

실시간 응용을 위한 웨이블릿 변환 기반의 얼굴 검출

(Wavelet Transform-based Face Detection for Real-time Applications)

송 해 진 [†] 고 병 철 ^{**} 변 혜 란 ^{***}
(Haejin Song) (ByoungChul Ko) (Hyeran Byun)

요 약 최근 화상 회의, 화상 전화, 모바일 환경에서의 화상 통신, 얼굴 인식을 이용한 보안 시스템 등의 상업화에 힘입어 비디오에서의 얼굴 검출 및 추적 기술은 눈부신 발전을 이룩하였다. 또한, 얼굴 요소 검출은 요소 그 자체뿐 아니라 정확한 얼굴 영역 검출을 위한 필수 단계로서 중요한 연구 주제가 되고 있다. 그러나 영상에 나타난 복잡한 배경과 카메라 조작 및 조명에 의한 색상 왜곡 그리고 다양한 조명 조건 등은 얼굴 검출 및 추적, 요소 검출에 있어 여전히 큰 장애가 되고 있다.

이에 따라, 본 논문에서는 실시간 화상 통신을 위한 새로운 얼굴 영역 검출 및 추적 알고리즘과 검출된 얼굴 영역에서 효과적으로 눈 영역을 검출할 수 있는 알고리즘을 제안한다. 제안하는 얼굴 검출 알고리즘은 복잡한 배경과 다양한 조명 조건에 관계없이 얼굴을 검출하고 추적하기 위해 웨이블릿 변환된 세 종류의 부 영역을 이용하여 얼굴 형판을 생성하고 웨이블릿 변환된 입력 영상과의 유사도를 측정하여 얼굴을 검출한다. 특히 다양한 조명 조건을 극복하기 위해 최소-최대 정규화와 히스토그램 평활화를 혼합 적용하여 매우 밝거나, 매우 어두운 영상에서의 얼굴 오 검출 및 놓침을 줄일 수 있었으며 세 가지 크기의 얼굴 형판을 이용함으로써 입력 영상에 존재하는 다양한 크기의 얼굴도 검출할 수 있었다. 또한 효과적인 얼굴 추적 알고리즘을 통해 다음 프레임에서의 얼굴 위치를 예측하고 예측된 얼굴 위치를 중심으로 탐색 영역을 정해 형판 정합을 수행함으로써 얼굴 검출률을 높이면서 수행 시간도 단축시킬 수 있었다. 수직, 수평 방향 투영을 이용한 합리적인 눈 검출 알고리즘은 어두운 조명이나 부정확한 얼굴 영역에서도 만족스러운 결과를 보여주었다.

키워드 : 얼굴 검출, 웨이블릿 얼굴 형판, 형판 정합, 얼굴 추적, 눈 검출

Abstract In this paper, we propose the new face detection and tracking method based on template matching for real-time applications such as, teleconference, telecommunication, front stage of surveillance system using face recognition, and video-phone applications. Since the main purpose of paper is to track a face regardless of various environments, we use template-based face tracking method. To generate robust face templates, we apply wavelet transform to the average face image and extract three types of wavelet template from transformed low-resolution average face. However template matching is generally sensitive to the change of illumination conditions, we apply Min-max normalization with histogram equalization according to the variation of intensity. Tracking method is also applied to reduce the computation time and predict precise face candidate region. Finally, facial components are also detected and from the relative distance of two eyes, we estimate the size of facial ellipse.

Key words : face detection, wavelet face template, template matching, face tracking, eye detection

· 본 논문은 학술진흥재단의 지원(2002-005-H20001)으로 연구되었음

[†] 비 회 원 : LG전자 Mobile Multimedia 연구소 연구원

haejins@lge.com

^{**} 비 회 원 : 연세대학교 컴퓨터과학과

soccer1@aipiri.yonsei.ac.kr

^{***} 종신회원 : 연세대학교 컴퓨터과학과 교수

hrbyun@aipiri.yonsei.ac.kr

논문접수 : 2003년 1월 7일

심사완료 : 2003년 6월 3일

1. 서 론

비디오로부터 자동으로 얼굴을 검출, 추적하는 기술은 화상 회의, 보안 통제 시스템, 내용 기반 비디오 색인, 얼굴 인식 시스템, 모바일 환경에서의 화상 통신 등과 같은 많은 상업적인 분야에서 중요한 연구 과제로 떠오르고 있다.

컴퓨터를 통해 개인의 고유한 생체 특징을 인식하고 분석하여 정보를 보호하고 신분을 확인하는 생체 인식(Biometrics)분야, 그 중에서 다른 생체에 비해 비 접촉성으로 거부감이 덜하며 대상자의 특별한 주의 없이 데이터를 획득할 수 있는 장점을 지닌 얼굴 인식에 대한 관심이 증대됨에 따라, 얼굴 인식 시스템의 전 처리 단계에 해당하는 얼굴 검출 및 얼굴 요소 검출에 대한 관심도 증가하고 있다[1]. 또한, 사용자 얼굴 및 표정 인식을 통해 사용자 맞춤의 친밀한 서비스를 제공하려는 사용자 인터페이스 분야뿐 아니라[1,2], 홍채 인식 시스템[3], 응시점 추적(gaze tracking) 시스템[4] 등이 일반화되는 등 얼굴 검출 및 추적, 얼굴 요소 검출의 응용 분야는 더욱 확대되고 있다. 특히 모바일 환경에서의 화상 통신을 구현하는데 있어 모바일 네트워크 상에서 객체 기반 압축이 가능한 MPEG-4를 적용하려고 하는 경우, 화상 통신에서의 가장 중요한 객체가 되는 얼굴을 실시간으로 검출하는 것은 화상 통신의 질을 결정짓는 것으로 매우 중요한 부분이 된다[5].

얼굴 검출에 있어서 가장 첫 번째 수행해야 할 작업은 영상 안에 얼굴이 존재하는지의 여부를 결정하는 것과 얼굴이 존재한다면 그 얼굴의 정확한 위치가 어디인가를 파악하는 것으로 매우 간단해 보이지만 정지 영상이나 동영상으로부터 정확하게 얼굴을 검출해 낸다는 것은 결코 쉬운 일이 아니다. 이것은 입력 영상에서의 얼굴이 다양한 크기를 갖고, 위치가 변하며, 카메라 앵글의 변화로 인해 각도가 변화하고, 얼굴 포즈(정면 얼굴, 45도 기울어진 얼굴, 측면 얼굴)의 변화와 얼굴 가림 현상, 얼굴 표정의 변화, 조명의 변화 등이 원인으로 작용한다[1]. 또한 눈 검출은 얼굴 영역을 검출한 후, 검출된 얼굴 영역에서 두 눈을 찾는 방법[1,4,8]과 입력 영상 전체를 대상으로 직접 두 눈을 검출하는 방법이 있는데[9], 실시간 응용을 위해 수행 시간을 고려할 때 입력 영상 전체를 대상으로 하는 방법은 거의 사용할 수 없게 되고, 얼굴 영역을 가정한다 할지라도 위에 언급한 여러 가지 다양한 조건들 하에서 정확하게 눈 영역을 검출하는 것은 생각만큼 간단하지 않다.

일반적으로 얼굴 검출을 위한 기존의 연구들은 계층적 지식 기반 방법을 이용한 연구[10], 특징 기반 방법을 이용한 연구[6,11,12,13], 형판(template) 정합을 이용한 연구[17,24], 신경망(neural network) 기반의 연구[1], 주성분 분석(Principle Component Analysis)을 이용한 연구[18] 등으로 분류할 수 있다. 첫 번째, 지식 기반 방법[10]은 인간의 기본 얼굴에 대한 연구자들의 지식으로부터 추론된 규칙들에 기반하여 얼굴을 검출하는 것으로, 얼굴을 구성하는 요소들 간의 상대적인 거리나 관계와 같은 몇 가지 단순한 규칙들을 정하고 이를 만

족하면 얼굴로 검출하는 것이다. 하지만 이 방법은 인간의 얼굴에 대한 지식을 정확하게 규칙으로 정의하기 어렵다는 점과 다양한 얼굴 포즈마다 규칙이 정의되어야 하는 문제점이 있다. 두 번째, 특징 기반 방법은 얼굴 검출을 위해 얼굴의 불변하는 특징(눈, 코, 입과 같은 얼굴 요소[11,12], 질감[13], 얼굴색[6])들을 이용하는 방법이다. 여러 가지 특징들 중 얼굴색은 얼굴의 이동, 회전, 크기 변화 등에 덜 민감한 특성을 가지므로 최근에 가장 많이 사용되고 있으나, 영상에 다양한 조명 조건이 존재하여 색상 왜곡이 발생한 경우 매우 비효율적이 되므로 최근 수행된 대부분의 연구들에서는 얼굴색을 단독으로 사용하기보다는 다른 방법(예지 검출을 통한 얼굴 타원[14], 얼굴 추적을 위한 움직임 정보[15], 얼굴 요소, 질감, 얼굴색을 모두 이용[16])과 결합하는 방식을 채택하고 있다. 세 번째로 형판(template) 정합 기반의 방법들[17,24]은 얼굴에 대한 몇 가지의 표준 패턴을 얼굴 형판으로 생성하여 저장하였다가 입력 영상의 탐색 영역 안에서 서로 비교하여 가장 유사도가 큰 지점을 얼굴로 검출하는 방법이다. 이 방법은 단순하다는 장점이 있으나, 정합 시 계산량이 많다는 단점을 가지고 있다. 네 번째, 신경망(neural network) 기반의 방법[1]은 다양한 영상들로부터 얼굴 영역과 비 얼굴 영역을 획득하여 신경망을 통해 얼굴과 비 얼굴로 학습한 뒤 입력 영상으로부터 얼굴을 검출하도록 하는 방법이다. 이러한 방법들은 정면 및 측면 얼굴 검출에 대해서는 만족할 만 하지만, 학습을 위한 비 얼굴 선택이 쉽지 않고 실수 위주의 복잡한 계산이 많으며 회전된 얼굴의 경우 좋은 결과를 보여주지 못하고 있다[1,12]. 마지막으로 주성분 분석(Principle Component Analysis)을 이용하는 방법[18]은 얼굴 영역을 추출하기 위해 복잡한 얼굴 패턴을 몇 개의 주성분(Principle Component) 값으로 표현하여 고유 얼굴 형판으로 생성하여 저장하였다가 입력 영상의 주성분 값과 비교하여 가장 유사도가 큰 지점을 얼굴로 검출하는 방법이다. 주성분 분석을 이용한 방법은 비교적 복잡한 배경에서도 효과적으로 얼굴을 검출하지만 얼굴의 위치와 조명 변화에 민감한 단점이 있다. 얼굴 검출을 위한 기존 연구들은 크게 위와 같이 분류할 수 있으나 최근에는 얼굴 얼굴색으로 걸러진 얼굴 후보에 형판 정합을 수행하는 방법[19], 얼굴 요소들로 얼굴 모형을 생성하여 이용하는 방법[20] 등 서로 다른 방법론들을 다양한 방식으로 결합시킨 연구들이 보고되고 있다.

이에 따라 본 논문에서는 화상 통신을 고려한 실시간 환경에서 복잡한 배경과 다양한 조명 변화 그리고 자연스러운 얼굴 회전과 다양한 크기의 얼굴이 존재하는 입력 영상으로부터 얼굴을 검출하고 추적하는 새로운 알

고리즘을 제안하고 얼굴 검출 알고리즘에 의해 검출된 얼굴 영역에서 두 눈을 검출하는 알고리즘을 소개한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 웨이블릿 변환 기반의 얼굴 검출 알고리즘과 추적 알고리즘에 대하여 설명하고 3장에서는 검출된 얼굴 영역에서 효과적으로 두 눈 영역을 검출하는 알고리즘을 소개한다. 4장에서는 제안한 알고리즘의 성능을 평가하기 위하여 수행한 다양한 실험들의 결과를 살펴보고, 마지막으로 5장에서는 본 논문의 결론 및 향후 연구 과제를 제시한다.

2. 웨이블릿 변환 기반의 얼굴 검출 및 추적 알고리즘

제안하는 얼굴 검출 알고리즘은 기본적으로 형판 정합 방법을 이용한다. 형판 정합과 관련된 여러 가지 문제점들을 해결하기 위하여 본 연구에서는 웨이블릿 변환된 얼굴 영상의 부 영역(sub-image)들을 평균하여 세 종류의 웨이블릿 얼굴 형판을 만들어 저장한 후, 웨이블릿 변환된 입력 영상과의 유사도를 측정하여 얼굴 영역을 검출하게 된다. 또한 이전 세 프레임에서의 얼굴 좌표는 다음 프레임에서의 얼굴 위치를 예측하는데 사용되며, 예측된 얼굴 위치를 중심으로 탐색 영역을 정해 주어 속도 및 정확성을 높이고자 하였다. 얼굴 검출 및 추적 알고리즘은 그림 1과 같이 크게 4 단계로 수행되어 진다.

2.1 전 처리

제안하는 알고리즘의 첫 번째 단계는 영상에 존재하는 다양한 조명을 보정해주기 위한 전 처리 단계이다. 일반적으로 영상이 디지털 카메라나 CCD 카메라로부터 얻어질 때 빛(스펙트럼, 빛의 강도)과 카메라의 특성(센서의 반응, 렌즈)으로 인해 영향을 받게 되므로 이러한

요인들로 인한 빛의 상태를 보정하는 작업은 잘못된 결과(false positive)를 줄이는데 필수적인 작업이다. 본 연구에서는 식 (1)의 최소-최대 정규화 방법을 적용한다. 최소-최대 정규화 방법은 원 영상을 새롭게 정의된 데이터 범위로 변형시켜주는 선형적인 변형(linear transformation) 방법이다.

$$y' = \left(\frac{y - \min_1}{\max_1 - \min_1} \right) (\max_2 - \min_2) + \min_2 \quad (1)$$

식 (1)에서, \min_1 과 \max_1 는 입력 영상의 최소, 최대 밝기 값이며 \min_2 와 \max_2 는 새로운 범위의 최소, 최대 밝기 값이다. 최소-최대 정규화를 위해, 본 연구에서는 밝기 성분의 상위 약 11.7%를 reference white(225~255)로, 하위 약 11.7%까지를 reference black(0~30)으로 간주하여 제거하여 주었다. 그러나 입력 영상이 매우 높은 밝기 값을 갖는 경우(gray 평균 밝기 > 180)나 매우 어두운 밝기 값을 갖는 경우(gray 평균 밝기 < 50)에는 최소-최대 정규화를 적용하기 전에 히스토그램 평활화를 적용하여 조명을 보정 하였다. 입력 영상이 매우 밝은 경우 영상의 밝고 어두움이 명확하지 않게 되어 영상의 질이 나빠지게 되는데, 이런 경우 히스토그램 평활화를 적용하여 영상의 밝고 어두움을 강조한 후 최소-최대 정규화를 적용하여 원하는 밝기 범위로 영상의 밝기를 조절하여 줌으로써 좋은 조명 보정 효과를 거둘 수 있었다.

2.2 색상 정보를 이용한 얼굴 후보 영역 추출

많은 연구들이 얼굴색 성분과 밝기 성분은 서로 독립적이라는 가정을 세우고 있지만, [8]에서는 실제 얼굴색은 비선형적으로 밝기 성분에 의존적이라는 이론을 실험과 함께 제시하였다. 따라서 [8]에서 제안한 변환식을 이용하여 입력 영상의 Cb, Cr성분을 입력 영상의 Y값에 따라 새로운 색상 모델인 NCb, NCr로 변형하였다.

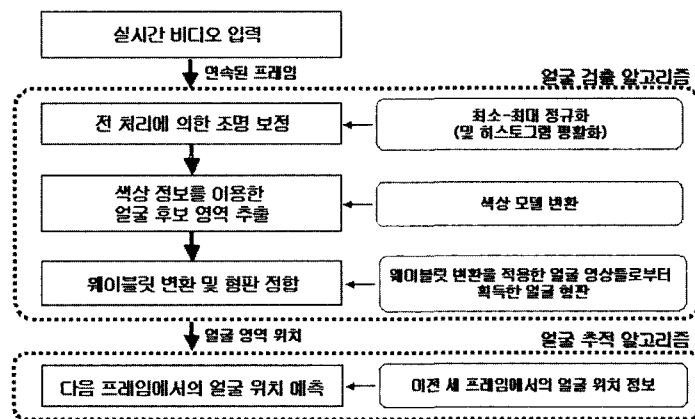


그림 1 웨이블릿 변환 기반의 얼굴 검출 및 추적 알고리즘

변형된 색상 성분에 임의의 임계값을 적용하여 얼굴 후보 영역을 검출하는데 이때 검출된 얼굴 후보 영역이 형판 정합이 수행되는 영역이 된다. 이때 적용되는 임계값은 [21]에 제시된 값을 기준으로 조명 상태에 따라 실제 얼굴 영역이 후보에서 제외되어 형판 정합조차 시도되지 못하는 경우를 막기 위해 보다 넓게 잡아 주었다.

$$Cf(x, y) = \begin{cases} 1 & \text{if } [NCb(x, y) \in T_{Cb}] \wedge [NCr(x, y) \in T_{Cr}] \\ 0 & \end{cases} \quad (2)$$

수식 (2)에서, $Cf(x, y)$ 는 색상 성분을 이용해 추출한 얼굴 후보 영역이며 얼굴 후보 영역 추출에 사용한 임계값인 T_{Cb} , T_{Cr} 은 [87, 167], [123, 173]으로 두 성분 모두 임계값을 만족한 경우에만 얼굴 후보로 추출하도록 하였다.

2.3 웨이블릿 변환 기반의 형판 정합을 이용한 얼굴 검출

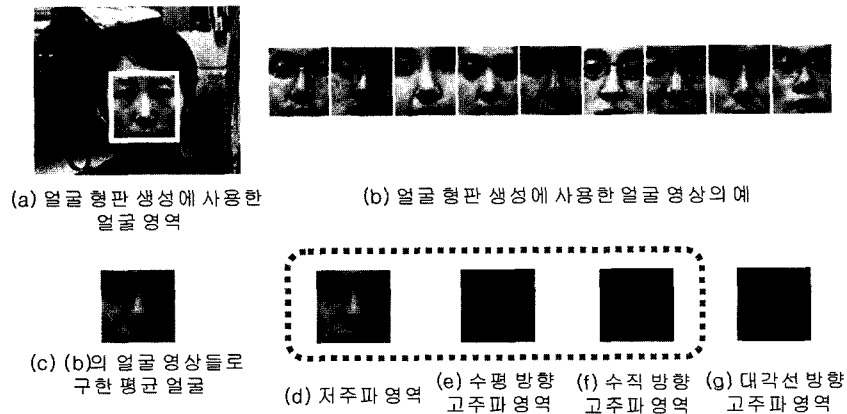
색상 정보를 이용하여 추출한 얼굴 후보 영역은 이번 단계에서 다해상도 얼굴 형판 정합의 후보 영역으로 사용된다. 형판 정합을 이용한 실시간 얼굴 검출 및 추적 방법의 문제점들 중 하나가 바로 탐색 시간이므로 본 연구에서는 얼굴 영상들에 이산 웨이블릿 변환(Discrete Wavelet Transform)을 적용하여 해상도가 낮아진 세 종류의 부 영역을 형판 정합에 이용함으로써 얼굴 영역을 검출하는 방법을 제안한다. 이때 얼굴 형판은 웨이블릿 변환으로 구한 저주파, 수평, 수직 고주파 성분을 이용하여 세 종류로 생성하였다. 또한 본 논문이 실시간 얼굴 검출을 기본으로 하기 때문에 메모리 소모가 크더라도 계산량을 줄이는 것이 더욱 중요하므로 영상에 존

재하는 다양한 크기의 얼굴을 검출하기 위해 여러 크기의 형판을 생성하였다가 얼굴 검출 시 이용하도록 하였다. 입력 영상이 화상 통신에서의 H.261 표준인 QCIF (176×144)이고 화상 통신의 특성상 상반신 이상의 영역이 영상에 존재한다는 점을 고려하여 얼굴 형판은 40×40, 60×60, 80×80의 크기로 정하였고 이 크기의 형판을 이용하면 저해상도에서의 형판 정합이라는 방법의 특성상 30×30~100×100의 얼굴까지 검출할 수 있다.

2.3.1 얼굴 형판 생성

그림 2(c)의 평균 얼굴은 연세대학교 컴퓨터과학과에서 수집한 얼굴 데이터베이스를 이용하여 만들어졌다 (<http://vip.yonsei.ac.kr/facetracking/>). 일반적으로 얼굴 형판은 인간 얼굴의 가장 일반적인 특징들을 가능한 한 모두 포함해야 하며 동시에 배경이나 개개인의 특징(예, 헤어 스타일, 수염의 모양)들에 민감하지 않아야 한다[17]. 이러한 사실에 입각해서, 본 연구에서는 68명으로부터 3장씩, 모두 204개의 얼굴을 형판 생성에 사용하였다. 형판은 위로 눈썹을 포함하고 아래로는 윗입술까지만 포함하도록 하였는데, 204개 얼굴을 그림 2(b)와 같이 눈썹과 윗입술이 포함되도록 잘라낸 다음, 40×40, 60×60, 80×80의 얼굴 형판을 생성하기 위해 각각의 크기로 정규화 한 후 정규화 된 얼굴들에 웨이블릿 변환을 적용하여 변환된 부 영역별로 평균하여 세 가지 형태의 웨이블릿 얼굴 형판을 생성하였다.(그림 2(d)~(f)).

40×40, 60×60, 80×80의 얼굴 영역에 2단계 웨이블릿 변환을 적용하면 얼굴 영역은 10×10, 15×15, 20×20으로 가로 세로 각각 4배씩, 16배 축소 되게 된다. 일반적으로



(a) 얼굴 형판 생성에 사용한 얼굴 영역 (b) 얼굴 형판 생성에 사용한 얼굴 영상의 예
(c) (b)의 얼굴 영상들로 구한 평균 얼굴 (d) 저주파 영역 (e) 수평 방향 고주파 영역 (f) 수직 방향 고주파 영역 (g) 대각선 방향 고주파 영역

그림 2 얼굴 형판 생성 과정

(a) 얼굴 형판 생성에 사용한 얼굴 영역, (b) 얼굴 형판 생성에 사용한 데이터베이스 내의 얼굴들의 예, (c) (b)를 평균한 얼굴, (d)~(g) (b)에 웨이블릿 변환을 적용하여 부 영역별로 평균 낸 영상

로 얼굴 영역은 수평, 수직 방향으로 에지가 강하게 나타나고 대각선 방향 에지는 두드러진 특징이 되지 못하기 때문에 수평(H), 수직 방향 고주파 성분(V)은 얼굴 영역 검출에 중요한 역할을 하는 반면 대각선 방향 고주파 성분(D)은 그렇지 못하게 된다. 따라서 본 연구에서는 오직 저주파 성분(L)과 수평 고주파 성분(H), 수직 고주파 성분(V)의 값을 각각 얼굴 형판으로 생성하였고 결국 각 얼굴 크기(40×40, 60×60, 80×80)별 세 종류(L, H, V)의 형판을 생성하므로 총 9개의 얼굴 형판이 정합을 위해 사용된다. 얼굴 형판 정합에 사용한 부 영역의 타당성을 평가하기 위하여 본 논문에서는 각 부 영역 및 부 영역을 조합한 얼굴 형판에 대한 얼굴 검출률을 비교하였다. 자세한 실험 결과는 4.1에 언급하였으며 본 논문이 제안한 세 가지 성분의 얼굴 형판을 이용하는 것이 얼굴 검출 성능을 높이는 데 효과적임을 확인할 수 있다.

2.3.2 형판 정합을 이용한 얼굴 검출

만약 $F_k(L, H, V)$ 가 $m \times n$ 크기의 얼굴 형판이고 L 이 $M \times N$ 크기의 웨이블릿 변환된 입력 영상의 저해상도 부 영역이라고 할 때, MAE(Mean Absolute Error)는 F_k 과 탐색 영역 S 사이에서 계산된다. 이때, 탐색 영역 S 의 크기는 $(m+6) \times (n+6)$ 이다. 형판 정합은 색상 정보를 이용하여 추출한 얼굴 후보 영역($CF(x, y) = 1$ 인 화소)에 대해서만 수행된다.

입력 영상으로부터 다해상도 얼굴 형판을 이용하여 얼굴 검출하는 방법을 아래의 형판 정합 과정을 통해 살펴본다.

- 형판 정합 과정

입력

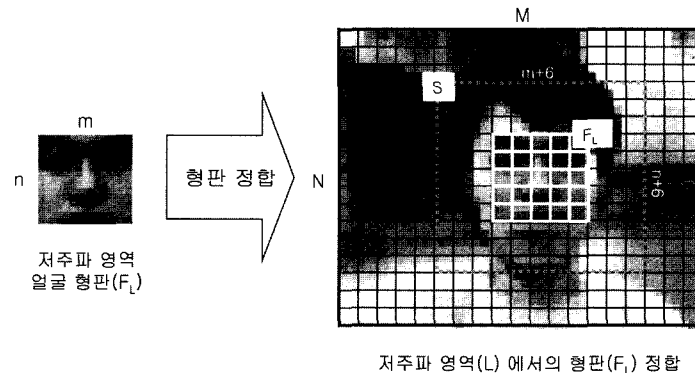
- 입력 영상에 조명 보정을 위해 전 처리를 적용한 그레이 영상
- 색상 모델 변환 후 수식 (2)를 이용하여 추출된 얼굴 후보 영역

1) 조명 보정된 그레이 영상에 웨이블릿 변환을 적용하여 저해상도의 저주파 영역(L), 수평 방향 고주파 영역(H), 수직 방향 고주파 영역(V)을 확보한다.

2) 1)에서 생성한 부 영역과 형판(40×40, 60×60, 80×80)의 유사도를 측정하는데, 형판 크기에 따라 모두 3번의 형판 정합이 수행된다. 유사도 측정의 대상이 되는 탐색 영역은 프레임 번호에 따라 다음과 같이 달라지는데, 형판의 크기에 따라 탐색 영역 또한 달라지게 된다.

- ① 첫 번째 프레임인 경우 영상 전체 영역을 대상으로 얼굴 후보 화소에 대해 형판 정합
- ② 두 번째 프레임부터는 이전 프레임에서의 얼굴 영역의 중심과 형판의 중심을 맞춘 상태에서 형판의 크기가 $m \times n$ 이라고 했을 때, $(m+6) \times (n+6)$ 의 탐색 영역을 정하고 이 안의 얼굴 후보 화소에 대해서만 형판 정합(그림 3)
- ③ 네 번째 프레임부터는 이전 세 프레임의 얼굴 좌표를 이용하여 현재 프레임에서의 얼굴 위치를 추정하고(얼굴 추적) 그 위치를 중심으로 ②와 같이 탐색 영역을 정함

형판 정합 시 부 영역과 형판의 유사도는 수식 (3)의 MAE(Mean Absolute Error)를 이용하여 측정하였다.



저주파 영역(L)에서의 형판(F_L) 정합

그림 3 얼굴 형판 정합 과정 : 저주파 영역으로부터 만들어진 형판 F_L 이 입력 영상의 저주파 영역(L) 탐색 영역(S) 안에서 $CF(x, y) = 1$ 인 화소(x, y)에 대해 정합되는 과정, 점선은 탐색 영역, 흰색 실선은 얼굴 형판을 의미

$$MAE(F, I)_k = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n (|F_k(x, y) - I_k|) \quad (3)$$

$$MAEs(F, I) = \sum_k^{L, H, V} (w_k \cdot MAE_k) \quad (4)$$

수식 (3), (4)에서 k 는 세 종류의 부 영역(L, H, V)를 나타내며, I 는 탐색 영역 안에서 현재 형판과 정합 되려는 중심이 (x, y) 인 한 영역을 나타낸다. w_k 는 각 부 영역 별 가중치를 의미하는데 본 논문에서는 $w_L : w_H : w_V = 1 : 1 : 1$ 로 정하였다. 수식 (4)에서 보여지듯 각 부 영역과 형판들의 $MAE(F, I)_k$ 값에 부 영역 별 가중치 곱하여 모두 더하면 $MAEs(F, I)$ 로 한 지점 (I)에서 형판(F)들과의 유사도가 된다.

3) 2)에서 형판 크기 별 $MAEs$ 가 구해지면 이 값들을 비교하여 최종 얼굴 영역을 검출하는데 수식 (5)와 (6)이 그 방법을 나타낸다. 이때 최소 MAE 값은 반드시 σ 보다 작아야 하는데, 이것은 얼굴과 비얼굴을 판단하기 위한 임계값으로 본 논문에서는 실험적으로 100으로 값을 정하였다.

$$CFace_{scale}(x, y) = \min(MAEs(F, I)) < \delta \quad (5)$$

$$RF_{scale}(x, y) = \min(CFace_{scale=10 \times 10, 15 \times 15, 20 \times 20}(x, y)) \quad (6)$$

수식 (6)은 최종 얼굴 영역 검출을 나타낸다. (5)에서 구한 $CFace_{scale}(x, y)$ 중 가장 작은 한 값을 골라 최종 얼굴 영역으로 결정하며, 이때 $scale$ 은 웨이블릿 얼굴 형판을 정합하여 검출한 얼굴의 크기이고 (x, y) 는 검출된 영역의 중심 좌표이다.

2.4 얼굴 추적 알고리즘

일반적으로 얼굴 추적에는 Kalman 필터가 많이 사용된다[22]. Kalman 필터는 측정된 데이터와 알려진 동적인 특성을 고려하여 시스템으로 하여금 비관측 상태를 추정하도록 설계되었지만[22], 정확성에 비해 계산 시간이 많이 걸리는 단점이 있다. Xu와 Sugimoto [6]는 계산 시간을 줄이기 위해 이전 3개 좌표의 지속적인 가속도와 속도를 이용하여 얼굴을 추적하는 방법을 제안하였다. 이 방법은 비교적 빨리 얼굴 위치를 추정해 주기는 하지만 움직임의 속도와 방향이 일정하지 않은 경우 잘못된 예측을 하게 되므로 본 연구에서는 [6]의 알고리즘을 개선하여 얼굴 움직임의 방향과 속도가 일정하지 않은 경우에도 효과적으로 얼굴을 추적할 수 있도록 하였다.

얼굴 위치가 t 시간에 $x(t) = [x(t), y(t)]^T$ 에 위치한다고 가정하자. 이제 다음 프레임에서의 얼굴 위치 $x(t+1)$ 를 $t-2$ 부터 t 까지의 평균 속도와 t 에서의 얼굴 위치의 좌표

를 결합하여 수식 (7)~(9)와 같이 구할 수 있다.

$$v(t-1) = x(t-1) - x(t-2) \quad (7)$$

$$v(t) = x(t) - x(t-1) \quad (8)$$

$$\bar{x}(t+1) = x(t) + \left[\frac{v(t-1) + v(t)}{2} \right] \quad (9)$$

수식 (7), (8), (9)에서 보여지는 $v(t-1)$ 는 $t-2$ 와 $t-1$ 사이의 변화가 되는데, 이전 프레임들에서 얼굴의 위치가 어떻게 변해왔는가를 살펴보고, 결국 다음 프레임에서의 얼굴은 변위 $v(t)$ 와 $v(t-1)$ 의 평균만큼 이동되어 나타날 것이라고 추정하는 것이다. 객체의 움직임은 속도 혹은 방향이 일정하지 않은 경우, 예를 들어, $t-2$ 에서의 얼굴 좌표가 4이고, $t-1$ 에서는 2, t 에서는 5인 경우, [6]의 방법을 적용하면 $t+1$ 에서의 얼굴 위치는 13인데 반해 제안하는 방법을 이용하면 얼굴의 위치는 5로, 제안한 알고리즘이 [6]의 방법보다 바람직한 예측을 하게 된다. 이에 본 연구는 처음 세 프레임의 얼굴 위치를 저장하였다가, 제안하는 추적 알고리즘을 이용하여 네 번째 프레임의 얼굴 위치를 예측하고, 예측된 얼굴 위치를 중심으로 탐색 영역을 정해주도록 하였다. 본 논문이 기반을 두고 있는 [6]의 방법과 본 논문이 개선한 방법에 대한 성능 평가는 4.3에 자세히 설명되어 있다.

3. 눈 검출 알고리즘

3장에서는 앞서 설명한 얼굴 검출 및 추적 알고리즘을 통해 검출된 얼굴 영역에서 두 눈을 검출하는 알고리즘을 소개한다. 영상에서 혹은 얼굴 영역에서 눈 영역을 검출하는 방법은 앞서 설명한 얼굴 검출 방법들과 마찬가지로 신경망 기반의 연구[9], 특징 기반의 연구[8,23,24], 형판 정합을 이용한 연구[25] 등으로 분류할 수 있다. 또한 접근 방법에 따라, 얼굴 영역을 가정하고 눈을 찾는 방법[1,4,8,23,24]과 입력 영상 전체에서 눈을 검출하는 방법[9]으로 나누어 볼 수 있으나 대부분의 연구는 대략적으로나마 얼굴 영역을 검출 한 후 그 안에서 얼굴 요소를 찾는 방식으로 진행되고 있다. 신경망 기반의 연구나 형판을 이용한 방법은 그 방법이 얼굴 검출 시 적용하는 방식과 유사하나, 특징 기반의 눈 검출 방법은 눈의 밝기 값이 인근 영역보다 어둡고 흰자와 검은자에 의한 밝기 대비가 크다는 특징을 기본 가정으로 하여 반복적인 이진화 만으로 눈을 검출하거나 [24], 그레이 모폴로지 연산자를 적용한 후 그 결과를 분석하여 눈을 검출하는 방법이다[8]. 또한 이진화 결과를 수직, 수평 방향으로 히스토그램을 그리거나 연결된 요소(connected component)를 분석하여 눈을 검출하는 것도 하나의 분석 방법이 되는데[23], 이러한 특징 기반의 눈 검출 방법은 다양한 응용이 가능하며 방식이 간

단한 장점이 있으나, 머리가 내려와 눈을 가린 경우나 상대적으로 눈썹이 짙은 경우 눈 검출은 사실상 어려워진다.

따라서 본 논문에서는 그림 4에서 도식화 한 것 같이, 검출된 얼굴 영역을 이진화하고, 이진화된 얼굴 영역을 수평, 수직 방향으로 투영시켜 그 결과를 분석하여 두 눈을 검출하는 알고리즘을 제안한다. 제안하는 눈 검출 알고리즘을 통해 찾아진 두 눈의 좌표는 얼굴 영역의 타원 추정을 위한 장축과 단축의 거리 계산에 사용된다.

3.1 적응적 임계값을 이용한 눈 후보 영역 검출

검출된 얼굴 영역이 입력으로 들어오면, 얼굴 요소들의 밝고 어두움을 강조하기 위하여 전 처리로서 히스토그램 평활화를 수행한다. 본 논문이 제안한 얼굴 검출 알고리즘에서 사용한 얼굴 형판이 눈썹부터 시작하여 윗입술까지 포함하도록 디자인되었기 때문에 정확한 눈 영역을 검출하기 위해 검출된 얼굴 영역의 중간 아래 부분은 눈 후보 영역에서 제거할 필요가 있다(그림 5(c)). 수식(10)를 통해 임계값을 구하고 검출된 얼굴 영역의 상위 부분에 대해서만 임계값 보다 낮은 밝기 값을 갖는 화소들을 눈 후보 영역으로 이진화 한다. 얼굴 요소 이진화를 위해 효과적인 임계값을 구하려는 연구들이 많이 있지만[23,24], 본 논문에서는 얼굴 영역의 밝기 평균과 표준 편차를 이용하여 비교적 간단하게 임계값을 구하도록 하였다.

$$Threshold = \mu - (\sigma * s) \tag{10}$$

수식 (10)의 μ 와 σ 는 전 처리가 적용된 얼굴 영역의 평균 밝기 값과 표준 편차이고 s는 표준 편차에 곱해지는 상수(scale factor)이다. 눈 영역이 얼굴의 다른 영역 보다 어둡다는 가정 하에 평균 밝기 값에서 표준 편차에 상수를 곱한 값을 감하여 임계값을 정하였는데, s가

작을수록 눈 후보 영역은 많아지고 s가 클수록 눈 후보 영역은 작아지게 된다. 상수 s는 여러 번의 실험에서 가장 좋은 이진화 결과를 보여주는 1.1로 정하였고, 수식 (10)에 의해 구해진 임계값이 너무 작아져 실제 눈 영역이 후보에서 제외되는 일을 방지하기 위해 최소 임계값은 30으로 잡았다.

3.2 수평 방향 투영 및 결과 분석

이진화된 영상을 수직, 수평 방향으로 투영시켜 눈 영역이 갖는 지형적 특성을 고려하여 투영 결과를 분석함으로써 두 눈의 위치를 추정한다. 우선 눈 후보 이진 영상에 대해 수직 방향 투영이 먼저 수행된다. 오른쪽 눈과 왼쪽 눈 영역을 분리하기 위하여 투영 결과 히스토그램에 대해 k-means 군집화($k = 2$)를 적용하고 나누어진 2개 군집의 중심 x 좌표를 두 눈의 x 좌표로 삼는다. 수직 방향 투영 결과 히스토그램 분석을 위해 k-means 군집화를 이용하는 것이 매우 복잡해 보일 수 있는데, 그림 5에서 사용된 영상과 같이 세로 방향 투영 결과가 코나 눈썹, 눈을 가린 앞머리 등에 의해 명확하게 두 개의 영역으로 구분되지 않고 하나로 연결되어 있거나 세 개 이상의 영역으로 나누어진 경우에는 얼굴 영역을 양쪽 눈이 속하는 두 영역으로 나눌 기준이 모호해지기 때문에 본 논문에서는 $k = 2$ 인 k-means 군집화를 이용하여 합리적인 기준을 세우고자 하였다.

3.3 수평 방향 투영 및 결과 분석

그림 5(e)에서와 같이 최종 구해진 두 군집의 중심 벡터의 x 좌표는 두 눈의 x 좌표가 된다. 두 눈의 y 좌표를 구하기 위하여, 두 군집을 분리하는 x 좌표 상의 d를 구하고, d를 중심으로 왼쪽 눈 영역과 오른쪽 눈 영역을 분리하여 각각 수평 방향 투영을 수행한다(그림 5(f)). 수평 방향 투영 결과 히스토그램 상에 존재하는

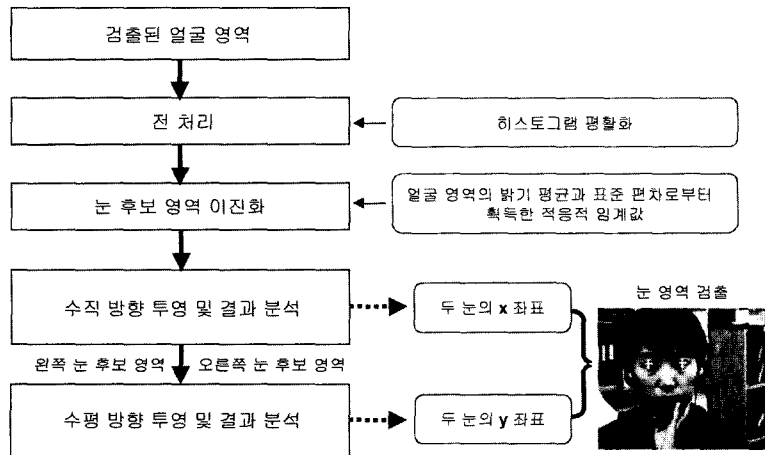


그림 4 제안하는 눈 검출 알고리즘

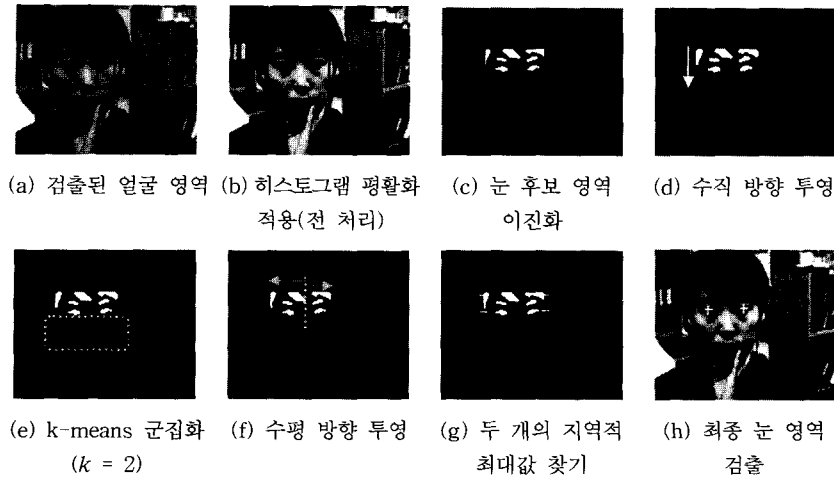


그림 5 눈 검출 알고리즘의 수행 과정

많은 지역적 최대값(local maxima)을 제거하기 위한 히스토그램 스무딩(smoothing)을 수행한 뒤, 오른쪽, 왼쪽 히스토그램에서 눈과 눈썹을 고려하여 2개의 최대값을 각각 찾는다(그림 5(g)). 지형적으로 눈은 눈썹보다 아래에 위치하므로 최대값을 갖는 y 값 큰 값(아래에 위치함)을 눈의 y 좌표로 선택한다. 따라서 그림 5(e)와 같이 수직 방향 투영 결과를 분석하여 구한 두 눈의 x 좌표와 수평 방향 투영 결과를 분석하여 구한 두 눈의 y 좌표(그림 5(g))를 조합하여 최종 두 눈 영역을 결정하게 된다. 그림 5는 눈 검출 알고리즘의 중간 결과 및 전체 수행 과정을 나타낸다.

눈 검출 알고리즘에 대한 성능 평가는 4.4의 표 2에 제시되어 있다.

3.4 두 눈의 거리를 이용한 얼굴 타원 추정

앞서 검출한 두 눈의 거리를 이용하여 입력 영상에 존재하는 얼굴 영역을 타원으로 추정한다. [8]에서는 얼굴 타원을 구하기 위해 타원 허프(Hough) 변환을 이용하였지만, 변환 자체의 복잡도가 높아 실시간 응용에는 적합하지 않다. 그리고 [14]에서는 얼굴 타원을 1.2의 비율을 갖는 수직 방향 타원으로 정하였지만 얼굴 영역은 얼굴의 회전된 정도, 기울임에 따라 그 타원의 모양도 변하게 되므로, 고정된 비율도 적당하지 않다. 따라서 본 연구에서는 검출된 두 눈 사이의 거리(e_{dis})를 이용하여 타원의 크기와 모양을 정해주었다. 일반적으로 얼굴의 형태는 타원이므로 단축(x-축)의 반경을 $e_{dis} \times 1.3$ 로 주고 장축(y-축)의 반경을 $e_{dis} \times 1.8$ 로 정하였고, 검출된 얼굴 영역 중심으로부터 장축과 단축 값을 이용하여 나머지 타원의 좌표를 계산하여 얼굴 타원을 추정한다.

4. 실험 및 결과 고찰

제안하는 알고리즘들은 운영체제 Windows 2000 Professional 상에서 Visual C++ 6.0 MFC를 이용하여 구현하였으며 실험은 Pentium-IV 1.8 GHz인 CPU와 512 MB 메모리 사양의 Windows 2000 Professional 환경에서 수행하였다. 실험에 사용된 데이터는 실생활에서 흔히 발생하는 6종류의 조명 조건을 기준으로 자체 촬영한 동영상 데이터이다. 영상 포맷은 화상 통신 표준인 H.261을 고려하여 QCIF(176×144)로 정하였고 영상에 존재하는 얼굴의 크기는 입력 영상과 사용 환경을 고려하여 30×30~100×100으로 가정하였다. 6종류의 조명 조건의 예제 영상은 그림 6에서 확인할 수 있으며 표 1은 6종류의 조명 조건 및 동영상 데이터의 설명이다. 실험에 사용한 데이터 및 실험 결과 동영상은 <http://vip.yonsei.ac.kr/facetracking/>에서 확인할 수 있다.

실험은 크게 네 부분으로 나누어 진행되었다. 첫 번째는 제안한 얼굴 검출 알고리즘에서 세 종류의 웨이블릿 부 영역을 이용하여 생성한 얼굴 형판에 대한 것이며, 두 번째는 얼굴 검출 시 조명 보정을 위해 사용한 전

표 1 실험에 사용된 동영상 데이터

동영상 데이터 (번호)	총 프레임 수	조 건
(1)	311	일정한 밝기의 영상
(2)	500	정면광 영상
(3)	500	측면광 영상
(4)	500	역광 영상
(5)	311	전체적으로 밝은 영상
(6)	290	전체적으로 어두운 영상



그림 6 6종류 조명 조건의 예제 영상 및 실험 결과

처리 방법에 대한 성능 평가이다. 세 번째는 탐색 영역 제한 방법 및 얼굴 추적 알고리즘의 성능 평가이고 마지막으로 네 번째는 눈 검출 알고리즘의 성능 평가 부분이다.

제안하는 알고리즘은 6종류의 실험 동영상 데이터에서 95.4%의 얼굴 검출률을, 91.9%의 눈 검출률을 보여 주었다. 평균 수행 시간은 얼굴 검출 시에는 0.091초/프레임, 눈 검출까지 수행하는 경우 0.11초/프레임으로 100프레임을 대상으로 한 경우 각각의 경우 9.08초, 11.02 초가 소요된다. 그림 7은 제안한 얼굴 검출 알고리즘의 6종류의 동영상별 얼굴 검출률과 수행 시간을 나타낸다.

그리고 앞서 설명한 6종류 이외의 동영상 데이터를 대상으로, 얼굴 가리워짐이나 2개 이상의 얼굴이 존재하는 경우에서의 얼굴 검출 및 눈 검출, 얼굴 영역 타윈

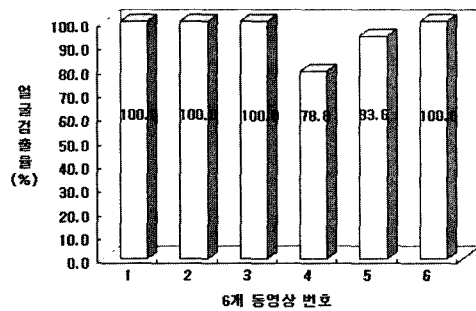


그림 7 제안한 얼굴 검출 알고리즘의 동영상 별 얼굴 검출률(%)

추정을 실험하였는데 결과는 그림 8에 나타나 있다. 특히 본 시스템은 실시간 응용 분야 중에서 화상 통신을 위해 설계되었으므로 여러 개의 얼굴이 존재할 경우 일

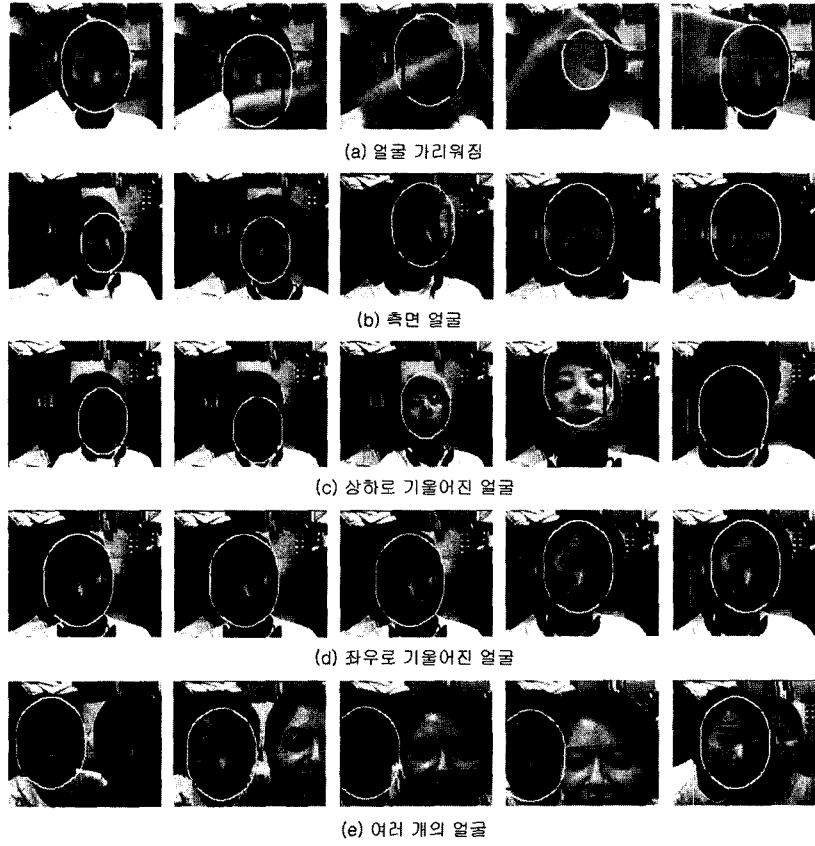


그림 8 다양한 조건 하에서의 얼굴 검출 결과

정한 크기를 갖고 가운데 위치하는 사람을 추적하도록 하였다. 따라서 그림 8(e)에서처럼 여러 개의 얼굴이 존재할 경우 모든 얼굴을 검출하는 것이 아니라 주가 되는 하나의 얼굴만을 검출하고 추적하게 된다.

4.1 얼굴 검출에 사용한 웨이블릿 부 영역

제안하는 얼굴 검출 알고리즘은 웨이블릿 변환된 얼굴 영역에서 저주파 영역(L), 수평(H), 수직 방향 고주파 영역(V)을 취해 얼굴 영역 검출 시 사용하고 있는데, 그림 9에서 보듯이 본 논문에서 사용한 세 종류의 부 영역(LHV) 얼굴 형판은 다른 종류의 부 영역을 이용한 경우에 비해 좋은 얼굴 검출률을 보여주었다. 특히 웨이블릿 변환을 통해 구할 수 있는 네 종류의 부 영역 전부를 얼굴 형판으로 생성한 경우(LHVD)를 보면 93.6%의 얼굴 검출률로 제안하는 세 종류의 부 영역을 사용한 경우와 거의 유사한 결과를 보여 주었지만 이러한 결과는, 제안하는 알고리즘에서 사용하지 않은 대각선 고주파 성분(D)이 사실상 얼굴 검출률에 큰 영향을 미치지 못한다는 것을 의미하며 같은 수준의 얼굴 검출률이라면 형판 저장을 위한 메모리 용량과 형판 정합

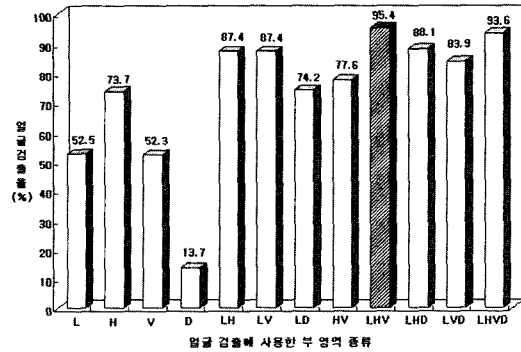


그림 9 얼굴 검출에 사용한 부 영역 종류에 따른 얼굴 검출률

L, H, V, D : 하나의 부 영역을 선택하여 얼굴 형판 생성, 정합
 LH, LV, HV : 두 종류의 부 영역을 선택하여 두 개의 얼굴 형판 생성, 정합

LHV, LHD, LVD : 세 종류의 부 영역을 선택하여 세 개의 얼굴 형판 생성, 정합

LHVD : 네 종류의 부 영역을 선택하여 네 개의 얼굴 형판 생성, 정합

에 있어서의 계산량을 고려할 때, 제안하는 방법과 같이 세 가지 성분을 이용하여 얼굴 검출의 효과를 극대화하는 것이 더욱 바람직하다고 보여진다.

4.2 전 처리

앞서 영상에 나타난 다양한 빛의 변화를 보정하기 위해 입력 영상의 평균 밝기 값에 따라 최소-최대 정규화 와 히스토그램 평활화를 조합하여 적용하는 전처리를 제안하였다. 6종류의 동영상 데이터에 아무런 전 처리를 적용하지 않은 경우의 평균 얼굴 검출률은 61.3%로 측정되었으나 제안한 전 처리 방법을 적용한 결과 약 95.4%로 매우 좋은 얼굴 검출률을 보여 주었다. 전 처리 방법의 성능 비교를 위해 일반적으로 많이 사용되는 다른 전 처리 방법(히스토그램 평활화, 명암 대비 스트

레칭(contrast stretching)(cut-off 8.5%), zero-mean 정규화, 감마 상관 관계 변환(gamma correction)($\gamma=1.5$), 최소-최대 정규화)을 적용하여 평균 얼굴 검출률을 비교하였는데, 그 결과는 그림 10에 보여지듯 92.5%, 55.7%, 80.6%, 63.9%, 86.0%로 제안한 전 처리 방법의 뛰어난 성능(95.4%)을 확인할 수 있었다.

4.3 탐색 영역 제한 방법 및 얼굴 추적 알고리즘

2.4.2에서 이전에 검출된 얼굴 위치를 이용하여 다음 프레임에서의 얼굴 위치를 추정하고 추정된 얼굴 위치를 중심으로 탐색 영역을 정하도록 하였다. 또한 정해진 탐색 영역 안에서도 색상 정보를 이용하여 추출된 얼굴 후보 영역에 대해서만 형판 정합이 수행되도록 하는 내용을 설명하였다. 본 논문에서 사용한 방법의 성능을 판단하기 위하여 ① 본 논문에서 제안하는 추적 방법과 ② 얼굴 추적 알고리즘을 사용하지 않고 프레임 전체를 탐색 영역으로 잡아 색상 정보를 이용하여 추출된 얼굴 후보 영역에 대해서만 형판 정합을 수행하는 방법(전역 탐색, full search), ③ 추적 알고리즘을 사용하여 탐색 영역을 정하긴 하되 색상 정보를 이용하지 않고 탐색 영역 전체에 형판 정합을 수행하는 방법등에 대해 얼굴 검출률 및 평균 수행 시간을 비교하여 보았다. 그림 11에서 보는 것처럼 탐색 영역을 제한해 주는 것(①)은 전역 탐색(②)과 비교하였을 때, 갑작스런 조명 변화에 의한 얼굴 오 검출 및 놓침을 예방하여 얼굴 검출률을 높이고 수행 시간을 줄여주는 효과가 있었다. 또한 제안한 바와 같이 탐색 영역 안에서도 색상 정보를 이용하여 검출된 얼굴 후보 영역에 대해서만 형판 정합을 수행하여 주는 것은 수행 시간을 줄이는데 효과가 있었다.

그림 12는 제안한 추적 알고리즘과 제안한 추적 알고리즘이 기반을 두고 있는 [7]의 추적 알고리즘의 얼굴

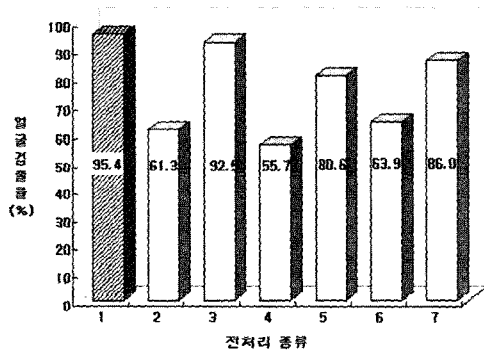


그림 10 전 처리 방법 별, 동영상 데이터에 대한 평균 얼굴 검출률

- 1 : 제안하는 전 처리, 2 : 전 처리를 적용하지 않은 경우,
- 3 : 히스토그램 평활화, 4 : 명암 대비 스트레칭,
- 5 : zero-mean 정규화, 6 : 감마 상관 관계 변환,
- 7 : 최소 최대 정규화

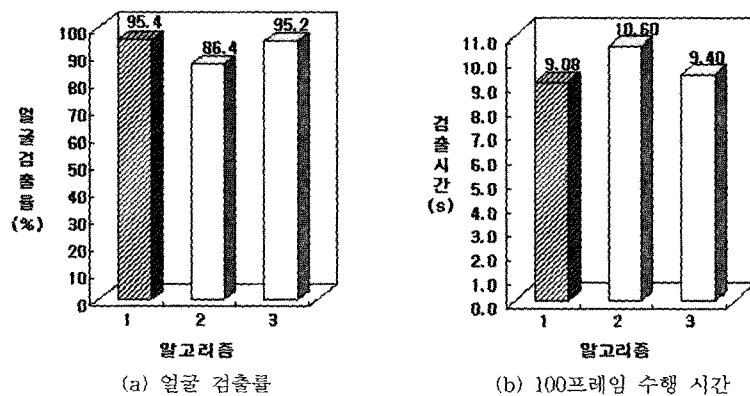
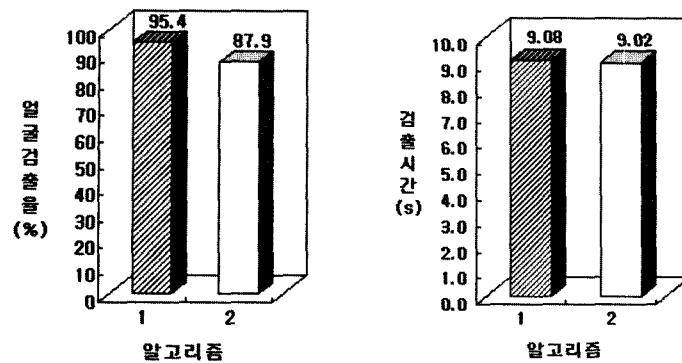


그림 11 탐색 영역 제한 방법에 따른 얼굴 검출률 및 수행 시간

알고리즘 ① : 본 논문이 사용한 탐색 영역 제한, 알고리즘 ② : 전역 탐색, 알고리즘 ③ : 탐색 영역 제한



(a) 얼굴 검출률 (b) 100프레임 수행 시간

그림 12 추적 알고리즘에 따른 얼굴 검출률 및 수행 시간

알고리즘 1 : 본 논문이 제안한 추적 알고리즘, 알고리즘 2 : [7]의 추적 알고리즘

표 2 제안한 눈 검출 알고리즘의 수행 결과

(*) : 얼굴 검출 알고리즘에 의해 성공적으로 검출된 얼굴 영역의 수(프레임 수)

동영상(*)	양쪽 눈 모두 검출	한쪽 눈만 검출	양쪽 눈 모두 검출 실패	검출률(%)
(1) (311)	311	0	0	100.0
(2) (500)	463	22	15	92.6
(3) (500)	404	95	1	80.8
(4) (394)	390	4	0	99.0
(5) (291)	261	24	6	89.7
(6) (290)	259	15	16	89.3
평균				91.9

검출률을 그래프로 표현한 것이다. [7]의 추적 알고리즘과 비교하였을 때, 제안한 추적 알고리즘을 이용하여 탐색 영역을 제한하는 것은 추적 알고리즘의 계산량에 따른 미세한 수행 시간 증가를 나타냈지만, 얼굴 검출률에 있어서는 긍정적인 영향을 확인할 수 있었다.

4.4 눈 검출 알고리즘

제안하는 눈 검출 알고리즘은 두 눈을 포함하고 있는 비교적 정확하게 검출된 얼굴 영역이 입력으로 들어온다는 가정을 세우고 있다. 그러므로 눈 검출 알고리즘의 성능 평가는 제안하는 얼굴 검출 알고리즘에 의해 검출된 얼굴 영역 중 성공적으로 검출된 얼굴 영역을 대상으로 하였고 눈 검출률은 정확하게 얼굴 영역이 검출된 프레임 중, 두 눈이 모두 검출된 프레임의 수로 계산해 주었다.

5. 결론 및 향후 연구 과제

본 논문에서는 화상 통신(telecommunication)을 위한 새로운 얼굴 검출 및 추적, 그리고 눈 검출 알고리즘을 제안하였다. 제안한 얼굴 검출 알고리즘은 복잡한 배경과 다양한 조명 조건에 관계없이 얼굴을 검출하기 위해

인간의 시각 체계와 유사한 웨이블릿 변환을 적용하여 얼굴 영역의 특성을 나타내기 적합한 세 종류의 주파수 영역으로 저해상도 얼굴 형판을 생성하고 웨이블릿 변환된 입력 영상과의 유사도를 비교하여 얼굴 영역을 검출한다. 웨이블릿 얼굴 형판의 낮은 해상도는 형판 정합의 복잡도를 낮추어 주었고 세 종류의 주파수 영역은 얼굴 특성을 효과적으로 표현하여 얼굴 영역 검출의 정확도를 높여 주었다. 세 가지 크기(40×40, 60×60, 80×80)로 생성한 얼굴 형판은 영상에 존재하는 다양한 크기의 얼굴을 검출할 수 있도록 해 주었는데, 알고리즘의 응용 분야를 고려할 때 미리 생성한 형판을 메모리에 저장해 이용하는 것은 처리기의 과부하를 줄여줄 수 있는 중요한 부분이라 판단된다. 다양한 조명 조건을 극복하기 위해 영상의 밝기 평균값에 따라 최소-최대 정규화와 히스토그램 평활화를 혼합하여 적용한 전 처리는 매우 밝거나, 매우 어두운 영상에서의 얼굴 오 검출 및 놓침을 감소시켜 주었다. 또한 제안한 얼굴 추적 알고리즘 이용하여 다음 프레임에서의 얼굴 위치를 예측하고 예측된 지점을 중심으로 탐색 영역을 정해 형판 정합을 수행하도록 함으로써 얼굴 검출의 정확도를 높

이고 수행 시간도 단축시킬 수 있었다. 눈 후보 영역의 수직, 수평 방향 투영 결과를 합리적으로 분석하는 눈 검출 알고리즘은 매우 낮은 밝기 값을 갖거나 앞머리가 내려온 얼굴 등 비교적 질이 나쁜 얼굴 영역에 대해서도 만족스러운 결과를 보여주었다.

그러나 아직 해결되어야 할 문제들은 남아 있다. 우선 제안한 얼굴 검출 알고리즘이 실시간 응용을 고려하여 앞단에 색상 정보를 이용하여 얼굴 후보 영역을 검출하고 검출된 얼굴 후보 영역에만 형판 정합을 수행하기 때문에 여타 연구들과 마찬가지로 빛의 스펙트럼에 의해 발생한 색상 왜곡에 대해서는 자유로워 질 수 없다. 그리고 제안한 얼굴 검출 알고리즘을 통해서도 상하-좌우 $-10^{\circ} \sim +10^{\circ}$ 정도 회전된 얼굴은 만족스럽게 검출이 가능하나 그 이상 회전된 경우엔 얼굴 검출을 보장할 수 없기 때문에 측면 얼굴, 상하-좌우로 기울어진 얼굴에 대한 알고리즘이 추가되어야 한다. 또한 본 논문에서 생성한 얼굴 형판은 한국의 20대 대학생 얼굴들을 대상으로 하였기 때문에, 인종, 나이, 성별 등에서 다양한 조건을 갖는 얼굴을 수집하여 새로운 얼굴 형판을 생성하고 다양한 조건의 실험 동영상을 대상으로 상당한 실험을 수행해야 할 것이다. 그리고 눈 검출에 있어서 현재 알고리즘은 이전 단계에서 얼굴 영역이 정확하게 검출되었다는 가정을 하고 있기 때문에 눈 검출은 검출된 얼굴 영역에 의존적이 된다. 따라서 눈 검출 알고리즘이 갖고 있는 가정을 완화시켜 실제 얼굴 보다 크게 검출된 얼굴 영역의 경우를 비롯하여 두 눈 영역이 검출된 얼굴 영역의 상단에 위치하지 않는 경우, 눈 영역이 한 쪽으로 치우쳐 있는 경우 등에서도 눈 영역을 검출할 수 있도록 지속적인 연구가 수행되어야 할 것이다.

참 고 문 헌

- [1] Sung, K.-K., Poggio, T. "Example-based learning for view-based human face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 20 Issue: 1, pp. 39-51, Jan. 1998.
- [2] Cohen, I., Sebe, N., Garg, A., Lew, M.S., Huang, T.S. "Facial expression recognition from video sequences," *IEEE International Conference on Multimedia and Expo*, Volume: 2, pp. 121-124, 2002.
- [3] Yamaguchi, T., Tominaga, M., Murakami, K., Koshimizu, H. "Re-generation of facial image eye contacting with partner on TV conference environment," *IEEE International Conference on Systems, Man, and Cybernetics*, Volume: 2, pp. 1169-1174, 2000.
- [4] Collet, C., Finkel, A., Gherbi, R. "CapRe: a gaze tracking system in man-machine interaction," *IEEE International Conference on Intelligent Engineering Systems*, pp. 577-581, 1997.
- [5] van der Schaar, M., Lin, Y.-T. "Content-based selective enhancement for streaming video," *International Conference on Image Processing*, Volume: 2, pp. 977-980 vol.2, 2001.
- [6] Xu, G., Sugimoto, T. "Rits Eye: a software-based system for real-time face detection and tracking using pan-tilt-zoom controllable camera," *International Conference on Pattern Recognition*, Volume: 2, pp. 1194-1197 vol.2, 1998.
- [7] Phillips, P.J., Hyeonjoon Moon, Rizvi, S.A., Rauss, P.J. "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 22 Issue: 10, pp. 1090-1104, Oct. 2000.
- [8] Rein-Lien Hsu, Abdel-Mottaleb, M., Jain, A.K. "Face detection in color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 24 Issue: 5, pp. 696-706, May 2002.
- [9] Jaeyoung Seol, Kwanghoon Sohn, Chulhee Lee "Selective coding of human faces using wavelet," *IEEE International Conference on Systems, Man, and Cybernetics*, Volume: 2, pp. 1384-1387 vol.2, 2000.
- [10] Kotropoulos, C., Pitas, I. "Rule-based face detection in frontal views," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume: 4, pp. 2537-2540, 1997.
- [11] S.A. Sirohey. "Human face segmentation and identification," *Technical Report CS-TR-3176*, University of Maryland, 1993.
- [12] Kin Choong Yow, Cipolla, R. "A probabilistic framework for perceptual grouping of features for human face detection," *Second International Conference on Automatic Face and Gesture Recognition*, pp. 16-21, 1996.
- [13] Huitao Luo "Algorithms for Video Object Detection and Segmentation with Application to Content-Based Multimedia Systems," *Dissertation of Ph. D. Columbia University*, 2000.
- [14] Birchfield, S. "Elliptical head tracking using intensity gradients and color histograms," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 232-237, 1998.
- [15] Graf, H.P., Cosatto, E., Gibbon, D., Kocheisen, M., Petajan, E. "Multi-modal system for locating heads and faces," *Second International Conference on Automatic Face and Gesture Recognition*, pp. 88-93, 1996.
- [16] Lixin Fan, Kah Kay Sung "Face detection and pose alignment using colour, shape and texture information," *Third IEEE International Workshop on Visual Surveillance*, pp. 19-25, 2000.
- [17] Liu, Z., Wang, Y. "Face detection and tracking in

- video using dynamic programming," International Conference on Image Processing, Volume: 1, pp. 53-56 vol.1, 2000.
- [18] Menser, B., Muller, F. "Face detection in color images using principal components analysis," Seventh International Conference on Image Processing and Its Applications, Volume: 2, pp. 620-624 vol.2, 1999.
- [19] Karlekar, J., Desai, U.B. "Finding faces in color images using wavelet transform," International Conference on Image Analysis and Processing, pp. 1085-1088, 1999.
- [20] Jun Miao, Baocai Yin, Kongqiao Wang, Lansun Shen, Xuecun Chen, "A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template," Pattern Recognition, Volume 32, Issue 7, pp. 1237-1248, July 1999.
- [21] Chai, D. Ngan, K.N. "Locating facial region of a head-and-shoulders color image," Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 124-129, 1998.
- [22] N. Oliver, A. Pentland, F. Berard "LAFTER: a real-time face and lips tracker with facial expression recognition," Pattern Recognition, 33, pp. 1369-1382, 2000.
- [23] Hyoung Woo Lee, Se Kee Kil, Younghwan Han, SeungHong Hong "Automatic face and facial features detection," IEEE International Symposium on Industrial Electronics, Volume: 1, pp. 254-259 vol.1, 2001.
- [24] J. Yang, R. Stiefelwagen, U. Meier, A. Waibel "Real-time face and facial feature tracking and applications," AVSP'98, Terrigal, Australia, pp. 79-84, 1998.
- [25] Haizhou Ai, Luhong Liang, Guangyou Xu "Face detection based on template matching and support vector machines," International Conference on Image Processing, Volume: 1, pp. 1006-1009 vol.1, 2001.

고 병 철

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 8 호 참조

변 혜 란

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 8 호 참조



송 해 진

1997년~2001년 2월 숙명여대 컴퓨터과
학과 졸업(이학사). 2001년 3월~2003년
2월 연세대학교 대학원 컴퓨터학과 졸업
(공학석사). 2003년 2월~현재 LG전자
Mobile Multimedia 연구소 연구원. 관심
분야는 얼굴검출, 인식, 컴퓨터 비전, 패

턴인식