

# 스테레오 영상의 정합값을 통한 얼굴특징 추출 방법

김상명<sup>\*</sup>, 박장한<sup>\*\*</sup>, 남궁재찬<sup>\*\*\*</sup>

## 요 약

본 논문에서는 스테레오 영상의 정합값(matching)을 통한 얼굴 특징추출 알고리즘을 제안한다. 제안된 알고리즘에서는 얼굴색상 정보의 RGB컬러공간을 YCbCr컬러공간으로 변환하여 얼굴영역 검출하였다. 추출된 얼굴영역으로부터 눈 형판(template)을 적용하여 눈 사이의 거리와 기울어짐, 코와 입에 대한 특징의 기하학적인 특징 벡터를 추출하였다. 또한 제안한 방법은 2차원 특징정보 뿐만 아니라 스테레오 영상의 정합을 통한 얼굴의 눈, 코, 입의 특징을 추출할 수 있었다. 실험을 통하여 약 1m이내 거리에서 73%의 일치율을 보였고, 약 1m이후 거리에선 52%의 일치율을 보였다.

## Face Feature Extraction Method Through Stereo Image's Matching Value

Sang-Myung Kim<sup>\*</sup>, Chang-Han Park<sup>\*\*</sup>, Jae-Chan Namkung<sup>\*\*\*</sup>

## ABSTRACT

In this paper, we propose face feature extraction algorithm through stereo image's matching value. The proposed algorithm detected face region by change the RGB color space of skin color information to the YCbCr color space. Applying eye-template from extracted face region geometrical feature vector of feature about distance and lean, nose and mouth between eye extracted. And, Proposed method could do feature of eyes, nose and mouth through stereo image's matching as well as 2D feature information extract. In the experiment, the proposed algorithm shows the consistency rate of 73% in distance within about 1m and the consistency rate of 52%in distance since about 1m.

**Key words:** Stereo Image(스테레오영상), Face Detection(얼굴검출), Face Feature Extraction(얼굴특징 추출), Eye-Template(눈형판), Matching Value(정합값)

## 1. 서 론

최근 몇 년간 인터넷을 통한 정보가 공유되면서 공유할 수 없는 개인정보에 대한 가치가 상승하고

있다. 더불어 개인정보 보호 문제가 대두되면서 생체 정보를 이용한 보안기술이 주목받고 있다. 생체정보는 타인에 의해 쉽게 복제될 수 없는 개인의 고유한 특징이다. 이러한 생체정보를 이용해 개인을 인식하는 기술에는 지문, 홍채, 서명검증, 정맥, 음성, 얼굴 인식 등이 있다[1]. 얼굴인식은 다른 생체인식 기술에 비해 경제성과 사용자 편리성이 높다. 또한 고가의 장비 대신 일반적인 PC카메라를 이용할 수 있으며, 사용자는 기기에 직접 접촉을 하지 않아도 인식이 이루어질 수 있다. 그리고 우리가 실제 눈으로 얻는 정보는 2차원 정보가 아닌 입체 영상이므로 입체 영상은 2차원 평면 영상과는 달리 사물에 대한 원근감, 즉 깊이(depth) 정보를 지니고 있기 때문에 이를 원격 작업에 이용하면 작업자는 마치 현장에 직접

※ 교신저자(Corresponding Author) : 박장한, 주소 : 서울시 노원구 월계동 447-1(139-701), 전화 : (02)940-5125, FAX : (02)941-2754, E-mail : parkch@kw.ac.kr

접수일 : 2004년 5월 17일, 완료일 : 2004년 10월 25일

<sup>\*</sup> 준회원, ㈜모빌링크

(E-mail : iic\_lab@hotmail.com)

<sup>\*\*</sup> 준회원, 중앙대학교 첨단영상대학원 Post-Doc.

<sup>\*\*\*</sup> 정회원, 광운대학교 컴퓨터공학과 교수

(E-mail : namjc@daisy.kw.ac.kr)

※ 이 논문은 2004년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.

투입되어 작업을 수행하는 것과 같은 현장감을 얻게 되어 원격 작업을 쉽게 수행할 수 있다. 영국 AEA 및 미국 ORNL 등에서 수행한 방사선 지역이나 우주에서 사용될 로봇의 시각 장치의 원격 조작성 실험에 관한 연구 보고서에서 입체 영상은 평면 영상에 비해 약 30%이상 원격 작업 능률을 향상시키는 것으로 보고하고 있다. 얼굴검출 및 특징 추출에서도 2차원을 이용한 방법의 한계를 극복하기 위하여 3차원 얼굴 데이터를 이용한 방법이 연구되고 있다.

얼굴검출 및 특징 추출의 일반적인 단계는 얼굴영상이 입력으로 주어지면 복잡한 배경으로부터 얼굴영역만을 분리해 내고, 분리된 얼굴영역으로부터 특징을 추출한다. 2차원 얼굴 영상에서 얼굴의 특징인 깊이 정보를 이용하지 못하는 점을 스테레오 영상을 통하여 얼굴의 깊이정보와 거리정보를 추출할 수 있다.

본 논문에서는 스테레오 영상의 정합값을 통한 얼굴영역에서 눈, 코, 입 등의 특징 추출방법을 제안한다. 제안된 알고리즘에서는 스테레오 영상의 시차(disparity) 보정처리 단계가 선행되어야 한다. 시차를 통한 특징영역의 깊이정보를 추출할 수 있다. 스테레오 영상에서 평면 영역은 주로 입체영상을 구성하는 좌·우 평면영상의 가장자리에 나타나고, 영상 획득 이전에 입체카메라의 캘리브레이션 작업이 선행되어야 한다. 또한 효율적인 얼굴영역 검출을 위해 RGB컬러영상을 YCbCr컬러공간[2,3]으로 변환하고, Cb, Cr성분을 가지고 일정한 범위 내에 있는 객체들을 제거(opening) 연산[4]을 통해 추출된 후보객체들 중에서 가장 큰 객체를 얼굴후보 영역으로 선정한다. 그리고 검출된 얼굴 영역에서 얼굴 특징(눈, 코, 입)들을 검출하기 위해 밝기 정보를 구별해 줄 수 있는 전처리 과정을 수행한 후 눈의 형판을 적용하여 눈의 검출과 각 특징들의 상대적인 크기, 위치, 구조적인 관계값에 대한 얼굴 특징 벡터를 추출하였다. 왼쪽과 오른쪽 영상에서 시차 정보의 특징을 분석하여 약 1m 이내의 영상, 약 1m 이후의 영상에 대한 일치 및 불일치의 오류를 검출할 수 있었다.

본 논문의 구성은 다음과 같다. 2장에서는 스테레오 영상과 얼굴영역 검출 및 특징 추출의 방법을 소개하고, 3장에서는 얼굴검출 및 특징 추출에 대한 알고리즘을 제안한다. 4장에서는 모의실험 결과를 분석하고, 5장에서 결론을 맺는다.

## 2. 관련 연구

### 2.1 스테레오 영상

스테레오 영상의 기술은 2차원 영상으로부터 3차원 정보를 추출하는 과정으로 2대의 카메라를 이용하는 방법이다[5,6]. 3차원 정보는 사물에 대한 원근감, 위치, 속도, 깊이 등의 정보를 지니고 있기 때문에 원격작업에 이용하면 마치 현장에 직접 투입되어 작업을 수행하는 것과 같은 현장감을 얻게 되어 원격작업을 쉽게 수행할 수 있다. 입체영상은 최소 좌·우로 구성된 2장 이상의 평면영상이 필요하며, 일반적으로 2대 이상의 영상센서를 이용하여 동시에 촬영하여 얻는다. 스테레오 영상을 이용한 입체카메라가 사람의 눈에서와 같이 자연스럽고 선명한 양질의 입체영상을 획득하기 위해서는 인간의 시각기능 중 최소한 세 가지 기능을 가지고 있어야 한다. 사람의 눈에 해당하는 카메라를 설치하는 방법에 따라 크게 순차식과 동시식으로 나눌 수 있다. 그림 1은 스테레오 카메라의 구성에 대한 그림을 나타내고 있는 것으로 순차식은 좌·우측의 카메라를 동시에 존재시키지 않고 하나의 카메라를 이용하여 순차적으로 영상을 입력하는 방식으로 이동에 따른 불편함과 정밀도 및 카메라 캘리브레이션에 대한 부정확 등의 이유로 사용되지 않는 방법이다.

동시식이란 좌·우측 카메라를 모두 설치하여 동시에 영상을 취득하는 것으로 현재의 스테레오 비전 시스템에서 많이 사용하고 있다. 한편 동시식 카메라 구성법에는 평행식과 교차식으로 나누어지는데, 카메라의 광축이 어떻게 구성하느냐에 따라 분류한다. 평행식은 두 카메라의 광축이 수직, 수평으로 평행한 구성법을 말하며, 교차식은 두 카메라의 광축이 일치하는 곳 즉, 주시점을 만드는 방법으로 사람의 눈 구조와 매우 가깝다고 할 수 있다. 일반적으로 평행식 카메라 구성법은 수평 시차만 존재하며 수직 시차는

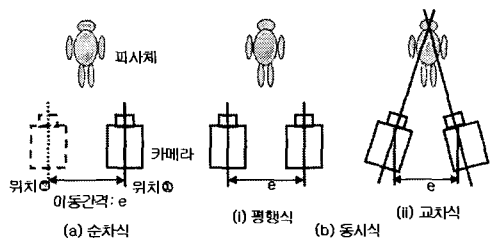


그림 1. 스테레오 카메라의 구성법

없기 때문에 취급하기에 용이하여 수학적인 모델링이 쉽다. 반면 교차식 구성방법은 광축을 일치시키는 문제가 있으며 특히 수직 시차의 변화에 따른 문제가 있기 때문에 수학적 모델링이 어려운 문제가 있다. 또한 평행식 카메라 구성법은 주시점이 없기 때문에 사람의 눈 구조와는 조금 다른 구조를 갖는다. 따라서 입체감을 표시하는 분야의 응용에는 사용하기 어려운 문제가 있지만, 거리 측정에는 수학적 표현이 쉽게 나타낼 수 있기 때문에 간단한 방법으로 거리를 측정할 수 있으며, 수직 시차를 고려할 필요가 없기 때문에 더욱 수월하다. 교차식 카메라 구성법은 주시점을 따라 카메라를 상·하, 좌·우로 항상 이동해야 하는 문제가 있으며, 두 대의 카메라가 각각 별도로 작동해야 하기 때문에 기계적 오차 및 움직임에 대한 이동속도 등이 문제가 된다. 이러한 문제는 카메라의 기하학적 수식을 만드는데 큰 제한점을 갖게 만들어 거리 측정에 어려운 면이 있다. 그러나 사람의 눈 구조와 매우 유사한 구조를 갖고 있어 스테레오 영상을 재현하는데 있어서 부드러운 입체 영상을 제공 가능하기 때문에 스테레오 영상의 표현에 많이 사용되고 있다. 본 논문에서는 거리 측정에 정밀도가 높고 제어의 편리성을 고려하여 평행식 카메라 구성법을 사용하며, 카메라의 설치 간격은 6.5(cm)로 설정하여 실험한다.

## 2.2 얼굴영역의 검출 및 추출 방법

얼굴영역 검출은 임의의 영상에서 얼굴의 위치를 자동으로 추출하는 기술로서 얼굴관련 연구 분야이다[7]. 초기 얼굴검출은 자동화된 얼굴인식을 위한 전 단계 기술로서 연구가 시작되었으나 문제의 어려움과 중요성으로 인해 최근에는 하나의 독립적인 분야로 연구가 활발히 진행되고 있다. 얼굴검출은 얼굴을 포함하는 최소영역을 찾아내는 것을 말하며, 응용에 따라 얼굴의 정확한 경계영역을 찾는 것을 포함할 수 있다. 얼굴영역 검출은 인종, 성별, 나이별, 얼굴 움직임, 표정 변화 등의 얼굴이 가지는 변화 요인뿐 아니라 주변 환경 즉, 복잡한 배경, 조명 변화, 장신구 착용 등에 의한 환경 요인에 무관하게 검출할 수 있는 기법을 개발 목표로 하고 있다[1]. 얼굴검출은 입력 데이터에 따라 정지영상에서의 검출과 동영상에서의 검출로 분류할 수 있다. 동영상에서의 연속적인 검출은 얼굴추적으로 구분되며, 큰 범주로는 동영상

에서의 검출로 분류한다. 동영상인 경우는 정지영상에서 사용할 수 있는 영역정보 이외에 시간정보를 추가적으로 사용할 수 있다. 이 방법은 컬러정보와 결합하여 실시간 검출과 같은 보다 넓은 응용범위를 가진다. 그러나 동영상은 카메라 움직임 보상 등과 같은 또 다른 문제점을 안고 있어 실제로 시간정보가 추가된 만큼 성능 향상이 이루어지지 않는다. 일반적인 얼굴검출 기법으로는 Kanade 연구팀이 제안한 신경망 기반 방법[8]과 Sung과 Poggio에 의해 제안된 예제 기반 학습 방법[9]이 있다.

얼굴특징의 추출은 얼굴영상 및 얼굴을 포함한 영상내의 분리된 얼굴영역에서 얼굴특징들, 즉 눈, 코, 입 등을 추출하는 것으로 다양한 방법에 의해 연구되어 왔다. 수염이 없는 얼굴 영상에 대해 KL변환을 적용하여 얼굴을 표현하는 방법[10], 얼굴특징 추출을 위해 Hough변환을 사용한 방법[4], 가변 형판을 사용하여 얼굴의 특징을 추출한 방법[11]. 눈을 검출하기 위해 형판을 사용한 방법[12], Gabor 웨이블릿 분해와 지역적 크기 변환의 상호 작용(local scale interaction)을 사용하여 곡률이 최대인 점에서 얼굴 특징을 추출하는 방법[13] 등이 있다. 얼굴 특징추출의 기술들은 매우 다양한 방법에 의해 연구되어 왔으며 최근에 웨이블릿 해석을 사용한 방법과 템플릿과 같은 영상 필터를 사용한 기술들이 선보이고 있다[14,15]. 본 논문에서는 얼굴영역을 검출하기 위해 RGB컬러공간에서 YCbCr컬러공간으로 변환하여 얼굴후보를 설정한다. 설정된 후보영역에서 눈 템플릿을 적용한 눈의 위치를 검출하고, 검출된 눈의 위치에 따라서 입과 코의 영역을 검출하고 특징을 추출한다.

## 3. 제안된 얼굴 영역 검출 및 특징 추출

### 3.1 제안된 얼굴 검출 블록도

본 논문에서 제안하는 정면 얼굴영역 검출 및 특징 추출 블록도는 그림 2와 같다. 제안된 블록도는 카메라에서 획득된 스테레오 영상의 영상위치 보정, 정면 얼굴후보 영역 추출, 정면 얼굴영역 추출 단계의 정규화과정을 거치며, 정규화된 영상의 크기는 128×128이다. 추출된 정면얼굴영역 내에서 눈 형판을 사용하여 눈을 검출하고, 눈, 입, 코의 기하학적인 관계를 통해 정면 얼굴의 특징을 추출한다. 그리고

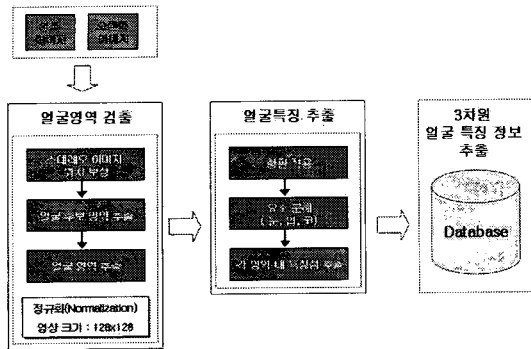


그림 2. 정면 얼굴영역 검출 및 특징 추출 블록도

추출된 정보는 데이터베이스에 정보를 저장한다.

### 3.2 스테레오 영상의 시차 보정 처리

그림 3의 스테레오 영상에서 평면 영역은 주로 입체영상을 구성하는 좌·우 평면영상의 가장자리에 나타난다. 그림 3에서 교차축에 대한 평면 영역을 표현하였으나 수평 이동축에서도 동일하게 평면 영역

이 존재한다. 평면 영역과 같은 비입체영역을 줄이기 위해서는 영상획득 이전에 입체카메라의 캘리브레이션 작업이 선행되어야 한다.

스테레오 영상에서 시차를 추출하는 방법 중, 본 논문에서는 고정된 크기의 블록정합 알고리즘을 사용한다. 블록정합에 사용되는 기저함수는 MAD (Mean Absolute Difference)와 MSD(Mean Square Difference)를 주로 사용하는데, 계산량을 줄이기 위해 SAD(Sum of Absolute Difference)를 기저함수로 사용하였다. 공간상의 한 점이 좌·우 영상에 투영되었을 때 두 점을 연결하는 가상의 선을 에피폴라선이라 한다[16]. 우측영상의 대응 블록은 좌측영상과 동일한 Y축의 좌표를 가지는 에피폴라선에서 정합된다. 식(1)은 스테레오 영상의 시차추출을 위해 사용된 SAD를 나타낸다.

$$SAD = \sum_i \sum_j^{N_x, N_y} |I_L(i, j) - I_R(i + k, j)| \quad (1)$$

여기서  $N_x, N_y$ 는 블록의 크기를 나타낸다.  $I_L$ 은 좌



그림 3. 스테레오 영상에서의 평면영역

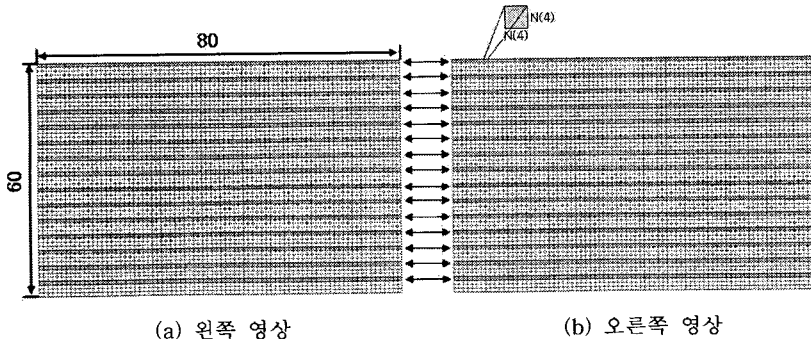


그림 4. 스테레오 영상에서의 시차 보정

측 영상의 블록을 나타낸다. 또한  $I_R$ 은  $I_L$ 에 대응되는 우측영상의 블록을 나타내며,  $k$ 는 왼쪽 영상에 대응되는 오른쪽 영상의 시차(disparity)이다.

스테레오 영상의 시차 보정은 스테레오 영상에 포함되어 있는 평면영역과 입체카메라의 에피폴라선이 정상적으로 정렬되지 않아 발생하는 불일치 영역을 최소화하는 과정이고, 평면영역은 입체영상을 구성하는 좌·우 평면영상의 좌·우 가장자리에 분포하며, 상·하 불일치 영역은 영상의 상·하 가장자리에 분포한다. 따라서 시차에 의해 발생하는 좌·우 영상의 불일치를 보정하는 과정이 선행되어야 한다. 이를 보정하기 위해 블록 크기( $N$ )를  $4 \times 4$ 로 하는 블록 정합 알고리즘을 이용하였다. 그림 4는 스테레오 영상의 블록 정합처리의 구성이다.

그림 4와 같이 시차 보정에서 블록 정합 알고리즘을 이용하여 많은 연산이 소요되는 불필요한 연산과정을 제거하면서도 전역 탐색과 비교해 거의 차이가 없는 결과를 얻을 수 있다.

먼저, 해당열의 SAD는 식(2)을 이용하여 각각 구하고,

$$SAD_k = \sum_i \sum_j |I_L(i, j) - I_R(i+k, j)| \quad (2)$$

각 동일 열에서의 최소 SAD는 식(3)을 이용하여 구하며,

$$SAD_{MIN}^P = \text{MIN}(\sum SAD_k) \quad (3)$$

마지막으로 전체 영상의 배경 시차인 최소 SAD는 식(4)을 이용하여 얻을 수 있다.

$$SAD_{MIN} = \text{MIN}(\sum SAD_{MIN}^P) \quad (4)$$

그림 5는 배경 시차를 이용하여 얻은 보상된 스테레오 영상을 나타내고, 보정된 영상을 식 5, 6과 같이 각각 쓸 수 있다[5].

$$\text{right}^* = \text{right}_{t-k} \quad (5)$$

$$\text{left}^* = \text{left}_{t+k} \quad (6)$$

### 3.3 정면 얼굴영역의 검출

카메라로부터 획득된 입력 영상은  $320 \times 240$  크기의 얼굴을 포함하는 RGB컬러영상이다. 먼저 획득된

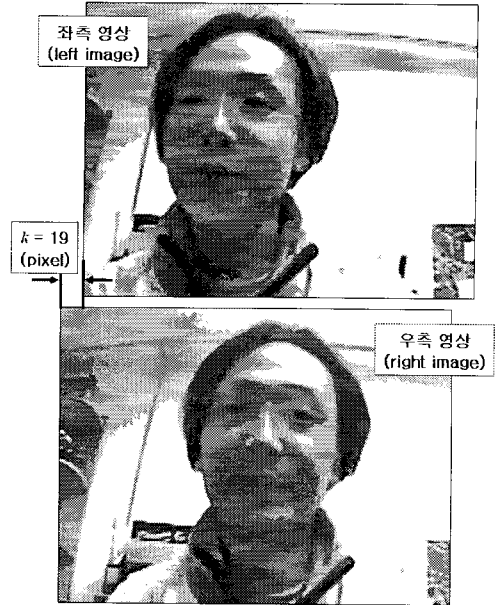


그림 5. 보정된 스테레오 영상

정면 얼굴 후보영역 영상에서 색상 정보만을 이용하여 조명에 대한 영향을 줄이기 위해 주어진 영상을 RGB컬러공간으로부터 YCbCr컬러공간으로 변환하게 된다. 입력 영상의 YCbCr 정보 중에서 Cb, Cr성분에 대해 동양인에 맞게 얼굴 피부 색상 범위를  $R_{Cb}=[77 \ 127]$ ,  $R_{Cr}=[133 \ 173]$ 으로 색상 영역분할과정을 거쳐 피부색 값을 갖는 영역만을 추출한다[17]. 식(7)을 이용하여  $R_{Cb}$ ,  $R_{Cr}$ 의 영역을 선택한다.

$$S(x, y) = \begin{cases} 1, & \text{if } [(Cb(x, y) \in R_{Cb}) \cap (Cr(x, y) \in R_{Cr})] \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

그림 6의 (a)는 입력된 영상을 색상영역 분할처리를 거친 결과이다. (a)의 영상을 보면 세밀한 잡음 형태가 존재하는 것을 볼 수 있는데, 이런 잡음의 형태를 제거하여 얼굴 영역만을 분리해 내기 위하여 형태학적 필터(morphological filter) 중에서 제거(opening) 연산을 수행하였다. 제거 연산은 원 영상을 침식 연산으로 영상의 최외각을 한 픽셀씩 없애고 다시 팽창 연산으로 최외각을 한 픽셀씩 확장시키는 방법이다. 따라서 제거 연산은 미세한 잡음을 제거하고, 다시 팽창 연산을 수행하므로 원 영상의 크기를 유지할 수 있다. 본 논문에서는 입력 차원을 결정하기 위해 영상을  $128 \times 128$ 로 얼굴 데이터베이스를 구축하였다.

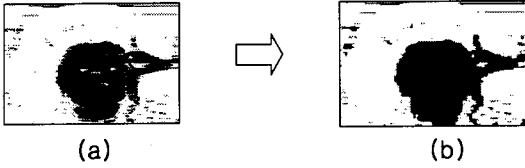


그림 6. 색상 영역을 이용한 얼굴영역 추출  
(a) 색상영역 분할 후의 영상 (b) 제거 연산을 적용한 후의 영상

다음 단계는 수직과 수평으로 스캔을 하게 되는데 수직·수평 방향으로 “255”의 값을 갖는 화소의 수를 센다. 전·후 5개의 화소값의 평균값을 취해 심하게 변하는 부분을 부드럽게 처리하여 결과의 신뢰도를 높였다. 그림 7의 영상에서 수직과 수평방향으로 “255”의 값을 갖는 화소 수를 합산하고 그래프 상에서 적응적 임계값 이상의 값(60)을 갖는 영역은 화소값을 모두 “255”로 처리한다. 최종적으로 화소값이 모두 조정된 결과는 그림 7의 (b)영상과 같으며, 얼굴영역이 추출된다.

### 3.4 정면 얼굴특징 추출

#### 3.4.1 정면 얼굴특징 추출을 위한 전처리 과정

정면 얼굴특징의 검출을 더욱 효율적으로 수행하기 위하여 분리된 얼굴영역 내에서 얼굴특징을 보다 명확한 명도값으로 대비시켜 전처리하게 된다. 보통의 경우에 얼굴특징을 추출하는 방법으로 경계선 검출 마스크가 일반적으로 쓰이지만 경계선 마스크는 가중치의 합이 “0”이기 때문에 명암도가 다른 영역들의 경계선은 검출이 가능 하지만 영역내부는 명암도에 상관없이 “0”의 값을 갖게 된다. 따라서 이런

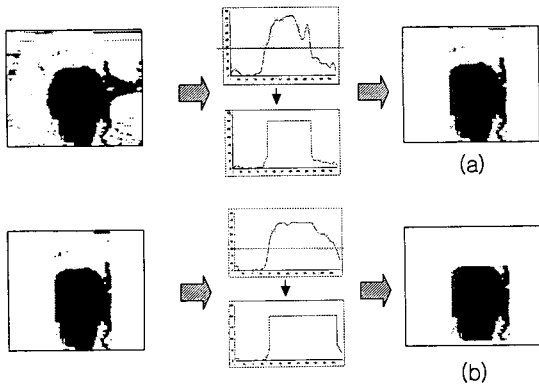


그림 7. 얼굴 영역 추출의 처리  
(a) 수직방향 화소값 조정 영상 (b) 수평방향 화소값 조정 영상

문제를 해결하고 얼굴색상과 얼굴특징들(눈, 코, 입)을 보다 명확히 대비시키기 위해서 그림 8의 3×3 마스크를 사용한다.

평균 조명과 동양인의 얼굴색 등을 고려하여 얼굴특징의 검출영역을 정하였지만, 보다 더 신뢰성 있는 결과를 얻기 위해서는 주변 조건의 변화에 따라 조금씩 틀러지는 입력 영상의 밝기 값을 조정한다. 따라서 얼굴 영역내의 화소값들의 차이만큼을 각 화소값에서 빼주거나 더해준다. 그림 9의 (a)는 얼굴영역의 밝기 값 조정 전의 영상이고, (b)는 얼굴영역의 명암값 70으로 조정 후의 결과 영상이다.

#### 3.4.2 스테레오 영상의 정합

그림 10의 (c)에서 (a)는 오른쪽 영상이고, (b)는

-1	-1	-1
-1	8	-1
-1	-1	-1

그림 8. 3×3 마스크

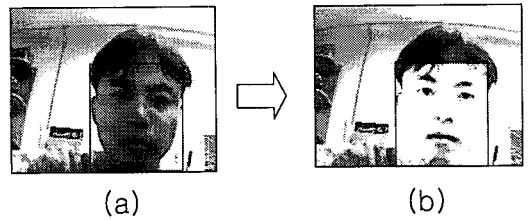


그림 9. 얼굴 영역 전처리  
(a) 밝기 값 조정 전 (b) 밝기 값 조정 후

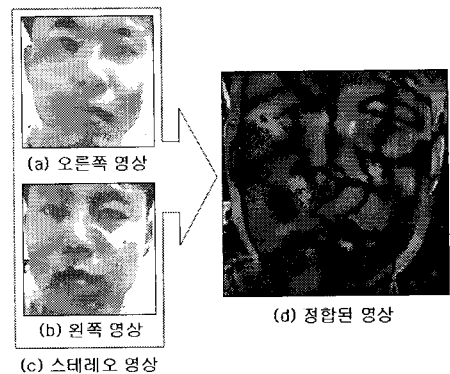


그림 10. 스테레오 영상 정합

왼쪽의 영상이며, 두 영상을 정합한 영상이 (d)이다. 본 논문에서 오른쪽과 왼쪽영상에 대해서 빼기 연산을 통한 모든 차 값의 평균인 특징 정보를 얻게 되는데, 정합과정에서 픽셀 단위의 명암값을 낮추는 빼기 연산은 영상 전체의 밝기값을 낮춰준다. 만약 빼 값이 음수 값이면 영상의 반전이 일어나기도 하지만 전처리 과정을 거친 경우는 원하는 정보를 얻을 수 있었다. 그림 11의 (a)는 정합 영상이고, (b)는 (a)의 정합된 얼굴을 반전시킨 영상이며, 얼굴의 특징 정보를 포함하고 있다. 또한 (c)은 (b)의 영상을 이진화시킨 영상이다. 본 논문에서는 이진화 영상의 특징 정보를 활용해 얼굴 영상의 특징을 추출한다.

3.4.3 정합된 영상에서의 얼굴특징 추출

정합된 영상에서 표 1과 같이 정합값, 거리, 각도 등의 세 가지 조건을 모두 만족해야만 최종적으로 정상적인 눈으로 판별하게 된다.

또한 얼굴영역에서 눈의 검출을 위해 최적의 형판을 설계하여 적용하였다. 입력 영상의 크기가 320×240이기 때문에 얼굴 전체가 영상 내에 존재한다는 가정 하에 입력 영상에서의 평균적 눈의 크기를 고려하여 크기가 20×16인 형판을 그림 12와 같이 설계하였다. 여기서 검은 부분의 화소는 “-5”, 하얀 부분의 화소는 “+5”의 값을 갖게 하였다. 이 값들은 모의실험에서 추출된 최적값을 선택한 것이다. 또한 눈의 가운데 좌표값(x, y) 추출은 식(8)을 통하여 구한다.

$$\begin{aligned} \text{눈의 가운데 좌표값} : x &= ((x_2 - x_1) / 2) + x_1 \\ y &= ((y_2 - y_1) / 2) + y_1 \end{aligned} \quad (8)$$

눈의 검출 척도로 사용하기 위해 정합값(matching value)을 제안하였는데, 0~255 사이의 화소값을 갖는 흑백 영상과 “-5”와 “+5”의 값을 갖는

표 1. 정합값, 거리, 각도를 사용한 눈 판단 조건

얼굴 판단 조건	
정합값	90 ≤ 정합값(Mv)
거리	10 ≤ 거리(distance) ≤ 20
각도	-10 ≤ 각도(angle) ≤ 10

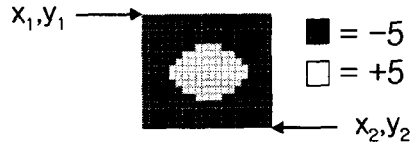


그림 12. 제안된 눈 형판 (20×16)

눈 형판과의 상관도(correlation) 값으로 볼 수 있다. 전처리 과정을 거친 얼굴영역을 눈 형판을 이용하여 검색하게 되는데, 이 때 화소단위로 형판과 검색영역과의 정합값을 식(9)과 같이 구하게 된다.

$$Mv(m, n) = \frac{1}{x_i \times y_i} \sum_{x=0}^{x_i-1} \sum_{y=0}^{y_i-1} (E[X][Y] \times \prod [X_i + (m-8)][Y_i + (n-10)]) \quad (9)$$

$$\text{단, } 8 \leq m < X_i - 8, \quad 10 \leq n < Y_i - 10$$

여기서 E[X][Y]는 사용된 눈 형판이고,  $x_i$  와  $y_i$  는 눈 형판의 수직 및 수평크기를 각각 나타낸다. 또한 E[X][Y]는 입력 영상의 얼굴영역이며,  $X_i$  와  $Y_i$  는 입력 영상에서 검출된 얼굴영역의 수평 및 수직의 크기이다. 이 정합값의 최대값은 검색영역 내에서 형판과 정확히 일치하는 패턴이 있을 때 형판의 “+5”와 영상의 “255”, 그리고 형판의 “-5”와 영상의 “0”이 일치할 때 구해진다. 그림 13은 제안된 눈 형판을 이용하여 왼쪽 눈과 오른쪽 눈의 중앙 좌표값을 표시한 것이다. 또한 식 10은 오른쪽 영상과 왼쪽

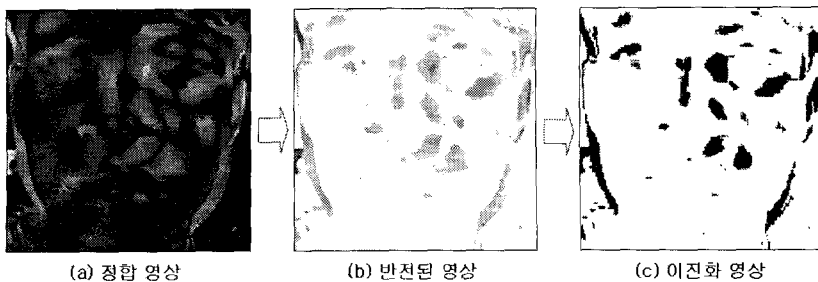


그림 11. 정합 영상 및 이진화 영상

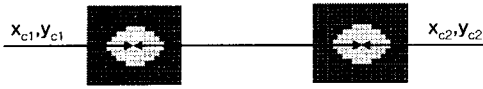


그림 13. 왼쪽과 오른쪽 눈의 중앙 좌표

영상에서 두 눈 사이의 거리(  $Dd_{eye}$  )를 구하는 식이다. 그림 14는 정합값을 갖는 눈의 위치를 표시한 결과 영상이며, 오른쪽 영상과 왼쪽 영상을 이용해 합성된 영상에서 불일치 영역을 추출한다.

$$Dd_{Eye} = \left| \sqrt{(x_{llc2} - x_{rlc1})^2 + (y_{llc2} - y_{rlc1})^2} - \sqrt{(x_{rlc2} - x_{rrc1})^2 + (y_{rlc2} - y_{rrc1})^2} \right| \quad (10)$$

정합값에 의해 결정된 두 눈 값의 신뢰도를 높이기 위해 두 개의 매개 변수를 제안하였다. 두 개의 매개변수는 눈 사이의 거리(distance)와 기울어짐을 나타내는 각도(angle)를 정의하였다. 그림 15는 두 매개변수의 기하학적인 관계를 보여준다.

얼굴영역의 특징정보로 오른쪽과 왼쪽 영상의 눈, 입 영역의 좌표 값을 얻기 위해 검출된 오른쪽과 왼쪽 눈의 차 좌표값, 입의 차 좌표값을 얻는다. 눈에 대한 추정 좌표값은 표 1의 조건에 이용하여 오른쪽( $Right_{eye}$ )/왼쪽( $Left_{eye}$ ) 영상의 눈 좌표값을 구할 수 있으며, 식(11)을 통해서 구한다.

$$\begin{aligned} Right_{eye} &= (x_{eye_n} + x_{eye_r})^2 + (y_{eye_n} + y_{eye_r})^2 \\ Left_{eye} &= (x_{eye_n} - x_{eye_r})^2 + (y_{eye_n} - y_{eye_r})^2 \end{aligned} \quad (11)$$

여기서 오른쪽 영상 좌표( $x_{eye_n}, x_{eye_r}, y_{eye_n}, y_{eye_r}$ )와 왼쪽 영상 좌표( $x_{eye_n}, x_{eye_r}, y_{eye_n}, y_{eye_r}$ )는 눈의 좌표를 나타낸다. 눈의 정보를 특징으로서 사용하기 위해서

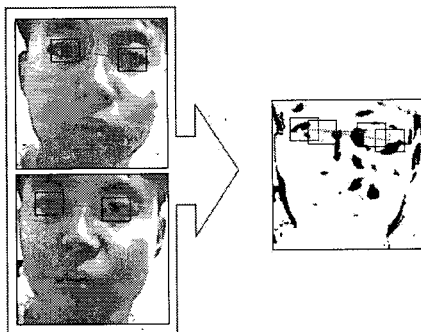


그림 14. 정합값으로 찾은 눈 영상

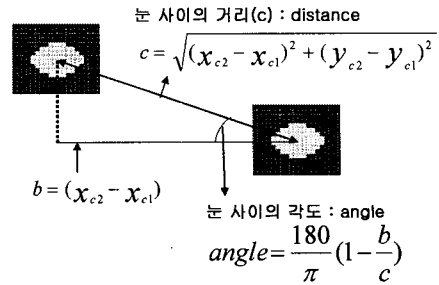


그림 15. 거리와 각도

눈 정보는 " $Eye_{info} = (Right_{eye} + Left_{eye})/2$ "로 획득한다. 입에 대한 추정값은 눈의 중앙 좌표값을 통하여 x, y의 좌표값을 구한다.  $x_c = ((x_2 - x_1)/2) + x_1$ 는 두 눈사이의 중앙 좌표값이며,  $y_c = x_c * 1.38$ 는 y축에 대한 좌표값으로 실험치에 의한 가중치값을 곱한다. 따라서 입에 대한 정보는 추정된 입 위치를 중심으로  $60 \times 32$ 의 입 영역을 설정하고, 식 12와 같이 구한다. 여기서 오른쪽 영상 좌표( $x_{mouth_n}, x_{mouth_r}, y_{mouth_n}, y_{mouth_r}$ )와 왼쪽 영상 좌표( $x_{mouth_n}, x_{mouth_r}, y_{mouth_n}, y_{mouth_r}$ )는 입의 좌표를 나타낸다. 이 크기는 눈 형판 및 얼굴 크기, 입 검출의 효율성 등을 고려하여 결정한 것이다. 그림 16은 검출된 눈과 입 영역을 보인 것이다. 입 영역 검출 판별 여부는 입 영역내에서 화소값들의 명암값으로 결정하였는데, 실험적으로 최적화된 값 60을 택하였다. 식 13은 추출된 눈의 중앙 좌표에서  $y_c = x_c * 0.75$ 값으로 코의 위치를 찾으며, 추정된 코 위치를 중심으로  $48 \times 25$ 의 코 영역을 설정하였다. 여기서 오른쪽 영상 좌표( $x_{nose_n}, x_{nose_r}, y_{nose_n}, y_{nose_r}$ )와 왼쪽 영상 좌표( $x_{nose_n}, x_{nose_r}, y_{nose_n}, y_{nose_r}$ )는 코의 좌표를 나타낸

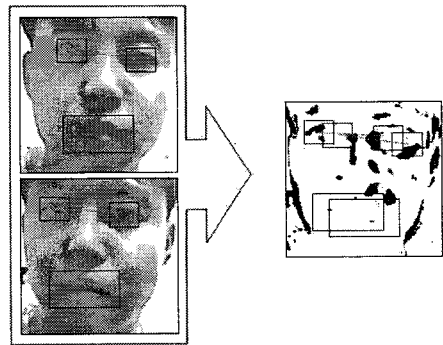


그림 16. 눈과 입의 검출



다. 그림 17은 눈, 코와 입의 위치를 찾은 영상이다.

$$\begin{aligned}
 &Right_{mouth} = \\
 &\quad (x_{mouth_{r1}} + x_{mouth_{r2}})^2 + (y_{mouth_{r1}} + y_{mouth_{r2}})^2 \\
 Left_{mouth} &= (x_{mouth_{l1}} + x_{mouth_{l2}})^2 + (y_{mouth_{l1}} + y_{mouth_{l2}})^2 \quad (12)
 \end{aligned}$$

$$\begin{aligned}
 &Right_{nose} = (x_{nose_{r1}} + x_{nose_{r2}})^2 + (y_{nose_{r1}} + y_{nose_{r2}})^2 \\
 Left_{nose} &= (x_{nose_{l1}} + x_{nose_{l2}})^2 + (y_{nose_{l1}} + y_{nose_{l2}})^2 \quad (13)
 \end{aligned}$$

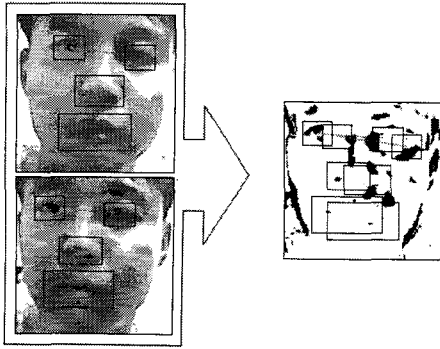


그림 17. 눈, 코와 입의 검출

### 4. 실험 및 고찰

#### 4.1 실험 결과 및 분석

본 논문의 실험은 얼굴특징 추출 조건으로 눈 검

표 2. 제안된 방법으로 찾아진 얼굴 조건

	얼굴 조건
눈 거리	10 ≤ 눈 ≤ 16
눈 기울기	0 ≤ 기울기 ≤ 10
눈 좌표 정보	10 ≤ 눈 좌표 정보 ≤ 20
코 좌표 정보	0 ≤ 코 좌표 정보 ≤ 10
입 좌표 정보	8 ≤ 입 좌표 정보 ≤ 12

출 과정에서 눈의 거리와 기울어짐, 눈, 입, 코 등의 좌표 추출 매개변수값과 왼쪽과 오른쪽 영상의 불일치(disparity) 정보를 추출한다. 매개변수들 중에서 눈의 기울기 변수는 특정한 조건 정보를 만들기 위해서 눈의 기울기 값에 100을 곱해서 사용한다. 또한 매개변수는 모의실험을 통한 최적화된 값들을 선택할 수 있는 기준을 표 2에 제시하였다.

그림 18은 모의실험에 사용된 스테레오 영상의 표본 입력영상이다. 본 논문에서 사용된 스테레오 영상은 특징 정보의 부분적인 특징이 강한 블록마다 명암값을 고려하여 20%미만의 임계값(60)을 설정해서 이진화값을 계산하게 된다. 그림 19는 제안된 알고리즘을 적용한 후의 결과 영상이다.

표 3은 제안된 알고리즘을 적용한 후에 각 표본 영상들에 대한 눈 영역(20×16), 코 영역(32×20), 입 영역60×32의 특징은 반전된 영상에서 각 픽셀수의 값을 더하여 평균을 나타낸 것이다. 입의 매개변수 값들과 각 매개변수 값을 모두 더해서 100으로 나누

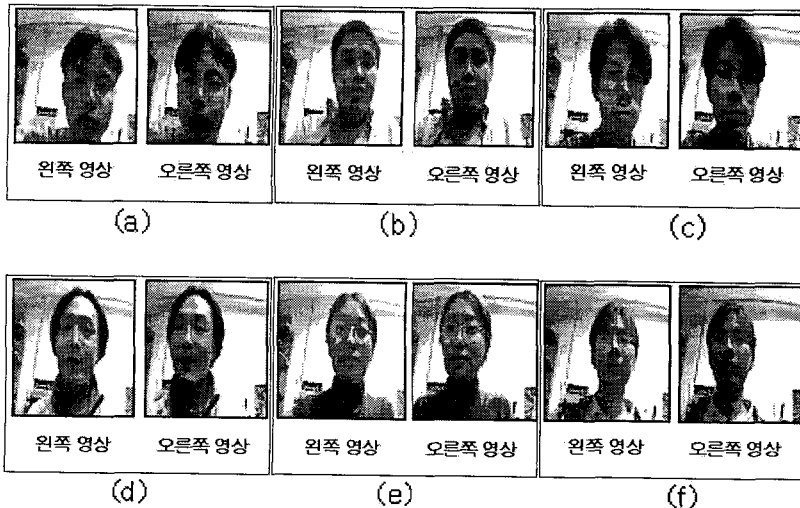


그림 18. 입력 스테레오 영상  
(a).(b).(c).(d).(e).(f) : 입력 영상

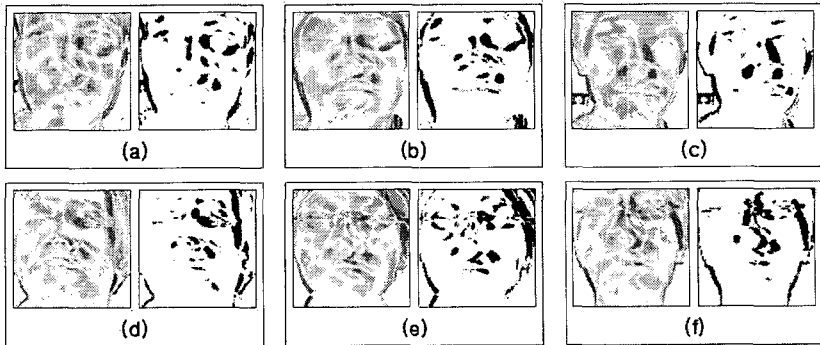


그림 19. 특징정보가 추출된 영상

어 처리된 영상의 결과 값들을 결과 (1)에 보이고, 결과 (1)와 특징 정보 값을 합한 최종값을 결과 (2)에 제시하였다.

표 4는 10명의 200개 영상의 얼굴영역에서 특징정보간의 거리값 통해 일치율을 나타낸 것이다. 표 4를 보면 비교적 가까운 약 1m 이내 거리 영상의 경우 약 73% 가량의 일치율을 보이며, 약 1m 이후의 거리가 멀어질수록 약 52% 가량으로 일치율이 급격히 저하됨을 볼 수 있다. 거리의 변화에 따른 일치율의 저하 요인으로는 거리가 멀어짐에 따라 스테레오 영상의 불일치가 로그 스케일(Log scale)로 감소하기 때문으로 볼 수 있으며, 그림 20에서 조명 변화, 특징정보의 사라짐, 눈을 검출하지 못한 경우에 발생하는 것을 볼 수 있다.

### 5. 결 론

본 논문에서는 스테레오 영상의 정합값을 통한 얼굴 특징추출 알고리즘을 제안하였다. 제안된 방법에서는 RGB컬러공간에서 YCbCr공간으로 변환하여

얼굴의 영역을 분리하고, 눈 형판을 적용하여 눈의 위치와 입, 코의 검출한다. 또한 얼굴의 특징정보를 이용함으로써 2차원 정보에서의 제약 조건 약화의 문제해결과 눈, 코, 입의 얼굴정보를 추출할 수 있었다. 결과적으로 얼굴판단에서 선글라스, 마스크 등으로 가려진 비정상적인 얼굴을 제외하고는 얼굴영상에서 눈, 코, 입의 기하학적인 특징정보를 찾을 수 있었다. 실험을 통하여 약 1m 이내와 1m 이후의 영상에 대한 일치율과 불일치율이 각각 73%, 52%을 얻을 수 있었다.

따라서 본 논문에서 찾아진 얼굴영상에 대해서는 얼굴특징의 강건함을 보일 수 있지만 입력영상을 획득할 때 정면얼굴의 크기나 영상의 밝기 등의 제약조건이 만족되어야 한다. 또한 형판의 크기가 가변적이지 아니기 때문에 카메라와 얼굴과의 거리 문제, 영상 내에서 얼굴의 크기가 작을 때와 크게 좌·우로 돌려진 얼굴 영상, 고개를 많이 숙이거나 올리는 것 등의 문제를 해결해야 한다. 그리고 스테레오 영상에서 카메라와 얼굴과의 거리를 계산해서 보정하는 방법 및 가변 형판 적용과 얼굴의 포즈에 대한 문제를 해결하

표 3. 각 표본 영상의 매개변수와 결과값

입력 영상	눈 거리	눈 각도	눈 특징	코 특징	입 특징	결과 (1)	특징 정보	결과 (2)
(a)	15.23	0.33	15.07	6.76	10.40	0.4779	0.7148	1.1927
(b)	14.52	×	15.34	7.09	10.57	0.4752	0.7701	1.2453
(c)	12.96	×	12.47	6.41	9.77	0.4161	0.7512	1.1673
(d)	16.00	×	14.99	9.81	5.97	0.4677	0.6164	1.0841
(e)	16.19	0.05	15.36	6.36	10.20	0.4816	0.8930	1.3746
(f)	16.16	×	16.57	6.71	10.04	0.4948	0.7848	1.2796

표 4. 추출된 얼굴 특징 정보의 결과

(a) 약 1m이내 거리 영상의 실험 결과

특징 정보	일치율(%)	일치시 평균오차(mm)	불일치시 평균 오차(mm)
왼쪽 눈 - 왼쪽 코	75.3	3.456712	10.546821
오른쪽 눈 - 오른쪽 코	76.1	3.184526	11.246578
왼쪽 눈 - 입 왼쪽	77.2	3.954126	8.428438
오른쪽 눈 - 입 오른쪽	77.9	3.673214	9.927242
입의 폭	70.0	3.175126	8.765184

(b) 약 1m이후 거리 영상의 실험 결과

특징 정보	일치율(%)	일치시 평균오차(mm)	불일치시 평균 오차(mm)
왼쪽 눈 - 왼쪽 코	62.2	4.845748	20.956127
오른쪽 눈 - 오른쪽 코	63.1	4.769852	22.248635
왼쪽 눈 - 입 왼쪽	64.1	4.875632	18.687950
오른쪽 눈 - 입 오른쪽	62.7	4.925435	19.965723
입의 폭	51.3	4.957829	33.278946

(c) (a)와 (b)의 평균

특징 정보	일치율(%)	일치시 평균오차(mm)	불일치시 평균 오차(mm)
왼쪽 눈 - 왼쪽 코	68.75	4.151230	16.747607
오른쪽 눈 - 오른쪽 코	69.6	3.977189	11.246578
왼쪽 눈 - 입 왼쪽	70.65	4.414879	13.558194
오른쪽 눈 - 입 오른쪽	70.3	4.299325	14.946483
입의 폭	60.65	4.066478	21.022065

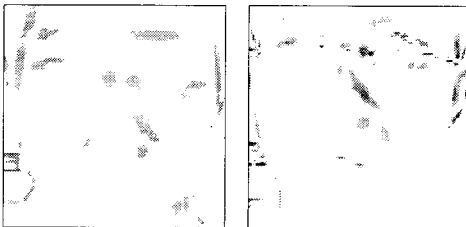


그림 20. 오류의 영상

면 향후 얼굴 인식에 3차원 정보를 이용할 수 있을 것이다.

참고 문헌

[1] H. Erik, L. B. Kee, "Face detection: a survey," *Computer Vision, Image Understanding*, vol. 83, no. 3, pp. 236-274, September 2001.  
 [2] C.H. Lin and W. Ja-Ling, "Automatic facial feature extraction by generic algorithms," *IEEE Trans. Image Processing*, vol. 8, no. 6, pp. 834-845, June 1999.  
 [3] G. Marcone, G. Martinelli, and P. D. Angelo, "Automatic face image recognition: an ac-

curate eye location procedure," *Proc. IEEE 13th Conf. Digital Signal Processing*, vol. 2, pp. 515-519, June 1997.  
 [4] B.A. McLindin, "Baselining illumination variables for improved facial recognition system performance," *Video/Image Processing and Multimedia Communications*, 2003. 4th EURASIP Conference, vol. 1, pp. 417-422, 2-5 July 2003.  
 [5] 이학찬, 박장한, 남궁연, 남궁재찬, "스테레오 영상을 이용한 물체 추적 방법", 대한전자공학회, 제39호 SP편 제5호, pp. 522-534, 11. 2002.  
 [6] W.A. IJsselsteijn, H. Ridder, and J. Vliegen, "Subjective evaluation of stereoscopic images: effects of camera parameters and display duration," *IEEE Trans., Circuits, Systems for Video Technology*, vol. 10, no. 2, pp. 225-233, March 2000.  
 [7] M.H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Trans., Pattern Analysis, Machine Intelligence*, vol. 24, no. 1, pp. 34-58, January 2002.

[8] G. Qian, S.Z. Li, "Combining feature optimization into neural network based face detection," *Pattern Recognition, 2000. Proc. 15th International Conference*, vol. 2, pp. 814-817, 3-7 September 2000.

[9] C. Nakajima, M. Pontil, T. Poggio, "People recognition and pose estimation in image sequences," *Neural Networks, 2000. IJCNN 2000*, vol. 4, pp. 189-194, 24-27 July 2000.

[10] Y. Hongxun, L. Mingbao, Z. Lizhuang, "Eigen features technique and its application," *Signal Processing Proc., 2000. WCCC-ICSP 2000*, vol. 2, pp. 1153-1158, 21-25 August 2000.

[11] Y. Zhong, A.K. Jain, D. Jolly, "Object tracking using deformable templates," *IEEE Trans., Pattern Analysis, Machine Intelligence*, vol. 22, no. 5, pp. 544-549, May 2000.

[12] T. Liang, H.K. Kwan, "Automatic localization of human eyes in complex background," *IEEE International Symposium, Circuits and Systems, 2002. ISCAS 2002*. vol. 5, pp. 669-672, 26-29 May 2002.

[13] L. Chengjun, H. Wechsler, "Independent component analysis of Gabor features for face recognition," *IEEE Trans., Neural Networks*, vol. 14, no. 4, pp. 919-928, July 2003.

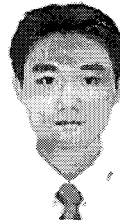
[14] R.S. Feris, T.E. Campos, R.M. Cesar, "Detection and tracking of facial features in video sequences," *Proc. Mexican International Conf. on Artificial Intelligence*, vol. 1793, pp. 197-206, April 2000.

[15] R.S. Feris, R.M. Cesar, "Tracking facial features using gabor wavelet networks," *Proc. 13th Brazilian Symposium on Computer Graphics and Image Processing*, pp. 22-27, September 2000.

[16] L.H. Chen and W.C. Lin, "Visual surface segmentation from stereo," *Image and Vision*

*Computing*, vol. 15, no.2 pp. 95-106, February 1997.

[17] D. Chai and K.N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Trans. Circuits, systems for video technology*, vol. 9, no. 4, pp. 551- 564. June 1999.



**김 상 명**

2001년 영동공과대학교 전자공학과 졸업  
 2003년 광운대학교 대학원 컴퓨터공학과 공학석사  
 2004년~현재 ㈜모빌링크 삼성 CDMA

관심분야 : 스테레오비전, MPEG2, 영상처리, 영상인식



**박 장 한**

1997년 방송통신대학교 전자계산학과 독학학위취득 (이학사)  
 1999년 광운대학교 전산대학원 소프트웨어공학과 이학석사  
 2005년 광운대학교 대학원 컴퓨터공학과 박사

2005년~현재 중앙대학교 첨단영상대학원 Post-Doc.  
 관심분야 : 스테레오비전, 3D 영상, 얼굴인식, 영상처리, 영상인식, 신경망, MPEG, Tracking



**남공재찬**

1970년 인하대학교 전기공학과 졸업  
 1976년 인하대학교 대학원 전자공학과 석사  
 1982년 인하대학교 대학원 전자공학과 박사  
 1979년~현재 광운대학교 컴퓨터

공학과 교수  
 관심분야 : 신경망, 영상처리, 문자인식, 패턴인식, 영상통신, MPEG, 3차원비전