

등고선 영역의 투영 벡터를 이용한 3차원 얼굴 인식

이영학[†] · 심재창^{**} · 이태홍^{***}

요 약

본 논문은 3차원 얼굴 영상으로부터 등고선 영역을 추출하여 얼굴의 지역적 특징이 잘 반영되는 투영 벡터를 이용한 얼굴 인식 알고리즘을 제안한다. 얼굴의 외곽 형상은 사람에 따라 비슷한 모양을 나타내므로 구분하는데 어려움이 많다. 그러나 3차원 얼굴 영상은 깊이 정보를 갖고 있으므로, 코로부터 일정 깊이 값에 대한 영역을 추출하면 사람마다 다른 형상이 추출 될 수 있다. 얼굴 내에서 가장 높은 코를 먼저 추출한 후, 이를 기준으로 깊이 값을 취하면, 코를 포함한 얼굴 내의 등고선 영역을 추출하였다. 이 영역 또한 비슷한 형상을 나타낼 수 있으므로, 논문에서는 영상을 투영한 후 투영 벡터의 국부화를 통하여 영상의 지역적 특징이 잘 반영되는 통계적 성질의 투영 벡터 방법을 사용하여 특징 벡터를 추출하였다. 제안된 방법을 이용한 유사도 비교는 입력과 데이터 베이스에 대하여 각각 두개의 깊이 데이터에 대해 유클리드 거리를 사용하였으며, 실험 결과 5위 이내의 인식률이 94.3%로 나타났다.

3D Face Recognition using Projection Vectors for the Area in Contour Lines

Yeunghak Lee[†], Jaechang Shim^{**} and Taihong Yi^{***}

ABSTRACT

This paper presents face recognition algorithm using projection vector reflecting local feature for the area in contour lines. The outline shape of a face has many difficulties to distinguish people because human has similar face shape. For 3 dimensional(3D) face images include depth information, we can extract different face shapes from the nose tip using some depth values for a face image. In this thesis deals with 3D face image, because the extraction of contour lines from 2 dimensional face images is hard work. After finding nose tip, we extract two areas in the contour lines from some depth values from 3D face image which is obtained by 3D laser scanner. And we propose a method of projection vector to localize the characteristics of image and reduce the number of index data in database. Euclidean distance is used to compare of similarity between two images. Proposed algorithm can be made recognition rate of 94.3% for face shapes using depth information.

Key words: 얼굴 인식

1. 서 론

일상생활에서 사람의 신체를 이용하여 사용자임을 확인 또는 인식하는 생체인식 연구가 지금까지 활발히 진행되고 있다. 생체인식은 자동으로 사용자임을 인식하기 위해 사람이 가지고 있는 생체정보인 - 지문, 얼

굴, 음성, 홍채, 망막, 손 모양, 귀 모양 그리고 정맥 등이 사용되고 있다[1,2]. 이러한 생체인식 기술은 외적인 영향에 의해 변형되지 않는다면, 분실, 원형의 변형 그리고 복제가 불가능하다는 장점을 가지고 있다. 얼굴인식 기술은 다른 생체 인식 분야에 비해 비접촉성과 비강제성의 장점을 가지고 있기 때문에 생체인식 가운데서도, 지문인식 다음으로 활발히 연구되고 있다. 얼굴인식이 사용되는 분야는 신분확인(주민등록증, 운전면허증), 현금자동인출기, 입후보자 등록, 보안, 범 집행과 감시 그리고 빌딩이나 사무실의 사용자

접수일 : 2002년 8월 29일, 완료일 : 2002년 11월 12일

[†] 대구과학대학교 겸임교수

^{**} (주)파미, (주)포디컬쳐, (주)페이스쓰리디 공동설립

^{***} 정희원, 서울대학교 대학원 전자공학과 졸업 (박사)

확인 등에 가능하다[3]. 그러나 얼굴인식이 앞에서 언급한 장점을 지니고 있는 반면, 실질적인 응용에서는 여러 가지 문제점을 가지고 있다. 가장 두드러진 문제점은 인식환경인 조명 변화와 카메라각도 그리고 얼굴 혹은 머리형태로서 안경, 머리모양, 표정 그리고 나이 등에 의해 크게 영향을 미친다는 것이다[4].

지난 수년간 얼굴인식은 밝기영상이나 컬러 영상을 대상으로 대부분 연구되어왔다[5]. 그러나 극히 일부만이 깊이 영상을 가지고 얼굴인식을 다루었다[6-9]. 밝기영상과 컬러영상들은 조명변화, 물체의 반사율 그리고 카메라의 각도에 대해 영상이 민감하게 반응하여 인식률이 크게 저하된다. 이러한 문제점을 해결하기 위하여 깊이가 있는 얼굴정보를 획득하여 얼굴인식에 사용한다.[10-12]. 3차원 얼굴 영상을 획득하는 방법에는 스테레오 정합법과 레이저 광 삼각법 등이 있다 [13-16]. 스테레오 정합법은 컴퓨터 비전의 기본원리인 두 눈간의 시차(binocular parallax)를 이용하여 깊이 정보를 측정하는 원리로서, 두 대의 카메라로 영상을 촬영하여 두 영상의 국부 영역의 매칭을 통해 변이(disparity)값을 구하여 3차원 정보를 추출하는 방법이다. 가까운 곳으로부터 먼 곳까지의 3차원 정보를 추출할 수 있는 장점이 있는 반면 정확도가 높지 않아 얼굴 인식에 적용하는 것이 어렵다. 반면, 레이저 광 삼각법은 레이저 쉬트빔(sheet beam)을 얼굴에 주사하면서 CCD카메라를 통해 촬영된 영상에서 광 삼각법을 이용하여 깊이 정보를 알아내는 방식이다. 레이저를 이용하는 또 다른 방법으로는 레이저를 보내고 받는 시간의 차이 값에 의해 깊이 정보를 알아내는 방식이다. 이러한 레이저 방식은 얼굴에 대해 깊이정보를 정확하게 획득 할 수 있으며, 필터와 레이저를 활용하므로 2차원 얼굴영상을 획득할 때 문제가 되는 조명의 영향과 카메라의 각도에 의한 영향을 거의 받지 않는 장점을 가지고 있다. 또한 3차원 얼굴영상은 얼굴과 카메라까지의 거리 차에 따라 생기는 축소 효과도 거리가 측정되므로 작게 할 수 있다[17].

3차원 얼굴영상은 2차원이 가지는 얼굴, 눈, 코 그리고 입에 대한 위치적 정보 이외에, 깊이정보를 가지고 있으므로, 지금까지의 연구가 얼굴의 주요 해석은 곡률을 이용하여 연구되어져 왔다. Lee와 Milios[6]은 얼굴의 각 픽셀에서의 평균 곡률과 가우시안 곡률의 부호를 이용하여 얼굴의 볼록한 부분을 추출한 후 확장 가우시안 영상(extended gaussian image)을 이용하여 두 영상의 얼굴 특징들을 비교하는 방법을 제시하였다.

Gordon[7]은 깊이와 곡률 특성을 이용하여 인식하는 방법을 제시하였다. 콧등과 눈의 특징들처럼, 얼굴 표면의 곡률들을 계산하여 얼굴을 정확하게 묘사하였다. Hiromi등[8]은 3차원의 단단한 물체의 표면에서 생기는 문제를 가지고 얼굴인식에 적용하였다. 입력과 데이터베이스 영상들에 대해 얼굴 표면의 주 곡률과 방향들로 구성된 확장 가우시안 영상(EGI)으로 표현하였고, 이를 Fisher의 EGI에 관한 구의 상관관계를 이용해 두 영상간의 유사도를 평가하여 인식하였다. Chua등[9]는 단단하지 않은 물체의 표면에 나타나는 포인트 시그니처(point signature)를 이용하여 얼굴인식에 적용하였다.

지금까지의 연구에서 대부분 논문들은 얼굴의 곡률을 이용하여 3차원 얼굴을 인식하는 알고리즘들을 제안하였다. 본 연구에서는 얼굴의 깊이에 해당되는 등고선을 이용한 것으로서, 사람마다 같은 깊이 값에 대하여 다른 얼굴 형상들이 나타난다는 점에 착안하여 등고선 영역을 이용하는 새로운 접근법을 제안한다. 또한 얼굴의 국부적인 특징을 잘 나타내고 계산량이 적은 통계적 성질의 투영 벡터 방법을 이용하여 특징 벡터들을 추출한다. 입력 영상과 데이터 베이스 영상 사이의 투영 벡터들을 유클리드 거리를 계산하여 유사도를 비교하고자 한다.

2. 등고선 영역의 투영 벡터를 이용한 얼굴 인식

2.1 3차원 얼굴 영상에서 코 위치 찾기

얼굴과 배경이 포함된 영상에 대해 전처리 과정으로써 배경영역과 얼굴 영역을 먼저 분리한다. 머리카락, 배경 그리고 옷은 코의 위치를 찾는 데 오류를 발생시킬 수 있는 가장 큰 요인이고, 불필요한 영역이므로 먼저 제거한다. 여기서 배경은 최저의 깊이 데이터를 갖는다. 입력 영상 그림 2 (a)를 식 (1)에 나타난 소벨(sobel) 연산자를 이용하여 그림 2 (b)처럼 얼굴, 배경 그리고 옷의 경계선을 추출한다. 여기서 사용된 σ 는 10이고 임계값(threshold)은 255로 하였으며, I 는 입력 영상이며 S 는 소벨 연산을 적용한 영상으로서, 모두 3차원 영상 값을 갖는다.

$$S_{xx} = d6 + 2d7 + d8 - d0 - 2d1 - d2$$

$$S_{yy} = d2 + 2d5 + d8 - d0 - 2d3 - d6$$

$$T = \sigma \cdot \sqrt{S_{xx}^2 + S_{yy}^2}$$

$$S(i, j) = \begin{cases} 255, & T > \text{Threshold} \\ I(i, j), & \text{else} \end{cases} \quad (1)$$

d0	d1	d2
i-1, j-1	i-1, j	i-1, j+1
d3	d4	d5
i, j-1	i, j	i, j+1
d6	d7	d8
i+1, j-1	i+1, j	i+1, j+1

그림 1. 3×3 마스크내의 9화소

배경으로 인한 잡음을 제거하기 위하여 네 모서리마다 5×5 영역을 취하여 이 영역들 가운데서 최고 값을 구하여 최고 값보다 작은 값을 255로 처리한다. 처리 결과는 그림 2 (c)와 같다. 네 모서리 영역을 취한 것은 얼굴보다 옷과 배경이 더 깊이 위치하기 때문이다. 배경을 제거한 3차원 영상에 대하여 그림 2 (c)처럼 얼굴 아래의 옷 부분은 아직 제거되지 않았으므로 식 (2)를 이용하여 평균값을 구한 후 이를 임계값으로 사용하여 이진 영상화시키면 그림 2 (d)와 같다. N은 3차원 값을 가지는 화소 수를 나타낸다. B는 이진 영상이다.

$$Avg = \frac{1}{N} \sum S(i, j) \quad \text{if, } S(i, j) \neq 255$$

$$B(i, j) = \begin{cases} 0, & S(i, j) > Avg \\ 255, & \text{else} \end{cases} \quad (2)$$

머리카락에 대한 잡음들을 제거하고 얼굴만을 추출하기 위해 얻어진 그림 2 (d)의 영상을 레이블링 (labeling)을 한 다음 가장 넓은 영역을 추출하면 그림 2 (e) 와 같이 배경, 머리카락 그리고 옷이 제거된 얼굴 영역만을 추출한 2진 영상을 얻을 수 있다.

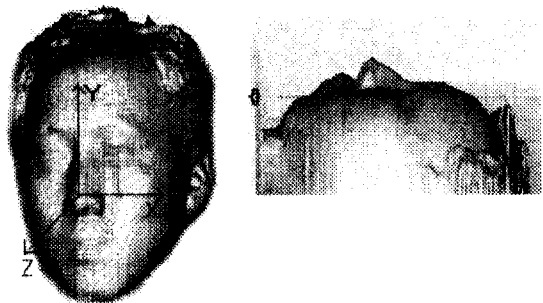


(a) 원 영상 (b) Sobel 처리 후 (c) 모서리 처리 후 (d) 임계 처리 (e) 배경 제거 영상

그림 2. 전처리 과정

사람의 얼굴은 전체적으로 곡면으로 이루어져 있으며, 정면의 얼굴에서 보면 크게 눈(좌, 우), 코 그리고 입의 3가지 부위로 나눌 수 있다. 입력된 영상은 3D 레이저 스캐너에 의해 획득된 3차원 깊이 정보를 가지고 있으므로 그림 3 (a)처럼 X, Y, Z 좌표를 가진다. 일반적으로 정면 얼굴에서의 최고점은 코끝이 된다. 그리고 얼굴이 심하게 회전된 경우와 특정인의 특별한 얼굴 형상을 제외하고는, 기울어진 영상과 회전된 영상에 대해서도 최고점은 코끝이 된다.

얼굴 영상을 옆으로 회전시켜 보면 그림 3 (b)에서처럼 최고점이 코끝이라는 것을 알 수 있다. 그림 3에 나타내는 것처럼 코는 얼굴의 중앙 영역에 위치하며, 가장 높은 정점을 가지고 있으므로 기준점으로써의 역할을 할뿐만 아니라 찾기가 용이한 특징이 있으므로, 코를 먼저 기준점으로 찾는다. 먼저 식 (3)을 이용하여 배경이 제거된 이진 영상 B를 이용하여 원 영상 I로부터 얼굴 영역을 나타내는 새로운 영상 P를 얻는다. 최고점은 식 (4)과 (5)를 이용한 반복 선택 방법을 이용하여 구하며, 임계값은 평균값으로 한다. 식 (4)을 이용하여, 배경이 제거된 얼굴 영역에서의 깊이 값에 대한 평균을 구하여 임계값을 선정한다. 식 (4)에서 구한 임계값을 이용하여 임계값 이상이 되는 영역만을 취한



(a) 3차원 좌표 (b) 옆에서 본 모습

그림 3. 3차원 얼굴 영상

것이 식 (5)과 같다. 여기서 M과 N은 가로 세로의 영상 크기이다.

$$P(i, j) = \begin{cases} I(i, j), & \text{if } B(i, j) = 0 \\ 0, & \text{else} \end{cases} \quad (3)$$

$$Avg = \frac{1}{N} \sum P(i, j), \quad P(i, j) > 0 \quad (4)$$

$$P(i, j) = \begin{cases} P(i, j), & \text{if } P(i, j) > Avg \\ 0, & \text{else} \end{cases} \quad (5)$$

식 (5)에서 구한 결과 영역 내에서 식 (4)을 이용하여 새로운 임계값을 구하고 식 (5)을 이용한다. 이 과정을 반복하면, 임계 평균값 이상이 되는 영역만이 추출되며, 최종적으로는 얼굴에서 가장 높은 값을 가지는 코의 정점이 추출된다. 사람에 따라서 코의 끝은 여러 가지의 형태를 가질 수 있으므로 최고점을 갖는 화소가 여러 개 있을 수 있다. 이 화소들로 이루어진 영역들의 무게 중심좌표를 구하여 최고점의 좌표 MAX(X, Y)를 구하면 그림 4와 같다.



그림 4. 3차원 영상에서의 최고점 추출

2.2 정규화 과정

사람의 얼굴을 입력받을 때 생기는 얼굴의 자세 변화에 대하여 정규화를 실시한다. 자세에 대한 변형은 크게 회전(panning), 기울어짐(rotation) 그리고 틸팅(tilting)의 3가지로 구분할 수 있다. 그림 5에 나타낸 것처럼 회전, 기울어짐 그리고 틸팅의 순서로 정규화를 실시한다. 먼저 입력된 영상에 대하여, -15° 부터 +15° 까지 최고점을 중심으로 영상을 회전시키면서 최고점의 좌, 우 거리가 50 픽셀 이내의 깊이 값의 평균의 차이가 가장 작게되는 지점의 각도를 찾아서 그 각도만큼 최고점을 중심으로 Y축을 중심으로 영상을 회전시킨다. 회전된 영상에 대하여 최고점으로부터 임의

의 깊이 값 이상이 되는 영역을 선택하고, 이 영역에 대하여 -15° 부터 +15° 까지 최고점을 중심으로 영상을 회전시켰을 경우, 수직 축을 중심으로 그 면적의 좌, 우 차이가 가장 작게되는 각도를 구하여 그 각도만큼 기울어진 영상을 Z축을 중심으로 회전시킨다. 이상의 두 가지 과정이 끝난 영상에 대하여 미간 사이의 중심점(nose bridge)과 코 밑점(nose base)을 찾아서, 두 지점의 높이 값의 차이가 10이 되도록 X축을 중심으로 영상을 회전시킨다. 여기서 각도 ±15° 는 정면 얼굴 영상인 경우 각도가 크게 벗어나지 않는다는 가정 하에 사용하였다.

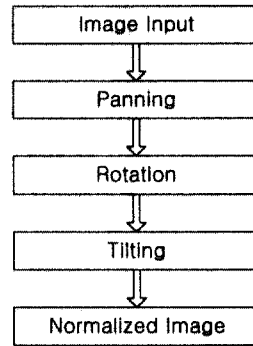


그림 5. 얼굴 자세 변화에 대한 정규화 과정

2.3 등고선 영역의 추출

배경이 제거된 3차원 영상에서 최고점으로부터의 일정 깊이에 있는 등고선 값을 취하여 그 내부영역과 외부영역으로 이진화 영상을 얻는다. 얼굴의 형태에 따라서 코 주변 이외의 불필요한 부분이 포함 될 수 있으며, 그림 6와 같다.

이를 제거하기 위하여 이진 영역에 대한 레이블링 작업을 실시하고 레이블링 영역에 대한 면적을 구한 후, 가장 넓은 면적을 가진 영역만을 취하여 최종적으로 선택한다. 레이블링 작업은 이진 영상의 시작부분부터 픽셀들을 읽어오면서 검은색 부분의 픽셀을 만나면 8방향 탐색을 시행하여 다른 픽셀의 연결 여부를 판단한다. 연결이 되어 있으면 자기 자신은 이전과 같은 레이블 값을 기록하고 다시 8방향 탐색을 한다. 만약 더 이상 연결 픽셀이 없을 경우 8방향 탐색을 멈추고 다음 픽셀로 이동하여 검은색 여부를 판단하고 검은색 픽셀이 존재하면 레이블 값을 증가시킨다.

등고선 값의 결정은 사용자가 임의로 설정을 할 수

있다. 등고선 값들에 대한 추출 결과가 그림 5와 같다. 아주 작은 면적을 나타내는 그림 7 (a)과 아주 큰 면적을 나타내는 그림 7 (b)은 사람마다 얼굴의 형상이 비슷한 형태를 나타내므로 얼굴을 구분하는데 많은 어려움이 존재한다. 그래서 본 논문에서 사용한 영역은 그림 7 (b)과 (c)이다. 각 영상의 최고점의 좌표가 영상의 중심 위치에 놓이게 하기 위하여 중심 좌표와 최고점 좌표와의 차이 값을 구하여 영상을 이동시킨다.

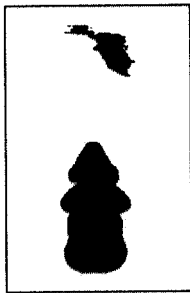


그림 6. 깊이 값 20에 대한 영역추출 결과

2.4 투영벡터

2.4.1 영역의 투영기법

일반적으로 투영은 n차원의 좌표 시스템의 점들을 n보다 작은 차원의 점들로 변환하는 역할을 수행한다. 투영을 영상에 적용할 경우 영상의 개략적인 형상 정보를 얻을 수 있으며, 1차원적으로 적용할 경우 행 방향으로의 투영을 통하여 수평 방향으로의 투영 벡터를 얻을 수 있고, 열 방향으로의 투영을 통해서는 수직 방향으로의 투영벡터를 얻을 수 있다. 이진 영상의 경우에 행 및 열 방향으로 투영을 적용한 경우를 그림 8과 같이 나타내었다. 그림 8에서와 같이 이진 영상에서의 투영벡터는 행 또는 열 방향으로의 화소 값들의 합으로 표시된다는 것을 알 수 있다[18].

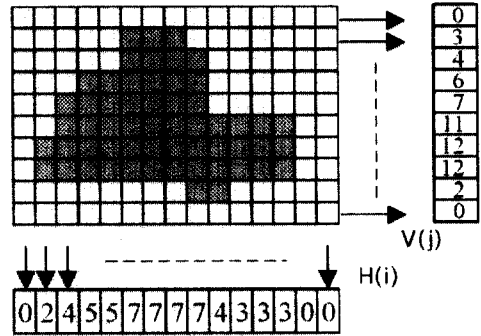


그림 8. 이진 영상에서의 투영예

크기가 M×N인 영상을 $I(i, j)$ 라 할 때 수평 및 수직 방향으로의 투영은 식 (6) 및 (7)과 같이 나타낼 수 있다.

$$H(i) = \sum_{j=0}^{N-1} I(i, j) \tag{6}$$

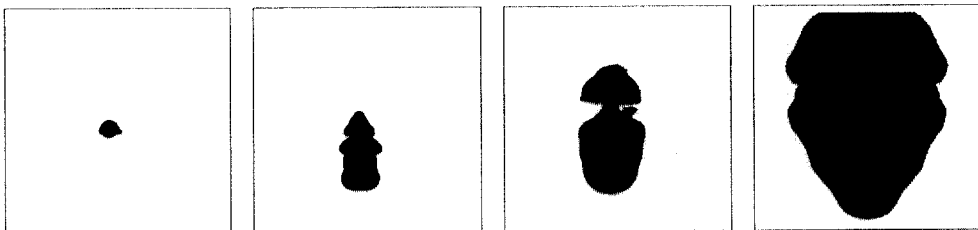
$$V(j) = \sum_{i=0}^{M-1} I(i, j) \tag{7}$$

영상에서의 1차 모멘트를 이용하여 물체의 위치를 알 수 있는 것과 마찬가지로, 영상에서의 1차 모멘트가 투영에서의 1차 모멘트로 나타난다는 성질을 이용하여 투영된 벡터에서 물체의 위치를 계산할 수 있다. 식 (8)을 이용하여 수평 및 수직 방향의 위치를 계산하면 식 (9)과 식 (10)과 같이 행 방향 및 열 방향의 위치를 계산할 수 있다[13].

그러나, 이와 같은 투영을 이용할 때 공간정보가 포함되어 인식 분야에 유용하게 사용된다는 장점이 있는

$$A = \sum_{i=0}^{N-1} H(i) = \sum_{j=0}^{M-1} V(j) \tag{8}$$

$$\bar{y} = \frac{\sum_{i=0}^{N-1} iH(i)}{A} \tag{9}$$



(a) 깊이등고선 값10 (b) 깊이등고선 값20 (c) 깊이등고선 값30 (d) 깊이등고선 값40

그림 7. 코로부터 등고선 깊이 값에 따른 추출된 이진 영상

$$\bar{x} = \frac{\sum_{j=0}^{M-1} jV(j)}{A} \quad (10)$$

반면 다양한 크기의 영상을 포함하는 데이터베이스 영상을 색인화 할 때 투영기법을 사용하기에는 문제가 있다. 즉 투영 기법을 사용할 때 일반적으로 영상의 크기에 따라 생성되는 투영 벡터의 크기가 달라지므로 동등한 길이의 벡터를 이용한 비교가 불가능하다는 문제가 있다. 또한 영상의 행 및 열의 수만큼의 벡터 요소가 색인데이터에 사용되기 때문에 검색을 위한 데이터베이스 색인 정보로 사용할 경우 색인 데이터량이 많아 대용량 데이터베이스에서는 결과를 얻기까지 많은 시간이 걸리는 단점이 있다.

2.4.2 투영벡터의 그룹 분산을 이용한 비교

투영 기법을 영상 색인화에 적용할 때의 문제점을 해소하기 위하여 식 (11)과 (12)에 나타낸 것처럼 투영된 투영벡터를 그룹화 하여 각 그룹별 합을 구하고 그 값을 이용하여 색인 정보로 사용하였다[19]. 여기서 G는 그룹 수를 나타낸다. 이때 구해진 그룹별 합은 투영 벡터를 그대로 사용하였을 경우에 비해 색인 데이터량을 많이 줄일 수 있으면서도, 투영 효과를 검색에 이용할 수 있고, 입력되는 영상의 크기의 변화에도 덜 민감하게 사용할 수 있다는 장점이 있다.

$$V(g) = n \left(\sum_{i=(g-1) \cdot MG}^{(g/MG)-1} \sum_{j=0}^{N} B(i, j) \right), \quad \text{if } B(i, j) = 0 \quad (11)$$

$$H(g) = n \left(\sum_{i=0}^{(g/NG)-1} \sum_{j=(g-1) \cdot NG}^{N} B(i, j) \right), \quad \text{if } B(i, j) = 0 \quad (12)$$

여기서, B는 등고선 값에 의해 추출된 2진 영상이며, $1 \leq g \leq G$ 이고, $n(\cdot)$ 는 개수를 나타낸다.

그림 9은 투영된 벡터의 합을 추출하기 위하여 그룹화를 행하는 과정을 보여준다. 이 때 구해진 합은 투영된 벡터에 대한 행 방향 및 열 방향으로의 그룹에 대한 크기 정도를 나타내는 값이기 때문에 영상을 블록으로 나누어 후 구한 국부 크기 값과는 달리 행 방향과 열 방향의 투영된 영상 특징이 나타나는 크기 값이라 할 수 있다.

$M \times N$ 의 크기를 갖는 영상에 대해 투영벡터는 $N+M$ 의 길이를 가진다. 반면 투영벡터의 분산을 이용할 경우 그 크기는 세 그룹씩으로 나눌 경우 6개가 된

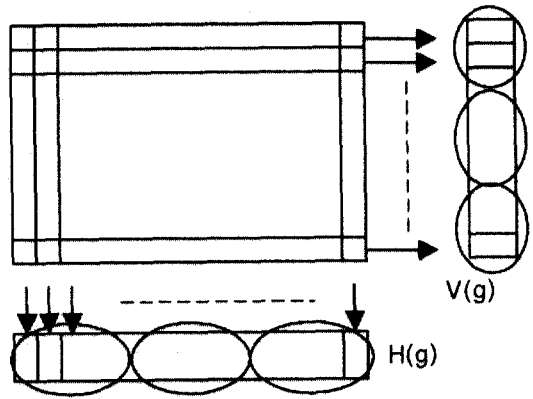


그림 9. 블록별 투영벡터 추출

다. 데이터베이스의 크기가 커질수록 이 차이는 더 커져 검색에 소요되는 시간 및 저장 장치의 비용 차이는 더 늘어난다.

투영벡터의 크기를 이용할 경우는 투영벡터를 그대로 사용한 경우에 비해 정밀도가 떨어지는 때에는, 적용하고자하는 대상의 정밀도에 적합한 크기 값을 구하는 블록의 수를 적절히 조절하여 원하는 블록의 수만큼을 생성하면 투영벡터를 그대로 이용하는 방법에 비해 색인 데이터량을 줄이면서도 원하는 결과를 얻을 수 있다. 그룹의 수는 영상의 국부적 특성이 잘 반영되고, 데이터량과 데이터베이스의 크기를 고려하여 7개의 그룹으로 나누어 14개의 투영벡터를 사용하였다.

3. 실험 및 결과

본 논문에서 3차원 얼굴 영상을 획득하기 위하여 4D컬처에서 제작된 3D Scanner를 이용하였다. 레이저 스캐너가 얼굴 영상을 획득하는 방법은 레이저 라인 빔을 얼굴에 주사하고 스캐너가 3초 동안 이동하면서 180개의 레이저 프로파일들이 포함된 영상을 얻는다. 여기서 3차원 라인 프로파일을 포함한 영상의 크기는 가로 세로 640×480 이며, 중심선 추출 알고리즘을 적용하여 중심선을 추출한다. 높이 값에 대해서 캘리브레이션(calibration) 과정을 거치고, 재표본화와 보간 작업을 수행하여 최종의 320×320 크기의 3차원 얼굴 영상이 추출된다. 또한 얼굴 스캐너에서 폐색(occludin)된 부분의 처리는, 폐색된 주변 경계를 연속하여 돌면서 주변화소들의 중간 값을 선택하는 방식을 사용하며, 머리카락 부분 등은 정확한 데이터를 중심으로

확장 보간하는 방법을 사용하였다. 인식에 사용된 데이터 영상은 정면을 향하고 안경을 착용하지 않은 일반 성인 35명에 대해 2장씩 70장의 영상을 사용하였으며 그림 10과 같다. 영상의 획득은 입력 영상의 스케일 문제를 배제하기 위해 카메라를 고정시킨 후 얼굴이 320×320에 들어 올 수 있는 거리에 의자를 놓고 동일인에 대하여 첫 번째 영상이 획득되어진 후 20분 후에 다시 같은 위치에 앉게 하여 2번째 영상이 획득되어졌다.

획득된 3차원 영상에서 코의 정점을 찾고 이를 기준점으로 정한 후 등고선 값이 20, 30 그리고 40 이상인 코 주변의 영역만을 취한 후 투영벡터의 블록별 크기 값을 이용하여 색인화한 시스템을 구축하여 질의 영상을 주었을 때의 인식 결과를 비교하였다. 두 영상간의 거리 값은 식 (13)과 같이 L_2 -norm을 사용하였다. c 는 코로부터의 등고선 값이 값(20~40)을 나타내며, G 는 등고선 값이 값에 따른 투영벡터의 블록 수이다. I 는 입력 영상 그리고 D 는 데이터 베이스영상이다. 식 (13)을 이용하여 두 영상 사이의 거리 값을 구한 후, 식 (14)을 이용하여 가장 작은 값을 나타내는 것이 유사도가 높은 것으로 비교하였다.

$$U_c(I, D) = \sum_{g=1}^G \sqrt{(I_{c,g} - D_{c,g})^2} \quad (13)$$

$$S(I, D) = \min \sqrt{(U_c(I, D))^2 + (\overline{U_c(I, D)})^2} \quad (14)$$

U 와 \overline{U} 는 서로 다른 영역에 대한 거리 결과이다. 각 영역을 식 (14)에 적용한 인식결과를 표 1과 그림 10에 나타내었다. 표 1과 그림 10에서 보면 등고선 값이 20과 40 그리고 30과 40일 경우 인식률이 낮게 나타났다. 이것은 등고선 40인 경우는 영역이 너무 넓게 되어 얼굴의 외형과 거의 유사한 형태를 나타내어 잘 구분이 되지 않는 것을 의미한다. 즉, 얼굴의 외곽은 사람마다 비슷하기 때문에 인식률이 낮게 나타났다. 반면, 얼굴의 내부 영역에 해당되는 등고선 값 20과

표 1. 영역에 대한 인식률

영역 \ 순위	5위 이내	10위 이내	15위 이내
20~30'	94.3%	97.1%	100%
20~40'	68.6%	85.7%	91.4%
30~40'	74.3%	85.7%	88.6%

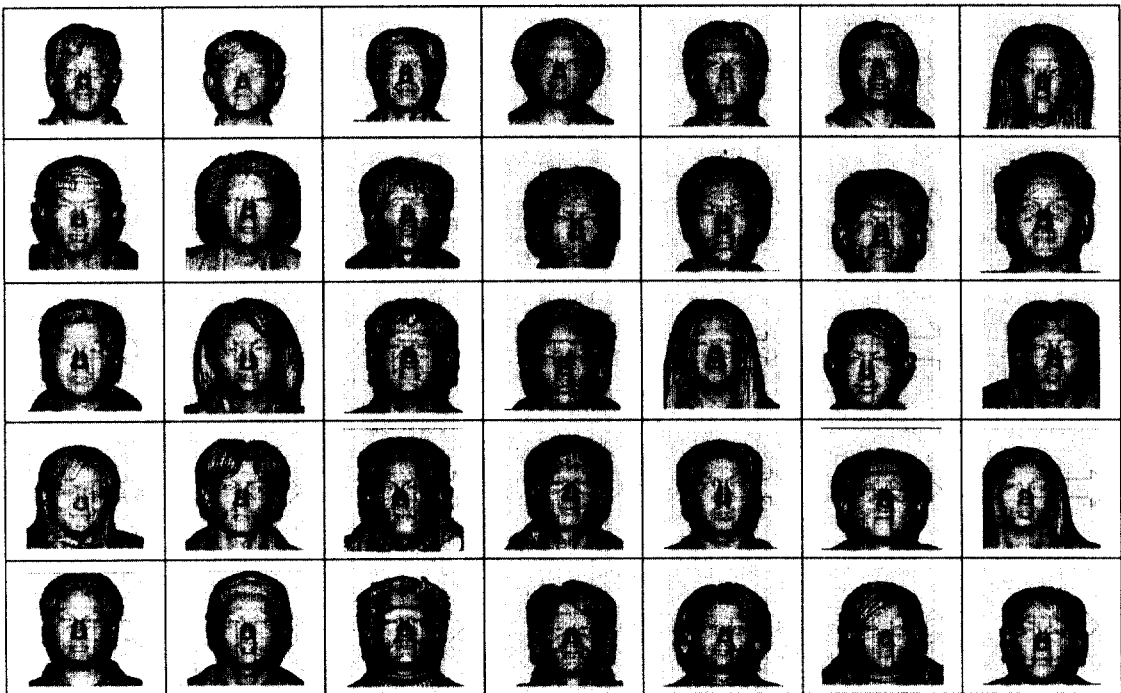


그림 10. 입력 영상

30에 대하여 순위 임계값을 5위, 10위 그리고 15위를 사용하여 실험을 하였다. 실험 결과, 5위 이내가 94.3%, 10위 이내가 97.1% 그리고 15위 이내는 100%의 인식을 나타내었다. 이를 순위 별로 나타낸 인식이 그림 12와 같다. 그림 13과 그림 14는 실험에 사용한 데이터의 인식결과 예와 실제로 두 영상에 대한 유클리드 거리를 계산한 결과를 나타내었다. 영상3이 순위가 높

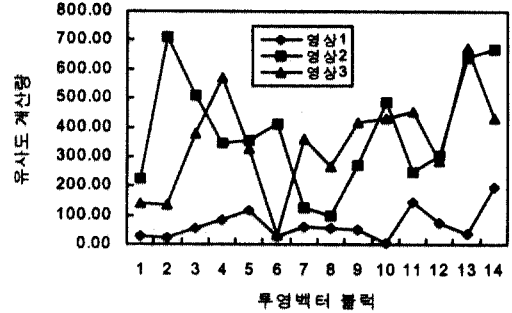


그림 14. 두 영상간의 유사도 비교 예

게 나타나는 이유는 영상 획득 과정에서 바른 자세를 취하지 않아 턱 부분이 약간 높게 되어 얼굴의 이마 영역을 포함하기 때문인 것으로 판단된다.

표 2는 가장 비슷한 인식 방법인 Takács and Wechsler[20]의 실험결과를 나타냈으며, 이는 영상 전체에 대하여 유사도를 비교하기 때문에 많은 계산량을 필요로 한다. 그러나 본 논문에서는 얼굴의 국부적인 영역에 대하여 투영벡터의 그룹 분산 값들을 인식 파라미터로 사용하므로 파라미터 개수가 적고 계산량을 많이 줄일 수 있었다. 또한 인식 결과도 Takács and Wechsler와 거의 비슷한 인식을 나타내었다. 표 3은 깊이 정보를 이용한 Gordon[7]에서 사용한 식을 이용하여 본 논문에서 사용한 데이터를 적용한 결과를 나

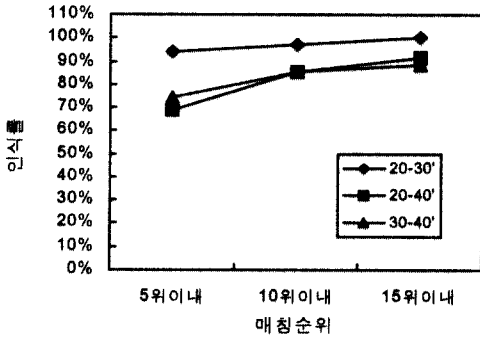


그림 11. 식 (13)을 사용한 유사도 결과

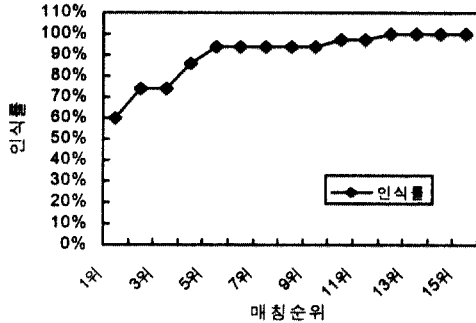


그림 12. 인식을

표 2. 참고문헌(20)의 결과 순위 비교

구분	5위 이내	10위 이내	30위 이내
Takács and Wechsler	97.1%	98.1%	99.7%

	영상1	등선 값 30	영상2	등선 값 30	영상3	등선 값 30
입력영상						
데이터베이스 영상						
검색순위	1		7		12	

그림 13. 인식 결과 예

표 3. 참고문헌(7)의 식을 이용한 순위 결과

구분	5위 이내	10위 이내	15위 이내
G. G. Gordon	85.7%	97.1%	100%

타내었다. 표 3은 얼굴 전체에 대하여 매칭을 수행한 결과이다. 표 1의 20-30'과 표 3을 비교해보면 5위 이내의 경우, 얼굴 전체에 대한 비교보다는 국부적으로 비교를 하는 것이 우수하다는 것을 알 수 있다. 또한, Gordon[7]은 얼굴에 대하여 위치에 대한 표준화를 거친 후 인식을 실험하였으나, 본 논문은 이러한 과정 없이 전처리만 수행하여 인식을 실험하였다.

데이터의 개수가 적었지만 데이터베이스를 구성하는 영상의 종류가 많아지고, 다양하더라도 영상의 투영벡터의 통계적 특징을 이용하기 때문에 우수한 결과를 얻을 수 있으리라 예상된다.

4. 결 론

본 논문은 3차원 얼굴 영상으로부터 일정 깊이 이상의 등고선 영역을 추출하여, 국부적 특징이 잘 반영된 투영벡터를 이용한 얼굴인식 알고리즘을 제안하였다. 얼굴은 사람마다 비슷한 외곽 모양을 가지고 있지만, 일정 깊이에 대한 얼굴의 형상은 많이 다르게 나타난다는 점을 이용하여, 2차원 얼굴영상에서 얻을 수 없는 깊이 정보를 이용한 얼굴 인식 실험을 수행하였다. 얼굴의 형상 특징을 얻기 위하여 광 삼각법을 이용한 레이저 스캐너를 사용하여 획득된 3차원 얼굴 영상 데이터를 사용하였다. 3차원 얼굴 영상 인식은 2차원 얼굴 인식에 있어서 가장 문제가 되는 조명의 변화와 카메라의 각도에 의한 영향이 적다. 또한 2차원 영상에서는 정확하게 추출할 수 없는 깊이 데이터를 3차원 영상이 가지고 있으므로 보다 쉽게 얼굴 내의 깊이에 대한 얼굴 형상을 추출할 수 있었다. 3차원 얼굴 영상에서 가장 두드러지게 나타나는 코를 찾는 것도 간단한 알고리즘을 이용하여 100% 찾을 수 있었다. 3차원 얼굴 영상으로부터 코 주변의 영역들을 추출하여, 영상을 투영한 후 투영 벡터의 국부화를 통하여 영상의 지역적 특성이 잘 반영되도록 하였다. 투영벡터의 통계학적인 특징을 적용하여 구현한 결과 35쌍의 실험 얼굴 영상 중 깊이 값이 20과 30인 경우 5위 이내는 94.3%, 10위 이내는 97.1%의 인식률을 나타내었다. 얼굴의 외

곽 형상은 구분이 잘 되지 않았으며, 얼굴 내부의 영역만을 취하는 경우 높은 인식률을 나타낸다는 것을 확인하였다. 또한, 투영벡터의 통계학적인 특징을 이용할 경우, 대용량 데이터베이스 구축과 색인 데이터량 및 계산상의 부담이 줄어 검색이 빠르다.

참 고 문 헌

[1] L. C. Jain, U. Halici, I. Hayashi, S. B. Lee, *Intelligent biometric techniques in fingerprint and face recognition*, CRC Press, 1999.

[2] Erik Hjeltnäs, "Biometric System: A Face Recognition Approach," *Proceedings of the Norwegian Conference on Informatics*, 2000.

[3] W. Zhao, R. Chellappa, A. Rosenfeld, J. Philips, "Face Recognition: A Literature Survey," *Technical Report, CFAR-TR-948*, Univ. of Maryland, 2000.

[4] 유명현, 박정선, 이상용, 최형철, "얼굴 기반 생체 인식 기술의 현황과 전망," *정보과학회지*, 제19권 제7호, pp. 22-31, 2001.

[5] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceeding of the IEEE*, Vol. 83, No. 5, pp. 705-740, 1995.

[6] J. C. Lee and E. Miliot. "Matching range image of human faces," *Third International Conference on Computer Vision*, pp. 722-726, 1990.

[7] G. G. Gordon, "Face Recognition based on depth and curvature feature," *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 808-810, 1992.

[8] H. T. Tanaka, M. Ikeda and Hchiaki, "Curvature-based face surface recognition using spherical correlation," *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 372-377, 1998.

[9] C. S. Chua, F. Han, Y. K. Ho, "3D Human Face Recognition Using Point Signature," *4th ICAFG*, 2000.

[10] (주)포디컬쳐, "http://www.4dculture.com".

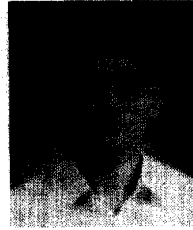
[11] Cyberware, "http://www.cyberware.com".
 [12] Minolta, "http://www.minolta.com".
 [13] P. L. Hallinan, G. G. Gordon, A. L. Yuille, P. Giblin, D. Mumford, *Two and three dimensional pattern of the face*, A K Peters. Ltd., 1999.
 [14] M. Grob, *Visual computing*, Springer-Verlag, 1994.
 [15] A. Nikolaidis, I. Pitas, "Facial feature extraction and pose determination," *Pattern Recognition*, pp. 1783-1791, 2000.
 [16] B. Moghaddam, T. Jebara, A. Pentland "Bayesian face recognition," *Pattern Recognition*, pp. 1771-1782, 2000.
 [17] 권용무, "디지털 문화재 기술," *물리학과 첨단기술*, 제9권 11호, 2000.
 [18] Ramesh Jain, *Machine Vision*, McGraw-Hill, 1995.
 [19] 권동현, 김용훈, 배성포, 이태홍, "투영벡터의 통계적 성질을 이용한 영상 검색," *한국통신학회지*, Vol. 25, No. 7A, 2000.
 [20] Barnabás Takács, Harry Wechsler, "Face Recognition Using Binary Image Metrics," *Proc. of the 3rd Int. Conference on Automatic Face and Recognition*, Nara, Japan, pp. 294-299, 1998.



이 영 학

1988년 2월 영남대학교 전자공학과 학사
 1991년 2월 영남대학교 전자공학과 대학원 석사
 1991년 1월 ~ 1995년 9월 LG정밀(주) 용인연구소 주임 연구원

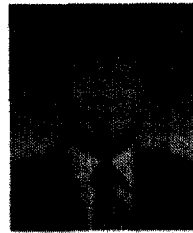
2001년 2월 영남대학교 전자공학과 박사과정 수료
 1998년 ~ 현재 대구과학대학교 겸임교수
 2002년 ~ 현재 (주)페이스쓰리디 공동설립
 관심분야 : 영상처리, 패턴인식, 통신시스템
 E-mail : annaturu@yumail.ac.kr



심 재 창

1987년 경북대학교 전자공학과 학사
 1990년 경북대학교 전자공학과 대학원 석사
 1993년 경북대학교 전자공학과 대학원 박사
 1997년 ~ 1999년 IBM T. J. Watson

Research Center, Researcher
 1994년 ~ 현재 안동대학교 컴퓨터전공 조교수
 1999년 ~ 현재 (주)파미, (주)포디컬처, (주)페이스쓰리디 공동설립
 관심분야 : 영상처리, 패턴인식, 컴퓨터비전
 E-mail : jcshim@andong.ac.kr



이 태 홍

1975년 서울대학교 공업교육학과 전자공학 전공 졸업 (학사)
 1976년 5월 ~ 1979년 10월 동양정밀(주) 중앙연구소 책임 연구원
 1980년 서울대학교 공업교육학과 전자공학 전공 졸업 (석사)

1980년 1월 ~ 1981년 2월 금성전기(주) 기술연구소 기좌
 1981년 3월 ~ 현재 영남대학교 전자정보공학부 교수
 1996년 10월 ~ 1997년 10월 한국통신학회 대구·경북 지부장
 1997년 서울대학교 대학원 전자공학과 졸업 (박사)
 관심분야 : 신호처리, 영상처리, 통신시스템
 E-mail : taehong@yumail.ac.kr

교신저자

이 영 학 712-749 경북 김천시 대동 214-1 영남대학교 전자정보공학부 영상통신 연구실