

# 모바일 영상통화 환경에 적합한 하이브리드 변환을 이용한 객체 기반 비디오 코덱 설계

전성혜<sup>†</sup>, 서용수<sup>\*\*</sup>, 권기룡<sup>\*\*\*</sup>

## 요 약

3G 통신 기술의 상용화로 모바일 단말기 상에서 영상통화 서비스가 활발히 제공되고 있다. 하지만, 단말기 상에서 비디오 시스템은 안정성과 성능 면에서 사용자들의 요구를 충족시키지 못하고 있으며, 이 때문에 영상통화 서비스에 대한 사용자들의 불만이 증가하는 것이 현 실정이다. 따라서 본 논문에서는 기존의 제한된 모바일 영상통화 환경에서 향상된 품질의 영상통화 서비스를 제공하기 위한 비디오 압축 방법을 제안한다. 제안한 방법에서는 매 프레임마다 중요 객체와 비중요 객체를 분리하며, 분리된 객체는 다른 비트율과 화질을 갖도록 각 객체에 적합한 변환 커널을 적용하여 부호화 절차를 달리한다. 중요 객체는 고화질을 가지도록 부호화하며, 비중요 객체에 대해서는 저화질을 가지도록 부호화함으로써 제한된 대역폭 내에서 중요 객체에 대한 화질을 보장하도록 한다. 실험 결과에서 동일 비트율에서 제안한 방법이 기존 방법보다 중요객체 영역에 대한 화질이 우수한 것을 확인할 수 있었다.

## Design of Object-based Video CODEC for the Mobile Video Telephony Using Hybrid Transform

Sung-Hye Jeon<sup>†</sup>, Yong-Su Seo<sup>\*\*</sup>, Ki-Ryong Kwon<sup>\*\*\*</sup>

## ABSTRACT

Recently, many people can easily contact video telephony service through the mobile terminal owing to the commercialization of 3G communication technology. However, the quality of the serviced video telephony has been not good yet by the actual mobile restrictions. For solving quality problems, this paper presents the design of the object-based video CODEC using hybrid transform in mobile video telephony. The proposed design firstly segment each frame into a significant object and an insignificant object. The proposed design is to improve the quality of a significant object by limiting the bit rate of a insignificant object. Thus, we compress a significant object with high quality and low compression ratio and compress an insignificant object with low quality and high compression ratio. Furthermore, we control the bit rate of the video stream in the limited bandwidth by adjusting the compression ratio of each object. From experimental results, we confirmed that our method has more higher quality than methods in the conventional CODECs at the significant region on the same bit rate.

**Key words:** Mobile Video Telephony(모바일 영상통화), Video CODEC(비디오 코덱), Object(객체), Hybrid Transform(하이브리드 변환)

\* 교신저자(Corresponding Author): 권기룡, 주소: 부산광역시 남구 대연 3동 599-1 1호관 1315호(608-737), 전화: 051)629-6257, FAX: 051)629-6410, E-mail: krkwon@pknu.ac.kr  
접수일: 2010년 3월 5일, 수정일: 2010년 4월 5일  
완료일: 2010년 4월 15일

<sup>†</sup> 준회원, 부경대학교 컴퓨터공학과  
(E-mail: shjeon@pknu.ac.kr)

<sup>\*\*</sup> 정회원, 동서대학교 정보시스템공학계열 전자공학과  
(E-mail: seoyong@dongseo.ac.kr)

<sup>\*\*\*</sup> 종신회원, 부경대학교 IT융합응용공학과

※ 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단 (KRF-2009-0075855) 및 한국산업기술진흥원의 지역혁신인력양성사업 지원으로 수행된 연구임

## 1. 서 론

최근 비동기식 IMT-2000 계열의 WCDMA(wideband code division multiple access)를 기반으로 한 HSDPA(high speed downlink packet access) 기술의 상용화로 본격적인 3세대 이동통신 서비스가 제공되고 있다[1]. HSDPA는 WCDMA를 확장한 고속 패킷 통신 규격으로서 이론적으로는 최대 14Mbps의 전송 속도와 실제 2~3Mbps의 WCDMA 보다 7배 정도 향상된 전송 속도로 데이터를 전송 받을 수 있다[2]. 이로 인해 동기식 IMT-2000 계열 기술에서는 제공하기 어려웠던 다양한 서비스가 가능해 졌다. 국내 이동통신사의 경우 SK 텔레콤과 KTF가 비동기식 IMT-2000의 사업자로 선정되어, 모바일 영상통화, 글로벌 로밍, 생활 인프라 서비스 그리고 USIM(universal subscriber identity module) 카드 기반 서비스 등 다양한 서비스를 제공하고 있다. 이중 가장 두각을 나타내는 서비스가 영상통화이다. 영상통화는 비디오 데이터의 수신 및 전송이 동시에 이루어지는 양방향 데이터 서비스이며, 비디오 데이터 자체가 방대한 처리량을 갖기 때문에 모바일 환경에서 영상 서비스를 제공하기 위하여 반드시 수반되어야 하는 기술이 바로 비디오 코덱 기술이다. 다양한 종류의 비디오 코덱이 사용목적과 환경을 고려하여 개발되었으며, 현재 모바일 영상통화를 위한 비디오 코덱으로는 H.261, H.263 코덱이 사용 중에 있다.

모바일 상에서 다양한 영상 서비스가 제공되며 빠르게 발전하고 있지만 실제로 비디오 시스템에 대한 안정성 및 성능은 만족스러운 수준에 이르지 못하고 있다. 원활한 영상 서비스를 제공하는데 제약이 되는 것은 크게 두 가지가 있으며, 첫 번째가 스트리밍 대역폭이고, 두 번째가 비디오 CODEC의 성능이다. 스트리밍 대역폭은 WCDMA, HSDPA의 상용화로 인하여 다소 해소되고 있으나, 무선 환경 자체가 유선에 비해 협대역임은 부정할 수 없는 현실이다. 비디오 코덱의 성능은 영상 서비스의 품질을 좌우하는 결정적인 요소로서, 연속된 방대한 비디오 스트림의 처리가 필요하므로 막대한 데이터 프로세싱을 요구한다. 하지만 단말기 배터리 용량의 문제로 인한 비디오 코덱의 저 전력 설계로 복잡도가 높은 고성능의 비디오 코덱을 적용하는 것이 현실적으로 어렵다. 특히, 영상통화 서비스는 부호화와 복호화가 동시에 이루어

여져야하며, 실시간 통화를 위해 매우 짧은 지연시간이 요구되는 더욱 제한적인 사항을 갖는다. 배터리 용량, 단말기 탑재 칩 기술의 저 전력화 등의 문제가 지속적으로 연구되고 있지만, 비약적 발전이 이루어지지 않는 한 현재의 문제점들을 해결하기는 어려울 것으로 보인다. 따라서 현재 제한된 모바일 환경 내에서 향상된 품질의 영상 서비스를 제공하기 위해서는 소프트웨어의 방법적인 개선, 즉 비디오 코덱의 개선이 필요한 실정이다.

현재 상용화된 모바일 영상통화 상에서 가지적으로 나타나는 문제점은 크게 두 가지를 들 수 있다. 첫 번째 문제는 배경조도에 의한 화질 열화 성분의 발생이며, 두 번째 문제는 프레임 내 큰 변화 성분에 의한 영상손실 발생이다. 본 논문에서는 이러한 문제점들을 향상시키기 위해 객체기반의 하이브리드 변환을 적용한 비디오 코덱을 제안한다. 압축될 모든 비디오 프레임은 중요 객체와 비중요 객체로 분리되어, 각 객체에 대해 압축 절차를 달리 수행한다. 모바일 영상통화 환경을 고려하여 중요객체는 얼굴 영역으로 설정하며, 얼굴 영역을 제외한 나머지 영역을 비중요 객체로 설정한다. 이와 같이 객체를 분리하여 코딩함으로써 중요객체에 대해 독립적인 AWB(auto white balance) 기술을 적용 가능하도록 하며, 배경조도에 의한 얼굴 영역의 화질 열화 성분을 향상할 수 있도록 한다. 그리고 각 객체에 비트율을 달리함으로써 비중요 객체에 해당하는 배경 영역에 많은 비트가 할당되는 문제점을 해결하도록 한다. 프레임 내에서 중요 객체에 해당하는 얼굴 영역에 대하여 많은 비트를 할당하도록 저 압축(low compression)을 수행한다. 반면, 비중요 객체에 해당하는 배경 영역은 고압축을 수행하여 기존 모바일 네트워크 환경의 제한된 대역폭 내에서 비트율을 만족하도록 한다. 결과적으로 복호된 비디오 영상의 중요 객체 영역은 고품질 영상이 제공되는 반면, 비중요 객체 영역은 저화질의 영상이 제공된다. 또한, 고압축이 수행 될 비중요 객체 영역에 기존 DCT(discrete cosine transform) 기반 블록 단위 압축방법을 적용할 경우, 블록 에러가 현저하게 나타날 것이다. 따라서 제안하는 방법에서는 DWT(discrete wavelet transform)을 적용한 프레임 단위 압축 절차를 적용한다. DWT 기반 프레임 단위 압축 기법을 적용함으로써 기존 영상통화 환경에서 배경영역의 급격한 변화 성분에 의한 블록 에러

를 해결하도록 한다.

실험에서 기존 영상통화 환경에서 사용되고 있는 비디오 코덱과 제안한 방법을 비교한다. 동일한 비트율 상에서 제안한 방법의 복호된 프레임과 기존 방법으로 복호된 프레임의 화질을 비교하였으며, 전체 프레임에 대한 화질은 기존 방법에 비해 낮게 나타났지만 중요 객체 영역의 화질은 기존 방법보다 우수함을 확인하였다.

본 논문의 구성으로 2장에서 현재 영상통화 상에서 사용되는 비디오 코덱과 그 문제점에 대하여 살펴본다. 3장에서는 제안한 비디오 코덱의 압축 절차에 대하여 단계별로 살펴본 후 부호화와 복호화를 위해 정의한 데이터 구조에 대해 기술하도록 한다. 4장의 실험 결과를 통하여 제안한 방법의 성능을 검토하여 고찰한 뒤, 5장의 결론으로서 제안한 방법에 대해 평가하고자 한다.

## 2. 모바일 영상통화 상에서의 비디오 코덱과 문제점

### 2.1 모바일 영상통화 상에서의 비디오 코덱

모바일 환경은 동일 대역폭에서 유선망에 비해 높은 비용이 소요되며, 모바일 전용 단말기는 저 전력 설계로 높은 성능을 확보하기 어려운 제한적 사항을 가지고 있다. 또한 영상통화와 같은 양방향 서비스는 부호화기와 복호화기의 동시 동작이 요구되며, 지연 시간이 매우 짧아야 한다. 이러한 사항을 고려하여 현재 영상통화 서비스는 3GPP(3rd generation partnership project)에서 표준화 한 3G-324M 규격을

따르고 있다[3,4]. 3G-324M 규격은 3G 모바일 네트워크에서 영상통화를 위한 프로토콜로서 기존 H.324 규격을 바탕으로 무선 환경에 맞도록 변경, 삭제, 보완된 규격이다. 3G-324M 규격 관련 3GPP TS(technical specification)를 그림 1에 나타내었다.

3G-324M 규격에서는 비디오 데이터의 코딩을 위해 그림 1에서와 같이 H.261(optional)[5]과 H.263(baseline)[6]코덱을 사용하도록 하며, 현재 대부분의 영상통화 상에서 사용되는 비디오 코덱은 H.263을 기반으로 한다.

### 2.2 기존 모바일 영상통화 서비스의 문제점

3세대 이동통신 서비스로 대표되는 영상통화 서비스에 대해 향상된 품질을 제공하기 위해 기존 모바일 영상통화 상에서 나타나는 대표적인 문제점을 분석하였다. 그림 2는 3G 휴대폰(IM-S240K)으로 영상통화 후 저장된 동영상을 캡처(capture)한 화면이다.

그림 2의 (a)에는 배경 조도에 의해 화질 열화 성분이 발생하는 것을 볼 수 있다. 영상통화 환경에서는 서로의 얼굴을 보고 대화하기 때문에 얼굴에 대한 화질이 보장되어야 한다. 하지만 그림 2의 (a)에서와 같이 배경 조도에 의해 얼굴의 형태를 알아보기 힘들 만큼 화질 열화가 발생하는 것을 확인할 수 있다. 이와 같은 문제를 개선하기 위하여 AWB 기술을 프레임에 적용할 경우, 화면의 전체적인 밝기가 증가하여 프레임 자체가 왜곡될 수 있다. 그림 2의 (b)에서는 빠른 변화 성분에 의해 블록 에러와 같은 영상 손실이 발생하는 것을 확인할 수 있다. 모바일 영상통화 사용자들은 단말기를 손으로 들고 통화하는 것이 일

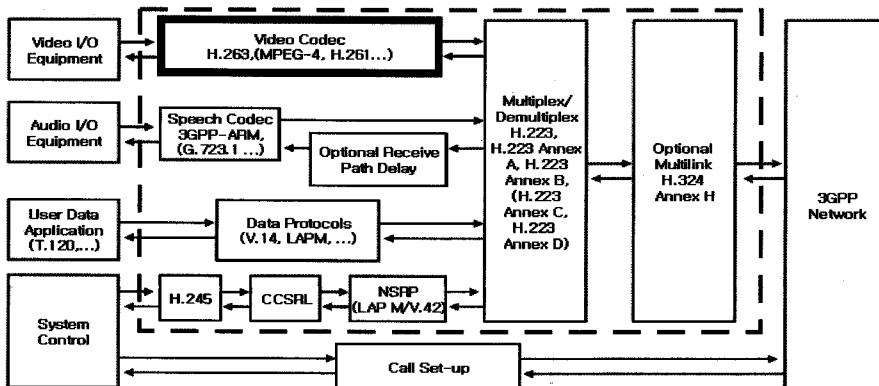


그림 1. 3GPP TS 26.111 규격 상에서 모바일 영상통화를 위한 비디오 코덱 영역

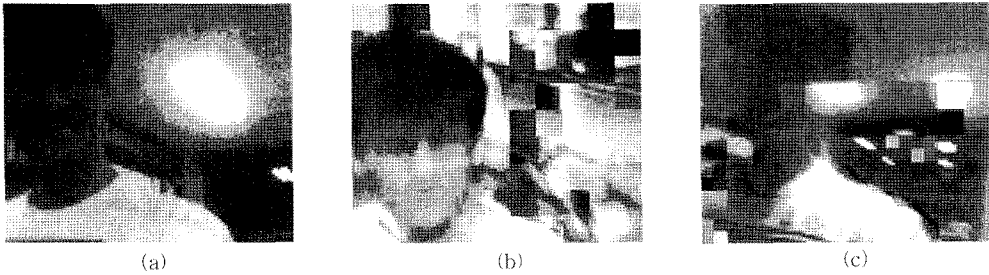


그림 2. 기존 서비스되고 있는 모바일 영상통화의 화질 열화 문제점 (a) 배경 조도에 의한 화질 열화 (b) 빠른 변화 성분에 의한 영상 손실 (c) 손떨림 현상에 의한 영상 손실

반적이다. 이럴 경우, 단말기와 가까운 거리에 있는 얼굴 영역이 프레임 내 전경(foreground)에 위치하게 되며, 상대적으로 거리가 먼 배경(background) 영역은 사소한 움직임에도 그 변화가 크게 나타나게 된다. 기존의 대다수의 DCT 기반 코덱은 프레임간의 상관성을 고려한 시간적 중복정보에 대한 압축을 수행한다. 즉, 이전에 코딩된 프레임과 현재 코딩될 프레임간의 중복되는 데이터 영역을 추정하여, 중복된 영역은 코딩을 수행하지 않고 변화된 영역에 대해서만 코딩을 수행한다. 따라서 기존 비디오 코덱은 제한된 대역폭 내에서 얼굴 영역에 비해 변화량이 큰 배경 영역에 대해 상대적으로 많은 비트를 할당하여 코딩을 수행한다. 또한, 배경 영역에 대해 갑작스런 움직임 등의 이유로 큰 변화량을 가질 때, 중복정보를 추정하기가 어렵게 되며, 그림 2의 (b)에서 볼 수 있듯이 블록 에러와 같은 화질 열화 성분들이 나타나게 된다. 영상통화 환경의 특성 상, 손 떨림 현상 등 프레임 전체에 대해 큰 변화 성분을 가질 수 있으며, 이러한 경우에는 그림 2의 (c)와 같이 프레임 전체에 걸쳐 블록 에러가 발생 할 수 있다. 따라서 본 논문에서는 모바일 영상통화 시 나타나는 이러한 문제점들을 향상시키기 위한 방법을 제안하며, 구체적인 절차와 방법은 3장에서 서술하도록 한다.

### 3. 제안한 모바일 영상통화 비디오 코덱

제안하는 비디오 코덱은 매 프레임마다 중요 객체와 비중요 객체를 분리하며, 중요 객체에 대해서는 저압축, 고화질을 제공하도록 부호화를 수행한다. 이에 반해, 비중요 객체에 대해서는 고압축, 저 화질을 제공하도록 부호화를 수행함으로써 기존의 제한된 대역폭 내에서 비트율을 만족하도록 조절하게 된다.

또한, 부호화 과정에서 각 객체에 적합한 변환 방법을 적용하여 압축 수행 절차를 달리 한다. 중요 객체는 압축효율을 고려하여 매크로 블록 단위의 DCT 기반의 압축 절차를 수행한다. 기존 블록 기반의 DCT를 이용한 비디오 방법은 고압축 수행 시, 심각한 블록 에러가 발생하는 문제점이 나타난다. 따라서 고압축이 수행될 비중요 객체에 대해서는 프레임 기반의 DWT를 이용한 부호화 방법을 적용한다. 그림 3에 제안하는 비디오 코덱의 부호화 과정을 도식한다.

#### 3.1 객체 분리

본 절에서는 프레임으로부터 각 객체를 분리하는 과정에 대하여 살펴보고자 한다. 모바일 영상통화 시, 얼굴 마주 보며 통화를 하는 것이 일반적이다. 따라서 제안하는 비디오 코덱에서는 영상통화 상황을 고려하여 얼굴이 포함된 사람 영역을 중요 객체로 설정하였고, 프레임 내에 중요 객체를 제외한 배경 영역을 비중요 객체로 설정하였다. 본 논문에서는 효율적인 압축 방법 및 절차에 초점을 두고 있으므로 객체의 분리는 실험의 편의 상 MPEG-4 코어 프로파일(core profile)에서 임의 형상 코딩(arbitrary shape coding)을 위해 사용된 이진 알파 마스크(binary alpha mask)를 이용하여 수행하였다[7,8]. 실제로 상용화된 휴대폰이나 디지털카메라의 경우, 얼굴 인식 LSI(large scale integration)를 이용하여 저성능의 CPU에서도 얼굴인식의 실시간성을 보장한다. 또한, 손 떨림 보정 및 초점 보정 등 다양하게 응용되고 있으며, 제안한 기법의 객체 분리 또한 얼굴 인식 LSI를 이용하면 적용가능하다.

객체 분리 시 매 프레임마다 해당 이진마스크를 가지고 있으며, 그림 4에 이진마스크와 분리된 객체에 대하여 나타내었다. 그림 4의 (b)에 첫 번째 프레

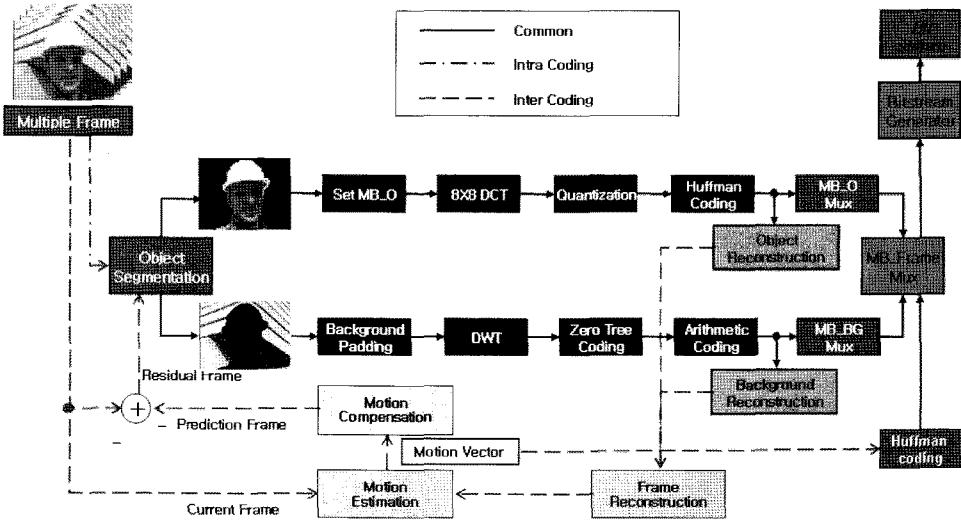


그림 3. 제안하는 비디오 코덱의 부호화 블록도

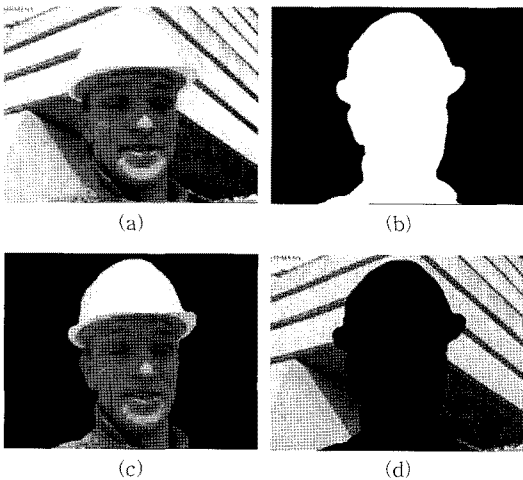


그림 4. 첫 번째 프레임에 대한 객체 분리 예 (a) 첫 번째 입력 프레임 (b) 첫 번째 프레임의 이진 마스크 (c) 분리된 중요 객체 (d) 분리된 비중요 객체

입에 대한 이진마스크와 그림 4의 (c)와 (d)에 분리된 각 객체를 나타내었다.

분리된 각 객체의 경계영역은 큰 고주파 성분을 포함하고 있다. 주파수 영역 상에서 손실 압축이 수행되기 때문에 경계영역에서 심각한 화질열화가 발생할 것이다. 따라서 경계 영역의 고주파 성분에 대한 처리가 필요하다. 중요 객체는 매크로 블록 단위로 DCT를 기반으로 압축이 수행되기 때문에 경계영역에 해당하는 매크로 블록에 대해서 패딩(padding)을 수행한다. 그림 5의 (a)와 같이 중요 객체의 경계

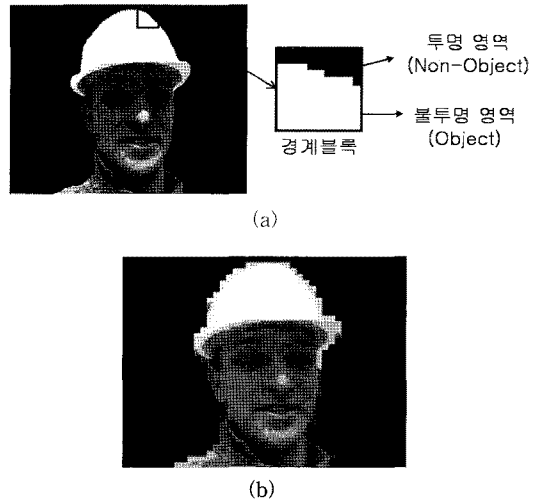


그림 5. 중요 객체에 대한 패딩 예 (a) 중요 객체의 경계 매크로 블록 (b) 패딩된 중요 객체

매크로 블록은 중요 객체에 해당하는 불투명 영역과 중요 객체에 해당하지 않는 투명 영역으로 구분할 수 있으며, 두 영역의 경계에 큰 고주파 성분이 포함되어 있다. 이를 해결하기 위하여 입력 프레임으로부터 투명 영역에 패딩을 수행한다. 그림 5의 (b)에 경계 매크로 블록에 대해 패딩을 수행한 결과를 나타내었다.

비중요 객체는 프레임 단위로 DWT를 수행한다. 분리된 비중요 객체의 투명영역, 즉 중요 객체가 있던 영역은 인접 픽셀로서 그림 6의 (a)와 같이 패딩이

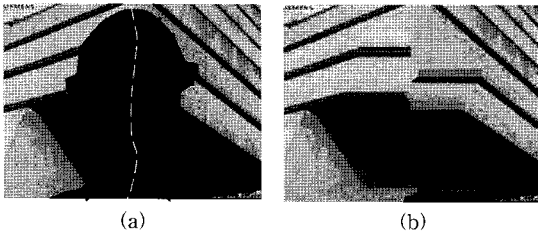


그림 6. 비중요 객체에 대한 패딩 예 (a) 인접 픽셀 패딩 수행 (b) 패딩된 비중요 객체

수행된다. 투명영역의 가운데를 기준으로 인접한 픽셀을 좌측과 우측으로부터 채워 나간다. 그림 6의 (b)에 첫 번째 프레임에 대해 패딩이 수행된 비중요 객체를 나타내었다. 패딩이 수행된 객체는 각각 다른 변환 방법을 적용한 부호화 과정을 거치게 된다.

### 3.2 중요 객체 부호화

시각적으로 중요하다고 간주되는 얼굴 영역의 객체에 대해서는 객체에 대한 정확한 움직임 추정 및 보상과 고화질의 양자화 파라미터를 적용하기 위해 기존의 잘 알려진 8x8 블록 단위의 DCT 변환을 이용하여 부호화한다. 중요 객체의 부호화 절차는 매크로 블록 재 정렬 및 중요 객체의 오프셋 값을 분리하는 매크로 블록 초기화 단계와 DCT 및 양자화 수행 단계, 그리고 허프만 부호화 단계까지 크게 4단계 구성된다. 각 단계별 중요 객체의 코딩 방법에 대하여 살펴해보도록 한다.

#### 3.2.1 매크로 블록 초기화

분리된 중요 객체는 효과적인 DCT 수행을 위해 2차원 버퍼에 재 정렬하게 된다. 재 정렬은 좌측 상단에서 우측 하단 방향으로 매크로 블록 단위로 버퍼에 저장되며, 15개의 매크로 블록이 버퍼의 수평 방향으로 위치하도록 정의하였다. 그림 7의 (a)에 실제 버퍼에 재 정렬된 매크로 블록을 나타내었다. 현 단계에서는 재 정렬 외에도 중요 객체의 복호를 위한 매크로 블록의 오프셋 값을 저장하게 된다. 본 코덱에서는 프레임이 내의 매크로 블록의 수평 라인을 하나의 슬라이스(slice)로 정의하였다. 그림 7의 (b)에 매크로 블록 단위로 분할된 프레임과 저장될 오프셋 포인트를 예를 들어 나타내었다. 사각형 하나는 매크로 블록에 해당하며, 음영이 있는 사각형은 중요 객체를 의미한다. 슬라이스 내의 첫 번째 매크로 블록의 시

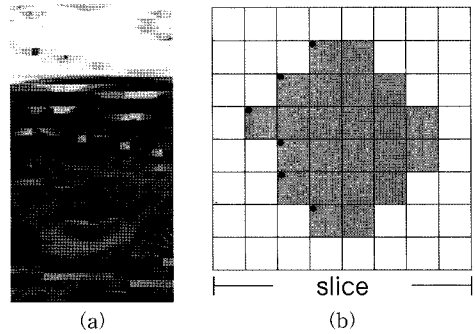


그림 7. 중요 객체의 재 정렬 (a) 재 정렬된 중요 객체에 대한 매크로 블록 (b) 중요 객체의 오프셋 값 저장 예

각 포인트와 매크로 블록의 개수를 저장 후 중요 객체의 복호를 위해 전송하게 된다.

#### 3.2.2 DCT 변환, 양자화, VLC(variable length coding)

중요 객체는 8x8 매크로 블록 단위로 DCT 기반의 압축절차가 된다. 먼저, 재 정렬된 각 매크로 블록은 DCT를 수행하여 공간영역에서 주파수 영역으로 변환된다. 이후 각 계수 값들은 사람의 눈에 영향이 적은 고주파 성분을 효과적으로 제거하기 위해 양자화가 수행된다. 양자화는 식 (1)을 이용하여 수행되며, 양자화 파라미터로서 중요 객체의 화질을 결정하게 된다. 양자화를 수행하면 크지 않은 DCT 계수들을 제거하여, '0' 값을 많이 포함하는 양자화된 계수들의 배열을 생성한다.

$$FQ = \text{Integer Round} \left( \frac{X}{QP} \right) \quad (1)$$

여기서  $FQ$ 는 순방향 양자화(forward quantiser)를 의미한다.  $QP$ 는 양자화 스텝 사이즈(step size)이며,  $X$ 는 DCT 계수 값이다.

양자화된 계수는 대부분 '0'으로 구성되어 있으므로, 보다 효과적인 엔트로피 코딩을 하기 위해 데이터를 정렬한다. 데이터 정렬은 보편적으로 계산 효율이 높은 DC 계수로부터 지그재그로 정렬하는 지그재그(zigzag) 스캔을 사용한다. 지그재그 스캔은 DCT 코딩된 계수들을 통계적인 기법으로 압축하기 위해 저주파 성분부터 고주파 성분으로 일정한 순서로 배열하기 위한 방법이다. 지그재그 스캔을 통하면 '0'의 계수들이 연속적으로 나열 될 수 있으므로 압축 효율을 극대화 할 수 있다. 다음으로 무손실 부호화 방법인 RLC(run length coding)이 수행된다. RLC는 동일한 수가 얼마나 반복되는지를 나타내어 압축하며,

반복된 연속된 수가 많을수록 압축률이 높아지게 된다. 따라서 RLC을 사용하면 연속된 '0'의 계수들을 간결하게 표현 할 수 있다. RLC이 수행된 후 엔트로피 코딩으로서 허프만 코딩이 수행된다. 허프만 코딩은 각 심벌에 대한 확률을 구하여 가장 큰 확률을 가지는 심벌에 대하여 짧은 코드 길이를 부여하는 가변 길이 코딩기법이다. 실제 코드에서는 MPEG 표준 권고안에서 정의한 VLC table을 사용하여 RLC와 허프만 코딩을 구현하였다.

### 3.3 비중요 객체 부호화

기존의 비디오 압축 기법들은 공간적 중복 정보(spatial redundancy), 시간적 중복 정보(temporal redundancy), 통계적 중복 정보(statistical redundancy)를 제거한다. 이 중에서 시간적 중복 정보 제거는 프레임간의 높은 상관관계를 이용하여 중복된 데이터를 제거함으로써 압축률을 높이는 방법을 사용한다. 영상통화 시, 비중요 객체에 해당하는 배경 영역은 중요객체인 얼굴영역보다 뒤에 위치하며, 그 변화가 크게 나타난다. 따라서 중요객체 보다 변화량이 많은 배경영역이 낮은 상관관계를 가지며, 이는 중요하지 않은 영상 정보에 상대적으로 많은 비트수를 할당하는 문제를 가지고 있다. 제안하는 비디오 코덱에서는 상대적으로 비중요 객체로 간주되는 배경영역에 대하여 저화질을 갖도록 고압축을 수행한다. 기존 대부분의 비디오 코딩 기법들은 블록 기반의 DCT 변환을 이용하여 압축을 수행한다. 따라서 기존의 방법으로 비중요 객체 대하여 고압축을 수행할 때 심각한 블록 에러가 발생한다. 이를 해결하기 위해 프레임 단위로 DWT 영역 상에서 제로트리 인코딩을 수행한다[9-11]. 프레임 단위로 인코딩을 수행함으로써 블록 에러와 같은 화질열화를 향상시킬 수 있다.

#### 3.3.1 웨이블릿 변환

분리된 비중요 객체는 DWT을 통해 주파수 영역으로 변환된다. 일반적으로 프레임 기반의 변환 방법은 블록 기반 변환 방법에 비해 요구되는 메모리 크기가 크다. 따라서 빠른 수행속도와 모바일 환경을 고려한 정수 연산 및 메모리 사용의 최소화를 위해 식 (2)의 Le Gall 5/3-tab 필터뱅크를 적용하여 리프팅 구조(lifting scheme)로 구현한다[12].

$$\begin{bmatrix} S(z) \\ D(z) \end{bmatrix} = \begin{bmatrix} -\frac{1}{8}z^{-1} + \frac{3}{4} - \frac{1}{8}z & \frac{1}{4} + \frac{1}{4}z \\ -\frac{1}{2}z^{-1} - \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} X_c(z) \\ X_o(z) \end{bmatrix} \quad (2)$$

1번째 프레임의 분리된 비중요 객체에 대한 웨이블릿 변환된 결과를 그림 8에 나타내었다. 웨이블릿 변환된 결과는 다해상도 분석이 가능하며, 부모 자손 관계의 자기 상관성(self-similarity)을 갖는다. 그림 8에서 볼 수 있듯이 영상의 거의 모든 정보가 저주파수 대역에 몰려 있으며, 그 외의 대역에서는 거의 '0'에 가까운 값을 갖고 있음을 알 수 있다. 따라서 비트 플레인 방법을 사용해서 부호화 할 경우에 저주파수 대역 외의 다른 대역들은 거의 '0'인 관계로 비트 플레인 전체를 부호화 하는 것은 상당히 비효율적이 되며, 따라서 각각의 비트 플레인을 부호화 할 때 모든 비트 플레인을 부호화 하지 않고, ZTR 방법을 사용해서 효율적으로 부호화 하게 된다.

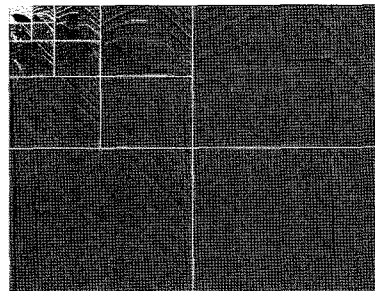


그림 8. 첫 번째 프레임의 비중요 객체에 대한 4-level 웨이블릿 변환 결과

#### 3.3.2 제로 트리 코딩

웨이블릿 변환된 계수 값들은 저주파 서브 밴드의 '뿌리(root)'에서 시작하는 '트리(tree)' 형태로 간주할 수 있다. 즉, 그림 9의 (a)에 나타낸 것과 같이 LL 밴드에 존재하는 단일 계수는 하위 레이어의 네 개의 계수와 대응하는 쿼드트리 구조의 부모-자손 관계를 갖는다. 또한, 하위 레이어에 해당하는 고주파 서브 밴드에 있는 많은 계수들이 대부분 작은 값을 가지고 있기 때문에 영상 화질의 큰 손실 없이 양자화 될 수 있다. 제로 트리 코딩은 이러한 특성을 이용하였으며, 비트율-양자화 왜곡 성능이 좋고 계수 값이 큰 중요 계수부터 먼저 부호화시켜 전송하는 임베딩(embedding) 특성을 가지고 있기 때문에 비트율 제

어가 가능하다. 제안하는 비디오 코덱에서는 구성파일(configure file) 내에 비중요 객체에 대한 비트율을 설정 가능하도록 하였으며, 식 (3)에 의해 비트율은 바이트 크기로 변환되어 정해진 크기만큼만 코딩하도록 한다.

$$ByteBudget = \frac{H_{frm} \times V_{frm} \times T_{bpp}}{8} \quad (3)$$

여기서  $H_{frm}$ 은 입력 프레임의 수평 해상도를 말하며,  $V_{frm}$ 은 입력 프레임의 수직 해상도를 말한다.  $T_{bpp}$ 은 미리 설정한 원하는 비트율을 의미한다.

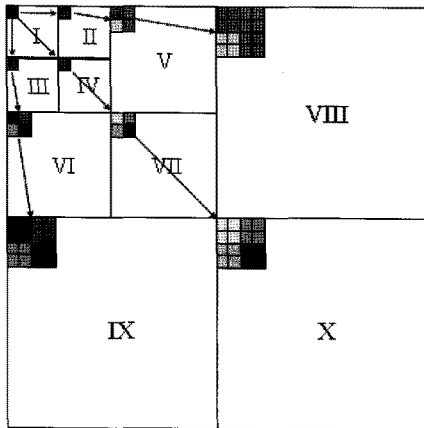
적용한 제로 트리 코딩의 자세한 부호화 과정은 다음과 같다. 부호화 과정은 크게 주부호화과정(dominant pass)과 종속부호화과정(subordinate pass)으로 나누어지게 되는데, 양자화 계수를 받으

로 줄여나가면서 이 두 과정을 반복하게 된다. 주부호화과정에서는 양자화 계수와 웨이블릿 계수를 비교해서 웨이블릿 계수의 중요도를 심별로 정의한다. 이때 중요도 판별 및 심별은 다음과 같이 결정한다. 부모계수의 값이 클 경우 자손계수의 값 또한 클 확률이 높으며, 부모계수의 값이 작을 경우 자손계수의 값 또한 작을 확률이 높다. 따라서 적용된 제로트리 코딩은 임계값을 설정하여 현재 부호화 하려는 계수의 절대값이 임계값 보다 큰지 작은지를 판단하고, 자손계수들 중에 절대값이 임계값 보다 큰 계수가 존재하는지를 판단한다. 다음으로 현재 부호화 하려는 계수가 임계값 보다 클 경우에 항상 자손계수를 부호화하고, 그렇지 않을 경우에는 자손계수의 상황에 따라서 자손계수를 부호화 할 것인지를 결정한다. 중요도 판별 후 계수들은 다음 4개의 심별을 사용해서 부호화 되며, 각 심별의 정의는 다음과 같다.

- POS - Positive Significant, 부호화 하려는 계수의 절대값이 임계값 보다 크며, 계수의 부호가 양수인 경우.
- NEG - Negative Significant, 부호화 하려는 계수의 절대값이 임계값 보다 크며, 계수의 부호가 음수인 경우.
- IZ - Isolated Zero, 부호화 하려는 계수의 절대값이 임계값 보다 작지만, 자손계수들 중에는 임계값 보다 큰 값이 존재할 경우.
- ZTR - Zero Tree Root, 부호화 하려는 계수의 절대값이 임계값 보다 작으며, 모든 자손계수들 역시 임계값 보다 작을 경우.

따라서 현재 부호화 하려는 계수가 POS, NEG와 같은 중요계수로 판단된 경우와, IZ로 판단될 경우에는 항상 자손계수 각각에 대해서 POS, NEG, IZ, ZTR 심별을 사용해서 다시 중요도를 판별하여 부호화하고, ZTR로 판단될 경우에는 더 이상 자손계수를 부호화하지 않는다. 즉 영상의 같은 위치를 복원하는데 관계된 부모, 자손계수 모두를 하나의 블록으로 생각할 때에, ZTR 심별은 현재 부호화하는 계수가 포함된 블록의 자손계수는 모두 0이라는 EOB(end of block)에 해당하는 심별이다.

종속부호화과정에서는 주부호화에서 중요 계수로 판단된 계수의 양자화의 정밀도를 높여 나간다. 이때 주부호화과정에서 부호화 하는 순서는 그림 9의



(a)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255
---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

(b)

그림 9. 3-Level DWT 영역 상에서의 (a) 대역 간 계수들의 상관관계와 (b) Raster scan order



(b)와 같이 가장 분해가 많이 이루어진 저주파 대역부터 고주파 대역으로 올라가는 순서로 부호화가 이루어진다. 종속부호화과정에서 부호화 하는 순서는 가장 먼저 중요계수로 판단된 계수부터 부호화가 이루어진다. 적용된 제로 트리 코딩의 부호화 과정을 요약하면 다음과 같다.

- ① 모든 계수를 조사해서 최대값을 찾아내어 이 값을 기준으로 양자화 계수를 설정한다.
- ② 주부호화과정에서 부호화가 이루어질 대상을 설정한다. 처음 시작일 경우 모든 계수가 부호화 대상이 된다.
- ③ 주부호화과정을 시작한다. 부호화 방법으로는 제로트리 Root 기법을 사용한다.
- ④ 종속부호화과정을 시작한다. 주부호화과정에서 판단된 중요계수들의 정밀도를 높여나간다.
- ⑤ 임계값을 반으로 낮춘다.
- ⑥ 주부호화대상에서 중요계수로 판단된 모든 계수를 제외시킨다.
- ⑦ ③번의 주부호화과정부터 반복 시행한다.

주부호화과정과 종속부호화과정을 통해서 중요 계수를 계속 발생시켜 나가며 양자화의 정밀도를 높여나감에 따라서 부호화나 복호화 시 임의의 위치 어느 곳에서나 멈출 수 있는 특성을 가지고 있어서 비트율을 조절할 수 있다.

### 3.3.3 산술 부호화

엔트로피 코더로서 산술 부호화(arithmetic coding)가 사용되는데, 이것은 주부호화과정에서 심벌이 하나 발생할 때마다 바로 부호화 해나갈 수 있으며, 마찬가지로 종속부호화과정에서도 부호가 하나 발생할 때마다 바로 부호화 할 수 있기 때문에 연속적인 근사 양자화의 특성을 그대로 살려줄 수 있는 장점이 있으면서도 허프만 부호화와 동등한 성능을 발휘하게 된다. 특히 산술 부호화는 각각의 심벌에 비트를 할당하는 방법이 아니라, 주어진 구간의 범위를 각각의 심벌에 해당하는 범위만큼 좁혀나가면서 부호화하는 방법으로 적용된 제로 트리 코딩 방법과 같이 부호화 하려는 심벌의 수가 적을 때에 아주 효과적으로 부호화 할 수 있다.

주부호화과정에서는 심벌 4개(POS, NEG, IZ, ZTR)로 구성된 히스토그램(histogram)을 사용해서

부호화하며, 종속부호화과정에서는 심벌 2개(1, 0)로 구성된 히스토그램을 사용해서 부호화 한다. 실제로 주부호화과정 시 부모 계수가 중요계수인지와, 바로 전에 부호화된 계수가 중요계수인지의 서로 다른 4가지 상황에 따라 다른 히스토그램을 사용하게 된다. 따라서 적용된 산술 부호화에서는 심벌 4개로 구성된 4개의 히스토그램과, 심벌 2개로 구성된 한 개의 히스토그램이 사용된다.

### 3.4 프레임 간 예측(inter-coding) 수행

제안하는 비디오 코덱은 압축률 향상을 위해 시간적 중복 정보(temporal redundancy)에 대한 압축을 수행한다. 연속된 프레임은 매우 높은 유사성을 가지므로 이들의 공통점을 찾아내어, 유사한 블록에 대하여 압축을 수행하지 않고 방향 정보 즉, 움직임 벡터(motion vector)만을 전송하는 방법이다. 시간적 중복 정보에 대한 압축을 수행하기 위하여 움직임 추정(motion estimation)과 움직임 보상(motion compensation)이 행해지며, 기존의 DCT 기반 압축방법에서 사용된 움직임 추정 및 보상 방법을 제안하는 비디오 코덱에 적용한다. 움직임 추정은 현재 프레임과 이전의 프레임간의 가장 유사도가 높은 블록을 구함으로서 수행된다. 저, 이전 프레임을 복원하기 위하여 각 객체가 복원되며, 복원된 각 객체로부터 프레임임을 복원 시킨다. 복원된 이전 프레임과 현재 입력 프레임간의 움직임의 변화를 비교하기 위해 16x16 블록별로 매칭을 수행한다. 제안한 비디오 코덱에서는 모바일 환경을 고려하였기 때문에 연산량이 적은 tree-step 탐색방식[13]과 SAE(sum of squared error)를 사용한 블록 매칭을 수행한다. 픽셀의 변화가 많을 경우, 비슷한 블록을 찾기 어려우므로 이미지를 2배 확대한 후 블록 매칭을 수행하는 1/2pel 정밀도 움직임 벡터 추정을 적용한다. 움직임 추정의 결과로서 생성되는 움직임 벡터는 가변길이 부호화로서 허프만 코딩 방법을 적용하여 프레임 복호를 위해 디코더로 전송하게 된다. 움직임 보상은 이전 프레임과 움직임 벡터로서 예측 프레임(predicted frame)을 생성하는 과정이다. 예측 프레임은 이전 프레임으로부터 움직임 벡터만큼 매크로 블록을 이동시켜 생성되며, 예측 프레임과 현재 프레임간의 차로서 오차 프레임을 생성한다. 생성된 오차 프레임은 다시 중요 객체 영역과 비중요 객체 영역으로 분리되며, 각 객

표 1. H.264 코덱의 NAL\_unit\_type 정의

NAL unit type	Contents of NAL unit and RBSP syntax structure	NAL unit type	Contents of NAL unit and RBSP syntax structure
0	Unspecified	8	PPS(picture parameter set)
1	Coded slice of a non-IDR picture	9	Access unit delimiter
2	Coded slice data partition A	10	End of sequence
3	Coded slice data partition B	11	End of stream
4	Coded slice data partition C	12	Filler data
5	Coded slice of a IDR picture	13-23	Reserved
6	SEI(supplemental enhancement information)	24-31	Unspecified
7	SPS(sequence parameter set)		

제별 압축이 수행된다.

### 3.6 데이터 구조 설계

부호화된 비트스트림의 복호를 위하여 데이터 구조를 정의하였다. 데이터 구조는 H.264 비디오 코덱에서 사용되고 있는 데이터 구조를 기반으로 설계하였다. H.264는 시작 단계부터 네트워크 전송을 목적으로 개발되었으며, 비디오 콘텐츠의 효율적 전송을 위해 NAL(network abstraction layer)로 구성된다. 제안하는 비디오 코덱 역시 네트워크 전송을 목적으로 하고 있으며, 객체에 독립적인 데이터 복호 및 향상 기법 적용을 위하여 NAL 구조를 기반으로 데이터 구조를 재정의 하였다. NAL은 그림 10의 (a)와 같이 NAL unit과 페이로드(payload)로 이루어지며, 페이로드에는 RBSP(raw byte sequence payload)가 위치하게 된다.

RBSP에는 SPS(sequence parameter set), PPS(picture parameter set)등의 정보를 알려주는 파라미터 셋과 VCL(video coding layer)에 해당하는 RBSP가 포함된다. VCL은 실제 H.264 압축된 데이터를 나타내며 VCL 데이터는 NAL unit를 통해 외부와 인터페이스가 이루어지게 된다. NAL unit은 4바이트 start pattern으로 시작하며 다음으로 1바이트

크기를 갖는 NAL header가 위치한다. NAL header는 그림 10의 (b)와 같이 Forbidden\_zero\_bit, NAL\_ref\_idc, NAL\_unit\_type으로 구성된다. Forbidden\_zero\_bit는 항상 '0'으로 설정되어야 하며, NAL\_ref\_idc에 정의된 값은 NAL header에 이어 나오는 페이로드가 프레임 간 예측 코딩 시 참조 픽처를 구성하는데 사용되는지의 여부를 확인한다. 다음에 위치하는 NAL\_unit\_type이 NAL header 뒤에 위치하는 페이로드가 실제로 어떤 RBSP를 포함하는지를 나타내며, 실제 H.264의 NAL\_unit\_type을 표 1에 나타내었다.

제안하는 비디오 코덱에 적합한 데이터 구조를 설계하기 위하여 복호화에 필수적인 정보들과 객체, 예측 모드, scalable 모드 등을 고려하여 NAL\_ref\_idc, NAL\_unit\_type을 표 2 및 표 3과 같이 재정의 하였다. Nal\_ref\_idc가 '0'으로 설정되면, 예측코딩이 적용되지 않고 모든 프레임에 대해 intra 코딩된 비디오 스트림이 페이로드에 위치함을 의미한다. '1'과 '2'는 예측 코딩이 적용된 intra/inter 코딩된 페이로드를 구분한다. '3'은 페이로드의 내용이 참조 픽처를 구성하는데 필요한 내용임을 의미하며, 복호화를 위한 각종 파라미터 셋과 중요 오브젝트의 오프셋 데이터,

표 2. 제안하는 방법에 적합한 데이터 구조를 위한 NAL\_ref\_idc 정의

NAL ref_idc	Contents of NAL_ref_idc
0	Only intra coding
1	Coded IDR object bitstream
2	Coded non IDR object bitstream
3	A content of the NAL unit is used to consist of reference picture.

<b>Starter code</b> 4Byte(0x00000001)	<b>NAL header</b> 1 Byte	<b>Payload</b>
--	-----------------------------	----------------

(a)

<b>Forbidden_zero</b> 1bit(0)	<b>Nal_ref_idc</b> 2bit	<b>Nal_unit_type</b> 5bit
----------------------------------	----------------------------	------------------------------

(b)

그림 10. H.264의 (a) NAL 구조와 (b) NAL header 구조

표 3. 제안하는 방법에 적합한 데이터 구조를 위한 NAL\_unit\_type 정의

NAL_unit_type	Contents of NAL unit and RBSP syntax structure
0	Unspecified
1	IDR bitstream for significant object
2	IDR bitstream for insignificant object
3	Non IDR bitstream for significant object
4	IDR bitstream for insignificant object
5	Residual data for significant object in the SNR scalability mode
6	Decoding information for a sequence
7	Offset data of significant object
8	Motion vector information
9	End of sequence
10	End of stream
11-12	Reserved

움직임 벡터가 여기에 해당된다. Nal\_unit\_type에는 중요/비중요 객체의 intra/inter 코딩 여부, 복호화를 위한 파라미터 셋, 중요 객체의 오프셋 데이터, 그리고 움직임 벡터 정보 등을 구분하였다. 추후 추가되는 기능을 위해 '11-31'를 'Reserved'로 설정하였다.

#### 4. 실험 결과 및 분석

본 장에서는 기존의 비디오 코덱과 비교하여 제안한 비디오 코덱의 성능 평가 및 그 결과를 고찰 해 본다. 기존 모바일 영상통화 환경에서 사용되고 있는 H.263과 성능을 비교하였으며, ITU-T TMN(test model near-term) 8 버전을 기반으로 하는 H.263+ 코더를 사용하였다. 실험은 입력 비디오의 휘도

(luminance) 성분에 한하여 수행하였으며, 제안한 방법으로 압축된 비디오와 H.263으로 압축된 비디오의 비트율을 같도록 하여 그 화질을 비교하도록 하였다. 화질 평가를 위해 원본 비디오와 복호화된 비디오 간의 PSNR(peak signal to noise ratio)을 사용하였다. PSNR은 식 (4)에 의해 구해진다.

$$PSNR[dB] = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad (4)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |O(i,j) - R(i,j)|^2 \quad (5)$$

실험 영상으로는 CIF(352×288) 해상도와 30fps의 프레임율을 갖는 Foreman 비디오를 사용하였다. Foreman 비디오 영상의 경우, 사람과 배경의 조합으로 프레임이 구성되어 있어 영상통화 상황으로 간주할 수 있다. 부호화 환경으로 GOP 구조를 '15'로 설정하였으며, 100 프레임에 대하여 비트를을 달리하여 객체별 화질을 비교하도록 하였다. 중요 객체에 대한 화질 변화를 비교를 위해 비중요 객체의 비트율을 intra, inter 프레임에 대하여 각각 0.2bpp, 0.005bpp로 고정하여 실험하였다. 중요 객체의 비트율은 양자화 파라미터 값으로 조절된다. 표 4에 동일 비트율상에서 제안한 방법과 H.263으로 압축된 비디오에 대한 평균 PSNR을 비교하였으며, 그림 11의 그래프에서는 비트율별 각 프레임에 대한 PSNR을 비교하였다.

표 4는 100 프레임에 대하여 압축을 수행 후 복호된 프레임에 대한 평균 PSNR 값을 나타낸 것이다. 표 4의 결과에서 동일한 비트율상에서 복호된 프레임에 대한 평균 PSNR은 H.263 방법이 높게 나타나지만, 중요 객체에 대한 평균 PSNR은 H.263 방법으로 복호된 중요 객체의 PSNR보다 높게 나타남을 확

표 4. 제안한 방법과 H.263방법에 대한 평균 PSNR 비교

CODEC		bpp(CR)	0.204 (1/39)	0.166 (1/48)	0.144 (1/55)	0.129 (1/61)
		제안한 OBVC PSNR (dB)	SO bpp(QP)	0.40(12)	0.30(16)	0.25(20)
Frame	31.13		31.20	30.99	30.83	
SO	37.70		36.16	35.02	34.11	
H.263 PSNR (dB)	Frame	34.33	33.76	32.82	32.60	
	SO	34.78	34.02	33.37	33.01	

※ OBVC - object based video coding, CR - compression ratio, QP - quantization parameter, bpp - bits per pixel, SO - significant object

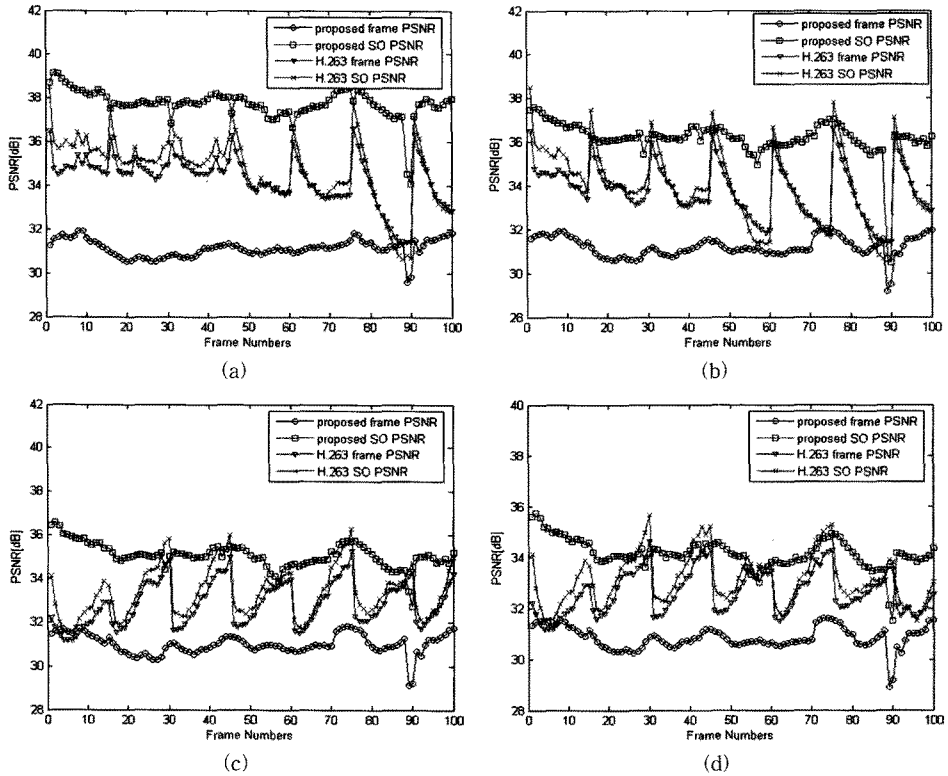


그림 11. 0.166 bpp에서 복호된 비디오의 프레임별 PSNR 비교

인 할 수 있었다. 그림11에 나타난 그래프를 보면 대체적으로 제안한 방법의 중요 객체에 대한 PSNR이 높게 나타나는 것을 확인할 수 있었다. H.263 방법의 intra 코딩된 프레임에 대한 중요 객체 PSNR이 제안한 방법보다 높게 나타날 때도 있지만 intra 코딩된 프레임과 inter 코딩된 프레임간의 화질 차이로 인하여 비디오 재생 시 화질 열화 성분이 가시적으로 드러났다. 90번째 프레임에서 급격한 움직임에 의해 화질이 저하되는 것을 제외하면, 제안된 방법의 중요 객체에 대한 화질변화가 크지 않음을 확인할 수 있었다. 제안한 방법과 H.263 방법으로 복호된 프레임에 대한 화질을 시각적으로 확인하기 위하여 그림 12부터 그림 15에 각 비트율별 복호된 프레임과 중요 객체의 임의 영역을 확대하여 나타내었다. 그림 12~15에 보이는 것처럼 동일한 비트율상에서 제안한 방법의 비중요 객체에 대한 화질은 고압축이 수행되어 H.263 방법의 비중요 객체 영역보다 화질이 좋지 않다. 하지만 높은 비트율을 부여하여 압축을 수행한 중요 객체 영역에서는 H.263보다 향상된

화질을 가진다. 그림 12, 14, 15에서 확대된 중요 객체 영역을 비교하면 H.263 방법의 중요 객체 영역에서 뚜렷한 블록화 현상을 볼 수 있었다. 또한, 그림 14의 (b)에서는 고압축을 수행하여 고주파 성분의 제거로 인해 H.263 방법의 확대된 중요 객체 영역에서 사람의 눈동자 부분이 제거된 것을 확인할 수 있었다. 하지만 그림 14의 (a)에는 눈동자 부분이 손실되지 않고 복호된 것을 확인할 수 있었다. 이와 같이 복호된 그림 12~15의 복호된 프레임을 비교함으로써 제안된 방법이 동일 비트율상에서 중요 객체영역에서 더 우수한 화질을 가지는 것을 시각적으로 확인할 수 있었다.

본 절에서는 제안한 비디오 코딩 방법과 기존의 H.263 방법에 의한 성능을 비교하기 위하여 객관적인 화질 평가로서 PSNR을 비교하였으며, 시각적 화질 평가로서 복호된 프레임을 비교하였다. 동일 비트율상에서 제안한 방법의 중요 객체 영역에 대한 화질이 기존 H.263 방법의 중요 객체에 대한 화질보다 우수함을 확인할 수 있었다.

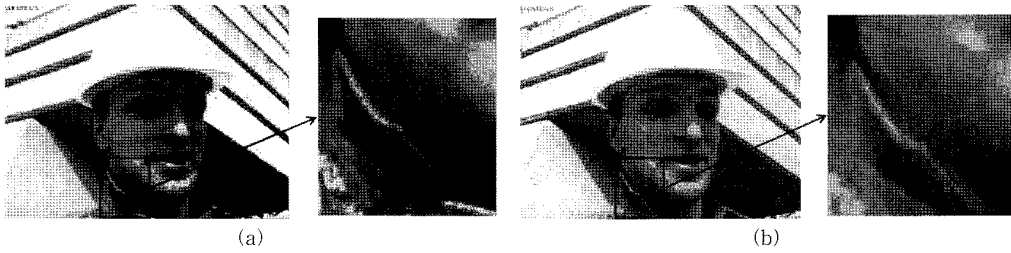


그림 12. 0.204 bpp에서 복호된 비디오의 3번째 프레임 (a) 제안한 방법 (b) H.263

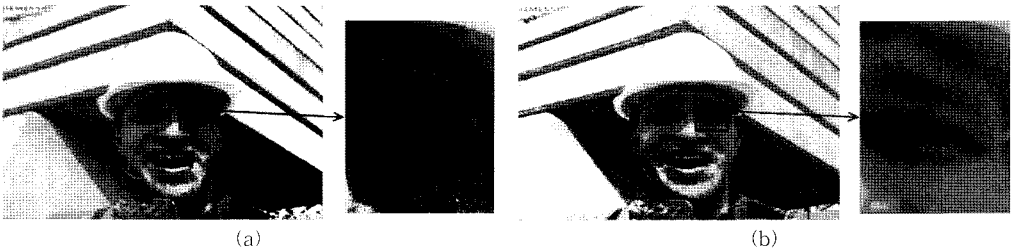


그림 13. 0.166 bpp에서 복호된 비디오의 25번째 프레임 (a) 제안한 방법 (b) H.263



그림 14. 0.144 bpp에서 복호된 비디오의 50번째 프레임 (a) 제안한 방법 (b) H.263



그림 15. 0.129 bpp에서 복호된 비디오의 90번째 프레임 (a) 제안한 방법 (b) H.263

## 5. 결 론

본 논문에서는 기존 3G 네트워크 환경에서 향상된 품질의 영상통화 서비스를 제공하기 위한 비디오 데이터 압축 방법을 제안하였다. 제안한 방법에서는 통화의 주체가 되는 사람의 얼굴 영역에 대한 화질 보장과 영상통화 시 발생하는 영상손실 성분의 최소화

에 목표를 두었다. 이를 위해 비디오 데이터 압축 시, 프레임 내에서 중요 객체와 비중요 객체를 분리하였다. 얼굴 영역이 포함된 사람 영역을 중요 객체로 설정하였으며, 중요 객체에 대해서는 많은 비트를 할당하는 저압축을 수행하여 고화질을 갖도록 하였다. 반면, 배경영역에 해당하는 비중요 객체는 고압축을 수행하여 저화질을 갖도록 하는 압축 절차를 적용하였

다. 중요 객체와 비중요 객체의 비트율은 상호보완적 관계를 가지며, 이로써 기존의 제한된 대역폭에서 비트율을 만족하도록 하였다. 비중요 객체 영역에 대한 코딩의 경우, 기존 방법으로 고압축을 수행하면 블록 에러와 같은 영상손실이 발생하기 때문에 이를 해결하기 위해 DWT 기반의 프레임 단위 압축을 수행하였다. 실험 결과에서 제안한 방법과 H.263의 비트율을 동일하게 두고 압축을 수행한 후에 복호된 중요객체 영역의 평균 PSNR을 비교하였으며, 제안한 기법이 H.263에 비해 중요 객체의 화질이 우수함을 확인하였다. 또한 주관적 화질 평가로서 복호된 각 프레임에 대한 시각적 평가에서도 제안한 방법의 중요 객체 영역이 우수한 화질을 가짐을 확인하였다.

참 고 문 헌

[ 1 ] 오규태, "IMT-2000(CDMA2000과 WCDMA)," 방송과 기술 통권 143호, pp. 109-126, 2007.  
 [ 2 ] 조경준, "HSDPA 서비스 동향," 방송과 기술 통권 140호, pp. 112-126, 2007.  
 [ 3 ] Codec for circuit switched multimedia telephony service; Modifications to H.324, 3GPP Technical Specification 26.111, Version 7.1.0: Jun. 2008. <http://www.3gpp.org/ftp/Specs/html-info/26111.htm>  
 [ 4 ] Terminal for low bit-rate multimedia communication, ITU-T Rec. H.324, Spt. 2005, <http://www.itu.int/rec/T-REC-H.324/en>  
 [ 5 ] Video Codec for Audiovisual Services at p×64 kbit/s, ITU-T Rec. H.261, ITU-T, Version 1: Nov. 1990, Version 2: Mar. 1993.

[ 6 ] Video Coding for Low Bit Rate communication, ITU-T Rec. H.263, ITU-T, Version 1: Nov. 1995, Version 2: Jan. 1998, Version 3: Nov. 2000.  
 [ 7 ] Coding of audio-visual objects-Part 2: Visual, ISO/IEC 14492-2 (MPEG-4 Visual), ISO/IEC JTC 1, Version 1: Apr. 1999, Version 2: Feb. 2000, Version 3: May 2004.  
 [ 8 ] M. Kamel and A. Campilho, "A New Object-Based Fractal Compression of Monocular and Stereo Video Sequences," Proc. ICIAR, pp. 103-115, 2009.  
 [ 9 ] J. M. Shapiro, "Embedded Image Coding Using Zerotree Wavelets Coefficients," *IEEE Trans. Signal Processing*, Vol. 41, pp. 3445-3462, 1993.  
 [10] A. Said and W. A. Pearlman, "A New Fast and Efficient Image Codec based on Set Partitioning in Hierarchical Trees," *IEEE Trans. Circuits and System for Video Technology*, Vol. 6, No. 3, pp. 243-250, 1996.  
 [11] W. FU, S. WANG, and D. KANG, "Hybrid Fractal Object-Based Zero-Tree Wavelet Video Coding Algorithm," Proc. ICSP, pp. 1147-1150, 2004.  
 [12] W. Sweldens, "The Lifting Scheme: A Construction of Second Generation Wavelets," *SIAM Journal on Mathematical Analysis*, Vol. 29, pp. 511-546, 1998.  
 [13] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 4, pp. 438-442, 1994.



전 성 혜

2008년 동서대학교 전자공학과 학사 졸업(공학사)  
2010년 부경대학교 컴퓨터공학과 석사 졸업(공학석사)  
관심분야: 비디오 코덱, 워터마킹, 영상신호처리



서 용 수

1975년 2월 경북대학교 전자공학과 졸업(공학사)  
1982년 2월 동아대학교 대학원 전자공학과 졸업(공학석사)  
1992년 2월 경북대학교 대학원 전자공학과 졸업(공학박사)  
1994년 3월~현재 동서대학교 정보시스템공학부 부교수.

관심분야: 영상처리, 패턴인식, Remote Sensing 등임.



권 기 룡

1986년 경북대학교 전자공학과 학사 졸업(공학사)  
1990년 경북대학교 전자공학과 석사 졸업(공학석사)  
1994년 경북대학교 전자공학과 박사 졸업(공학박사)  
2000년~2001년 Univ. of Minnesota, Post-Doc.

1996년~2005년 부산외국어대학교 디지털정보공학부 부교수

2006년~현재 부경대학교 IT융합응용공학과 교수

2009년~현재 한국멀티미디어학회 총무이사

관심분야: 멀티미디어 정보보호, 영상처리, 웨이블릿 변환