

인간-로봇 상호작용을 위한 자세가 변하는 사용자 얼굴검출 및 얼굴요소 위치추정

Face and Facial Feature Detection under Pose Variation of User Face for Human-Robot Interaction

이 태 근*, 박 성 기, 박 민 용
(Taigun Lee, Sung-Kee Park, and Mignon Park)

Abstract : We present a simple and effective method of face and facial feature detection under pose variation of user face in complex background for the human-robot interaction. Our approach is a flexible method that can be performed in both color and gray facial image and is also feasible for detecting facial features in quasi real-time. Based on the characteristics of the intensity of neighborhood area of facial features, new directional template for facial feature is defined. From applying this template to input facial image, novel edge-like blob map (EBM) with multiple intensity strengths is constructed. Regardless of color information of input image, using this map and conditions for facial characteristics, we show that the locations of face and its features - i.e., two eyes and a mouth - can be successfully estimated. Without the information of facial area boundary, final candidate face region is determined by both obtained locations of facial features and weighted correlation values with standard facial templates. Experimental results from many color images and well-known gray level face database images authorize the usefulness of proposed algorithm.

Keywords : face and facial features detection, pose variation, edge-like blob map, directional template, eye pairs

1. 서론

최근 로봇 연구에 있어 한가지 연구경향은 지능화된 서비스 로봇이다. 이러한 지능형 로봇에서 매우 중요한 기능 중 하나는 인간-로봇 상호작용(Human-Robot interaction, HRI)기술의 구현이다. HRI에 있어서 카메라를 이용한 얼굴영상 인터페이스(Fig.1)는 최소한의 사용자 협조만이 요구되며, 폭넓은 정보 획득이 가능한 인터페이스로 인간과 로봇간의 감성정보 교류와 같은 수준높은 상호작용 구현에 적합하다. 이를 위해서는 사용자 얼굴 및 표정인식이 중요한 요소기술이며, 여기에서는 필수적으로 얼굴검출 및 얼굴요소의 정확한 추정방법이 선행적으로 요구된다. 본 연구에서는 이러한 HRI 관점에 적합한 새로운 얼굴 위치검출 및 얼굴요소 위치 추정 방법을 제안한다.

기존의 일반적인 얼굴검출은 정지 카메라에 기반한 연구가 대부분[1,2,3,4,5]이나 HRI를 위한 얼굴검출 기술에서는 그 구현에 있어 또 다른 특성이 존재한다. 우선, 획득된 영상내의 얼굴은 로봇 플랫폼의 이동성(mobility)으로 인해 정면이 아닌 상당한 자세변화를 가진다. 또한, 로봇 주위 공간에 위치하는 사용자 얼굴영상에는 임의의 복잡한 뒷배경을 가지게 되며 상당한 조명변화가 포함된다. 뿐만 아니라, 연속적으로 변화하는 상황에서의 사용자 인지가 요구되므로 검출작업은 되도록 실시간으로 이루어져야 한다. 본 논문에서는 이러한 검출상의 한계조건을 극복하기 위한 새로운 얼굴 및 요

소검출 방법을 제시한다. 먼저, 주요 얼굴요소인 눈 영역쌍(pair) 및 입 영역 위치의 명도특성(intensity characteristics)에 근거한 방향성 템플릿을 정의하고, 이를 영상에 적용하여 만들어지는 유사 에지 영역맵(edge-like blob map)을 새롭게 제안한다. 이를 통해서 조명변화 및 얼굴의 자세변화에 비교적 강인하며 대강의 얼굴영역 정보가 없이도 세부적인 얼굴요소의 효과적 위치추정이 가능함을 보인다. 또한, 본 제안방법은 카메라에서 얻어지는 얼굴영상의 컬러특성 여부에 상관없이 얼굴검출이 가능하여 입력영상 조건에 유연한 대응이 가능한 장점을 갖는다. 이 방법을 널리 알려진 얼굴영상 데이터베이스(그레이 영상)와 다양한 컬러영상에서 수행된 결과를 통해 그 타당성을 보이도록 한다.

본 논문의 구성은 2장에서는 얼굴 및 요소검출에 관한 기존연구를 소개하고, 3장에서는 제안하는 구체적 방안을 설명한다. 4장에서 실험을 통한 수행결과를 보이고 5장에서 결론을 제시하기로 한다.

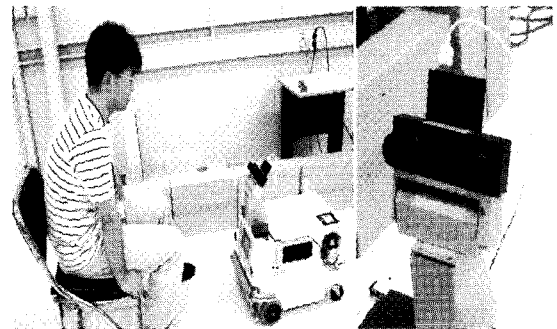


그림 1. 인간-로봇 상호작용의 예 : 실내용 이동로봇에서의 사용자 얼굴검출 시스템.

Fig. 1. Human-robot interaction application-face detection system with indoor mobile robot.

* 책임저자(Corresponding Author)

논문접수 : 2004. 6. 28., 채택확정 : 2004. 10. 8.

이태근, 박민용 : 연세대학교 전기전자공학부

(dungding@kist.re.kr/mignpark@yonsei.ac.kr)

박성기 : 한국과학기술연구원 지능로봇연구센터(sksee@kist.re.kr)

※ 본 연구는 산업자원부 지원으로 수행하는 21세기 프론티어 연구 개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다.

II. 기존 연구

기존의 얼굴 검출 방안[1,2,6]은 영상내 얼굴 위치검출 성능에 한정시켜, 완결된 개별 방법으로서의 의미만을 가진다. 그러나 본 연구에서의 구현하려는 얼굴검출 작업은, 전체적인 HRI 큰 관점에서 볼 때, 얼굴 및 표정인식을 위한 사전단계로 고려된다. 따라서, 이러한 특성상 임의의 비제한적 배경을 갖는 영상에서 얼굴의 효과적 검출 뿐만 아니라 얼굴내의 주요 요소위치에 대한 정확한 추정이 동시에 요구된다. 한편, 얼굴요소 검출에 관련된 많은 연구 중에서 포즈가 변화하는 얼굴내 위치검출에 대한 최근의 가장 대표적인 연구로는 얼굴색 컬러 영상에 기반한 방법[7]이 있다. 여기서는 조명에 따른 컬러보정(compensation)을 수행시킨 컬러영상을 이용하여 각 얼굴요소의 컬러 색도 맵을 작성, 이에 따른 포즈 변화하는 얼굴 내의 요소를 찾는다. 그러나 실제 수행결과에서 카메라의 색도 특성이 좋지 않은 경우에는 컬러 색도 맵에서 얼굴 요소의 위치가 명확하게 구분하기 어렵고, 더욱이 영상내에 얼굴 크기가 작은 경우이나, 눈 또는 입의 크기가 작을 경우 이러한 구분이 더욱 어렵다.

이러한 경우에서 얼굴요소의 특징은 이 얼굴영상에 마스크를 씌워 구성된 에지(edge)영상에서 보다 잘 표현된다. 이에 관련된 최근의 연구방법으로는 edge oriented map을 사용하여 얼굴내 요소들을 비롯한 에지 성분의 전반적 특성을 파악 미터화 하는 접근방안[8], line edge map을 정의하여 눈을 비롯한 얼굴 요소 부근의 에지 성분 특징을 얻어내어 얼굴검출 및 인식을 수행[9]하는 연구등이 있어왔다. 하지만 이러한 방안들은 대체로 정면영상 관점에서 전체적 에지 맵을 구성하고, 이를 바탕으로 유사도를 측정하기 때문에 본 논문의 경우와 같이 HRI 응용에서 포즈가 변화하는 얼굴 영상의 경우에 적용할 경우 얼굴요소 검출율의 상당한 저하를 보였다.

III. 얼굴 및 요소검출 방안

본 논문의 얼굴 및 요소검출 알고리즘의 전체구성은 Fig. 2와 같다. 본 논문에서 제안하는 방법의 특징 중 하나는 컬러영상 뿐만 아니라 그레이영상에서도 얼굴검출 수행이 가능하다는 점이다. 따라서 영상 입력특성에 따라 추가적인 작업 단계를 이용, 효과적으로 얼굴 및 요소검출을 수행하도록 한다.

1. 얼굴가능 영역범위의 빠른 추정

입력영상이 컬러이미지일 경우에는 얼굴색 컬러정보를 이용한 얼굴영역 추정단계를 통해 보다 빠른 얼굴 및 요소검출이 가능하도록 한다. 즉, 이 단계는 얼굴색 컬러특성을 통해 영상내의 얼굴영역 검색범위를 효과적으로 줄이게 되어, 전체 얼굴 및 요소 검출과정을 실시간에 가깝게 수행하는 것이 가능하게 한다. 사람의 얼굴색 컬러특성은 인종차이, 다양한 조명조건 및 얼굴색의 밝기 차에 따라 달라지며, 이를 컬러공간에서 표시할 경우 매우 다양한 형태로 분포된다. 얼굴색 특성을 나타내는 입력영상내 픽셀 컬러성분의 보다 효과적인 구분(segmentation)을 위해서 본 논문에서는 YCbCr 컬러공간 변환[4,10]을 사용하기로 한다. 일반적인 RGB컬러공간에 비해 YCbCr 공간에 분포하는 얼굴색 특성픽셀은 보다 밀집된 형태[4,10]로 나타나며, 이로인해 간단한 컬러공간 경계조건 만으로 얼굴색/비얼굴색 픽셀을 구분한 얼굴색 이진영상

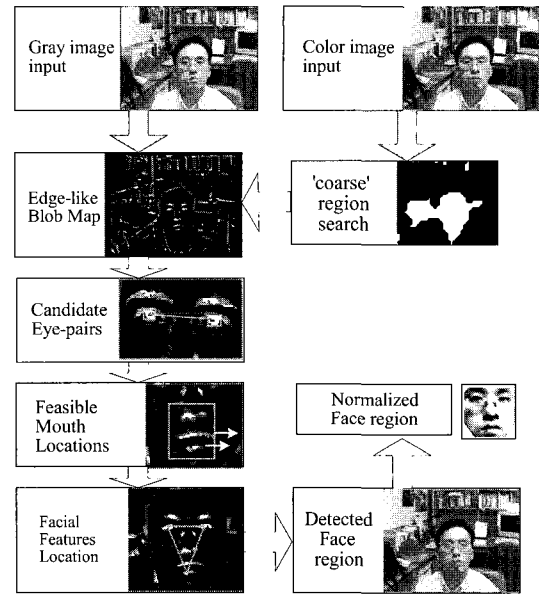


그림 2. 얼굴 및 요소검출 알고리즘의 개요.

Fig. 2. Summary of face detection and facial feature localization algorithm.

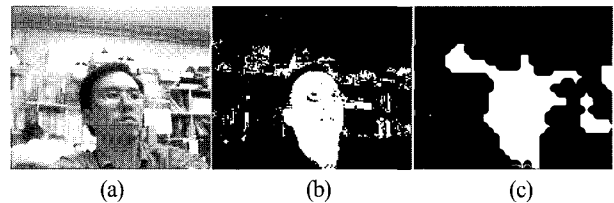


그림 3. 컬러입력에서 얼굴가능 영역의 전처리 작업 (a) 원 영상, (b) 이진화된 영상, (c) 변형 모폴로지 작업후 얼굴가능 범위 이진 영상.

Fig. 3. Preprocessed region of face candidate area (a) original image (b) binary image by facial color (c) tolerant region of candidate face location by modified closing operation.

(Fig. 3(b) 참조)을 얻을 수 있게 된다[10].

얻어진 얼굴색 이진영상에 대해 웨이블릿 변환기법을 이용한 계층적 피라미드 영상과 새로운 유연한 모폴로지 영역 처리기법을 사용하여 얼굴이 존재가능한 영역범위를 빠르게 찾아내도록 한다.

먼저, 가장 간단한 웨이블릿 형태인 Haar wavelet을 이용하여 이진영상 크기를 효과적으로 줄인 계층적 피라미드 영상 [10]을 만든다. 이를 이용하면 원 영상크기에서 찾는 것보다 적은 개수의 스케일 변화로 적절한 크기의 얼굴영역 범위를 빠르게 찾아낼 수 있다.

본 논문에서 제안하는 방법의 특성상 이 단계에서는 정확한 얼굴영역 결정을 필요로 하지 않는다. 따라서, 대강의 유연한 영역으로 얼굴의 존재가능범위를 빠르게 결정하기 위해 Eq.(4)와 같은 변형된 모폴로지 닫기연산(modified closing operation)을 사용한다. (모폴로지 수행식에 대한 보다 자세한 내용은 [11] 참조) 수행결과 피부색 컬러부근의 유사한 컬러영역까지 비교적 폭넓게 포함하게 되는 형태 (Fig. 3(c)참조)의 얼굴가능 영역범위가 선택된다.

$$\begin{aligned}
 & \mathbf{X}^{\text{mod}CL} = (\mathbf{X} \oplus^m \mathbf{B}) \otimes^n \mathbf{A} \quad (\text{modified closing operation}) \\
 & (\mathbf{X}: \text{facial image}, \oplus: \text{dilation}, \otimes: \text{erosion}) \quad (4) \\
 & \mathbf{X} \oplus^m \mathbf{B} \equiv \mathbf{X} \oplus \mathbf{B} \text{ only if } (\{x: B_x \subset X\} \text{ AND } \{m: B_m \subset X, m \geq 5\}) \\
 & \mathbf{X} \otimes^n \mathbf{A} \equiv \mathbf{X} \otimes \mathbf{A} \text{ only if } (\{x: A_x \cap X \neq \emptyset\} \text{ AND } \{x \geq n = 7\}) \\
 & (\mathbf{A}, \mathbf{B}: 3 \times 3 \text{ masks}, \mathbf{A} = \{255, 255, \dots, 255\}, \mathbf{B} = \{0, 0, \dots, 0\})
 \end{aligned}$$

2. 얼굴 요소 검출을 위한 전처리 작업

본 절에서는 얼굴요소를 효과적으로 검출하기 위한 전처리 단계로서 새롭게 제안하는 유사에지 영역맵(edge-like blob map)의 구성과정을 설명하기로 한다.

얼굴 내에는 그 특성을 나타내는 여러가지 요소들(예: 두 눈, 눈썹, 코, 입, 코 끝, 양쪽 귀 등)이 있다. 그 중에서 두 눈(eye pair)은 매우 두드러진 특징 요소로서 여러 연구[12,13]에서 가장 중요한 얼굴검출 단서로 사용되었다. 입(mouth) 역시 매우 두드러지나 얼굴요소들 중에서 형태상의 가변성이 가장 큰 요소로 일반적인 특성기반 모델링 방법으로는 표현이 어렵다. 입 위치와 관련된 연구로는 비교적 고 해상도인 영상에서 입 양쪽 끝의 특징점을 찾아 얼굴인식에 적용한 예 [5], 입술 부근의 적색이 많이 검지되는 컬러특성을 이용한 위치 추정연구[7] 정도가 있어왔다. 본 논문에서는 입력영상의 컬러특성 여부에 상관없이 얼굴요소의 명도특성(gray intensity)만을 이용한 위치추정방법을 제시한다. 두 눈, 입과 같은 얼굴요소 부근 픽셀의 명도값은 주변 얼굴영역보다 어두운 특성을 가지며, 대개의 경우 이러한 어두운 영역은 수평방향으로 보다 길게 나타나는 흡사 수평에지와 유사한 'blob'형태의 영역으로 얻어진다. 이러한 명도특성을 고려하여 효과적인 얼굴요소 표현을 위한 방향성 템플릿(directional template)을 정의한다. 그리고 이를 이용하여 유사에지 영역맵을 구성한다. 먼저, 얼굴요소 방향성 템플릿을 정의한다. 검출하려는 얼굴크기를 고려하여 크기를 결정하되, 눈 영역과 비슷한 크기로 가로가 더 길게 정한다(Fig. 4참조). 이 템플릿을 입력영상에 적용하여 모든 픽셀에 대한 명도편차를 다음과 같이 계산한다. 크기가 $W \times H$ 인 영상 안의 픽셀 $P(x, y)$ 에서 템플릿 중심이 놓이는 픽셀 $P_{cent}(x_c, y_c)$ 기준으로 방향성 템플릿(크기: $h_{FF} \times w_{FF}$)의 8방향으로 주변픽셀들의 평균밝기 \bar{I}_i (Eq.5)를 계산하고, 중심픽셀 P_{cent} 의 밝기 I_{cent} 와의 명도편차 ΔI_j 를 Eq.6와 같이 얻는다. 이러한 방향성 템플릿을 통한 명도편차를 계산하는 예를 Fig. 4에 나타내었다. 그 결과 가장 밝기편차가 큰 쪽의 명도편차크기 I_{pr} 를 얻게된다.

$$\begin{aligned}
 & \bar{I}_i \quad (i=1, 2, \dots, 8) \\
 & ; \text{avg. intensity of neighborhood pixels in template for each direction} \quad (5) \\
 & (\text{directions: } i=1 \text{ (left), } 2 \text{ (right), } 3 \text{ (top), } 4 \text{ (bottom),} \\
 & \quad \quad \quad 5 \text{ (NW), } 6 \text{ (SE), } 7 \text{ (NE), } 8 \text{ (SW)})
 \end{aligned}$$

$$\begin{aligned}
 & \text{Ex) } \bar{I}_1 = \left\{ \sum_{i=-w_{FF}/2}^{i=w_{FF}/2} \sum_{j=y_c}^{j=y_c+k} P(x+i, y) \right\} / (w_{FF} / 2) \\
 & \Delta I_1 = \text{avg} [I_{cent} | \bar{I}_1, \bar{I}_2] = (|I_{cent} - \bar{I}_1| + |I_{cent} - \bar{I}_2|) / 2 \\
 & \Delta I_2 = \text{avg} [I_{cent} | \bar{I}_3, \bar{I}_4], \Delta I_3 = \text{avg} [I_{cent} | \bar{I}_5, \bar{I}_6], \Delta I_4 = \text{avg} [I_{cent} | \bar{I}_7, \bar{I}_8] \quad (6)
 \end{aligned}$$

$$I_{pr} = \text{Max}_{j=1, \dots, 4} \{ \Delta I_j \} \quad (7)$$

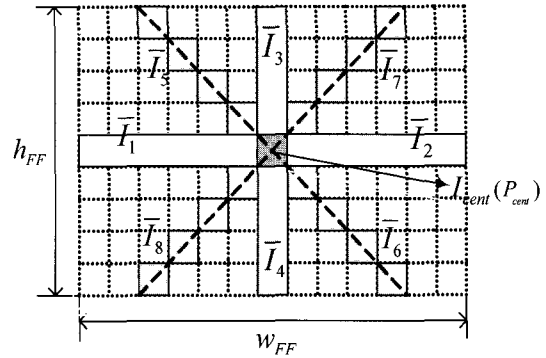


그림 4. 방향성 템플릿과 각 방향의 명도편차.
Fig. 4. Directional intensity template and the intensity differences of 8-directions from center pixel of template.

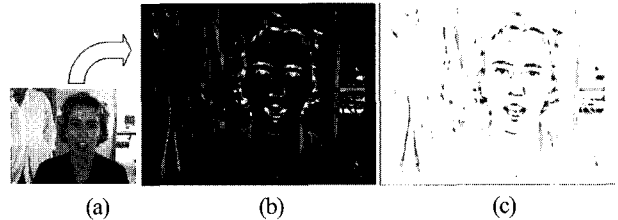


그림 5. 그레이 원 영상에서 작성된 유사에지 영역맵.
Fig. 5. Edge-like blob map from original gray image.



그림 6. 언어진eye-pair의 예(원래 크기 및 확대).
Fig. 6. Eye-pair detection example(original and enlarged eye patches with histogram equalization).

이제 현 위치에서 각 8방향의 밝기편차에 따라 해당 픽셀 위치에 다양한 밝기강도(multiple intensity strengths)가 부여되는, 유사에지 영역맵(edge-like blob map, EBM) 영상을 다음과 같은 방법으로 구성한다. 먼저 I_{pr} (Eq. 7)를 이용하여 영상내 해당 위치에서 수평방향으로 양쪽의 주위의 평균 명도편차 ΔI_j 가 해당가중치가 할당된 I_{pr} 의 크기(Eq. 8참조) 보다 클 경우, 이 픽셀위치에 +1레벨의 밝기강도를 할당한다. 이와 같이 나머지 세 방향으로도 계산된 평균명도차가 Eq.8과 같은 조건을 만족할 경우 각각 +1씩의 밝기강도를 부여한다. 이러한 과정을 통해서 전체 입력영상은 총5단계까지 다른 명도값 밝기강도를 가지는 영역맵(+4의 밝기강도 위치에 가장 밝은 명도값 부여)이 작성되며, 이는 에지영상과 유사한 형태의 gray-level 영역맵으로 만들어진다. Fig. 5(b)에서 표시되는 영역맵에서 매 단계의 밝기강도 명도값은 각각 0, 40, 80, 120, 200이다. Fig. 5(c)는 반전된 유사에지 영역맵 영상으로, 원 영역맵에서 쉽게 구분하기 힘든 +0 ~ +4단계의 다른 밝기강도가 보다 잘 나타나는 그림이다.

For each pixel $p(x,y)$ in input image,

if $\Delta I_{1,p(x,y)} > \alpha_{pr} |I_{pr}|$ then add (+1) level strength intensity at $p(x,y)$
 also if $\Delta I_{2,p(x,y)} > \beta_{pr} |I_{pr}|$ then add (+1) level strength intensity at $p(x,y)$
 also if $\Delta I_{3(4),p(x,y)} > \gamma_{pr} |I_{pr}|$ then add (+1) level strength intensity, each other
 (where $\alpha_{pr} = 1.0, \beta_{pr} = 0.9, \gamma_{pr} = 0.8$)
 (8)

3. 얼굴요소 검출1; 가능한 두 눈의 쌍(Eye-pair) 찾기

이 절에서는 얻어진 유사에지 영역맵에서 눈 영역으로 볼 수 있는 에지 형태의 작은영역(blob)을 찾아, 얼굴의 기하학적 조건을 만족시키는 두 눈의 쌍(pair)을 결정하는 과정을 수행한다. 두 눈의 경우 주변보다 상대적으로 어두운 특성이 보다 강하므로, 영역맵에서 밝기강도 레벨 +4인 픽셀로 구성되는 blob에 대해서만 눈 가능 위치로 선택한다. 먼저 레벨+4인 영역과 그 이하인 레벨 영역을 구분하고 component labeling 기법[10]을 통해 각 영역blob의 기하학적 특성정보를 얻는다. 먼저, 모든 blob에 대하여 기본 영역특성 조건(Eq.9)을 적용하여 적합한 눈 가능 영역blob들만 선택한다. 여기서, 얻어진 blob영역에 대한 윤곽 사각형(bounding rectangle)을 작성한 후 그 크기가 템플릿 수평크기 및 수직크기의 1.5배 이상일 경우이거나 너무 작은 noisy 영역일 경우에는 그 해당 blob영역을 제거하도록 한다.

select certain blob as 'candidate eye' blob (E.B.).
 only if $\{width_{E.B} < c \cdot w_H\} \cap \{height_{E.B} < c \cdot h_H\} \cap \{area_{E.B} > \varepsilon\}$ (9)
 (where $c = 1.5, \varepsilon = 6$)

이렇게 선택된 눈 가능 후보영역 blob들에 대해 짝지을 수 있는 모든 eye-pair를 구성하고, 만들어진 pair의 영역특성이 얼굴의 기하학적 특성 및 명도분포 특성을 만족하는지에 따라 유력한 eye-pair만을 선택한다. 기하학적 특성조건으로는 (1) 만들어진 eye-pair의 수평길이의 얼굴과의 적합 여부, (2) eye-pair의 방향특성, (3) 두 눈 blob영역의 넓이비 등이 고려되며, 명도 분포 특성으로는 (4) 두 눈 영역의 명도값 비율, (5) 두 눈과 그 미간의 명도패턴 특성, (6) 두 눈 영역blob 주위의 명도패턴 특성등을 고려하여 선택한다. Fig.6은 다층 에지맵을 통해 선택된 유력한 두 눈 가능 영역과 이의 확대된 모습을 보여준다.

4. 얼굴요소 검출2; 입 가능 위치의 추정

얼굴자세가 변하는 경우에 얼굴범위를 정확히 포함하는 영역은 가변적 직사각형 형태를 갖게된다. 두 눈 사이의 수평길이를 기준으로 얼굴영역의 수직방향 길이는 대개 더 길고 오히려 짧을 수도 있다. 이때, 두 눈과 함께 주요한 얼굴요소인 입의 위치를 정확하게 추정하는 것은 정확한 얼굴영역 범위를 결정하기 위해 매우 중요하다. 이러한 표준화된 형태로 얼굴영역을 구성하는 과정은 이후의 얼굴인식과정 등에서 얼굴의 인식률을 높이기 위해서 필수적으로 요구된다. 앞에서 제시된 유사에지 영역맵은 얼굴 요소중 가장 두드러지는 두 눈 영역 뿐만 아니라, 비교적 가변적인 입 영역 위치를 추정하는 데에도 효과적이다. 영역맵 내에서 눈 영역과 비슷하게 입 영역에도 주변보다 어두운 blob영역이 상당히 존재한다. 그러나, 입 형태의 가변적인 특성 때문에 대개 눈 영역과 같은 강한 밝기강도를 보이지 못하고 다양한 강도

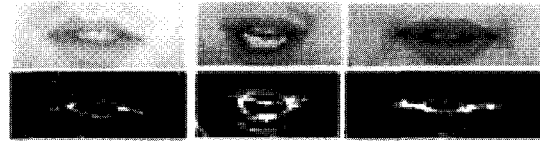
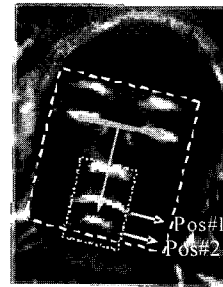


그림 7. 영역맵상에서 입영역 주위의 다양한 에지형태. Fig. 7. Various strength and shape of blobs near mouth area.



* Thick arrow - eye pair vector
 * Thin arrow - its normal vector
 * Dashed box - candidate area of face from eye and normal vector lengths
 * Dotted box - boundary of possible locations of mouth
 * position #1, position #2 - candidate positions of mouth (from below)

그림 8. 입 가능위치 추정과정의 도식. Fig. 8. Representation of estimating probable mouth locations.

- (1) Eye-pair의 위치 및 방향을 기준으로 입 방향의 수직벡터 결정
- (2) 입 존재 가능한 부분 영역 설정 (Dashed Box, in Fig. 8)
- (3) 입 가능 영역을 명도 표준화 및 수평회전 하여 구성
- (4) 영역 수평방향으로 에지강도 +1 이상인 픽셀이 일정 문턱치 (threshold) 이상일 경우, 입 가능한 후보위치로 선택
- (5) 영역 수직방향으로 눈 사이길이를 기준으로 정해진 두께 이하인 얇고 긴 형태의 입 후보위치 만을 선택



그림 9. 변형상관도 계산을 위한 표준얼굴 템플릿. Fig. 9. Some face templates for modified correlation.



그림 10. 검출된 얼굴영역 및 얼굴요소 위치 결과. Fig. 10. Detection result of face area and feature locations.

로 분포된다. 한편 입 영역에서는 대개의 경우 blob영역이 눈 부근에서 보다 더 선명하게 수평에지 형태로 나타난다. (Fig. 7참조) 유사에지 영역맵을 이용한 본 논문에서의 입 가능위치 추정과정은 이러한 수평에지 특성이 잘 나타나는 위치를 복수개로 얻고 이에 따른 얼굴영역을 선택하여, 이 영역을 이용한 최종적인 영역의 명도패턴 분석작업을 수행하도록 한다. 앞 절에서 눈의 경우 밝기강도 레벨 +4 이상인 경우의 에지형태blob만을 눈 가능영역으로 선택하였다. 그러나, 입 영역은 레벨 +3이하의 밝기강도를 가진 얇고 긴 형태의 에

지형blob이 있는 영역이 eye-pair방향과 평행하게 자주 나타난다(Fig. 7,8참조). 따라서, 입 가능 위치는 눈과 수평방향으로 상당한 밝기강도가 존재하는 위치로 추정한다. 아래에 입 가능위치 추정을 위한 세부적인 수행과정을 제시하였다.

5. 최종 얼굴영역 및 요소위치 결정

이 절에서는 마지막 단계로서 앞에서 결정된 두 눈 쌍의 위치와 입 위치를 이용하여 결정된 복수개의 얼굴 가능영역에 대해 변형된 상관계수(correlation coefficient)식을 사용하여 미리 작성된 표준 얼굴영상 템플릿과의 유사도(similarity)를 계산, 이를 통해 최종적인 얼굴범위 및 요소위치를 결정하는 과정을 설명한다. 먼저 얻어진 두 눈 사이의 길이를 기준으로 얼굴영역 폭을 정하고, 입까지의 거리를 나타내는 수직백터를 고려하여 세로방향 영역크기를 설정한다. 이러한 직사각형 형태의 얼굴영역을 가로, 세로방향으로 두 단계의 스케일 변형으로 정사각형 형태의 표준화 영역(patch)으로 만든다. 복수개의 입 영역 위치에 따라 한 eye-pair에 최대 3개 까지 가능한 얼굴영역으로 얻어지며 이 patch를 이용하여 표준 얼굴 템플릿과의 유사도를 계산한다. 유사도 계산을 위한 얼굴 영상 patch 과 얼굴 템플릿과의 상관계수 [14]은 Eq.10를 기본으로 사용한다. 단, 얼굴 주요요소인 두 눈과 입이 위치하는 위치부근 영역에 대해서는 상관도 계산시에 가중치를 부여하는 변형된 형태로 수행한다.

한편, 검출된 얼굴영역과 표준 얼굴영상 템플릿의 영상의 정사각형 영역 안에는 비얼굴부분이 불가피하게 포함된다(예: 뒷 배경, 얼굴주변부). 본 연구에서는 주변부를 잘라낸 'mask'형태로 patch영역을 설정하고 히스토그램 평활화 작업을 수행하여 표준화된 얼굴 patch 를 얻었다(Fig. 9,10참조). 이렇게 만들어진 patch영역에 대해 얼굴 템플릿과의 변형 상관도 값이 가장 높은 영역을 최종 얼굴영역으로 선택하고, 해당 patch 영역에서의 얼굴요소 위치를 원 입력영상에 삼각형(triangle) 형태로 표시하였다. Fig. 10은 최종적인 입력영상에서 검출된 얼굴영역 및 요소위치를 보인다.

$$\rho_{I_{FD}, I_{tmpl}} = \frac{E[I_{FD}I_{tmpl}] - E[I_{FD}] \cdot E[I_{tmpl}]}{\sigma_{I_{FD}} \sigma_{I_{tmpl}}} \quad (10)$$

(where I_{FD} : obtained facial area, I_{tmpl} : face templates)

IV. 실험 결과

폭넓은 수행실험을 통해 보다 객관적인 검출성능 평가를 위해 얼굴검출 실험세트를 세 종류로 구성하였다. 첫번째 세트는 실제 HRI구현상황에 적합한 상하좌우 임의의 방향으로 30° 내외의 자연스러운 얼굴 pose변화를 갖는 영상을 10사람에 대해 실험실 조명환경에서 얻었다. 이러한 pose 변화는 실제 실험에서 사용자의 두 눈과 입이 적절하게 영상에 나타날수 있는 최대 범위이다. 이때, 카메라에서부터의 관측거리를 0.6m, 0.9m, 1.2m로 바꾸어가며 10사람 * 9방향 * 3거리변화 = 총270장의 얼굴영상을 구성하였다. 가장 먼거리(1.2m)에서 관측되는 얼굴영역 내 얼굴요소간의 거리는 평균 22픽셀 내외이다. (얼굴영역 최소 크기: 평균 가로 40*세로46). 두번째 세트는 최근 얼굴검출 연구[12,13]에서 많이 사용되는

BioID face database[15]를 사용하였다. 이 DB는 총1521장의 단일인물 영상(384*288 pixel, grayscale)으로 구성되어 있으며, 23명의 사람에 대해 몇 개의 다른 장소에서 각각 다른시간에 획득한 영상들로 상당한 수준의 조명, 배경, 그리고 얼굴크기의 변화를 포함한다. 세번째 세트로는 널리 사용가능한 컬러 기반 얼굴검출 DB가 없는 관계로 CMU face images[16], Internet news sites[17]에서 선택적으로 얻은 얼굴영상으로 구성하였다. HRI의 작업환경과 유사한 단일인물 중심으로 자연스러운 포즈변화 및 조명변화가 포함된 총120장의 영상으로 만들어진다. 먼저 첫번째 테스트 세트에서 수행한 검출작업의 몇가지 예를 Fig. 13에 나타내었다. 카메라부터 얼굴까지 거리를 다양하게 변화시켜가며 얼굴검출 및 요소추정을 수행하였다. 두번째 테스트 세트에서 다단계 에지 맵을 통한 얼굴요소 위치를 찾는 결과는 Fig. 14에 나타내었다. 세번째 세트에서 다양한 크기의 영상에 대해 수행된 결과를 Fig.15에 나타내었다. 본 논문에서는 요소검출 수행결과의 객관적 제시를 위해 [13]에서 제시하는 상대적 에러측도(relative error measure)를 기본으로 사용하였다. 추정된 두 눈의 위치와 실제 위치간의 추정오차를 두 눈 사이거리를 기준으로 Eq.11과 같이 에러측도값 d_{eye} 를 정의한다.

$$d_{eye} = \frac{\max(d_l, d_r)}{\|C_l - C_r\|}$$

(C_l, C_r : true eye centers in facial image) (11)

(d_l, d_r : distance between C_l, C_r and estimated eye positions)

Table.1에서 세가지 테스트 세트에 대해 수행한 두 종류의 얼굴 및 요소검출 수행결과를 제시하였다.

기존 얼굴검출 연구[13]에서 사용한 에러측도값(=0.25)을 적용한 얼굴영역 검출률('DR')과 측도값=0.1을 채택한 얼굴요소 검출률('FDR')의 결과를 각각 나타내었다. DR은 일반적 얼굴검출 성공률을, FDR은 보다 정확히 얼굴요소 위치를 찾아 삼각형 형태로 나타낸 비율을 의미한다. 또한, 평균 수행시간을 함께 제시하였다 (P4-2.8G PC). 이러한 수행시간은 영상내에 얼굴요소로 추정가능한 에지 형태의 작은영역(3,4절 참조)의 개수에 비례한다. 컬러 영상의 경우 얼굴색 컬러특성을 통해 영역(blob) 검색범위를 줄이게 되어 수행시간이 실시간 처리에 적합한 수준(본 논문에서는 1초내외로 정한다.)으로 구현되며, 검출 얼굴영역의 크기 및 카메라와의 거리와는 무관하다. 두번째 테스트 세트의 경우 gray-level 기반 영상으로서 (3.1)절의 얼굴가능 영역범위의 빠른 추정과정을 수행할 수 없으므로 보다 많은 수행시간이 요구된다.

수행결과 얼굴요소 검출상에 오류가 발생하는 경우는 주로 눈이나 입이 명확히 드러나지 않는 영상에서 나타났다. 즉, 눈이 주변보다 어둡게 보이지 않는 경우로 안경의 반사효과(glaring)등으로 눈이 발견되기 힘든 경우와 수염 등으로 인한 입 영역 위치가 모호할 경우가 해당된다. 두번째 테스트 세트에 대해서는 이러한 경우의 영상을 배제한 검출 수행결과를 별도로 제시하였다. (제외한 구체적 영상범위: 320-417, 594-602, 1174-1213, 총 117 영상) 이 경우에 약 97% 정도의 얼굴 검출률과 94%의 얼굴요소 검출 성공률을 나타내었다.

표 1. 세 종류의 테스트 세트에서의 수행결과 요약.

Table 1. Summary of detection results for three test sets.

Test Set	#1 (Lab.)	#2 (Bio ID face database)		#3 (internet)
image origin	our Lab.	(all cases)	(Non-glaring case, only)	CMU faces, Internet news images,
No. of images	270 images	Total 1521 images	Total 1404 images	120 images
FDR (%)	96.67 (261/270)	90.53 (1377/1521)	94.02 (1320/1404)	89.17 (107/120)
DR (%)	98.52 (266/270)	94.74 (1441/1521)	96.94 (1361/1404)	93.33 (112/120)
avg.time (sec)	0.8 ± 0.2	2.08 ± 1.5	2.16 ± 1.4	1.0 ± 0.3

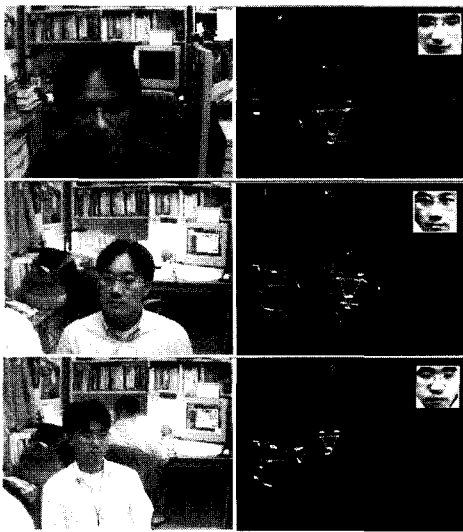


그림 13. 테스트 세트1에서의 수행결과 의 예.
Fig. 13. Some results of 'test set #1' (pose-variant faces).



그림 14. 테스트 세트2에서의 수행결과 의 예.
Fig. 14. Some results of 'test set #2'-BioID face database.



그림 15. 테스트 세트3에서의 수행결과 의 예.
Fig. 15. Some examples of 'test set #3'.

표 2. 기존 연구와의 검출성능 비교(for BioID DB).

Table 2. Comparison with detection result of previous studies.

Test Set	Jesorsky, et al. [13]	Fröba, et al. [8]	Proposed method
Relative error measures (Eq.12)	0.25 / -	0.30 / -	0.25 / 0.1
DR / FDR (%)	91.8 / -	95.67 / -	94.74 / 90.53

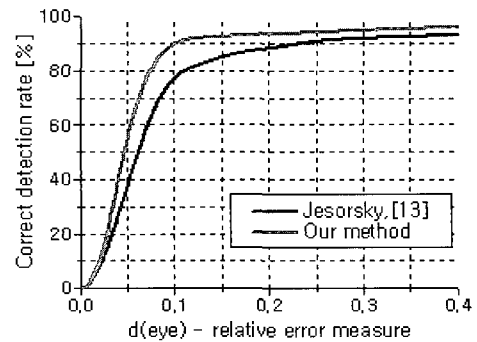


그림 16. 상대 에러측도와 검출률의 분포함수.
Fig. 16. Distribution function of relative error measure value against detection rates for test set #2 (BioID).

다음으로는 제안된 방법의 검출성능 평가를 위해 BioID face database에 대한 얼굴검출을 수행한 기존 연구[8,13]과 비교한 결과를 Table. 2에 나타내었다. 얼굴검출 수행결과를 살펴보면 기존 방법과의 상대적 에러측도의 차이(허용 에러측도값이 클수록 검출률 향상을 감안하여도 비슷하거나 보다 높은 검출 성공율을 보였다. 한편, 본 제안된 방안은 비교적 낮은 에러측도 만이 허용되는 강인한 검출조건(에러측도값 0.1이하)에서도 상대적으로 높은 검출성능 결과를 나타낸다 (Fig. 16참조). 또한, 기존 연구에 비해 얼굴검출 뿐만 아니라, 얼굴요소의 위치까지 정확한 추정기 가능하다는 상대적 장점을 갖는다.

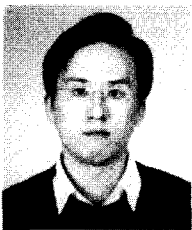
V. 결론

본 논문에서는 이동로봇에서 적합한 HRI관점의 사용자 얼굴 검출 및 얼굴요소 검출방안을 제시하였다. 주요 얼굴 요소인 두 눈 과 입 영역의 명도특성을 이용한 방향성 얼굴요소 템플릿을 정의하고, 이를 통해 구성되는 유사에지 영역맵을 새롭게 제안한다. 이러한 영역맵과 얼굴특성 조건을 이용하면 얼굴검출과 얼굴요소 위치의 강인한 추정이 가능하며, 또한 향후에 사용자 얼굴인식 및 표정인식 등의 응용에 적합하다. 사용자가 로봇에 장착된 카메라에 비교적 멀리 떨어져 있는 경우에도 비교적 높은 검출성공률을 보였으며, 컬러영상뿐만 아니라 그레이 영상에서도 얼굴 요소검출 수행이 가능한 유연한 알고리즘으로 다양한 얼굴영상 입력에 대한 성능평가가 가능하였다. 보다 높은 검출 성능을 보장을 위해서는 향후에 조명조건 변화를 보상하는 얼굴영상 명도분포의 고려에 대한 연구를 필요로 한다.

참고문헌

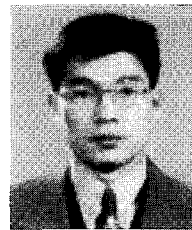
- [1] Nikolaidis, I. Pitas, "Facial feature extraction and pose determination", *Pattern Recognition*, vol. 33, pp. 1783-1791, 2000.
- [2] M. H. Yang, N. Ahuja, and D. Kriegman, "Mixtures of linear subspaces for face detection", *Proc. Fourth Int'l Conf. Automatic Face and Gesture Recognition*, pp. 70-76, 2000.
- [3] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images", *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 24, no. 1, pp. 34-58, Jan. 2002.

- [4] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis", *IEEE E Trans. Multimedia*, vol. 1, no. 3, Sep. pp. 264-277, 1999.
- [5] Y. J. Wang, C.-S. Chua, Y.-K. Ho, "Facial feature detection and face recognition from 2D and 3D images", *Pattern Recognition Letter*, vol. 23, pp. 119-183, 2002.
- [6] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection", *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 20, no. 1, pp. 23-38, Jan. 1998.
- [7] R. L. Hsu, Mohamed Abdel-Mottaleb, Anil K. Jain, "Face detection in color images", *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 24, pp. 696-706, may 2002.
- [8] B. Fröba and C. Küblbeck, "Robust face detection at video frame rate based on edge orientation features", *Proc. 5th Int'l Conf. Automatic Face and Gesture Recognition*, pp. 327-332, 2002.
- [9] Y. S. Gao, M. K. H. Leung, "Face recognition using line edge map", *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 24, no. 6, June, pp. 764-779, 2002.
- [10] T. G. Lee, S.-K. Park and M. S. Kim, "Development of pose-invariant face recognition system for mobile robot applications", *Proceedings of Int'l Conf. of Control, Automation and Systems Engineers (ICCAS)*, pp. 783-788, 2003.
- [11] A. K. Jain, "Fundamentals of digital image processing", Prentice Hall, 1989.
- [12] Z. H. Zhou, X. Geng, "Projection functions for eye detection", *Pattern Recognition*, pp. 1049-1056, May 2004.
- [13] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance", *Proc. 3rd Int'l Conf. on Audio-and Video-based biometric person Authentication(AVBPA)*, pp. 90-95, 2001.
- [14] A. Leon-Garcia, "Probability and random processes for electrical engineering", Addison Wesley, 1994.
- [15] The BioID face database; <http://www.bioid.com/downloads/facedb/facedatabase.html>
- [16] Carnegie Mellon University, Image database, face; <http://vasc.ri.cmu.edu/idb/html/face/index.html>
- [17] Yahoo! News, News photos; <http://news.yahoo.com/news?tmpl=index2&cid=702>



이 태 군

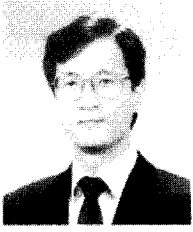
1997년 고려대학교 기계공학과 졸업(공학사). 1999년 동대학원 기계공학과 졸업(공학석사). 1999년~현재 연세대학교 대학원 전기전자공학과 박사과정 재학중. 관심분야는 컴퓨터 비전, HCI, 패턴인식.



박 성 기

1987년 서울대학교 기계설계학과 졸업(공학사). 1989년 동대학원 기계설계학과 졸업(공학석사). 2000년 KAIST 자동화 및 설계 공학과 졸업(공학박사) 2000년~현재 한국과학기술연구원 선임연구원 재직중. 관심분야는 컴퓨터

비전, 로봇 비전, 지능 로봇.



박 민 용

1973년 연세대학교 전자공학과 졸업(공학사). 1977년 동대학원 전자공학과 졸업(공학석사). 1982년 동경대 전자공학과 졸업(공학박사) 1982년~현재 연세대학교 전자공학과 교수. 관심분야는 퍼지 제어 및 응용, 로봇틱스, 퍼지 바이

오메디칼 시스템.