

모바일 아바타 전송을 위한 얼굴 영역 검출 및 압축에 관한 연구

최재영^{*}, 황승호^{**}, 황보택근^{***}, 양영규^{****}

요 약

기존의 음성 통신 도구로 사용되던 휴대 전화는 최근 들어 데이터 통신 기기로서의 그 역할이 증대되고 있다. 그러나 늘어난 대역폭으로 인하여 멀티미디어 콘텐츠를 전송할 만한 충분한 능력이 되었지만 여전히 여러 콘텐츠들을 이용하기에는 이용자의 통신 및 정보 이용료의 부담이 크다. 이에 본 연구에서는 모바일 아바타 생성을 위한 얼굴 영상의 전송이라는 특수한 상황을 적용하여 일반적인 압축방법을 통하여 얻을 수 있는 압축률보다 높은 성능의 압축 방법을 제안함으로써 콘텐츠 이용자의 정보 이용 부담 및 통신 트래픽을 줄이고자 한다. 제안한 방법은 웨이블릿으로 분해된 영상의 대역간 공간적 상관관계뿐만 아니라 얼굴 영역 검출을 통하여 얻은 웨이블릿 영상의 중요계수 위치를 참조함으로써 모바일 아바타 생성에 적합한 주요영역별 압축율을 자동화 하고자 하였다. 실험의 결과를 통하여 기존의 EZW 압축 방법보다 입력 영상의 중요계수 위치 정보를 사전에 파악하여 차등 압축하는 방법이 모바일 아바타 생성을 위한 시스템에 효율적임을 확인하였다.

A Study on Facial Region Detection and Compression for Transmission of Mobile Avatar

Jae-Young Choi^{*}, Seung-Ho Hwang^{**}, Taeg-Keun Whangbo^{***},
Young-Kyu Yang^{****}

ABSTRACT

In recent times, mobile phones that play an important role in voice communication have become increasingly exciting due to the ability to transmit multimedia contents. However, user's communication cost and service usage fees for digital context exchange are still prohibitive. In this paper, we consider the task of transmitting a facial image in the context of mobile avatar generation. We present image compression as a useful technique to reduce service usage costs of data exchange in mobile communications. In view of this we propose an image compression algorithm that performs better than existing techniques. The suggested algorithm in this paper uses the location of significant coefficient on the wavelet image by detecting facial region as well as the spatial correlation between decomposed bands by wavelet transformation. The algorithm perform well and archives an efficient compression ratio. The proposed method recognizes location information of significant coefficient in advance, thus it is more suitable than existing general compression methods for mobile avatar generation systems.

Key words: Face Detection(얼굴 검출), Avatar(아바타), EZW(임베디드 제로트리 웨이블릿), Wavelet(웨이블릿), Compression(압축)

※ 교신저자(Corresponding Author): 황보택근, 주소: 경기도 성남시 수정구 복정동 산65번지(461-701), 전화: 031) 750-5417, FAX: 031)757-9508

E-mail: tkwhangbo@kyungwon.ac.kr

접수일: 2004년 10월 28일, 완료일: 2005년 2월 7일

^{*} 준회원, UCLA Post-Doc Researcher

(E-mail: jy1001@kyungwon.ac.kr)

^{**} 준회원, 경원대학교 소프트웨어 연구소 연구원
(hsh1052@kyungwon.ac.kr)

^{***} 종신회원, 경원대학교 소프트웨어 대학 부교수
(tkwhangbo@kyungwon.ac.kr)

^{****} 경원대학교 소프트웨어 대학 교수

※ 이 논문은 2004년도 대학 IT 연구센터 육성지원사업비의 지원을 받아 연구되었음.

1. 서론

오늘날 모바일의 폭 넓은 사용과 함께 이를 지원하기 위한 모바일 디지털 콘텐츠들에 대한 연구가 급속히 증가 되어지고 있다. 특히, 개인의 익명성을 보장하면서도 개성과 느낌을 표현해 줄 수 있는 모바일 상에서의 아바타나 캐리커처 시스템에 관한 연구가 활발히 진행되고 있으며, 이에 필요한 기반기술이 필수적으로 병행되어 지고 있다. 하지만 아직까지 휴대폰의 성능 제약으로 인하여 휴대장치에서 모든 과정을 처리하는 것은 불가능하며 복잡한 연산은 서버로 전송하여 처리하는 것이 효율적이다. 이에 본 논문에서는 단말기와 서버 양단간의 통신 정보량을 줄임으로써 사용자의 정보 이용부담을 줄이고자 한다. 본 논문의 구성은 2장에서 효율적인 아바타 압축을 위한 얼굴 영역 검출 및 본 연구의 영상 변환으로 사용할 웨이블릿 변환에 대하여 소개하고, 이를 개선한 압축 전송 방법을 제안한다. 3장에서는 제안한 알고리즘의 성능을 평가하고 이를 기존의 압축 알고리즘과 비교 분석한 후 4장에서 결론을 맺는다.

2. 얼굴 중요 영역 설정 및 압축

일반적으로 아바타는 그 사용자를 대표하는 신체 전체가 될 수도 있겠으나, 디스플레이 크기가 작은 핸드폰을 일차적인 어플리케이션으로 고려해볼 때 주로 상반신이나 얼굴을 사용한 아이템이 될 수 있겠으며, 이에 본 연구에서는 중요 영역을 얼굴에 제한하였다. 또한 모바일의 특성상 적은 패킷과 과금을 통한 데이터의 전송이 이루어져야 하며, 이를 위해 적은 비트율을 통한 높은 품질의 영상을 만족시키고자 한다면 결국 중요영상에 우선순위의 품질의 차등을 두어 전송하는 방법이 해결책이 될 수 있을 것이다. 이에 본장에서는 이러한 모바일 아바타 전송의 문제점을 효율적으로 극복하고자 차등 압축전송을 위한 중요영역 설정 및 압축방법을 제시하고자 한다.

2.1 TSL 색상 공간을 이용한 얼굴 영역 탐색

본 논문에서는 얼굴 영역 검출을 위하여 일반적인 방법들이 많이 적용하는 변환 색상모델을 사용하는데, 특히 J. C. Terrillon이 얼굴색 검출용으로 좋은 성능임을 증명한 TSL 색상 공간을 본 논문의 얼굴색

검출 모델로 사용하였다. TSL 색상 공간 중 T-S 공간은 조명과 독립적으로 하나의 군집을 이루고 있고, 그 분포가 밀집되어 있어 사람의 얼굴색 추출을 위한 방법으로 적합하다[1].

그림 1은 얼굴 영역 추출을 위한 본 연구의 흐름도이다. 본 연구에서처럼 RGB 색상 공간을 선형이나 비선형의 적절한 2차원 공간으로 변환하는 것은 공간적 감소 및 연산의 효율성을 위하여 매우 중요한 일이며, 잡음을 배제한 정확한 얼굴 영역을 획득하기 위하여 거친(coarse) 단계와 상세(fine) 단계의 2단계를 사용함으로써 첫 단계에서는 잡음에 강한 얼굴의 대략 영역을 확보할 수 있었고, 두 번째 단계에서는 검출 색상의 범위를 확장 시켜 1단계에서 미처 추출하지 못한 얼굴부위를 검출할 수 있었다.

입력 영상은 핸드폰 카메라에 의해서 얻어진 영상으로서 얼굴 아바타의 특징상 얼굴이 중앙에 위치하는 영상만으로 제한하여 실험하였다. 먼저 입력영상이 들어오면 핸드폰 연산의 한계성을 고려해 입력 영상들은 다운 샘플링을 하여 사용함으로써 연산량을 감소시키고, 화질의 열화와 조명의 영향으로 생기는 국부적 잡음을 제거하여 미 검출되는 영역을 최소화 하였다.

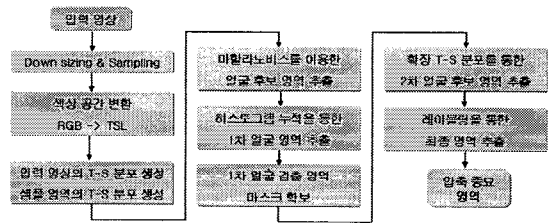


그림 1. 얼굴 영역 추출

다운 샘플링은 세로를 일정크기로 고정하고 가로는 입력영상의 가로 : 세로 비율에 의해 자동으로 설정되게 하였다. 이러한 입력 영상은 RGB 색상으로 입력되므로 식 (1)을 이용하여 TSL 색상 공간으로 변환하였다.

$$T = \begin{cases} \arctan(r'/g')/2\pi + 1/4, & g' > 0 \\ \arctan(r'/g')/2\pi + 3/4, & g' < 0 \\ 0, & g' = 0 \end{cases}, \quad (1)$$

$$S = [9/5(r'^2 + g'^2)]^{1/2},$$

$$L = 0.299R + 0.587G + 0.114B$$

$$\text{where } r' = (r-1/3), g' = (g-1/3), \\ r = R/(R+G+B), \text{ and } g = G/(R+G+B)$$

식 1에서 T는 색조(tint), S는 선명도(saturation), 그리고 L은 휘도(luminance)를 나타내며, R,G,B는 각각 입력영상의 컬러 채널값을 나타낸다. 본 연구에서 T-S 정보는 얼굴색 검출에 이용하며, 필요에 따라서는 L정보를 이용하여 조명을 분석하고 보정할 수 있다. 또한, 얼굴 색상 공간이 T-S 색상 공간에서 타원형의 구조를 가지고 있는 특징을 이용하여 얼굴 중심 영역의 T-S값을 근간으로 식 (2)와 같은 마할라노비스 거리에 의한 얼굴 영역 화소 추출 방법을 사용하였다.

$$[\lambda_K(i,j)]^2 = [X(i,j) - m_K]^T C_K^{-1} [X(i,j) - m_K] \quad (2)$$

$X(i,j)$ 를 좌표 i, j 에서의 $[T(i,j) S(i,j)]^T$ 벡터로 표현 가능하다고 할 때, 얼굴색 공간 K 에 대하여 C_K 를 얼굴색 공간 K 의 공분산이라 하고, m_K 를 $[m_{T_K} m_{S_K}]^T$ 로 표현되는 평균 벡터라 가정하면, 마할라노비스 거리 $\lambda_K(i,j)$ 는 식 (2)와 같다. 여기서 중심은 m_K 가 되며, 중심축은 공분산 C_K 에 의하여 결정된다[2].

그림 2는 입력 영상의 중심 영역에서 해당 영상의 얼굴색을 샘플링한 후 추출한 색상의 분포를 근간으로 얼굴 화소에 해당하는 영역을 추정하는 과정으로써 그림 2의 (b)는 전체 입력 영상의 T-S 분포 상에서 얼굴에 해당하는 화소값 및 식 (2)에 구해진 마할라노비스 거리분포를 통한 얼굴색 분포를 표시한 것

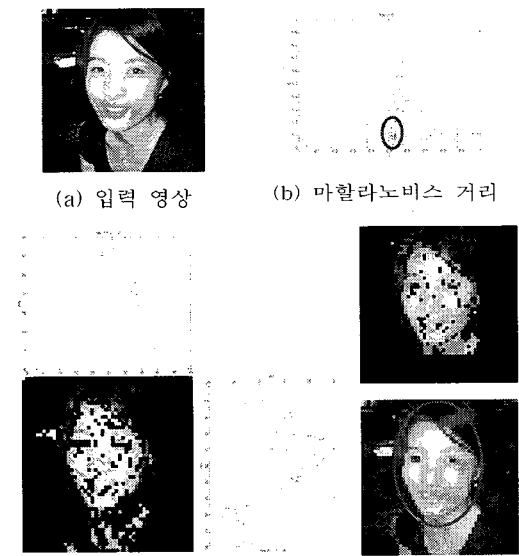
이며, 그림 2의 (c)는 이렇게 추출된 얼굴 영역을 가로, 세로 누적 히스토그램을 분석하여 임계값보다 작은 영역이라면 얼굴 후보 영역에서 제거한다.

하지만, 1차 분석에 의한 영상은 그림 2의 (c)와 같이 얼굴 영역 내의 많은 부분들이 추출되지 못할 경우도 있고, 우연한 오류로 인하여 얼굴이 아닌 잡음들이 가로 세로 누적 히스토그램에 포함되어 함께 검출되는 경우도 있다. 따라서 이러한 잡음 제거를 위한 추가적인 알고리즘이 필요하며, 본 연구에서는 레이블링을 이용하여 얼굴 후보 영역을 추출하였다. 이때, 얼굴내부의 미 추출된 요소에 의해서 레이블링 시 얼굴 영역이 분할되는 문제를 보완하기 위하여 얼굴 내부의 요소들을 이용해 마할라노비스 거리를 1차 추출보다 20% 추가하여 확장하는 방법을 그림 2의 (d)와 같이 사용하였다. 20%의 확장 적용은 확보된 많은 실험 영상들의 다양한 확장 비율을 실험한 결과 약 20%의 추가비율을 적용할 때 가장 효과적인 얼굴색 추출이 이루어졌으며, 그보다 적을 때에는 여전히 미검출 얼굴영역이 발생하였으며, 그보다 많을 경우에는 얼굴이외의 추가 잡음 영역이 발생하였다. 따라서, 다양한 실험치를 적용해 본 결과 고정치 20%를 사용하여도 실험 결과에는 영향을 미치지 않으며, 고정계수 사용으로 인하여 연산소요시간의 절약을 가져올 수 있었다. 이러한 연산을 한 후 얼굴 후보 영역을 레이블링 연산하면 보다 정확한 얼굴 후보 영역의 추출이 가능하다[6].

2.2 웨이블릿 변환

휴대폰 카메라로 입력되어진 영상이 인덱스 칼라로 가정하더라도 만약 크기가 352×288인 영상을 처리하고자 할 때, 약 800k 비트의 정보량이 필요하며 이를 그대로 통신에 적용하여 전송한다면 이 모든 비용은 고스란히 이용자의 몫으로 돌아가게 되며 전송 시간 및 처리시간에 있어서도 매우 큰 손실이 아닐 수 없다. 이렇듯 영상 처리에 있어서 압축은 필수 불가결한 상태이며 아무리 채널의 대역폭이 커지고 저장 매체의 양이 늘어나도 그에 비례하여 원하는 데이터의 양은 증가하기 때문에 압축은 필요하게 된다.

웨이블릿을 이용한 영상 압축은 기존에 가장 널리 이용되고 있는 정지 영상의 압축 방식인 JPEG의 블록화(blocking) 현상을 극복하였으며, 점진적인 전송



(a) 입력 영상 (b) 마할라노비스 거리 (c) 얼굴 후보 영역의 누적히스토그램 (d) 얼굴영역 추출

그림 2. 얼굴 후보 영역 추출

방식(progressive transmission)을 사용할 수가 있어서 사람의 심미적 만족도를 더욱 충족시킬 수가 있다.

그림 3은 2차원 영상의 웨이블릿 변환을 나타낸 것으로서, 본 연구에서 이용하는 방식은 옥타브 밴드 분해 방식이다. 즉 계속해서 저역 부분을 분해해 가는 것으로서 그림 4에서 그 결과를 보여주고 있다.

웨이블릿 변환을 하게 되면 각 부대역은 수평, 수직, 대각선 방향으로 상관성이 높다. 즉 비 중요 계수로 판정된 주위의 계수들은 역시 비 중요 계수가 많다고 추측할 수 있으며, 이러한 특징을 EZW (Embedded Zerotree Wavelet)에 적용한 것이 오늘날 많이 쓰이고 있다[3].

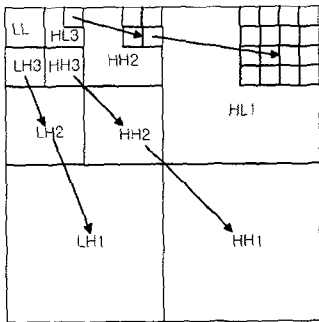


그림 3. 웨이블릿에서 부대역 간의 위치적 연관성



(a) R 채널 (b) G 채널 (c) B 채널

그림 4. RGB 채널별 웨이블릿 변환 결과

2.3 제안하는 EZW 방식

영상 압축이라 하는 것은, 영상 데이터의 중복성을 제거하는 과정으로써 공간적, 시간적, 대역간, 그리고 통계적 중복성을 고려하게 된다.

본 논문에서 제안하는 압축방식은 그림 5와 같이 기존의 EZW방식에서 사용되는 임계값을 공간적 중요도, 즉 세밀하게 표현하고자 하는 얼굴영역의 여부에 따라 두 가지 형태로 분리하여 적용하는 방법을 제안하였다. 또한, 웨이블릿 변환에서 대역간 상관관계를 분석함으로써 얼굴 영상의 특성에 맞는 수평대

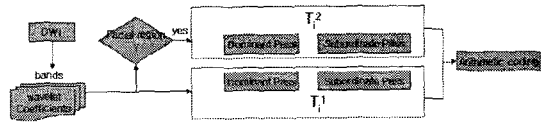


그림 5. 제안한 EZW 압축 흐름도

역의 중요성을 강조하여 계수 탐색의 우선순위를 두었다. 최종적으로 통계적 중복성은 추출한 데이터에 내재한 상관성을 이용한 산술연산 부호화 방법을 사용하였다[4]. 시간적 중복성은 일반적으로 동영상에서 사용하는 것으로써 본 논문에서는 적용하지 않았다.

먼저, 그림 5는 제안한 EZW 압축 알고리즘의 흐름도로서 입력 영상은 그림 3 및 4와 같이 다해상도의 웨이블릿 변환을 거쳐 여러 개의 부대역으로 분해되며, 분해된 각 대역의 계수값은 두개의 채널로 분리되어 양자화 과정을 거치게 된다.

첫 번째 채널은 임계값 T_i^1 를 기준으로 하여 각 대역의 모든 변환 계수들의 크기가 설정된 임계값 이상이면 유효계수(significant coefficient)로, 임계값 이하이면 무효계수(insignificant coefficient)로 분류되며, 유효계수에 대해서만 양자화가 적용된다. 두 번째 채널은 본 연구가 얼굴 영상의 압축이라는 조건을 전제로 각 대역상에서 얼굴의 위치적 공간관계를 고려하여 2.1절에서 획득한 얼굴 공간에 해당하는 변환계수는 임계값 T_i^2 를 적용하여 양자화 과정을 거치게 된다.

일반적으로 양자화는 시각에 미치는 영향에 따라 가중치를 두어 적용하며, EZW의 경우는 내장형 부호화를 위하여 양자화 과정에서 연속근사 양자화를 사용한다. 즉, 고정된 양자화 간격으로 한번만 양자화 하는 것이 아니라 간격을 좁혀가면서 연속적으로 양자화하는 기법인데 크게 주 단계(dominant pass)와 부 단계(subordinate pass)의 두 과정을 반복한다. 주로 주 단계에서는 계수의 유효성을 판별한 후 부 단계에서 유효계수의 크기를 세밀화 한다. 이때, 그림 6과 같이 유효계수의 실제 크기가 해당 양자화 범위의 상위인지 하위부분인지 구분한 후 양자화 범위를 반으로 줄이게 되는데, 본 연구에서는 MDC (Multiple Descriptin Coding)기법과 엄일권 외 1명이 제안한 공유 임계값을 이용한 EZW 코딩 알고리즘을 응용하여 각 조건에 맞는 임계값을 사용하고자 하였다[5-7].

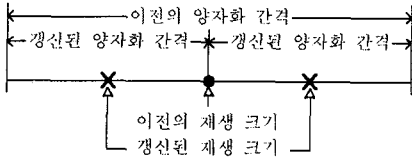


그림 6. EZW의 양자화 범위 보정

그림 7의 (a)는 EZW에서 쓰이는 일반적인 임계값 범위이며, (b)와 (c)는 기존의 임계값 범위 진행단계를 두 단계씩 묶어서 사용하는 방법으로서 (b)에서 설정된 임계값의 범위는 (c)에서 설정된 임계값의 범위와 중첩이 되도록 함으로써, 서로간의 상관관계를 높일 수 있다.

식 (3)은 i 번째 단계에서 쓰이는 일반적인 임계값 범위의 계산식으로서 그림 7의 (a)에 해당한다. 또한, 식 (4) 및 식 (5)는 변형된 EZW의 임계값을 표현한 식으로서 각각 그림 7의 (b)와 (c)의 범위를 나타내고 있다.

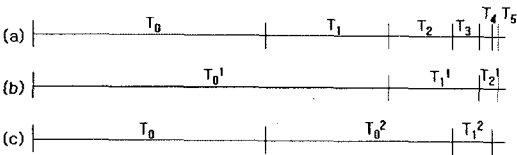


그림 7. 제한한 EZW의 각 채널별 임계값 범위

$$T_i = 2^{-i} T_0 \quad (3)$$

$$T_i^1 = 2^{-2i} T_0 \quad (4)$$

$$T_i^2 = 2^{-(2i+1)} T_0 \quad (5)$$

이렇게 정해진 임계값을 기준으로 계수를 판단하여 유효계수가 결정되며, 근사적으로 정해지는 유효계수의 크기를 세밀화하기 위하여 양자화 범위를 줄이게 되는데 표 1은 이러한 양자화 범위의 보정값을 나타낸 것이다. 표 1에서 $E_{d,i}$ 는 주 단계에서 사용하는 세밀화 보정값이며, $E_{s,i}$ 는 부 단계에서 사용하는 세밀화 보정값으로서 양자화 단계에서 발생하는 오차의 범위를 줄여주는 역할을 한다.

그림 8의 (a)는 레벨 3의 옥타브 밴드로 구성된 웨이블릿 변환 계수들의 전체 스캐닝 순서를 나타낸 것이고, (b)는 저대역, 수평/수직, 대각선의 각 밴드 별로 스캐닝하는 순서를 보여준 것이다. 일반적으로

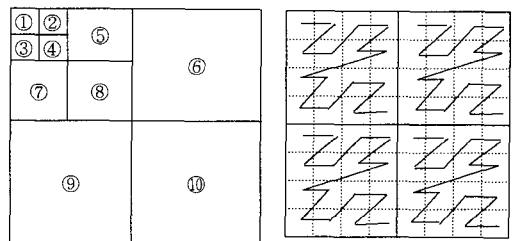
표 1. 양자화 오차에 따른 범위 보정

	식(3)의 양자화 범위보정	식(4)와 식(5)의 양자화 범위보정
$E_{d,i}$	$\frac{1}{2} T_i$	$\frac{3}{2} T_i$
$E_{s,i}$	$\frac{1}{4} T_i$	$\frac{3}{8} T_i$

계수들의 유효 여부를 판단할 때에는 저해상도 부대역에서 고해상도 부대역의 순으로 주사하는 것이 원칙이나 본 논문에서는 그림 8의 (a)에서 볼 수 있듯이 ①-⑤까지는 일반적인 주사 순서를 지키되, ⑥번부터는 수평 성분의 대역을 우선순위로 주사하였다.

즉 그림 3과 같이 3레벨인 경우, LL3→HL3→LH3→HH3→HL2→LH2→HH2→HL1→LH1→HH1의 순으로 하는 것이 일반적인 순서이나 본 논문에서는 수평대역 HL1의 순서를 우선순위로 앞당겨서 LL3→HL3→LH3→HH3→HL2→HL1→LH2→HH2→LH1→HH1의 순으로 하였다. 만약 레벨이 증가한 경우라도 수평대역을 모두 처리한 후 나머지 대역을 처리한다. 이러한 이유는 현재 압축하고자하는 입력 영상이 얼굴을 주로 다루고 있으며 사람의 눈썹, 눈, 콧망울, 입에 대한 정보는 대부분 수평성분을 가졌으며 ②, ⑤, ⑥번 대역에 분포하고 있다는 정보를 이용한 것이다.

따라서, 모바일 상에서 원하는 비트율을 만족시키기 위하여 일부 정보만을 필요로 하다면 수직이나 대각선보다 수평 성분의 정보들이 우선적으로 전송되도록 하는 것이다. 따라서 전체 비트의 인코딩이 이루어지지 않을 때에는 주요 성분으로 설정한 눈썹, 눈, 코 밑, 입술 등의 정보 위주로 전송이 될 수 있도록 유도한다. 또한 그림 8의 (b)는 수평, 수직, 대각선 각각의 대역에서 사용되는 주사 방향을 나타낸 것



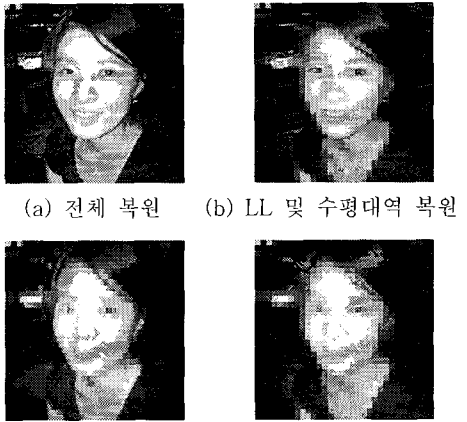
(a) 전체 스캐닝 순서 (b) 밴드별 스캐닝 순서

그림 8. 제한한 EZW 스캐닝 순서

로서 일반적으로 많이 쓰이는 morton 방식을 사용하였다.

그림 9는 레벨 3의 옥타브 밴드를 구성해 놓은 상태에서 수평, 수직 대각선 등의 해당 대역만을 이용하여 복원한 결과이다. 그림 9의 (a)는 전체 대역을 모두 복원한 결과이고, (b)는 수평성분의 대역만을 복원한 결과이다. 그림 9의 (b)와 (c), (d)를 비교해보면 얼굴 영상에서의 압축 및 복원은 눈썹, 눈, 입 등의 특징을 잘 살릴 수 있는 수평성분에 중요도가 있음을 알 수 있다. 그림 10은 실제로 높은 압축율로서 동일한 비트를 전송할 때 기존의 일반 스캔 방식과 수평대역 우선 스캔 및 전송 방식을 비교한 영상이다.

그림 10의 (a)는 전체 대역을 고루 복원하였으나 전체적인 블러링 현상이 나타나며, 그림 10의 (b)는 수평성분의 우선적 전송으로 인하여 눈썹, 눈, 입 부분이 좀 더 선명하게 복원됨을 확인할 수 있다.



(a) 전체 복원 (b) LL 및 수평대역 복원

(c) LL과 수직대역 복원 (d) LL과 대각대역 복원

그림 9. 각 대역별 복원 결과



(a) 일반적인 순차스캔 (b) 수평성분 우선스캔

그림 10. 스캔 방식에 따른 복원 영상

3. 실험 및 성능평가

본 장에서는 지금까지 설명하였던 제안한 알고리

즘을 바탕으로 구현 및 테스트한 결과를 보이고, 이를 웨이블릿 압축 방법인 EZW 방법과 PSNR, 처리 속도면에서 비교, 평가하여 각 알고리즘의 장단점에 대하여 논하고자 한다.

3.1 알고리즘의 실험

제안한 알고리즘의 실험 및 고찰을 위하여 사용한 실험 영상 데이터는 휴대폰 카메라에서 획득한 얼굴 영상으로 하였으며, 본 논문에서는 알고리즘의 실험 및 비교 평가를 위하여 20장의 얼굴 사진을 대상으로 제안한 알고리즘에 적용하였다. 웨이블릿 필터로는 HAAR 필터가 이용되었으며, 식 (6)으로 계산된 레벨의 옥타브 밴드 분해를 이용하였다. 즉, 예를 들어 256×256 영상이 들어오면 약 8단계의 레벨을 가지게 된다.

$$Level = \log_2(n) \tag{6}$$

n : 정방형 영상의 가로 크기

각 실험 영상은 제안한 압축방법으로 압축을 한 후에 복원하여 원 영상과의 차이를 측정하며, 정량적인 실험 결과를 위하여 복원 영상에 대한 손실 측정은 식 (7)과 같은 PSNR을 사용하였다.

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{N} \sum_{i=0}^N (f(x_i) - f'(x_i))^2} [dB] \tag{7}$$

N : 영상의 전체 화소수

$f(x_i)$: 원 영상의 x_i 에서의 화소값

$f'(x_i)$: 복원 영상의 x_i 에서의 화소값

3.2 실험 결과

그림 11은 각각 기존의 EZW 방식과 제안한 방식을 사용하여 압축한 후 복원한 영상들이다.



(a) 기존의 EZW 복원 (b) 제안한 EZW 복원

그림 11. 기존방식과 제안방식에 따른 복원 영상

두 영상을 비교하여 볼 때의 큰 특징으로서 배경 부분과 얼굴 영역을 먼저 확인 할 수 있는데, 그림 11의 (a)영상의 배경(얼굴 영역을 제외한 나머지 부분)은 우측 영상의 배경과 비교해 볼 때 좀 더 뚜렷하며 블록화 현상이 적은 반면, 중요 영역(얼굴 영역)의 부분은 열화된 것을 확일 할 수 있다. 반면, 그림 11의 (b)영상은 비록 배경 부분의 블러링 및 화질 열화 효과가 있지만 얼굴 영역 부분의 색상이나 블록화 현상이 현저히 줄어들어 중요 영역의 효과적인 복원이 되었음을 확인할 수 있다.

그림 12는 그림 11의 얼굴 영상을 포함한 20장의 얼굴 영상을 테스트하여 0.01에서 0.1까지의 각 비트율(bpp)별로 평균 PSNR을 그래프로 나타낸 것이다. 여기서 비트율이라고 하는 것은 한 화소당 적용되는 비트수로서, 원 영상에서의 비트율은 한 화소당 8비트 3채널로서 24의 비트율을 갖는다.

그림 12에서 알 수 있듯이 제안한 알고리즘의 전체 영상에 대한 PSNR은 상당히 낮은 편이지만 상대적으로 얼굴 영역에 대한 복원값은 상당히 증가함을 알 수 있다. 또한 기존의 EZW 방법과 비교해 볼 때, 전체 영상에 대한 PSNR 차이는 비트율에 따라 약간씩 상이하지만 약 1~3dB 정도의 화질 열화가 발생하였으나 얼굴 영역에 있어서는 약 1~1.5dB 정도 화질이 기존의 EZW 복원 방법에 비해 향상되어 복원되었다.

본 논문의 실험에 사용한 영상은 대부분 얼굴과 배경이 복잡하지 않은 사진을 대상으로 하였고 영상에서의 중요부분과 그렇지 않은 부분의 차지하는 비율이 비슷한 데이터를 사용하였으나, 압축 전송에 있어서 중요하지 않는 부분이 좀 더 복잡한 형태를 가

지고 있는 영상이라면 본 논문에서 제안하고자 하였던 영역 중요도별 부호화방법이 훨씬 효율적으로 적용될 수 있다. 또한, 본 연구에서는 다루지 않았으나 불필요한 배경부분을 영역 축소(down size)하여 압축한다면 더욱 높은 압축율을 얻을 수 있을 것이다.

표 2는 제안한 방법과 기존의 방법과의 처리속도를 추가적으로 비교하였다. 실험에 사용된 컴퓨터는 Pentium-III 1GHz 노트북이며, 실험의 분석을 위하여 전체 알고리즘은 MATLAB으로 수행하였고, 그 중 제안한 부분인 주(dominant) 단계와 부(subordinate) 단계는 EZW와 제안한 방법 모두 DLL로 구현하여 연동하였다.

결과에서 알 수 있듯이 압축과정에 있어서는 두 알고리즘 모두 별 차이를 나타내지 않았으나 복원과정에 있어서는 비트율이 높아질수록 최고 2배 가까운 처리속도를 보였다. 이는 압축과정에서 사용한 DLL의 속도차이이며, 본 연구의 목표가 압축율을 높이기 위한 제안 방법이므로 수행 속도를 고려한 좀 더 최적화된 코드화작업이 보완된다면 훨씬 높은 성능을 얻을 수 있다.

표 2. 비트율에 따른 처리속도 (단위:초)

알고리즘		비트율				
		0.02	0.04	0.06	0.08	0.1
압축	EZW	4.3	5.4	7.5	9.3	10.2
	M-EZW	5.5	7.1	8.9	10.3	11.8
복원	EZW	4.0	6.1	9.4	13.1	16.7
	M-EZW	4.8	8.0	11.7	15.7	21.1

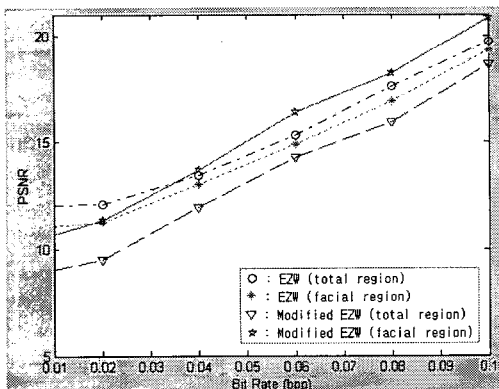


그림 12. 비트율에 따른 PSNR

4. 결 론

본 논문에서는 전송에 있어서 중요한 부위를 2단계 MDC를 사용하여 차등 부호화함으로써 적은 비트율 혹은 동일한 비트율을 가진 전송에 있어서 원하는 영역의 품질을 향상시킬 수 있으며, 필요시 비트율에 따라 두 개의 채널 중 한 채널만을 전송하여도 무방하다는 장점을 가지고 있다. 또한 제안한 방법은 특히 배경부분이 복잡한 영상에서의 불필요한 코딩 생산을 억제함으로써 더욱 높은 효율을 얻을 수 있다.

향후 연구 과제로는 현재 얼굴에 국한되어 있는 중요영역의 대상 확장 및 한 영상에서 중요영역이 여러 영역으로 산재해 있을 때의 처리에 대한 것이

다. 또한 현재의 압축시간은 기존의 임계값 단계와 전체 단계의 횟수는 같기 때문에 크게 차이는 없으나, 각 채널에 대한 병렬처리 방법을 적용한다면 좀더 빠른 수행시간의 개선이 가능할 것이다.

참 고 문 헌

[1] J. C. Terrillon and S. Akamatsu, "Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images," *Proc. of the 4th IEEE International conference on Automatic Face and Gesture Recognition*, pp. 54-60, 2000.

[2] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," *IEEE Transactions on Multimedia*, Vol. MM-1, No. 3, pp. 264-277, 1999.

[3] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Transactions on signal Processing*, Vol. 41, No. 12, pp. 3445-3462, 1993.

[4] K. Sayood, *Introduction to Data Compression*, Morgan kaufmann Publishers, pp. 77-116, 2000.

[5] I. K. Eom and Y. S. Kim, "Robust EZW Coding with Shared Threshold," *IEEE LETTERS*, Vol. 39, No. 21, 2003.

[6] S. D. Servetto, K. Ramchandran, V.A. Vaishampayan, and K. Nahrstedt, "Multiple Description Wavelet Based Image Coding," *IEEE Transactions on Image Processing*, Vol. 9, pp. 813-826, 2000.

[7] A. I. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded Multiple Description Scalar Quantizers for Progressive Image Transmission," *IEEE ICASSP*, pp. 736-739, 2003.



최 재 영

1995년 경원대학교 수학과 학사
 1999년 경원대학교 전산학과 석사
 2004년 경원대학교 전산학과 박사
 2004년 3월~10월 경원대학교 소프트웨어연구소 책임연구원

2004년 11월~현재 UCLA Vision Lab. Post-doc researcher

관심분야: 영상처리, 패턴인식, 멀티미디어



황 승 호

2001년 경원대학교 수학과 학사
 2003년 경원대학교 전산학과 석사
 2004년~현재 경원대학교 소프트웨어연구소 연구원

관심분야: HCI, 얼굴인식, 영상처리



황보택근

1983년 고려대학교 공과대학 학사
 1987년 CUNY 전산학과 석사
 1995년 S.I.T. 전산학과 박사
 1988년~1993년 Q-Systems, Senior Technical Staff
 1995년~1997년 삼성종합기술원 선임연구원

1997년~현재 경원대학교 소프트웨어대학 소프트웨어학부 부교수

관심분야: 컴퓨터 그래픽스, 영상처리, GIS



양 영 규

1972년 서울대학교 학사
 1974년 서울대학교 석사
 1985년 Texas A&M Univ. Remote Sensing 박사
 1973년~1996년 시스템공학연구소 책임연구원

1996년~2003년 한국전자통신연구원 책임연구원

2002년~2003년 University of California at Irvine 객원연구원

2003년~현재 경원대학교 소프트웨어대학 교수, 전산정보원장

관심분야: LBS, GIS/RS, 텔레매틱스