에 따라 유연한 순서로 코딩이 가능하다. 본 논문에서는 여러 slice group map type들 중에 ROI 구현을 위하여 type 2(Foreground and background)를 사용하였다. 이 형식은 프레임 내에서 마지막 슬라이스 그룹을 제외한 모든 슬라이스 그룹을 직사각형 형식의 영역으로 정의하며, 마지막 슬라이스 그룹은 어느 그룹에도 속하지 않은 매크로블럭들(background)을 모두 포함한다.

그림 3은 slice group map type 2를 사용하여 4개의 슬라이스 그룹을 표현한 예이다.

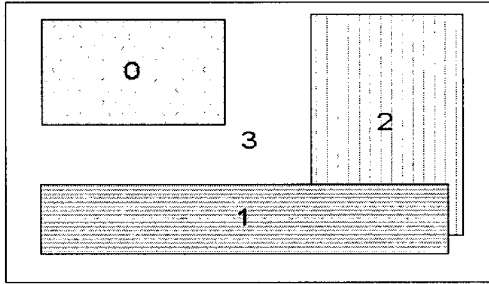


그림 3. FMO의 slice group map type 2에 의해 정의된 슬라이스 그룹
Fig. 3 slice groups defined by FMO with slice group map type 2

2. FMO를 이용한 ROI의 부호화 및 복호화

SVC에서 ROI의 구현은 다음과 같은 방식을 통해 이루어진다. 먼저 앞 절에서 언급한 slice group map type 2 형식에 따라 FMO를 이용하여 ROI를 하나의 슬라이스 그룹으로 정의한다. 해당 프레임에서 ROI 영역 외에 존재하는 매크로블럭들은 또 다른 하나의 슬라이스 그룹으로 정의한다. 만약 ROI가 한 개 존재한다면, ROI가 슬라이스 그룹1이 되고, 나머지 배경이 슬라이스 그룹2가 되는 것이다. 프레임 내의 매크로블럭들은 자신이 속한 슬라이스 그룹의 ID를 배정받는다. 각 매크로블럭은 단 한 개의 슬라이스 그룹 ID만을 받을 수 있다. SVC에서 프레임 부호화는 슬라이스 그룹 단위로 이루어진다. 이때, 특정 슬라이스 그룹이 부호화 될 때에는 해당 슬라이스 그룹의 ID와 동일한 ID를 가지는 프레임 내의 매크로블럭들만이 부호화 된다. 그리고 부호화할 때 picture parameter set에서 전체 슬라이스 그룹의 개수와 각 슬라이스 그룹의 ID를 표시하고, 각 슬라이스 그룹의 좌측상단과 우측하단 매크로블럭 주소를 기록한다. 이들 정보는 슬라이스 단위로 복호화를 수행할 때 사용된다.

3. FMO를 통한 다중 ROI 구현

한 프레임 내에 하나의 ROI가 존재할 때와는 달리, 여러 개의 ROI가 존재하는 다중 ROI를 구현할 때에는 ROI간

중복영역에 대한 고려가 필요하다. 그림 4는 중복영역을 가지는 두 개의 ROI가 정의된 경우를 나타내고 있다. 그림 4에서와 같이 ROI 1과 ROI 2가 정의되어 있는 경우, FMO를 통해 ROI1과 ROI2를 각각 하나의 슬라이스 그룹으로 정의하면, 중복영역은 ROI1 또는 ROI2가 포함되는 슬라이스 그룹들중 하나의 슬라이스 그룹에 포함된다. 즉 ROI 1 또는 ROI 2 중에서 slice group id가 낮은 값을 가지는 슬라이스 그룹에 중복영역이 포함된다. 만약 ROI 1의 slice group id가 0, ROI 2의 slice group id가 1인 경우, 공통된 영역의 매크로블럭들은 ROI 1에 속하게 된다. 결과적으로 부호화 결과를 보면, ROI 1은 원래의 해당 영역의 범위를 유지한 채 부호화가 되지만, ROI 2는 중복 영역을 제외한 영역에 대해서만 ROI 2의 영역으로 할당되어 부호화 된다. 이와 같은 상황에서 만약 사용자가 부호화된 비트스트림에서 ROI 2 영역만을 복호화 한다면, 복호화기에서는 중복 영역이 빠진 ROI 2 부분만을 복호화해서 보여주게 된다. 이 같은 문제를 해결하기 위해 ROI 1과 ROI 2 사이의 중복영역을 독립된 하나의 슬라이스 그룹으로 정의하여 사용하는 방법을 사용할 수 있다. 이를 통해 중복영역(Overlapped Region)이 독립적으로 복호화가 가능하게 됨으로써, 만약 사용자가 ROI 1을 보려고 하면, 중복영역과 ROI 1을 복호화하고, ROI 2를 보려고 한다면, 중복영역과 ROI 2를 복호화 함으로써 올바른 결과를 볼 수 있게 되는 것이다. 단 이때, FMO에서 정의하는 규칙을 따르기 위해, 중복영역의 slice group id는 ROI1, ROI2의 slice group id보다 작아야

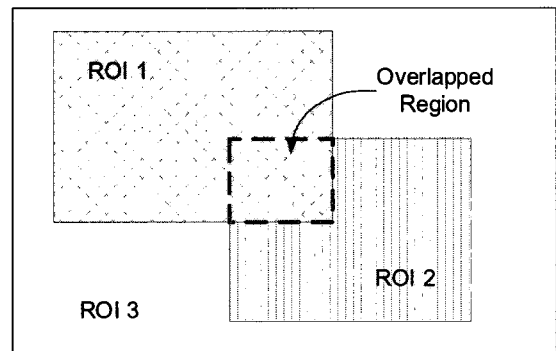


그림 4. FMO를 이용한 중복영역이 존재하는 다중 ROI 표현
Fig. 4 Description of multiple ROI with overlapped region by FMO

한다. 그리고 필요한 메타데이터를 SEI에서 표현해 줌으로써 ROI의 추출을 위해 필요한 정보를 비트스트림 추출기나 복호화기에게 전달할 수 있다.

4. ROI 비트스트림 추출을 위한 메타데이터

중복영역이 존재하는 다중 ROI를 FMO를 통해 제안한 방법으로 영상 내에 정의하는 경우, 하나의 ROI가 하나 이상의 슬라이스 그룹으로 구성될 수 있다. 그러므로 특정 ROI만을 포함하는 비트스트림을 추출하기 위해서는 각 슬라이스 그룹이 어떤 ROI에 포함되는지를 알려주는 정보가 필요하다. 이러한 메타데이터는 SVC 또는 H.264에서 SEI 형태로 존재할 수 있다^[6,7]. 현재 SVC에서는 3 Scalable Info. SEI, Sub-picture Scalable Layer Info. SEI, Motion-Constrained Slice Group Set SEI 를 이용하여 ROI 관련 정보를 표현하고자 한다.

Scalability Info. SEI 에는 모든 스케일러블 레이어를 기술하는데, ROI를 위해 sub_pic_layer_flag를 두고 이 플래그가 TRUE이면 해당 레이어가 ROI를 표현하는 레이어임을 의미한다. ROI와 관련된 각 슬라이스 그룹은 하나의 레이어로 기술되고 layer_id는 개별 슬라이스를 구분할 수 있는 ID가 된다. roi_id는 해당 슬라이스의 ROI ID를 나타내며, ROI가 위치하는 공간해상도에 관계없이 동일한 위치에 존재하는 ROI는 동일한 ID를 가지며, 비트스트림 추출기에서 ROI를 선택하는 입력 변수가 된다. 만약 해당 슬라이스

가 공통 영역을 의미하는 슬라이스일 경우의 roi_id값은 이미 정의되어 사용 중인 ROI ID 값 이외의 어떤 값이든 가능하다. 해당 슬라이스와 관계된 ROI 가 2개 이상일 경우에는 layer_dependency_info_present_flag 가 TRUE가 되고, num_directly_dependent_layers에는 관계된 슬라이스 그룹의 수가 기술된다. 그리고 directly_dependent_layer_id_delta 에는 관계된 슬라이스 그룹의 layer_id와 현재 슬라이스 그룹의 layer_id값과의 차이를 기술한다.

Sub-picture Scalable Layer Info. SEI와 Motion-Constrained Slice Group Set SEI는 Scalability Info. SEI에서 기술된 슬라이스 그룹정보와 실제 슬라이스 그룹 id를 연결하는 역할을 한다. Sub-picture Scalable Layer Info. SEI 는 항상 Motion-Constrained Slice Group Set SEI바로 앞에 위치한다. Sub-picture Scalable Layer Info. SEI는 scalable_layer_id 값을 가지며, 이 값은 Scalability Info. SEI에서 기술된 layer id를 의미한다. 그리고 Motion-Constrained Slice Group Set SEI 에서는 슬라이스 ID 값을 나타내는 slice_group_id값을 가진다. 따라서 기술된 layer id와 slice group id로부터 Scalability Info. SEI에서 기술된 ROI와 관련된 실제 슬라이스 그룹을 연결할 수 있다.

ROI 비트스트림 추출시에는 SEI 메타데이터를 읽고 분석하여, ROI ID와 슬라이스 ID사이에 관계를 찾고, 부호화된 SVC 비트스트림으로부터 특정 ROI만 포함된 비트스트림을 추출하게 된다. 그림 5는 ROI관련 SEI 정보에 대한 내용과 서로의 상관 관계를 보여주고 있다.

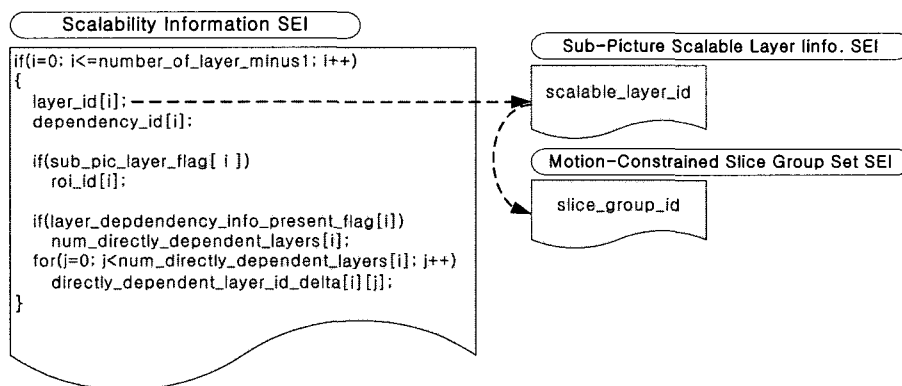


그림 5. SEI를 통한 ROI의 표현
Fig. 5 ROI description by SEIs

IV. 독립적으로 복호화 가능한 ROI를 위한 고려 사항

부호화후 특정 ROI만 복호화 했을 때 영상이 예러없이 올바르게 복호화되기 위해서는 복호화시 슬라이스 그룹 외부 화소값을 이용하는 영상처리과정과 예측 부호화시 슬라이스 그룹 외부에 있는 정보를 사용하는 경우에 대해 ROI 외부 정보에 대한 의존성을 갖지 않도록 하는 처리가 필요하다. 본 논문에서는 ROI가 하나 이상의 슬라이스 그룹으로 구성되어 있으므로, 각각의 슬라이스 그룹이 독립적으로 복호화될 수 있도록 함으로써 ROI의 독립적 복호화하는 기능을 가능토록 하였다.

FMO에서는 슬라이스 그룹의 부호화시 슬라이스 그룹간의 복호화와 관련된 의존관계가 없도록 부호화하도록 하고 있다. 그러나 FMO의 목적이 네트워크 에러에 강인한 복호화를 목적으로 하기 때문에, 의존적인 복호화에 관련해서 고려되는 것은 현재 프레임내에서의 슬라이스 그룹간 의존 관계로써, 정확하게는 매크로 블록의 인트라 부호화 모드(*intra coding mode*)에 한정되어 있다. 즉 슬라이스 그룹의 경계에 존재하는 매크로블록에 대해, 예측부호화를 통한 인트라 부호화시 슬라이스 그룹 밖의 매크로블록 정보를 사용하지 못하도록 함으로써, 현재 프레임에 존재하는 슬라이스 그룹간의 복호화시의 의존관계를 없앤다. 만약 하나의 프레임을 구성하는 슬라이스 그룹 가운데서 손실되는 슬라이스 그룹이 있더라도 복호화기는 남은 슬라이스 그룹에 대해 복호화를 수행할 수 있는 것이다. FMO를 통한 ROI 부호화를 하는 경우, FMO에서 제공하는 하나의 프레임내 슬라이스 그룹간의 독립적인 복호화기능은 충분하지 못하다. 특히 슬라이스 그룹의 경계에 위치하는 매크로블록의 경우, 움직임 보상과 SVC에서 새로이 지원하는 *Inta_BL*모드에서의 독립적인 복호화를 위해서는 추가적인 고려가 필요하다. 또한 독립적인 ROI 복호화를 지원하기 위해서는 시간상의 여러 프레임간에서도 슬라이스 그룹의 독립적인 복호화가 되도록 할 필요가 있다.

1. 움직임 벡터의 검색 영역에 대한 제약

ROI가 독립적으로 부호화와 복호화가 가능하기 위해서

는 움직임 검색 영역을 ROI 내로 국한시켜야 할 필요가 있다. 이것은 여러 프레임간에서도 다른 슬라이스 그룹의 정보를 사용하지 않도록 함으로써 슬라이스 그룹의 독립적인 복호화가 되도록 하기 위한 것이다. 즉, 만약 부호화 시에 ROI 내에 있는 매크로블럭이 움직임 추정시에 ROI 밖에 있는 데이터를 참조하여 부호화가 이루어 지는 경우, 그림 6에서와 같이 ROI영역만을 복호화하면 ROI영역밖의 데이터를 필요로 하게 됨으로써 정확한 복호화가 수행되지 않는다. 따라서 움직임 벡터를 ROI영역 안으로 한정함으로써, 움직임에 의한 영상 예측시에 슬라이스 그룹밖의 영상 정보를 사용하는 경우를 회피할 수 있다. 단 중복영역의 경우 검색영역에 포함될 수 있으며 이를 위해서는 중복영역이 우선적으로 복호화 되어야 한다.

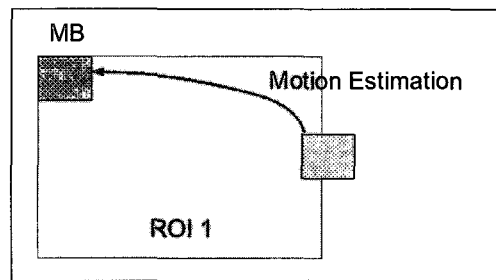


그림 6. 복호화 시에 ROI 영역 밖의 데이터를 필요로 하는 예
Fig. 6 Motion estimation/compensation that refers pixels outside of ROI

2. 슬라이스 그룹의 경계영역 처리

FMO를 통한 ROI 부호화를 하는 경우, FMO에서 제공하는 하나의 프레임내 슬라이스 그룹간의 독립적인 복호화기능은 충분하지 못하다. 특히 슬라이스 그룹의 경계에 위치하는 매크로블록의 경우, 움직임 보상과 SVC에서 새로이 지원하는 *Inta_BL*모드에서 독립적인 복호화를 위해서는 추가적인 고려가 필요하다.

2.1 1/2 샘플 보간 수정

SVC에서는 1/4 샘플 움직임 추정을 수행한다. 이를 위해 1/2 샘플은 기존의 정수 위치 샘플로부터 보간법을 사용하여 생성된다. 1/2 샘플 보간은 MPEG-4 AVC 에서 제공하

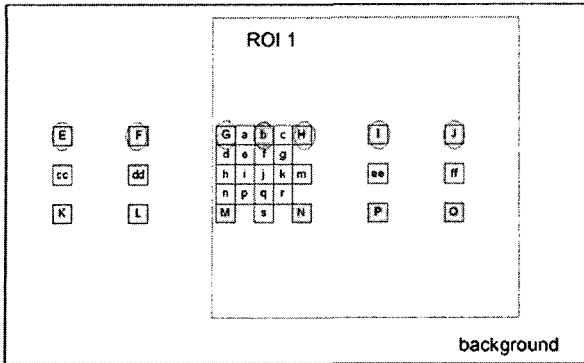


그림 7. ROI 경계 부분에서 1/2 픽셀 보간
Fig. 7 Half-pel interpolation in the ROI boundary

는 6-tap FIR 필터를 사용하여 보간 되는데, 그림 7은 1/2 샘플 보간을 보여준다^[8]. 그림 7에서 회색의 박스에 적힌 알파벳 대문자가 정수 위치 샘플이다. 식 (1)은 b(또는 h,s,m) 위치의 화소값을 계산하는 FIR 필터의 수식이다.

$$b = \text{round}((E - 5F + 20G + 20H - 5I + J)/32) \quad (1)$$

그리고, a,c,d,n 위치의 1/4 샘플 보간은 1/2 샘플 보간이 이루어진 다음, 식(2)에 의해 계산된다.

$$a = \text{round}((G + b)/2) \quad (2)$$

이와 같은 샘플 보간 방법에 의해 만약 1/2 샘플 보간에

서 에러가 발생하면 1/4 샘플 보간에 영향을 준다. 그림 8에서 ROI 경계 부분에서 이루어지는 1/2 샘플 보간은, ROI 내의 'b' 샘플을 보간할 때 ROI 밖에 존재하는 E와 F 같은 정수 위치 샘플들을 사용하게 된다. 이럴 경우에 주변 데이터가 존재하는 전체영상을 부호화/복호화하는 경우에는 문제가 발생하지 않으나, ROI 혼자 복호화가 되어야 하는 상황에서는 부호화 시에 사용했던 ROI 밖에 존재한 정수 위치 샘플들의 값이 존재하지 않기 때문에 복호화 결과에서 경계부분에 깨짐 현상이 나타난다. 그러므로, 부호화 시에 ROI 경계 부분에서의 1/2 샘플 보간을 할 때에는 값을 결정할 수 없는 ROI 외부의 화소값을 특정한 값으로 결정하는 과정이 필요하다. 현재 H.264와 SVC에서는 영상의 경계밖의 화소를 참조하는 경우, 영상의 경계값을 영차 외삽을 한 값을 사용하도록 하고 있다. ROI 또한 하나의 작은 영상으로 고려 할 수 있으므로 동일한 방법을 적용할 수 있다. 1/2 샘플보간에서 참조하는 외부 화소는 최대 ROI 경계밖의 2화소거리에 있는 값이므로 영차 외삽은 2화소 거리의 화소에 대해서만 수행을 하도록 하였다. H.263의 경우 2-tab 양선형(bilinear) 필터가 1/2 샘플 보간을 위해 사용되므로 SVC에서와 같은 문제는 발생하지 않으며, MPEG-4의 객체기반부호화는 움직임 벡터가 영상 밖으로 나가는 것을 허용하기 때문에 문제가 발생하지 않는다. 또 다른 접근방법으로는 움직임 벡터의 검색 영역에 대한 제약을 ROI 영역에서 보다 작게 ROI 안쪽으로 2화소안으로 제약하는 것이다. 그러나 움직임 영역을 ROI 안

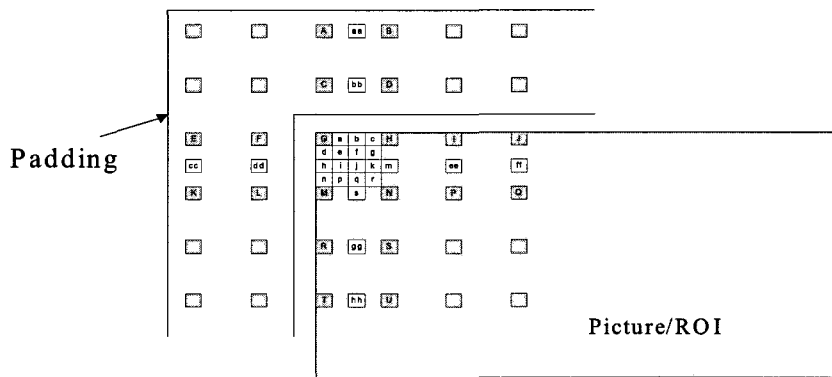


그림 8. ROI 경계 부분에서 1/2 픽셀 보간
Fig. 8 Half-pel interpolation in the ROI boundary

쪽으로 2화소나 제약하는 것은 부호화효율을 떨어뜨리는 원인이 된다.

2.2 IntraBL모드를 위한 업샘플링 수정

IntraBL모드는 Inter-layer intra-texture prediction을 가리키는 것으로 상위 레이어에서 intra-texture prediction을 수행할 때 하위 레이어의 영상정보를 사용하는 것을 말한다. 이를 위해서는 하위 레이어의 texture 정보를 업샘플링해야 하는데, 이때 업샘플링을 위해 참조영상의 1/2 픽셀 보간 필터를 사용한다. 이때, 슬라이스 그룹의 경계 영역에 위치하는 매크로블록을 업샘플링하는 경우, 원래의 슬라이스그룹 영역 밖에 있는 데이터가 사용되어 경계 부분에서 에러가 발생할 수 있다. 이를 방지하기 위해 슬라이스 그룹경계에 위치한 매크로블록이 IntraBL모드로 부호화할 때, upsampling을 위한 1/2 픽셀 보간필터가 매크로블록 외부의 화소를 참조할 때에는 외부 화소값 대신 매크로블록의 가장자리 화소값을 참조하도록 한다. 이것을 구현하는 한 방법으로는 그림 9과 같이 슬라이스 그룹 경계에 존재하는 매크로블록의 가장자리를 패딩하는 것이다. H.263의 경우, (Annex O)에서 공간적인 확장성을 제공하고 있다. H.263은 업샘플링을 위해 2-tab의 bilinear 필터를 사용하고 있는데, 영상의 경계에서는 경계밖의 화소를 참조하지 않는 필터를 따로 제공하고 있고, ROI의 경계를 영상의 경계로 고려함으로써 SVC에서 발생하는 문제를 해결하고 있다.

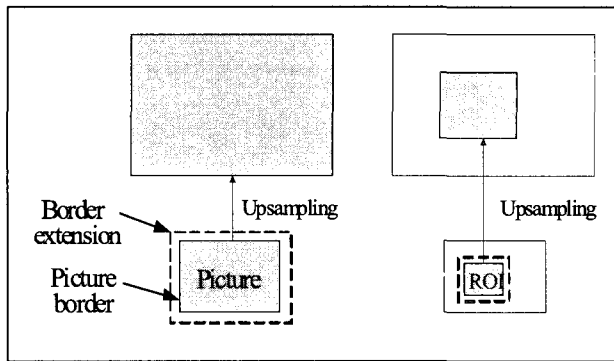


그림 9. 업샘플링을 위한 ROI와 영상의 경계영역 확장
Fig. 9 Border extension of picture and ROI for upsampling

2.3 디블럭킹 필터 동작의 제한

H.263의 ISD모드에서는 디블럭킹 필터모드가 사용되면 영상의 경계에서 디블럭킹 필터의 동작을 제한한다. 동일한 이유로 SVC에서도 ROI의 독립적인 복호화를 위해 동일한 적용이 필요하다. 디블럭킹필터는 블록 사이의 왜곡 현상을 줄이기 위해 사용된다. 이와 같은 필터링은 블록 단위로 이루어지는데, 매크로블럭 내의 4x4 블록의 수직 또는 수평 가장자리에 대해 수행된다. 그러나 디블럭킹 필터의 목적은 두 매크로블럭의 부호화시 손실차에 의해 생기는 블록화현상을 줄이기 위한 것이므로 슬라이스 그룹의 경계면에서는 비교되는 매크로블록이 없으므로 필터링을 수행할 필요가 없다. H.264와 SVC에서는 'disable_deblocking_filter_idc' 값이 2인 경우 슬라이스 그룹의 경계에서 디블럭킹 필터링의 동작을 제한하도록 이미 동일한 기능을 제안하고 있다. 그러므로 ROI 부호화와 복호화시에는 'disable_deblocking_filter_idc' 값을 2로 변경하도록 할 필요가 있다.

V. 실험 및 분석

제안한 다중 ROI의 독립적인 부호화/복호화 기능은 JSVM1.0에서 구현하였다. 먼저 JSVM1.0에 FMO와 독립적인 ROI 복호화를 위한 처리를 구현하였으며, ROI 관련정보를 SEI로 추가적으로 기술하도록 하였다. 제안한 방법에 의해 FMO를 통해 영상에 다중 ROI를 정의하고, 독립적인 복호화를 위한 경계영역 처리유무에 따른 부복호화 결과를 확인하는 실험을 수행하였다. 각1/2 샘플 보간과 업샘플링 필터링시의 경계처리에 대해 고려한 결과와 고려하지 않은 경우에 대한 부복호화를 수행하였다. 실험은 SVC테스트 영상인 "BUS", "ICE" 영상을 사용하였으며, "BUS"영상의 경우 QCIF, 15fps, CIF, 30fps의 2개의 공간 레이어를 가지도록 부호화하였으며, "ICE" 영상은 QCIF, 15fps, CIF, 30fps, 4CIF, 30fps의 3개의 공간 레이어를 가지도록 부호화하였다. 그림 10은 SVC 테스트 영상인 "BUS" 영상에 대해 두 개의 ROI영역을 정의한 것을 나타낸다. 두 개의 ROI는 중복영역을 가지고 있으며, 따라서 중복영역과 배경을 포함한

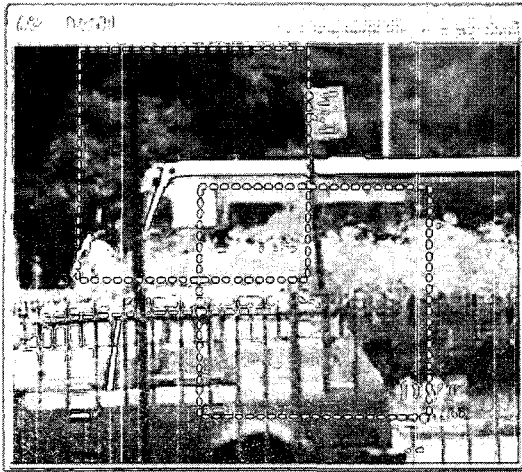


그림 10. 영상내 2개의 ROI의 구성
Fig. 10 Defined two ROIs in the video sequence

4개의 슬라이스 그룹을 정의할 필요가 있다. "ICE" 영상에 대해서는 영상 내에 하나의 ROI를 정의하였다.

그림 11은 "BUS"영상에 대해, 움직임 추정시 1/2 샘플 보간에 대한 경계영역 처리를 하지 않고 부복호화한 결과이다. 그림 11의(b)는 (a)의 영상에서 ROI영역만을 확대한 것이다. 그림 11의 (b)에서, 세로로 그어진 줄 같은 에러가 있음을 확인할 수가 있다. 이것은 ROI 좌측 경계 부분에 위치한 움직임 참조영상의 1/2 샘플 보간시 ROI외부의 값을 이용하여 보간함으로써 에러 값을 가진 픽셀이 발생하는 결과를 보여주는 것이다. 그림 12은 IntraBL 모드를 위한 업샘플링시 슬라이스 그룹의 경계영역처리를 수정하지 않고 ROI만을 부복호화한 결과이다. 그림 12의 (b)에서 ROI의 우측과 하단 경계 부분에서, 경계를 따라 에러가 발생하는 것을 확인할 수 있다. 이것은 IntraBL 모드 복호화시 하위 레이어

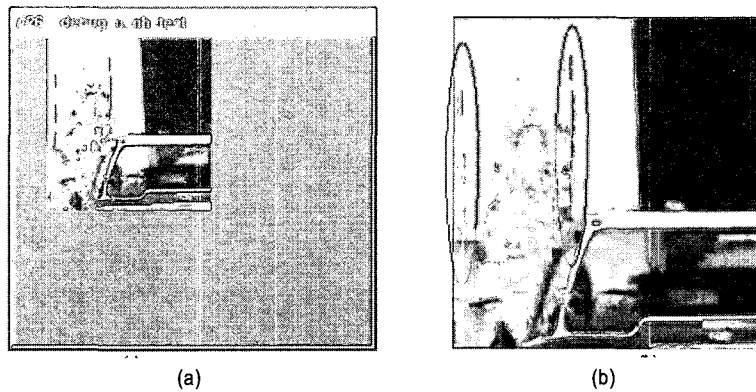


그림 11. 움직임 추정시1/2 샘플 보간을 수정하지 않고 ROI를 부복호화한 결과 (2 Layer 구성) : (a) 전체영상, (b)ROI
Fig. 11 Decoded result when boundary handling for half-pel interpolation is not handled (2layer configuration): (a) whole picture, (b) ROI

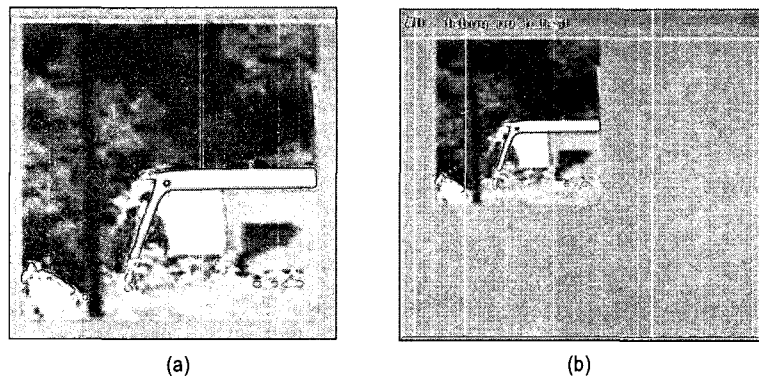


그림 12. IntraBL모드로 부호화시 업샘플링을 수정하지 않고 ROI를 부복호화한 결과 : (a) 전체영상, (b)ROI
Fig. 12 Decoded result when boundaryhandling for upsampling is not handled: (a) whole picture, (b) ROI

의 영상 데이터가 업샘플링 과정에서 ROI 영역 밖의 픽셀을 참조되면서 발생한 에러이다. 그림 13는 움직임 추정시1/2 샘플 보간 및 IntraBL 모드의 업샘플링 과정에서의 경계영역을 처리한 후 부복호한 결과이다. 그림 13의 (a)는 ROI 1만을 부복호한 결과이고 (b)는 ROI 1과 ROI 2를 동시에

부복호한 결과이다. 그림 11, 12와는 달리 부복호한 결과영상의ROI 경계 부분에서 에러가 보이지 않는다.

그림 14, 15, 16는 "ICE"영상을 공간 해상도에 대해 QCIF, CIF, 4CIF의 3 레이어로 부복호한 결과이다. 그림 14는 움직임 추정시1/2 샘플 보간에 대한 경계영역 처리를

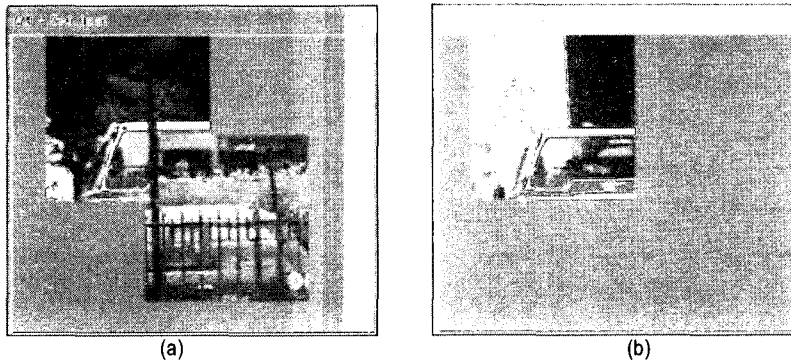


그림 13. ROI 구현과 관련한 에러 요인을 수정하고 한 후 ROI 비트스트림 결과: (a) ROI 1, (b) ROI 1과ROI 2
 Fig. 13 Decoded result when boundaryhandling for both half-pel interpolation and upsampling are handled : (a) ROI 1, (b) ROI 1 and ROI 2



그림 14. 3 Layer로 부호화된 영상의 움직임 추정시1/2 샘플 보간을 수정하지 않고 ROI를 부복호한 결과
 Fig. 14 Decoded result when boundaryhandling for half-pel interpolation is not handled: 2 layer configuration



그림 15. 3 Layer로 부호화된 영상의 IntraBL모드로 부호화시 업샘플링을 수정하지 않고 ROI를 부복호한 결과: (a) 전체 영상, (b) ROI 영역
 Fig. 15 Decoded result when boundaryhandling for upsampling is not handled (3 spatial layer configuration): (a) whole picture, (b) ROI

하지 않고 부복호화한 결과로써 에러가 발생하여 영상의 움직임에 따라 영상 안에서 에러가 전파되는 것을 볼 수 있다. 그림 15는 IntraBL 모드를 위한 업샘플링시 슬라이스 그룹의 경계영역처리를 수정하지 않고 ROI만을 부복호화한 결과로써, 두 개의 공간 레이어로 부호화한 그림 12에 비해 에러가 더 심각함을 알 수 있다. 이것은 하위 레이어의 에러가 상위 레이어로 전파됨으로 해서 에러 정도가 커지기 때문이다. 그림 16는 움직임 추정시 1/2 샘플 보간 및 IntraBL 모드의 업샘플링 과정에서의 경계영역을 전혀 고려하지 않고 부복호화한 결과이다. 그림 14, 15에서 확인할 수 있는 에러가 복합적으로 영상에 나타나는 것을 알 수 있다.



그림 16. 3 Layer로 부호화된 영상의 1/3 샘플보간과 업샘플링을 수정하지 않고 ROI를 부·복호화한 결과

Fig. 16. Decoded result when boundary handling for both half-pel interpolation and upsampling are not handled

VI. 결 론

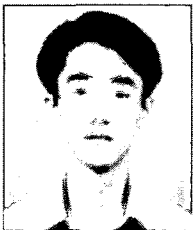
본 논문에서는 SVC에 독립적으로 복호화 가능한 다중 ROI를 구현하기 위한 요구사항에 대해 살펴보고, 그에 대

한 해결 방법을 제시하였다. H.264의 FMO를 이용하여 JSVM에서 다중 ROI를 정의할 수 있는 syntax와 방법을 제시하고, 실제로 부호화 및 복호화하는 방법을 제시하였다. FMO를 통해 ROI를 정의하는 경우, ROI가 겹치는 중복영역을 처리하기 위해 슬라이스 그룹의 집합으로 ROI를 정의하도록 하였다. 또한 부복호화시 슬라이스 그룹의 경계에서 발생하는 문제를 해결하기 위해 움직임 벡터의 검색 영역에 대한 제약 뿐만 아니라, 1/2 보간 필터와 업샘플링 필터의 경계영역에서의 처리방법을 제안하였다. 또한 이를 기반으로 ROI만을 포함하는 비트스트림 추출을 위한 signaling 방법과 추출방법을 제안한다. 제안한 방법은 JSVM1.0에 구현하였으며, 실험을 통해 제안한 방법의 필요성과 유효성을 확인할 수 있었다.

참 고 문 헌

- [1] ITU-T and ISO/IEC JTC1, "Scalable Video Coding - Working Draft1", N9601, Jan. 2005
- [2] ISO/IEC JTC 1/SC 29/WG 11, "Joint Scalable Video Model (JSVM) 1.0 Reference Encoding Algorithm Description," N6899, Jan. 2005, Hong Kong.
- [3] "Scalable Video Coding Applications and Requirements", N6880, Hongkong, Jan. 2005.
- [4] ITU-T, "Video Coding for Low Bitrate Communication, " ITU-T Recommendation H.263, Version 1: Nov. 1995, Version 2: Jan. 1998.
- [5] ISO/IEC JTC1, "Information technology Coding of audio-visual objects Part 2: Visual," ISO/IEC 14496-2 (MPEG-4), 1998.
- [6] ISO/IEC JTC 1/SC 29/WG 11, "Text of ISO/IEC FDIS 14496-10: Advanced Video Coding 3rd Edition," N6540, Jul. 2004, Redmond.
- [7] Ye-Kui Wang and Truong Cong Thang, "Revised signaling of scalability information", JVT-P113, Poznan, July 2005.
- [8] Iain E.G Richardson, "H.264 and MPEG-4 Video Compression," John Wiley, 2003.

저 자 소 개



배 태 면

- 1996년 2월 : 경북대학교 전자공학과 졸업
- 1998년 2월 : 경북대학교 전자공학과 석사
- 2000년 8월 : 경북대학교 박사과정
- 2000년 9월~2001년 3월 : Togabi 기술 연구원
- 2001년 3월~현재 : 서울대학교 컴퓨터기술연구소 연구원
- 2002년 9월~현재 : 한국정보통신대학교 박사과정
- 주관심분야 : semantic video analysis, Scalable Video Coding, Video Adaptation, Watermarking

저 자 소 개



김 덕 연

- 2003년 2월 : 중앙대학교 전기전자공학부 졸업
- 2006년 2월 : 한국정보통신대학교 석사
- 2006년 3월~현재 : 삼성전자
- 주관심분야 : Scalable Video Coding, Video Adaptation, MPEG-21



T. C. Thang

- 1997년~2000년 : project engineer in Vietnam Datacommunications Company
- 2000년 : MS degrees from Hanoi University of Technology, Vietnam
- 2001년~2005년 : a satellite engineer of Vietnam Telecom International.
- 2006년 2월 : 한국정보통신대학교 박사
- 2006년 3월~현재 : 한국정보통신대학교 연구교수
- 주관심분야 : content adaptation, Video abstraction, MPEG-21



노 용 만

- 1985년 2월 : 연세대학교 전자공학과 졸업
- 1987년 2월 : KAIST 전기공학과 석사
- 1992년 2월 : KAIST 전기공학과 박사
- 1992년~1995년 : Dept. of Radiological Science, University of California, Irvine, 초빙 연구원
- 1996년 : Dept. of Electrical Eng. and Computer Science, University of California, Berkeley 연구원
- 1997년~현재 : 한국정보통신 대학교 정교수
- 주관심분야 : 이미지/비디오 처리 및 분석, MPEG-7, 특징인식, 이미지/비디오 인덱싱



강 정 원

- 1993년 : 한국항공대학교 항공전자공학과 (학사)
- 1995년 : 한국항공대학교 항공전자공학과 신호처리전공 (석사)
- 2003년 : Georgia Institute of Technology ECE (공학박사)
- 2003년~현재 : 한국전자통신연구원 방송미디어연구그룹 선임연구원
- 주관심분야 : 비디오 신호처리, 비디오 적응, MPEG-7/MPEG-21



김 재 곤

- 1990년 2월 : 경북대학교 전자공학과(학사)
- 1992년 2월 : KAIST 전기 및 전자공학과(석사)
- 2005년 2월 : KAIST 전기 및 전자공학과(박사)
- 2001년 9월~2002년 11월 : 뉴욕 콜롬비아대학교 연구원
- 1992년~현재 : ETRI 방송미디어연구그룹 방통융합미디어연구팀장/선임연구원
- 주관심분야 : 영상통신, 비디오 신호처리, 디지털 방송, 멀티미디어 프레임워크, TV-Anytime/MPEG-7/MPEG-2