

## 임의의 ROI를 포함하는 JPEG2000 이미지의 ROI 코딩 기법

홍석원\*, 김상복\*, 서영건\*\*

### An ROI Coding Technique of JPEG2000 Image Including Some Arbitrary ROI

Seok Won Hong\*, Sang Bok Kim\*, Yeong Geon Seo\*\*

#### 요 약

이미지를 사용하는 영상처리 시스템이나, 단순히 특정 이미지를 보기를 원하는 일반 사용자에게 한 이미지 내에 특정 부분을 타 영역보다 높은 품질을 갖도록 한다면, 더 좋은 서비스를 제공할 수 있을 것이다. 특히 모바일 환경에서는 화면의 크기가 작으므로 우선적으로 보여야 될 부분이 필요하게 되었다. JPEG2000에서는 이러한 기능을 지원하고 있다. 하지만 구체적인 영역 추출 과정이나 서비스 기능은 없지만, 추가로 기능을 넣을 수 있도록 제공하고 있는데, 이것을 ROI(Region-of-Interest) 기법이라 한다. 본 논문에서는 인물이 포함된 이미지를 이용하여, 얼굴 영역이 가장 우선적으로 처리되고 높은 품질로 압축되도록 부가 서비스를 제공하는 것이다. 하나의 이미지가 사용자에게 서비스되기 전에 압축되고 저장되어야 하는데, 얼굴 영역은 배경 영역 보다 더 좋은 품질로 압축되고 배경은 품질을 상대적으로 떨어뜨리는 것이다. 전송될 때도 얼굴 영역은 우선적으로 전송하여 사용자에게는 훨씬 좋은 서비스를 제공할 수 있다. 또한 압축은 일반적인 방법으로 하고, 전송될 때 얼굴영역을 찾아 우선적으로 전송해도 된다. 얼굴영역 추출은 신경망 기반의 얼굴 검출 기법을 사용하고, 우선 처리는 JPEG2000의 EBCOT 기법을 사용한다. 실험은 여러 개의 얼굴이 포함된 이미지를 사용하며, 객관적인 평가와 주관적인 평가를 실시하며, 충분히 좋은 기법으로 증명되었다.

#### Abstract

In some image processing system or the users who want to see a specific region of image simply, if a part of the image has higher quality than other regions, it would be a nice service. Specifically in mobile environments, preferential service was needed, as the screen size is small. So, JPEG2000 supplies this function. But this doesn't support the process to extract specific regions or service and does the functions to add some techniques. It is called by ROI(Region-of-Interest). In this paper, we use images including human faces, which are processed most preferentially and compressed with high quality. Before an image is served to the users, it is compressed and saved. Here, the face parts are compressed with higher quality than the background which are relatively

• 제1저자 : 홍석원    교신저자 : 김상복

• 투고일 : 2010. 09. 07, 심사일 : 2010. 09. 10, 게재확정일 : 2010. 09. 13.

\* 경상대학교 컴퓨터학과, 컴·정보통신연구원    \*\* 경상대학교 컴퓨터교육과, 컴·정보통신연구원

with lower quality. This technique can offer better service with preferential transferring of the faces, too. Besides, whole regions of the image are compressed with same quality and after searching the faces, they can be preferentially transferred. In this paper, we use a face extraction approach based on neural network and the preferential processing with EBCOT of JPEG2000. For experimentation, we use images having several human faces and evaluate objectively and subjectively, and proved that this approach is a nice one.

▶ Keyword : ROI, Automatic ROI extraction, JPEG2000, ROI Coding, Neural Network

## I. 서론

여러 장의 이미지에서 사용자가 일일이 모든 이미지를 눈으로 보면서 특정 작업을 해야 할 일이 생긴다면, 사용자는 아주 성가신 일이 될 것이다. 이런 경우에 이미지 내에 특정 이미지를 포함하는 공통적인 특성이 있다면 그 영역을 자동적으로 찾아 주도록 하고, 그 영역이 중요한 영역이므로 이미지 압축 시에도 높은 품질로 저장되거나 전송되어야 할 필요성이 있다. 특히 큰 이미지나, 모바일과 같은 작은 화면에서처럼, 이미지의 일부가 다른 영역에 비해 좋은 품질로 보여야 할 경우가 허다하다. 이런 서비스를 위해 JPEG2000에서는 ROI (Region-of-Interest) 기법으로 제공하고 있다.

이와 같이 최근에는 영상의 물리적인 특성을 이용하여 사용자의 도움 없이 자동으로 영상의 주제가 되는 관심 영역을 추출하기 위한 논문이 활발히 진행되고 있다[1]-[5]. 특히 영상에서 촬영자가 자신의 의도를 포커스 정보를 이용하여 영상의 주제가 되는 관심영역을 표현한, 낮은 피사계 심도 영상의 경우 포커스 유무만 판단할 수 있다면 사용자의 도움 없이 주제가 되는 관심영역을 추출할 수 있다[5].

일반적으로 인간의 관심영역은 중앙에 분포한다는 조건으로 관심영역을 자동 추출하는 기법은 간단하면서도 충분한 서비스를 제공한다. 하지만 인간의 관심영역이 특정 객체에 기반하고 있다면 그 객체를 찾아서 높은 품질을 제공해야 한다. 인간의 관심을 특정 도메인에 한정한다면 특정 객체를 찾아서 우선적으로 처리하거나 배경보다 높은 품질로 압축하는 기법이 좋은 서비스가 될 수 있다. 이를 위해 본 논문에서는 사람의 얼굴이 포함된 이미지를 대상으로 하여, 얼굴을 배경보다 높은 품질로 서비스를 하는 기법을 제안하고 실험하여 충분히 좋은 기법이라는 것을 보인다.

얼굴 영역의 추출은 먼저 얼굴의 윤곽 마스크를 이용하여 얼굴의 개략적인 영역을 찾고, 그 속에 눈의 위치를 탐색한다. 눈은 얼굴의 위쪽 양쪽에 나란히 위치하는 특징이 있으므로

이 조건으로만 검색해도 충분히 얼굴의 위치를 찾을 수 있다. 이렇게 검색된 얼굴은, 얼굴의 위치에 대하여 ROI 마스크가 생성되고 마스크된 영역은 ROI 코딩된다. ROI 코딩이란 ROI 마스크된 영역을 배경보다 높은 품질로 압축하거나 우선적으로 전송하는 것이다. 한 이미지에 여러 개의 얼굴이 포함되어도 멀티 ROI로 제공한다.

ROI 코딩 방법[6]은 Maxshift 방법, Implicit 방법, 수정된 Implicit 방법, 적응적 방법[7] 등이 있으나, 본 논문에서는 Maxshift 방법을 이용한다. 실험을 통해, 제안된 기법을 이용하면, 객체를 고려하지 않고 압축한 이미지에 비해 훨씬 더 좋은 서비스가 될 수 있음을 증명해 보인다.

## II. ROI 코딩과 ROI 자동 추출

### 2.1 ROI 코딩

영상 처리에서 사람이 사물을 바라볼 때 먼저 전체적인 윤곽을 파악하고 차츰 자세한 부분에 집중한다는 사실에 기반을 두어 웨이블릿 변환은 아주 적합한 변환 방법이라 할 수 있다. 웨이블릿 변환은 영상을 다해상으로 분할 가능하게 하며, 분해된 각 대역은 원 영상에 비하여 분산 값이 크게 줄어들기 때문에 영상 부호화가 용이하다[6]. 또한 JPEG2000에서도 ROI를 표준으로 지원해 주기 때문에 ROI의 추출과 마스크 생성의 과정을 통하여 사용자에게 훨씬 더 좋은 서비스를 제공하는 논문도 진행되고 있다[7]-[8]. 영상의 2차원 웨이블릿 변환은 1개의 평균 영상(LL)과 가로(LH), 세로(HL), 대각선(HH) 에지 정보를 가진 고주파 서브밴드를 만들어 낸다. 변환이 수행되는 횟수는 요구되는 압축률, 원 이미지의 크기, 필터의 길이 그리고 실시간 문제에 따라 달라질 수 있다.

다양한 에지 검출방법에 비해 DWT 변환된 고주파 서브밴드들의 에지 정보를 이용하면 낮은 피사계 심도 영상의 포커스된 영역은 디포커스된 영역보다 고주파 정보를 더욱 많이 가지고 있다[5]. 따라서 관심 객체의 에지들을 찾기가 수월하다는 것을 알 수 있다[9]. ROI 코딩은 JPEG2000의 주요

특징 중의 하나로서 전체 영상을 전송 및 복원하기보다는 ROI를 배경보다 우선적 처리하는 기술이다. 이것은 점진적 이미지 전송 방법보다 압축률과 전송시간을 단축시킬 수 있을 뿐만 아니라 효율적인 메모리 관리 등으로 인해 보다 빠른 이미지 서비스가 가능하다.

계수 스케일링 기반의 ROI 코딩 방법은 대부분 정적 ROI에 속하며, 대부분의 응용 분야에서는 ROI 서비스를 위해 정적 ROI 코딩 방법으로도 충분히 가능하다. 하지만 인코딩 과정에서 ROI를 모른다면, 디코딩 과정에서 ROI 모양 정보를 받아서 ROI 코딩을 한다. 이 처럼 디코딩 과정에서 ROI 모양을 정의한 후 ROI 코딩하는 방법을 동적 ROI 코딩이라고 한다. 본 논문은 정적 ROI 코딩 기법에 속하며, 그림 1과 같은 과정으로 코딩한다. 그림 1에서 ROI 추출은 컬러 변환 후에 실행된다.

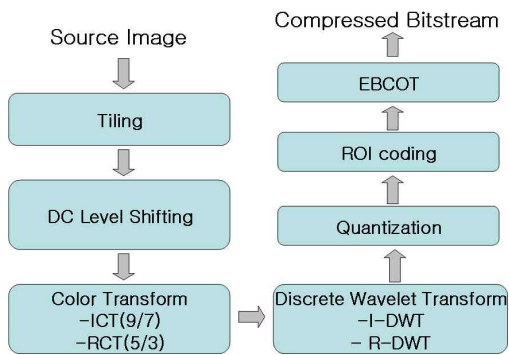


그림 1. JPEG2000 인코딩 과정  
Fig. 1. JPEG2000 Encoding Process

### 2.2 Maxshift 코딩

Maxshift방법[6]은 양자화된 계수 중에 ROI 계수와 배경 계수를 구분하여, 배경 계수 중에서 가장 큰 계수 값,  $s$ 를 구한 다음, ROI 계수를  $s$ 보다 높은 비트-평면에 이동시키는 방법으로서 JPEG2000 Part1 표준이다. 식(1)은  $s$ 를 구하는 수식이고, 식(2)는 ROI 처리 후의 계수,  $a'(u, v)$ 를 구하는 수식이다.

$$s \geq \max(M_b) \quad \dots\dots\dots (1)$$

$$a'(u, v) = \begin{cases} a(u, v), & M(u, v) = 0 \\ 2^s a(u, v), & M(u, v) = 1 \end{cases} \quad \dots\dots\dots (2)$$

식(1)에서  $\max(M_b)$ 은 각 서브밴드에서 양자화된 배경 계수 중에 가장 큰 값을 의미한다. 식(2)에서  $M(u, v)$ 는 ROI 마스크 정보로서 계수가 ROI에 속하는 좌표인 경우는 1,

배경에 속하는 좌표인 경우는 0을 의미한다. 장점은 인코딩 시에 ROI 모양 정보 대신  $s$  값만 가지게 되므로 코딩 효율이 좋다. 단점은 ROI가 모두 복원될 때까지 배경을 얻을 수 없다. 즉 ROI 중요도 조절이 불가능하다. 또한 다수 ROI 지원이 어렵다.

### 2.3 ROI 자동 추출

하나의 이미지를 저장하거나 전송하기 위하여 인간의 인지를 이용하여 ROI를 지정하고 ROI 코딩을 하는 것은 별로 문제가 되지 않는다. 처리해야 할 이미지가 수천 장이라면 자동으로 ROI를 추출하지 않으면 너무 많은 작업 시간이 필요하다. 뿐만 아니라, 특정 객체가 관심을 갖는 부분이라면, 처리할 모든 이미지에 객체가 들어가 있을 것이다. 이것을 자동적으로 찾아내어 특별한 처리를 하는 것이 ROI 자동 추출 및 ROI 코딩이다.

기존에는 ROI의 추출을 위하여 이미지의 에지를 이용하여 일반적으로 추출하며[5], 웨이블릿 변환을 이용하므로 주로 LH, HL, HH 밴드의 정보를 이용한다. [1]에서는 세 밴드의 계수를 OR하여 계수의 값이 특정 임계치를 초과하면 1로 설정하고 그렇지 않으면 0으로 설정하여 에지를 찾아낸다. 다음으로 잡음제거 과정을 거친다. 추출된 ROI의 형태는 임의 모양, 네모, 원을 가질 수 있고, 한 개 이상의 ROI를 가지는 기법도 있다. 일반적으로 관심영역은 영상의 가장자리가 아닌 중앙에 위치할 가능성이 크다. 그림 2는 lena 이미지를 1차 웨이블릿 변환한 것이다. LL 밴드를 제외한 나머지 3 밴드에는 이미지의 에지들이 분포되어 있음을 볼 수 있다.



그림 2 lena 1차 웨이블릿 변환  
Fig. 2 lena wavelet 1-decomposition

기존의 방법은 단지 에지의 분포에 의존하므로 인간의 시각 인식을 전혀 고려하지 않은 처리 방법이다. 즉, 에지의 분

포가 골고루 퍼져 있다거나, 에지가 잘 나타나지 않는 이미지에서는 결코 좋은 ROI를 얻지 못하는 단점이 있다. 그래서 본 논문에서는 인간의 이미지 인식을 고려하여 객체를 찾아내어 그 객체를 ROI로 간주하고, ROI 코딩한다. 본 논문에서 사용하는 객체는 인간의 얼굴이다. 일반적인 인물 사진인 경우에는 인물이 우선적으로 고려되어야 하며, 특히 범죄에 관련된 인물 사진인 경우에는 특히 얼굴 부분을 중요한 객체로 처리해야 한다. 그림 3에서는 피사계 심도가 낮은 영상에서 ROI 자동 추출후의 결과를 보이고 있다. (a) 이미지는 1.2bpp로 압축하여 추출한 것이며, (b)는 0.2bpp 압축한 후에 얻은 ROI 영역을 표시하고 있다. 그림에서와 같이 인간의 눈으로는 같은 그림처럼 보이지만 압축을 많이 하면 에지도 더불어 사라지므로 ROI를 구하는 데 문제를 일으키는 것을 보이고 있다.

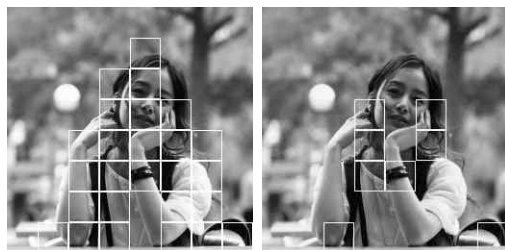


그림 3. [5] 기법의 ROI 추출  
Fig. 3. Edge Distributions of [5]

2.4 ROI 마스크

ROI 마스크 테이블은 영상의 해상도와 같은 크기로 할당되며 각 픽셀이 ROI에 속하는 픽셀이면 ROI 마스크에 있는 그 값도 1이 되며, 배경이면 0으로 설정된다. ROI 블록 내에 ROI 픽셀과 배경 픽셀이 같이 있을 수도 있지만, 모두 하나의 특성으로 분류된다. 우선 처리를 위해서는 EBCOT의 각 품질 레이어에서 코드블록의 임베디드 비트 스트림으로부터 임의의 공헌도를 포함하게 한다. 전체 손실 최소화에 의해서 코드블록 공헌도를 할당하기 때문에, ROI는 손실 감소와 ROI가 일치하도록 코드블록 공헌도를 할당한다. 식(3)은 손실 계산방법을 나타낸다.

$$D_j^{n_i} = \begin{cases} W_{ROI} w_b \sum_{u,v \in B_j} (\hat{a}^{n_i}(u,v) - a(u,v))^2, & ROI \text{ codeblock} \\ w_b \sum_{u,v \in B_j} (\hat{a}^{n_i}(u,v) - a(u,v))^2, & \text{그외} \end{cases} \quad (3)$$

식(3)에서  $W_{ROI}$ 는 ROI 가중치이고,  $D_j^{n_i}$ 는  $n_i$ 에서의 가

중 MSE 손실이고,  $w_b$ 은  $B_j$ 를 포함하는 서브밴드의 가중치이고,  $B_j$ 는  $j$ 번째 코드블록이고,  $a(u,v)$ 는 웨이블릿 계수이고,  $\hat{a}(u,v)$ 는  $n_i$ 으로 양자화된 웨이블릿 계수이고,  $n_i$ 는 절단점을 의미한다.

III. 얼굴 영역 추출과 우선적 처리

본 장에서는 이미지에서 자동으로 의미 있는 관심 객체, 얼굴을 추출하고 압축 또는 저장하는 알고리즘을 제안하고 ROI 코딩에서 우선 처리하는 기법을 기술한다.

3.1 신경망 기반의 얼굴 추출

본 논문에서는 흑백 이미지를 이용하여 이미지 내의 사람의 얼굴을 추출하고, 추출된 영역을 ROI로 간주한다. 얼굴은 정면 얼굴로 한정하고, 많이 기울어진 얼굴은 추출 안 될 수도 있다. 알고리즘은 크게 두 부분으로 나뉜다. 이미지에 여러 개의 신경망 기반의 필터를 적용하는 첫 번째 단계와 첫 번째 단계의 결과를 조합하기 위해 조정자를 사용하는 것이 두 번째 단계이다[10]-[12].

먼저 20x20 픽셀 영역을 갖는 필터를 실행한 결과는 -1 ~ 1 사이의 값을 갖는다. 1은 100% 얼굴을 갖는다는 의미이며 -1은 100% 얼굴을 갖지 않는다는 의미이다. 그림 4는 20x20 픽셀 영역의 필터를 나타낸다. 그림에서 1은 255의 값을 의미하고 0은 0을 의미한다.

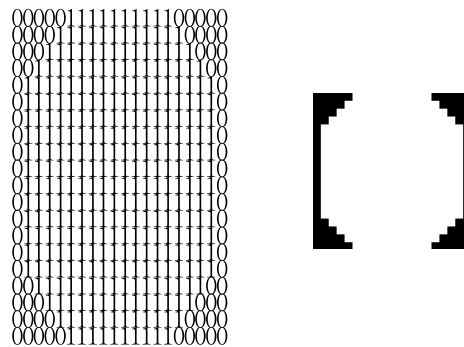


그림 4. 20x20 얼굴 마스크  
Fig. 4. 20x20 Face mask

얼굴 이미지가 마스크 보다 크면, 이미지를 단계적으로 줄이면서 정합을 하고, 이미지가 마스크보다 작다면 그 얼굴은 찾지 못한다. 정합은 한 이미지 내의 모든 픽셀에 대해 실행하고, 이미지의 스케일을 줄이는 것은 1.2배씩 감소시킨다.

필터링 알고리즘은 그림 5에 나타내었다. 먼저 전처리 과정으로써, 이미지의 윈도우에 적용된다. 윈도우는 그 윈도우 내에 얼굴이 포함되어 있는지를 알기 위해 신경망을 통과한다. 선형함수는 윈도우 내의 강도(intensity value)에 맞춘 후, 어떤 극단적인 밝기 상태에서는 수정을 하고, 차이를 구한다. 다음으로 콘트라스트를 개선하고 다양한 camera gain을 수정하기 위해 히스토그램 평준화가 적용된다. 이 각각의 단계에 대해서 그림 4의 타원 마스크 내의 픽셀에 기반을 두어 정합이 계산되고 이후에 전체 윈도우에 적용된다.

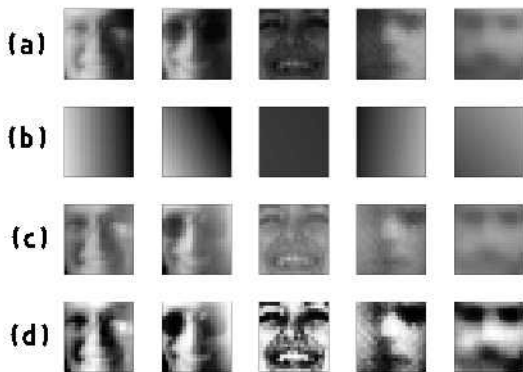


그림 5. (a) 원 윈도우 (b) 최적 선형함수 (c) 수정된 윈도우 조명 효과 (d) 평준화시킨 윈도우의 히스토그램  
Fig. 5. (a) Original window (b) Best fit linear function (c) Lighting corrected window (d) Histogram equalized window

다음으로, 전처리된 윈도우가 신경망을 통과하는 것이다. 신경망으로 20x20 픽셀이 입력되면 수용영역을 거쳐 은닉 단위를 통과하여 결과를 출력한다. 은닉 단위는 3 가지로 구성되는데, 10x10 픽셀 부영역을 찾는 4 은닉 단위, 5x5 픽셀 부영역을 찾는 16 은닉 단위, 겹치는 20x5 픽셀 수평 무늬 픽셀을 찾는 6 은닉 단위로 구성된다. 이 모든 은닉 단위는 얼굴 추출에 중요한 지역적 특징을 찾는 데 사용된다. 특히, 수평 무늬는 입이나 눈, 코, 입의 끝 등을 찾는 데 사용된다.

### 3.2 웨이블릿 변환 후의 얼굴 추출

얼굴 영역을 추출하는 시점은 전체 이미지에서도 가능하지만, 웨이블릿 변환을 한 후에도 가능하다. 전체 이미지에서 찾을 때는 이미지 도메인이 크기 때문에 탐색 시간이 많이 걸리지만, 얼굴을 정확하게 찾을 수 있다. 웨이블릿 변환 후에 추출을 하면 이미지가 25%로 줄어들기 때문에 20x20 마스크보다 더 얼굴 영역이 줄어들 수도 있다. 이 경우에는 추출이 불가능하다. 그림 6은 원 이미지와 웨이블릿 분해 후의 얼

굴 영역 추출 결과를 보이고 있다. (a)는 원 이미지에서, (b)는 1차 웨이블릿 변환 후, (c)는 2차 웨이블릿 변환 후, (d)는 3차 웨이블릿 변환 후이다. 샘플 이미지는 768x768 크기이며, 3차 웨이블릿 변환 후의 LL 밴드의 크기는 96x96이 된다. 그림 상의 흰색 네모로 된 영역은 ROI의 precinct이다. 압축을 위해서는 LL 영역 내의 ROI만 압축하는 것이 아니라, 나머지 고주파 밴드 내의 precinct, 즉, 이미지 도메인 내의 같은 위치에 있는 계수까지 압축해야 한다.

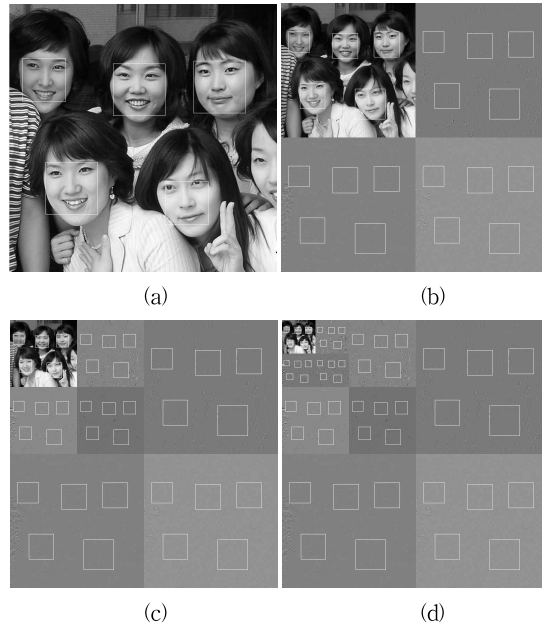


그림 6. 웨이블릿 변환 분해 레벨 별 ROI precinct  
Fig. 6. ROI precincts by wavelet decomposition levels

ROI 추출 후의 밴드별 precinct를 모두 압축을 하기 위해 각 밴드 별로 마스크가 설정되어야 하며, 이 마스크는 이미지 도메인에서의 마스크가 웨이블릿 도메인에서의 마스크로 정의되어야 한다. 웨이블릿 도메인에서의 ROI 마스크 정보는 이미지 도메인에서의 ROI 마스크 정보와 IDWT를 이용하면 구할 수 있다. 즉, 웨이블릿 도메인에서의 ROI 마스크 생성 과정은 우선 이미지 도메인에서의 ROI 마스크와 마지막 IDWT에 의해 2개의 서브밴드 안에 어떤 위치의 계수가 ROI 계수로 재구성하기 위해 사용되었는지를 추적하게 된다. 그런 후에, 두 서브밴드에서의 ROI 마스크와 마지막 이전의 IDWT에 의해 각각 2개의 서브밴드 안에 어떤 위치의 계수가 ROI 계수로 재구성하기 위해 사용되었는지를 추적하게 된다. 이렇게 하여 각 분해레벨에서 모든 서브밴드 내에 어떤 위치의 계수가 ROI 계수로 재구성하기 위해 사용되었는지를 추적

하게 된다. 그림 7은 이미지 도메인에서의 ROI 마스크를 웨이블릿 도메인에서의 ROI 마스크로 변환되는 과정이다. 웨이블릿 도메인에서의 ROI 마스크는 각 분해 레벨에서 서브밴드 안에 ROI 코드블록, 배경 블록, 부분 배경 계수를 포함하는 ROI 블록으로 구성된다.

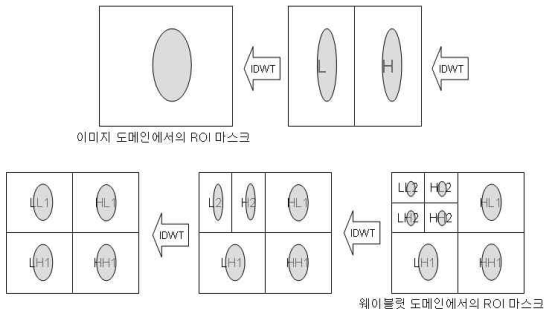


그림 7. 관심영역 마스크 변환  
Fig. 7. Conversion of ROI mask

### 3.3 ROI 코드블록의 우선처리

관심영역은 코드블록 단위로 하나의 이미지를 나누고 각 코드 블록에 해당하는 ROI 마스크를 EBCOT에 입력시킨다. 본 논문에서 주어진 ROI 영역은 모두 네모로 구성되고 코드 블록은 다음 세 가지의 마스크 영역으로 구성된다.

- 모든 픽셀이 ROI인 코드블록
- 일부 픽셀만 ROI인 코드블록
- 모든 픽셀이 배경인 코드블록

ROI 블록에 해당하는 꼭짓점 (x1, y1), (x2, y2)가 구하여 졌으므로 이 공간 내에 속하면 ROI로 지정하고 속하지 않으면 배경으로 지정한다. EBCOT에서는 각 서브밴드를 코드 블록으로 나누어서 코드블록에 대해 독립적으로 코딩된 임베디드 비트스트림으로부터 최종 비트스트림이 만들어 진다. 블록 전체가 ROI인 경우는 우선권 조절이 필요 없이 그냥 보내면 되지만, 일부가 ROI인 경우는 배경 영역을 k만큼 절단하여 우선권 조절을 한다. 일부가 ROI인 코드블록의 우선권 조절은 식(4)에 의해 얻어진다.

$$\bar{a}(u, v) = \begin{cases} \text{sigm}[a(u, v)] \left[ \frac{|a(u, v)|}{2^k} \right] 2^k, & \text{배경계수} \\ a(u, v), & \text{ROI 계수} \end{cases} \dots\dots\dots (4)$$

매개 변수 k는 임의로 지정 가능하며 본 논문에서는 Maxshift 방법을 사용하였으므로 k는 모든 배경 영역을 ROI 영역 보다 낮은 우선순위를 갖도록 k를 결정한다.



그림 8. 실험에 사용된 이미지  
Fig. 8. Images used in experimentation  
(5girls, 8persons, hyesun)

## IV. 실험 및 평가

이 장에서는 제안한 방법을 적용한 결과를 보이기 위하여 하나 이상의 얼굴이 포함된 이미지를 이용하여 얼굴 검출 후 ROI를 지정하고 ROI가 포함된 이미지를 압축 후의 PSNR을 여러 기법과 비교한다. 비교는 압축률과 웨이블릿 분해 레벨에 따라 어떻게 다르게 나타나는 지도 보인다.

### 4.1 실험 방법

실험에 사용된 이미지는 그림 8과 같으며, 입력 이미지의 해상도는 (a)와 (b)는 768×768이며, (c)는 1024x1024이다. 웨이블릿 변환 시 사용된 필터는 D-9/7필터를 사용하였으며, 실험은 Maxshift 방법을 이용하며, PSNR을 측정한다. 사용된 이미지는 자연 이미지이며, ROI의 모양은 네모이며, 분해 레벨(l)은 1~3 레벨까지 다양하게 실험하였다. 실험에 사용된 모든 이미지는 흑백 영상이고, 5girls 이미지는 2 분해 레벨, 8persons 이미지는 1 분해 레벨, hyesun 이미지는 3 분해 레벨에서 얼굴 추출 후 ROI를 결정하였다. ROI를 추출하는 데 LL 밴드 계수만 사용하였으며, LH, HL, HH 밴드의 계수는 사용되지 않았다.

실험에 사용된 모든 이미지는 0.1bpp, 0.2bpp, 0.4bpp, 0.8bpp로 각각 압축되었으며, 비교를 위하여 같은 이미지에서 같은 분해 레벨에서 얼굴을 추출하였다. bpp는 원 이미지를 압축할 때의 비트율을 의미한다. 비트율이 낮으면 원 이미지에

손상이 많이 가해지므로 얼굴 영역을 추출하는 것이 불가능할 수도 있다. 뿐만 아니라 분해 레벨이 클수록 원 이미지를 축소 한 상태에서 추출을 하므로 얼굴 추출이 더 어려워진다.



그림 9. 각 이미지의 특성에 따른 ROI위치  
Fig. 9. Each image's ROI regions by its characteristics

4.2 ROI 추출 위치

실험에 사용된 이미지에서 얼굴 추출이 웨이블릿 변환을 한 후에도 잘 되는 지를 실험을 통해 보인다. 사용된 3 개의 이미지를 압축률에 따라, 분해 레벨에 따라 어떻게 ROI가 결정되는지를 그림 9에 나타내었다. 5girls에서 각각 다른 압축률에 똑같이 한 사람의 얼굴을 찾을 수 없었다. 머리칼이 이마를 가리고 있기 때문이다. 실제 눈으로 구분하기는 힘들지만 (a)의 이미지가 (b)의 이미지보다 약간 화질이 떨어지는데도 불구하고 같은 영역에서 얼굴이 추출되었다. 8persons는 1 분해 레벨에서 ROI를 추출하였는데, 0.1bpp 환경에서는 하나의 얼굴을 검출하지 못하였다. (c)와 (d) 모두 한 사람의 얼굴을 추

출하지 못하였는데, 이것도 (a) (b)와 마찬가지로 머리칼이 얼굴을 많이 가리고 있기 때문이다. (e)와 (f) 이미지는 3 분해 레벨에서 추출하였음에도 성공적으로 실행되었다.

비교를 위해, [5]에서 제안한 기법을 사용하여 ROI추출한 결과를 그림 10에 나타내었다. [5]에서는 객체의 특징을 이용하지 않고 에지의 분포를 이용하여 ROI를 추출했기 때문에 사용자의 의도와는 다른 부분이 추출됨을 볼 수 있다. 그림 10의 (a)는 768x768 이미지를 0.8bpp로 압축한 후, 2 레벨 웨이블릿 변환 후에 [5] 기법에 의해 ROI를 추출한 것이다. [5] 방법은 압축을 전혀 하지 않아도 에지 분포로만 찾기 때문에, 에지의 분포 특성이 ROI 분포 특성과 일치하지 않으면 그림과 같은 결과가 나타난다. (b)는 1024x1024 이미지를 0.2bpp로 압축한 후, 1레벨 웨이블릿 변환 후에 [5] 기법에 의해 ROI를 추출한 것이다. 에지의 분포가 사용자가 원하는 ROI를 일부 포함하지만, 이 이미지는 배경이 모두 같은 패턴이므로 가능한 결과이다.

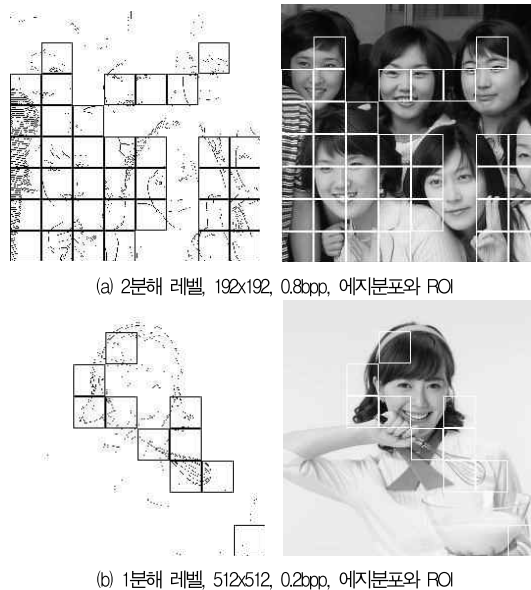


그림 10. [5] 기법을 이용한 ROI 위치  
Fig. 10. ROI positions of [5] approach

4.3 화질 평가

본 논문은 객관적인 화질 평가를 위해 PSNR을 사용한다. 평가는 식(5)와 같고, 샘플 당 n 비트의 길이를 가진다.

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \dots\dots\dots (5)$$

MSE는 식(6)과 같다.

$$MSE = \frac{\sum (\hat{x} - x)^2}{A} \dots\dots\dots (6)$$

여기서  $x$ 는 원 화소 값을 의미하고,  $\hat{x}$ 은 재구성된 화소 값을 의미하고,  $A$ 는 ROI를 의미한다.

표 1. 각 이미지의 비트율에 따른 PSNR  
Table 1. PSNRs by each bit rate of 3 images

영상		5girls	8persons	hyesun
비트율(bpp)				
0.1	no ROI	26.740	27.090	29.001
	with ROI	26.009	26.811	28.789
0.2	no ROI	28.800	29.892	32.089
	with ROI	28.451	29.589	31.456
0.4	no ROI	32.833	33.834	36.610
	with ROI	32.552	33.680	36.534
0.8	no ROI	38.616	39.340	41.802
	with ROI	38.348	38.934	41.000

ROI를 포함하지 않은 이미지의 화질은 모든 이미지 영역에서 같고, 화질은 상대적으로 ROI를 포함한 이미지보다는 낮다. 하지만 ROI를 포함한 이미지는 전체적으로 화질은 상대적으로 조금 떨어지지만, Maxshift를 사용하여 ROI 코딩했기 때문에 ROI 부분은 최고의 화질을 제공한다. ROI가 포함되었을 때와 포함되지 않았을 때의 화질 차이는 표 1에서 보이는 것처럼 아주 미미하기 때문에 ROI가 사용자의 의도에 따라 잘 추출된다면 훌륭한 서비스를 제공할 수 있음을 보이고 있다. 이미지의 화질을 타 기법과의 비교했을 때 우수하다는 것은 큰 의미가 없다. 대부분의 기법이 ROI 코딩을 했을 때와 하지 않았을 때의 화질을 비교해 보면 큰 차이가 없기 때문이며, 제안한 기법과 타 기법과의 비교도 화질 면에서는 큰 차이가 없으므로 생략한다.

주관적 화질 평가를 위해서 그림 10에서와 같이 ROI 우선 전송을 했을 때, 화면에 보이는 이미지를 나타내었다. 전체 이미지가 완전히 전송되었을 때는 그림 9에서와 같이 보이겠지만, 완전히 전송되지 않고 ROI와 배경의 일부만 전송되었을 때는 그림 11과 같이 보이게 된다. hyesun의 이미지를 2분해 레벨 0.1bpp에서 압축하여 전송 중에 보이는 일련의 이미지를 나타낸 것이다. 2분해를 했으므로 ROI는 이미지가 256x256 크기에서 구했다. (a)는 ROI만 나타난 것이며, (b)는 ROI와 배경의 첫 비트만 전송된 후, (c)는 ROI와 배경의 두 번째 비트 전송된 후, (d)는 ROI와 배경의 세 번째 비트까지 전송된 후를 보이고 있다.



그림 11. hyesun의 우선 처리된 사용자 화면  
Fig. 11. User view with preferred processing of hyesun

### V. 결론 및 향후과제

자동으로 인간의 관심영역을 추출하는 것은 사용자의 다양한 요구와 방대한 데이터에서 원하는 데이터를 빠른 시간 안에 처리하려는 의지에서 비롯된 것이다. 물론 사람이 직접 처리하는 것과는 정확성 면에서는 비교도 안 되겠지만, 유사한 종류의 데이터에서 같은 종류의 객체를 추출한다면 충분히 유용하다는 것을 본 논문을 통해 보였다. 본 논문은 인간의 얼굴을 관심영역으로 하는 응용에 유용한 자동 객체 추출과 그것의 우선 처리에 대해 제안한 방법이 우수함을 보였고, 관심영역을 처리하지 않은 이미지 처리와 비교해서 질적으로 우수함을 보였다. 본 논문에서는 앞면만 추출할 수 있는 데, 향후에는 옆면도 추출할 수 있는 ROI 코딩을 논문할 것이며, 나아가서 다양한 객체를 추출할 수 있고, 다중 객체가 포함된 이미지에 대해서도 ROI 코딩을 적용하는 논문을 할 것이다.



## 참고문헌

[1] O. T. Chen and C. C. Chen, "Automatically Determined Region of Interest in JPEG 2000," *Multimedia, IEEE Trans*, Vol. 9-7, pp.1333-1345, Nov. 2007.

[2] C. Kim, "Segmenting a low-depth-of-field Image Using Morphological Filters and Region Merging," *IEEE Trans, Image Processing*, Vol.14-10, pp. 1503-1511, Oct. 2005.

[3] C. S. Won, K. Pyun and R. M. Gray, "Automatic Object Segmentation in Images with low depth of Field," *2002 International Conf. Image Processing*, Vol. 3-24-28, pp. 805-808, June 2002.

[4] 박재홍 외, "JPEG2000에서 ROI의 자동 추출과 우선적 처리," *한국컴퓨터정보학회 논문지*, 제 13권, 제 5호, 127-136쪽, 2008년 11월.

[5] 박순화 외, "피사체 심도가 낮은 이미지에서 웨이블릿 기반의 자동 ROI 추출 및 마스크 생성," *한국컴퓨터정보학회 논문지*, 제 14권, 제 3호, 2009년 3월.

[6] A. Graps, "An Introduction to Wavelets," *Computational Science & Engineering, IEEE* Vol. 2-2, pp. 50-61, Summer 1995.

[7] 강기준 외, "JPEG2000의 관심영역 부호화를 위한 적응적인 관심영역 마스크 생성 방법," *한국컴퓨터정보학회 논문지*, 제 12권, 제 5호, 39-47쪽, 2007년 11월.

[8] 강기준 외, "A Performance Evaluation of Factors Influencing the ROI Coding Quality in JPEG2000," *한국컴퓨터정보학회 논문지*, 제11권 4호, 197-206쪽, 2006년 1월.

[9] J. Z. Wang, J. Li and R. M. Gray, "Unsupervised Multiresolution Segmentation for Images with low depth of Field," *Pattern Analysis and Machine Intelligence, IEEE Trans.*, Vol. 23, pp. 85-90, Jan. 2001.

[10] B. Zoran and S. Andreja, "Face Detection Approach in Neural Network Based Method for Video Surveillance," *Neural Network Applications in Electrical Engineering, NEUREL* 2006. pp. 44-47, Sept. 2006.

[11] Rowley, Baluja, and Kanade, "Neural Network-Based

Face Detection," *PAMI*, Jan. 1998.

[12] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images : A Survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp.34-58, Jan. 2002.

## 저자 소개



홍 석 원

2001년 : 경상대학교  
멀티미디어과 석사  
2003년~현재 : 경상대학교 컴퓨터  
과학과 박사과정  
관심분야 : 영상처리, 데이터 압축



김 상 복

1989년 : 중앙대학교 전자과 박사  
1984년~현재 :  
경상대학교 컴퓨터과학과 교수  
2001년~현재 :  
경상대학교 컴퓨터정보통신연구소원  
관심분야 : 멀티미디어통신, 영상인식,  
보안



서 영 건

1987년 : 경상대학교 전산과 학사  
1997년 : 숭실대학교 전산과 박사  
1989년~1992년 : 삼보컴퓨터  
1997년~현재 : 경상대학교  
컴퓨터교육과 교수  
2001년~현재 : 경상대학교 컴퓨터  
정보통신연구소원  
관심분야 : 알고리즘, 영상인식,  
창의적 알고리즘교육