

다중 관심영역의 자동 추출 및 부호화 방법

서영건*, 홍도순**, 박재홍***

요 약

JPEG2000에서는 영상에서 원하는 영역을 타 영역(배경)보다 고화질로 압축하는 기법인, 관심영역 부호화 방법을 제공하고 있는데, 본 연구에서는 얼굴이 포함된 영상을 이용하여, 얼굴 영역이 가장 우선적으로 처리되고 높은 품질로 압축되도록 부가 서비스를 제공한다. 제안 기법은 크게 두 단계로 구성된다. 첫 번째는 얼굴 추출 단계이고, 두 번째는 관심영역 부호화 단계이다. 얼굴 추출은 영상의 모든 화소에 대해 20x20 윈도우 화소 크기로 자르거나 축소하여 전처리 과정을 거친 후 신경망을 이용하여 인식한다. 추출된 각 영역은 관심영역 마스크로 표시되고, Maxshift 방식을 이용하여 부호화된다. 이후에 EBCOT 과정을 거쳐 압축 및 저장된다. 기존의 방법은 고주파 성분의 분포에 의해 관심영역을 찾은 후 부호화하는 방법이 많이 연구되었다. 반면에 본 연구는 인간의 인지 능력을 이용하여, 여러 개의 얼굴이 포함된 영상에서 충분히 유용한 기법임을 보인다.

Automatic Extraction and Coding of Multi-ROI

Yeong-Geon Seo*, Do-Soon Hong**, Jae-Heung Park***

Abstract

JPEG2000 offers the technique which compresses the interested regions with higher quality than the background. It is called by an ROI(Region-of-Interest) coding method. In this paper, we use images including the human faces, which are processed uppermost and compressed with high quality. The proposed method consists of 2 steps. The first step extracts some faces and the second one is ROI coding. To extract the faces, the method cuts or scale-downs some regions with 20x20 window pixels for all the pixels of the image, and after preprocessing, recognizes the faces using neural networks. Each extracted region is identified by ROI mask and then ROI-coded using Maxshift method. After then, the image is compressed and saved using EBCOT. The existing methods searched the ROI by edge distributions. On the contrary, the proposed method uses human intellect. And the experiment shows that the method is sufficiently useful with images having several human faces.

Key Words : 관심영역(ROI), 관심영역 추출(ROI Extraction), JPEG2000

1. 서론

최근에는 영상의 물리적인 특성을 이용하여 사용자의 도움 없이 자동으로 영상의 주제가 되는 관심 영역을 추출하기 위한 연구가 활발히

진행되고 있으며[1]-[5]. 특히 영상에서 촬영자가 자신의 의도를 포커스 정보를 이용하여 영상의 주제가 되는 낮은 피사계 심도 영상의 경우 포커스 유무만 판단할 수 있다면, 사용자의 도움 없이 주제가 되는 관심영역을 추출할 수 있다 [5]. 여러 장의 영상에서 사용자가 일일이 모든 영상을 눈으로 보면서 특정 작업을 해야 할 일이 생긴다면, 사용자는 아주 성가신 일이 될 것이다. 이런 경우에 영상 내에 특정 영상을 포함하는 공통적인 특성이 있다면 그 영역을 자동적으로 찾도록 할 수 있다. 또한, 그 영역이 중요한 영역이므로 영상 압축 시에도 높은 품질로 저장되거나 전송되어야 할 필요성이 있다. 특히

※ 제일저자(First Author) : 서영건

※ 교신저자(Corresponding Author) : 홍도순

접수일:2010년 10월 03일, 수정일:2010년 12월 06일,

완료일:2011년 03월 30일

* 경상대학교 컴퓨터과학과, 컴퓨터정보통신연구원

young@gnu.ac.kr

** 경상대학교 컴퓨터과학과

*** 경상대학교 컴퓨터과학과

큰 영상이나, 모바일과 같은 작은 화면에서처럼, 영상의 일부가 다른 영역에 비해 좋은 품질로 보여야 할 경우가 허다하다. 이런 서비스를 위해 JPEG2000에서는 관심영역(ROI : Region-of-Interest) 기법으로 제공하고 있다.

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 연구에서는 사람의 얼굴이 포함된 영상을 대상으로 하여, 얼굴을 배경보다 높은 품질로 서비스를 하는 기법을 제안하고 실험하여 충분히 좋은 기법이라는 것을 보인다.

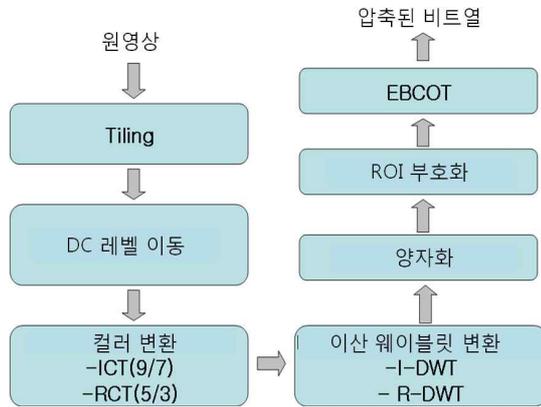
관심 영역의 추출은 먼저 얼굴의 윤곽 마스크를 이용하여 얼굴의 개략적인 영역을 찾고, 그 속에 눈의 위치를 탐색한다. 눈은 얼굴의 위쪽 양쪽에 나란히 위치하는 특징이 있으므로 이 조건으로만 검색해도 충분히 얼굴의 위치를 찾을 수 있다. 이렇게 검색된 얼굴은, 얼굴의 위치에 대하여 관심영역 마스크가 생성되고 마스크된 영역은 관심영역 부호화된다. 관심영역 부호화란 마스크된 영역을 배경보다 높은 품질로 압축하거나 우선적으로 전송하는 것이다. 한 영상에 여러 개의 얼굴이 포함되어도 다중 관심영역으로 제공한다. 관심영역 부호화 방법[6]은 Maxshift 방법, Implicit 방법, 수정된 Implicit 방법, 적응적 방법[7] 등이 있으나, 본 연구에서는 Maxshift 방법을 이용한다. 실험을 통해, 제안된 기법을 이용하면, 객체를 고려하지 않고 압축한 영상에 비해 훨씬 더 좋은 서비스가 될 수 있음을 증명해 보인다.

본 논문의 구성은 2장에서 JPEG2000과 관심영역 부호화에 대해 기술하고, 3장에서는 얼굴추출 및 그 부호화에 대해 기술하고, 4장에는 실험을 통하여 본 연구의 유효성을 보이고, 5장에는 결론을 기술한다.

2. JPEG2000과 관심영역 부호화

2.1 JPEG2000

웨이블릿 변환은 사람이 사물을 바라볼 때 먼저 전체적인 윤곽을 파악하고 차츰 자세한 부분에 집중한다는 사실에 기반을 두기 때문에 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 관심영역을 표준으로 지원해 주기 때문에 관심영역의 추출과 마스크 생성의 과정을 통하여 사용자에게 훨씬 더 좋은 서비스를 제공하는 연구도 진행되고 있다[7]. 영상의 2차원 웨이블릿 변환은 1개의 평균 영상(LL)과 가로(LH), 세로(HL), 대각선(HH) 가장자리 정보를 가진 고주파 서브밴드를 만들어 낸다. 변환이 수행되는 횟수는 요구되는 압축률, 원 영상의 크기, 필터의 길이 그리고 실시간 문제에 따라 달라질 수 있다.



(그림 1) JPEG2000 부호화 과정

그림 1에서처럼 관심영역 추출은 컬러 변환 후에 실행되며, 본 연구는 정적 기법에 속한다. 계수 스케일링 기반의 관심영역 부호화 방법은 대부분 정적 기법에 속하며, 대부분의 응용 분야에서는 관심영역 서비스를 위해 정적 부호화 방법으로도 충분히 가능하다. 하지만 부호화 과정에서 관심영역을 모른다면, 복호화 과정에서 관심영역 모양 정보를 받아서 관심영역 부호화를 한다. 이 처럼 복호화 과정에서 관심영역 모양을 정의한 후 부호화하는 방법을 동적 부호화라고 한다.

2.2 관심영역 부호화

관심영역 부호화 방법은 다양하게 있지만, 그 중에서 본 연구에 적용하는 Maxshift방법[6]은 양자화된 계수 중에 관심영역 계수와 배경 계수를 구분하여, 배경 계수 중에서 가장 큰 계수값, s 를 구한 다음, 관심영역 계수를 s 보다 높은 비트-평면에 이동시키는 방법으로서 JPEG2000 Part1 표준이다. 식(1)은 s 를 구하는 수식이고, 식(2)는 관심영역 처리 후의 계수, $a'(u, v)$ 를 구하는 수식이다.

$$s \geq \max(M_b) \tag{1}$$

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

식(1)에서 $\max(M_b)$ 은 각 서브밴드에서 양자화된 배경 계수 중에 가장 큰 값을 의미한다. 식(2)에서 $M(u, v)$ 는 관심영역 마스크 정보로서 계수가 관심영역에 속하는 좌표인 경우는 1, 배경에 속하는 좌표인 경우는 0을 의미한다. 장점은 부호화 시에 관심영역 모양 정보 대신 s 값만 가지게 되므로 부호화 효율이 좋다. 단점은 관심영역이 모두 복원될 때까지 배경을 얻을 수 없다. 즉 관심영역 중요도 조절이 불가능하다. 또한 다수 관심영역 지원이 어렵다.

2.3 관심영역 자동 추출

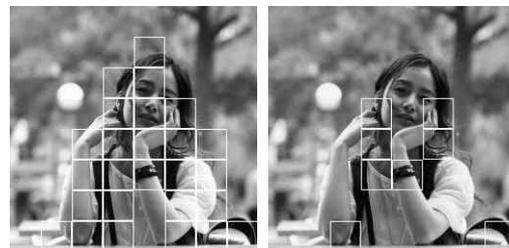
특정 부분을 관심영역 처리를 위해, 영상이 수천 장이라면 자동으로 관심영역을 추출하지 않으면 너무 많은 작업 시간이 필요하다. 뿐만 아니라, 특정 객체가 관심을 갖는 부분이라면, 처리할 모든 영상에 객체가 들어가 있을 것이다. 이것을 자동적으로 찾아내어 특별한 처리를 하는 것이 관심영역 자동 추출 및 관심영역 부호화이다. 기존에는 관심영역의 추출을 위하여 영상의 가장자리를 이용하여 일반적으로 추출하며 [5][6][8], 웨이블릿 변환을 이용하므로 주로 LH, HL, HH 밴드의 정보를 이용한다. [1]에서는 세 밴드의 계수를 OR하여 계수의 값이 특정 임계치를 초과하면 1로 설정하고 그렇지 않으면 0으로 설정하여 가장자리를 찾아낸다. 다음으로 잡음제거 과정을 거친다. 추출된 관심영역의 형태는 임의의 모양, 네모, 원을 가질 수 있고, 한 개 이상의 관심영역을 가지는 기법도 있다. 일반적으로 관심영역은 영상의 가장자리가 아닌 중앙에 위치할 가능성이 크다. 그림 2는 child3 영상

을 1차 웨이블릿 변환한 것이다. LL 밴드를 제외한 나머지 3 밴드에는 영상의 가장자리들이 분포되어 있음을 볼 수 있다.



(그림 2) 1차 웨이블릿 변환

가장자리의 분포에 의존하는 방법은 인간의 시각 인식을 전혀 고려하지 않은 처리 방법이다. 즉, 가장자리의 분포가 골고루 퍼져 있다거나, 가장자리가 잘 나타나지 않는 영상에서는 결코 좋은 관심영역을 얻지 못하는 단점이 있다. 그래서 본 연구에서는 인간의 영상 인식을 고려하여 객체를 찾아내어 그 객체를 관심영역으로 간주하고, 관심영역 부호화한다. 그림 3에서는 피사체 심도가 낮은 영상에서 관심영역 자동 추출 후의 결과[5]를 보이고 있다. (a) 영상은 1.2bpp로 압축하여 추출한 것이며, (b)는 0.2bpp 압축한 후에 얻은 관심영역을 표시하고 있다. 그림에서와 같이 인간의 눈으로는 같은 그림처럼 보이지만 압축을 많이 하면 가장자리도 더불어 사라지므로 정확한 관심영역을 구하기는 어렵다. [6]에서도 가장자리의 분포에 의해 관심영역을 찾는 데, [5]와의 다른 점은 한 블록 내에서 가장자리의 분포가 어떤 임계치를 초과하면 그 블록은 관심영역 블록으로 간주한다. [5]에서는 상하좌우 4방향에서 안쪽으로 탐색하여 고주파 성분의 가장자리가 나타나면 관심영역 블록으로 간주한다.



(a) 1.2bpp (b) 0.2bpp

(그림 3) [5] 기법의 관심영역 추출

2.4 관심영역 마스크와 우선처리

관심영역 마스크 테이블은 영상의 해상도와 같은 크기로 할당되며 각 화소가 관심영역에 속하는 화소이면 관심영역 마스크에 있는 그 값도 1이 되며, 배경이면 0으로 설정된다. 관심영역 블록 내에 관심영역 화소와 배경 화소가 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, 관심영역은 손실 감소와 관심영역이 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 관심영역 가중치이고, $D_j^{n_i}$ 는 n_i 에서의 가중 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.

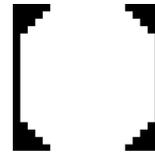
3. 얼굴 추출 후 관심영역 부호화

본 장에서는 영상에서 자동으로 의미 있는 관심 객체를 추출하고 압축 또는 저장하는 알고리즘을 제안하고 관심영역 부호화에서 우선 처리

를 기술한다.

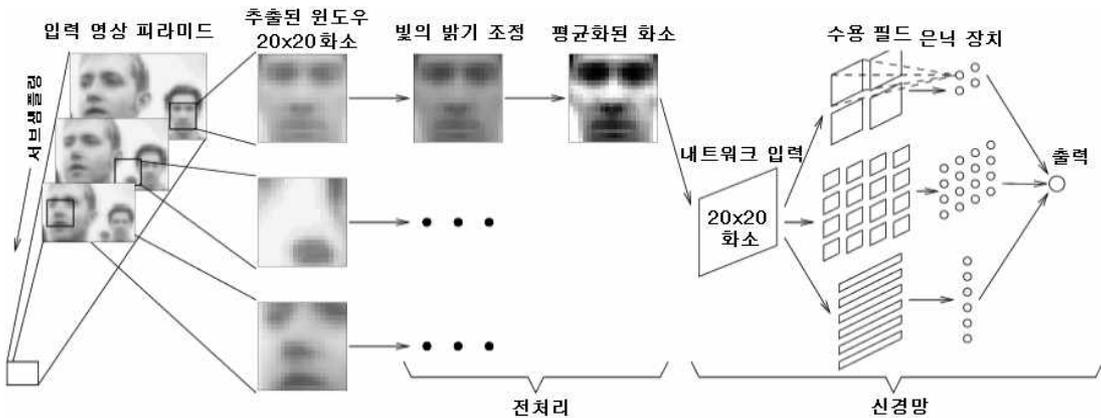
3.1 제안한 얼굴 추출 방법

추출된 얼굴 영역을 관심영역으로 간주하고, 얼굴은 정면 얼굴로 한정한다. 또한 많이 기울어진 얼굴은 추출 안 될 수도 있다. 알고리즘은 크게 두 부분으로 나뉜다. 영상에 여러 개의 신경망 기반의 필터를 적용하는 첫 번째 단계와 첫 번째 단계의 결과를 조합하기 위해 조정자를 사용하는 것이 두 번째 단계이다[9]-[11]. 먼저 20x20 화소 영역을 갖는 필터를 실행한 결과는 -1 ~ 1 사이의 값을 갖는다. 1은 100% 얼굴을 갖는다는 의미이며 -1은 100% 얼굴을 갖지 않는다는 의미이다. 그림 4는 20x20 화소 영역의 필터를 나타낸다.



(그림 4) 20X20 얼굴 마스크

얼굴 영상이 마스크보다 크면, 영상을 단계적으로 줄이면서 정합을 하고, 영상이 마스크보다 작다면 그 얼굴은 찾지 못한다. 정합은 한 영상 내의 모든 화소에 대해 실행하고, 영상의 스케일을 줄이는 것은 1.2배씩 감소시킨다. 얼굴 탐지 알고리즘은 그림 5에 나타내었다. 먼저 전처리 과정으로써, 영상의 윈도우에 적용된다. 윈도우는 그 윈도우 내에 얼굴이 포함되어 있는지를 알기 위해 신경망을 통과한다. 선형함수는 윈도우



(그림 5) 얼굴 탐지에 사용되는 기본 알고리즘

우 내의 강도(intensity value)에 맞춘 후, 어떤 극단적인 밝기 상태에서는 수정을 하고, 차이를 구한다.

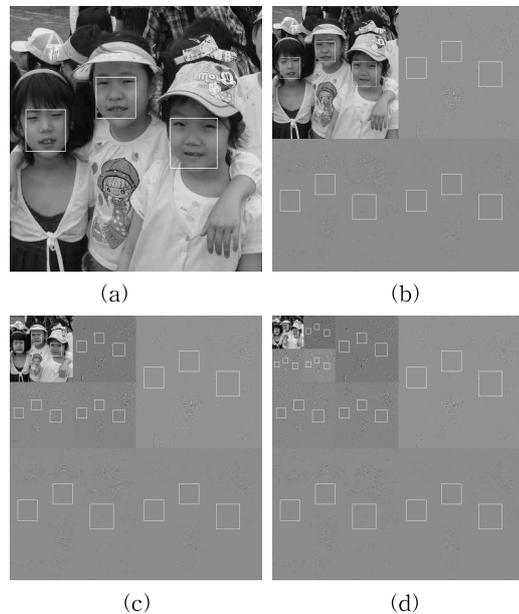
다음으로 색의 대비를 개선하고 다양한 카메라 계인을 수정하기 위해 히스토그램 평균화가 적용된다. 이 각각의 단계에 대해서 그림 4의 타원 마스크 내의 화소에 기반을 두어 정합이 계산되고 이후에 전체 윈도우에 적용된 후, 전 처리된 윈도우가 신경망을 통과하는 것이다. 신경망으로 20x20 화소가 입력되면 수용영역을 거처 은닉 단위를 통과하여 결과를 출력한다. 은닉 단위는 3 가지로 구성되는데, 10x10 화소 부영역을 찾는 4 은닉 단위, 5x5 화소 부영역을 찾는 16 은닉 단위, 겹치는 20x5 화소 수평 무늬 화소를 찾는 6 은닉 단위로 구성된다. 이 모든 은닉 단위는 얼굴 추출에 중요한 지역적 특징을 찾는데 사용된다. 특히, 수평 무늬는 입이나 눈, 코, 입의 끝 등을 찾는 데 사용된다.

3.2 웨이블릿 기반의 얼굴 추출

전체 영상에서 얼굴을 찾을 때는 영상 도메인이 크기 때문에 탐색 시간이 많이 걸리지만, 얼굴을 정확하게 찾을 수 있다. 하지만, 웨이블릿 변환 후에 추출을 하면 영상이 25%로 줄어들기 때문에 일부 얼굴을 찾지 못할 수도 있다. 20x20 마스크보다 더 얼굴 영역이 줄어들 경우에는 추출이 불가능하다. 그림 6은 원 영상과 웨이블릿 분해 후의 얼굴 영역 추출 결과를 보이고 있다. (a)는 원 영상에서, (b)는 1차 웨이블릿 변환 후, (c)는 2차 웨이블릿 변환 후, (d)는 3차 웨이블릿 변환 후이다. 샘플 영상은 768x768 크기이며, 3차 웨이블릿 변환 후의 LL 밴드의 크기는 96x96이 된다. 그림 상의 흰색 네모로 된 영역은 관심영역의 프리시트이다. 압축을 위해서는 LL 영역 내의 관심영역만 압축하는 것이 아니라, 나머지 고주파 밴드 내의 프리시트, 즉, 영상 도메인 내의 같은 위치에 있는 계수까지 압축해야 한다.

관심영역 추출 후의 밴드별 위치를 모두 압축하기 위해 각 밴드 별로 마스크가 설정되어야 하며, 이 마스크는 영상 도메인에서의 마스크가 웨이블릿 도메인에서의 마스크로 정의되어야 한다. 웨이블릿 도메인에서의 마스크 정보는 영상 도메인에서의 관심영역 마스크 정보와 역웨이블릿

릿 변환을 이용하면 구할 수 있다. 즉, 웨이블릿 도메인에서의 관심영역 마스크 생성 과정은 다음 두 단계로 이루어진다. 먼저 영상 도메인에서의 관심영역 마스크를 추적하고, 다음으로 마지막 역웨이블릿 변환에 의해 2개의 서브밴드 안에 어떤 위치의 계수가 관심영역 계수로 재구성하기 위해 사용되었는지를 추적하게 된다. 이렇게 하여 각 분해레벨에서 모든 서브밴드 내에 어떤 위치의 계수가 관심영역 계수로 재구성하기 위해 사용되었는지를 추적하게 된다. 웨이블릿 도메인에서의 관심영역 마스크는 각 분해 레벨에서 서브밴드 안에 관심영역 코드블록, 배경블록, 부분 배경 계수를 포함하는 관심영역 블록으로 구성된다.



(그림 6) 웨이블릿 변환 분해 레벨 별 관심영역 위치

3.3 관심영역 코드블록의 우선처리

본 연구에서 주어진 관심영역 영역은 모두 네모로 구성되고 코드블록은, 모든 화소가 관심영역인 코드블록, 일부 화소만 관심영역인 코드블록, 모든 화소가 배경인 코드블록으로 나뉜다. 관심영역은 코드블록 단위로 하나의 영상을 나누고 각 코드블록에 해당하는 관심영역 마스크를 EBCOT에 입력시킨다. 관심영역 블록에 해당하는 꼭짓점 (x1, y1), (x2, y2)가 구하여 졌으므로 이 공간 내에 속하면 관심영역으로 지정하고

속하지 않으면 배경으로 지정한다. EBCOT에서는 각 서브밴드를 코드블록으로 나누어서 코드블록에 대해 독립적으로 부호화된 임베디드 비트스트림으로부터 최종 비트스트림이 만들어진다. 블록 전체가 관심영역인 경우는 우선권 조절이 필요 없이 그냥 보내면 되지만, 일부가 관심영역인 경우는 배경 영역을 k 만큼 절단하여 우선권 조절을 한다.

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

일부가 관심영역인 코드블록의 우선권 조절은 식(4)에 의해 얻어진다. 매개 변수 k 는 임의로 지정 가능하며 본 연구에서는 Maxshift 방법을 사용하였으므로 k 는 모든 배경 영역을 관심영역 영역 보다 낮은 우선순위를 갖도록 k 를 결정한다.

4. 실험 및 평가

실험을 위해 여러 얼굴이 포함된 영상을 이용하여 얼굴 검출 후 관심영역을 지정하고 관심영역이 포함된 영상을 압축 후의 PSNR을 여러 기법과 비교한다.



(a) child3

(b) child12

(그림 7) 실험에 사용된 영상

표 1은 실험에 사용된 영상의 특징을 보이고 있다. 비트율은 원 영상을 압축할 때의 비트율을 의미한다. 비트율이 낮으면 원 영상에 손상이 많이 가해지므로 얼굴 영역을 추출하는 것이 불가능할 수도 있다. 뿐만 아니라 분해 레벨이 클수록 원 영상을 축소된 상태에서 추출을 하므로 얼굴 추출이 더 어려워진다.

4.1 실험 방법

실험을 위해 웨이블릿 변환 시 사용된 필터는 D-9/7필터를 사용하였으며, 실험은 Maxshift 방법을 이용하며, PSNR을 측정한다. 사용된 영상은 자연 영상이며, 그림 7과 같고, 해상도는 모두 768×768이다. 관심영역의 모양은 네모이며, 분해 레벨(l)은 1~2 레벨로 실험하였다. 실험에 사용된 모든 영상은 흑백 영상이고, child3 영상은 2 분해 레벨, child12 영상은 1 분해 레벨에서 얼굴 추출 후 관심영역을 결정하였다. 관심영역을 추출하는 데 LL 밴드 계수만 사용된다.

<표 1> 실험에 사용된 각 영상의 특성

영상 비트율(bpp)	child3	child12
0.1	2 분해 레벨	1 분해 레벨
0.2	2 분해 레벨	1 분해 레벨
0.4	2 분해 레벨	1 분해 레벨
0.8	2 분해 레벨	1 분해 레벨

4.2 웨이블릿 변환 후의 얼굴 추출

실험에 사용된 영상에서 얼굴 추출이 웨이블릿 변환을 한 후에도 잘 되는 지를 실험을 통해 보인다. 사용된 2 개의 영상을 압축률에 따라, 분해 레벨에 따라 어떻게 관심영역이 결정되는 지를 그림 8에 나타내었다. child3에서 압축률에 따라 얼굴 추출이 완벽하지 않음을 볼 수 있다. 이는 압축률이 높아 원 영상에 손상이 많이 가해졌기 때문이다. 실제 눈으로 구분하기는 힘들지만 (a)의 영상이 (b)의 영상보다 약간 화질이 떨어져서 하나의 얼굴 밖에 찾지 못했다. 반면에 0.4bpp 이상으로 압축을 하면 모두 잘 추출되었다. child12는 1 분해 레벨에서 관심영역을 추출하였는데, 0.1bpp 환경에서는 몇 개의 얼굴을 검출하지 못하였다. 화질이 나쁜 탓도 있지만, 얼굴이 안경이나, 머리카락으로 가려져 있기 때문이다.

본 실험 영상으로 기존의 연구 방법인 [5]와 [6]의 기법으로 추출한 관심영역 영역은 그림 9와 같이 나타난다. 이 방법은 가장자리의 분포에 의해 관심영역을 추출하므로 고주파 성분이 관심영역 주위에 충분히 분포되어 있어야 하는데, 본 실험 영상은 그렇지 않기 때문이다. 제안된

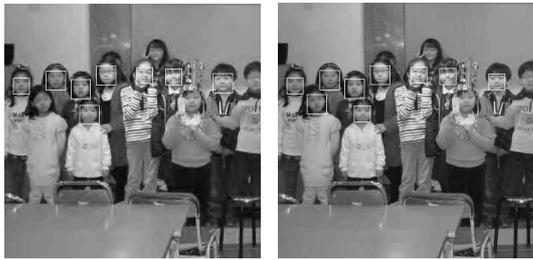
방법은 주파수 성분의 분포에 의해 관심영역을 찾는 것이 아니라 인간의 인지 능력에 의해 찾는 것이므로 그림 8과 같이 훌륭한 결과를 얻을 수 있었다.



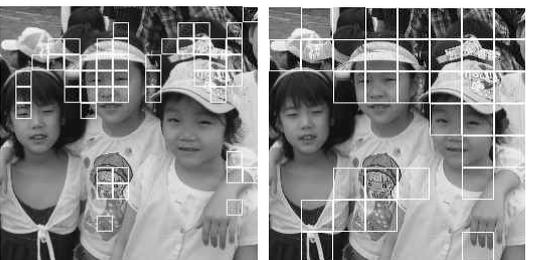
(a) 2분해 레벨, 192x192, 0.1bpp (b) 2분해 레벨, 192x192, 0.2bpp



(c) 2분해 레벨, 192x192, 0.4bpp (d) 2분해 레벨, 192x192, 0.8bpp



(e) 1분해 레벨, 384x384, 0.1bpp (f) 1분해 레벨, 384x384, 0.4bpp
(그림 8) 각 영상의 특성에 따른 관심영역 위치



(그림 9) [5]와 [6]의 방법에 의한 관심영역 위치
(2 분해 레벨, 192x192 0.8bpp)

4.3 객관적 화질 평가

관심영역 부호화된 영상과 관심영역 부호화가

안 된 영상을 비교하여 전체 화질의 차이가 있는지를 보이고자 한다. 객관적인 화질 평가를 위해 PSNR을 사용하고, 평가는 식(5)와 같고, 샘플 당 n 비트의 길이를 가진다.

$$PSNR = 10\log_{10}((2^n - 1)^2 / (\frac{\sum(\hat{x} - x)^2}{A})) \quad (5)$$

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

<표 2> 각 영상의 비트율에 따른 PSNR

영상		child3	child12
비트율(bpp)		(2레벨 분해)	(1레벨 분해)
0.1	관심영역 비포함	27.678	27.345
	관심영역 포함	27.233	26.990
0.2	관심영역 비포함	29.230	29.442
	관심영역 포함	28.901	29.129
0.4	관심영역 비포함	33.833	33.459
	관심영역 포함	33.457	33.122
0.8	관심영역 비포함	39.234	39.001
	관심영역 포함	38.988	38.776

표 2에서 보이는 것처럼, 관심영역을 포함한 것이나 하지 않는 것이나 화질은 거의 차이가 나지 않는다. 다만 분해를 많이 하면 관심영역을 찾지 못할 수도 있으므로 서비스 질이 떨어질 수도 있다. 화질이 거의 차이를 보이지 않는 것은 전체 영상 중에서 관심영역 영역이 차지하는 부분이 적기 때문이다. 만약 관심영역이 커다면 배경 영역에 많은 손상을 가해야 하므로 PSNR 화질은 떨어지게 된다. 관심영역을 포함한 영상은 전체적으로 화질은 상대적으로 조금 떨어지지만, Maxshift를 사용하여 관심영역 부호화했기 때문에 관심영역 부분은 최고의 화질을 제공한다. 관심영역이 포함되었을 때와 포함되지 않았을 때의 화질 차이는 표 2에서 보이는 것처럼 아주 미미하다. 여기서, 관심영역이 사용자의 의도에 따라 잘 추출된다면 훌륭한 서비스를 제공할 수 있음을 보이고 있다.

전체 영상이 완전히 전송되었을 때는 그림 8과 같이 보이지만, 완전히 전송되지 않고 관심영역과 배경의 일부만 전송되었을 때는 그림 9와 같이 보인다. 주관적 화질 평가를 위해서 그림 10과 같이 관심영역 우선 전송을 했을 때, 화면

에 보이는 영상을 나타내었다. girl3의 영상을 2분해 레벨 0.4bpp에서 압축하여 전송 중에 보이는 일련의 영상을 나타낸 것이다. 2분해를 했으므로 관심영역은 영상이 192x192 크기에서 구했다. 보이는 순서대로 첫 번째는 관심영역 블록만 나타난 것이며, 두 번째는 관심영역 블록 전부와 배경의 첫 비트만 전송된 후, 이 후는 관심영역 블록과 각각 배경의 두 번째, 세 번째, 네 번째, 다섯 번째 비트까지 전송된 후를 보이고 있다.



(그림 10) girl3의 0.4bpp, 2레벨 분해 후 관심영역 부호화된 영상의 사용자 화면

5. 결론 및 향후과제

제안한 방법은 자동으로 인간의 관심영역을 추출하고 그것을 압축 저장 전송하여 사용자에게 보다 나은 서비스를 제공하는 것이다. 얼굴 추출은 압축을 0.1bpp 또는 0.2bpp로 하거나, 웨이블릿 변환을 3 레벨 분해하면 충분한 결과를 얻을 수 없었다. 객관적인 화질 면에서 제안된 관심영역 부호화 방법을 적용한 것과 관심영역 부호화를 적용하지 않은 것의 차이는 거의 미미

했다. 하지만, 사용자 입장에서 우선처리 서비스를 제공받을 수 있기 때문에 충분히 유용한 기법이라 할 수 있다. 본 연구에서는 앞면만 추출할 수 있는 데, 향후에는 옆면도 추출할 수 있는 관심영역 부호화를 연구할 것이며, 나아가서 다양한 객체를 추출할 수 있는 연구를 할 것이다.

참 고 문 헌

- [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, 2007.11.
- [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, 2005.10.
- [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, 2002.6.
- [4] 박재홍, 서영건, 김상복, 강기준, 김호용, "JPEG2000에서 ROI의 자동 추출과 우선적 처리", 한국컴퓨터정보학회 논문지, 제13권, 5호, pp. 127-136, 2008년 11월.
- [5] 박순화 이부권, 강기준, 김호용, 김형준, 김상복, "피사계 심도가 낮은 이미지에서 웨이블릿 기반의 자동 ROI 추출 및 마스크 생성", 한국컴퓨터정보학회 논문지, 제14권, 3호, pp. 93-101, 2009년 3월.
- [6] 서준성, "JPEG2000 이미지에서 코드블록 내에서 예지 분포를 이용한 자동 관심 영역 추출", 경상대학교 교육대학원 석사학위 논문, 2010. 2.
- [7] 이부권, 서영건, 박순화, 김호용, 김형준, "실시간 원격 강의에서 영상 인코딩 기법을 적용하기 위한 요소", 한국디지털콘텐츠학회 논문지, 제9권, 4호, pp. 687-695, 2008년 12월.
- [8] 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, 2001.1.
- [9] 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, 2006.9.

- [10] H. A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection", IEEE Transactions on PAMI, Vol. 20, No. 1, pp. 1-27, 1998.1.
- [11] 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, 2002.1.



서 영 건

1987년 : 경상대학교 전산과 (이학사)
1997년 : 숭실대학교 전산과 (공학박사)
1989년~1992년 삼보컴퓨터
1997년~현재 경상대학교 컴퓨터과
학과 교수

2001년~현재 : 경상대학교 컴퓨터정보통신연구소원
관심분야 : 의료영상처리, 영상분할, 웨이블릿기반영상처리



홍 도 순

1988년 : 경상대학교 전산과(이학사)
1993년 : 경상대학교 전산과(이학석사)
1993년~현재 충청여자고등학교 교사

2006년~현재 : 경상대학교 전산과 박사 수료
관심분야 : MPEG, JPEG2000, 웨이블릿기반영상처리



박 재 홍

1978년 : 충북대 수학교육과 (이학사)
1989년 : 중앙대학교 전산과 (공학박사)

1983년~현재 : 경상대학교 컴퓨터과학과 교수
관심분야 : 소프트웨어 공학, 소프트웨어 신뢰성