

# 웨이블릿을 이용한 3차원 얼굴인식

서윤식, 이영학, 배기역, 이태홍  
영남대학교 전자공학과

## 3D Face Recognition using Wavelet

Yunsik Seo, Yeunghak Lee, Kiup Bae, Taihong Yi  
Department of Electronic Engineering, Yeungnam University  
E-mail : yunsiks2002@ymail.ac.kr

### 요약

본 논문에서는 표면의 에지를 잘 나타내는 웨이블릿을 이용한 3차원 얼굴 인식 알고리듬을 제안한다. 먼저 얼굴영역을 추출하고 정규화과정을 수행한다. 코는 얼굴에서 가장 높고 기준점의 역할을 하므로 반복 선택방법을 이용해서 코끝을 찾는다. 코끝 최고점을 기준으로 깊이값 20, 30, 40인 영역에 대해 웨이블릿 변환을 수행하여 얼굴마다 저주파와 고주파들을 생성하는데, 저주파를 제외한 고주파들에 대하여 히스토그램을 특징벡터로 사용하였다. 유사도의 비교는  $L_1$ 거리함수를 사용하여 수평, 수직, 대각고주파, 그리고 이 고주파들의 유사도 비교치를 합한 합성의 경우 각각에 대하여 실험하였다. 깊이값에 따른 영역에서 고주파별로 실험한 결과, 순위 임계값이 10 위를 기준으로 깊이값 30 대각고주파에서 91%가 나타났고 합성에서는 93%의 인식률이 나왔다.

### 1. 서론

오늘날 컴퓨터를 이용한 정보보호와 신분확인을 위해 다양한 분야의 생체인식이 연구되고 있다. 생체인식 기술로는 지문, 홍채, 얼굴, 손모양, 혈관, 음성, 서명 등 다양한 종류가 있다. 그 중에 얼굴인식은 비접촉식으로 거부감이 없는 자연스럽고 간편한 기술로 인정 받고 있다[1]. 기존에 많이 연구되어 온 2차원 얼굴인식에 비해 3차원 얼굴인식은 외부 조명변화에 대한 영향을 받지 않고 접근거리에 따라 일어나는 크기 변화에 장인한 장점이 있다[2].

본 논문에서는 정규화된 3차원 얼굴 영상을 이용하여 얼굴 형태학적 특성을 웨이블릿 변환을 통해 인식하는 방법을 제안하다. 코끝을 기준으로 깊이값 20, 30, 40이내의 영역을 실험 대상으로 하였다. 각 고주파성분에 대한 히스토그램을 색인정보로 구축하여 유사도를 측정하여 인식률을 구하였다..

구한다. 그 다음 최고값을 기준으로 그 보다 작은 값을 모두 0으로 두고 레이블링 과정에서 가장 넓은 영역을 추출하여 그림 1.(b)와 같은 순수 얼굴 영역을 구한다.

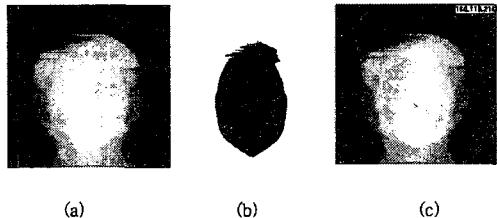


그림 1. 전처리 및 코의 최고점 추출영상

코의 최고점을 찾는 과정은 배경이 제거된 얼굴영역에서 배경소거법인 식 (1)과 식 (2)를 반복 수행하여 구한다.[3]

$$Avg = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} D(i, j) \quad D(i, j) > 0 \quad (1)$$

$$D(i, j) = \begin{cases} D(i, j), & D(i, j) > Avg \\ 0, & Otherwise \end{cases} \quad (2)$$

### 2. 전처리 과정 및 정규화

#### (1) 전처리

전처리 과정에서 하는 일은 배경에서 얼굴을 추출하는 과정이다. 먼저 그림 1.(a)와 같은 3차원 얼굴 원영상의 네 모서리에  $5 \times 5$ 영역을 설정하여 최고값을

여기서  $D(i, j) > 0$ 는  $i, j$  위치의 깊이정보이고,  $M, N$ 은 영상의 가로, 세로의 크기이다.

관성 능률 축에 의해 만들어지며 그 각도의 값은 식 (6)에 의해 구해진다.

## (2) 정규화

3차원 얼굴 영상인식에서 고려되어져야 할 한 가지는 획득된 영상의 정면 자세이다. 그래서 정규화 과정이 필요하다.

본 논문에서의 정규화는 다음과 같다. 전체 보상각에 크게 영향을 주는 평면회전(Z축 회전)과 깊이회전(Y축 회전)의 보상각도를 구해 비교를 한 다음 더 큰 보상각을 가진 회전부터 먼저 처리하게 되며 상하회전(X축 회전)은 단면 패턴의 대칭성을 그대로 유지하기 때문에 평면회전(Z축 회전)과 깊이회전(Y축 회전)에 의해 보상된 후에 수행한다.

### 1) 깊이 회전보상(Y축 회전)

얼굴은 형태학적으로 좌우 대칭인 형태이므로 좌우의 깊이값을 비교함으로 회전 각도를 구할 수 있다.

코의 최고점 좌표  $(x_p, y_p)$ 에 해당하는 높이값을  $z_p$ 라고 할 때,  $y_p$ 를 기준으로 평행한 평면을  $X-Z$  공간으로 나타낼 때 원영상의 깊이 회전 보상각  $\theta_y$ 를 구하기 위해 식 (3)을 이용한다.  $D_1$  값이 최소가 되는 각을 회전 보상각으로 결정한다.

$$D_1 = \frac{1}{M \times N} \sum_{i=0}^M \sum_{j=0}^N |Rseg(x_p + i + 15, y_p + j) - Lseg(x_p + i - 45, y_p + j)| \quad (3)$$

여기서  $Rseg(i, j)$ ,  $Lseg(i, j)$ 은 좌표  $(i, j)$  위치의 깊이값이고  $M, N$ 은 각각 30이다.

### 2) 평면 회전보상(Z축 회전)

얼굴이 좌, 우 대칭이므로 코의 최고점을 기준으로 일정한 깊이 내에 있는 영역들에 대하여 무게중심파 모멘트인 식 (4)와 식 (5)을 이용하여 회전 보상각  $\theta_z$ 를 찾아 낼 수 있다[4].

$$x_{q,p} = \frac{1}{n} \sum_{i=1}^n x_{i,p}, \quad y_{q,p} = \frac{1}{n} \sum_{i=1}^n y_{i,p} \quad (4)$$

$$\mu_p^{k,l} = \sum_{i=1}^n (x_{i,p} - x_{q,p})^k (y_{i,p} - y_{q,p})^l \quad (5)$$

여기서  $q = (x_q, y_q)$ 은 무게중심 좌표이고,  $p$ 는 영역들의 깊이 값들이며,  $p = 15, \dots, 20$ 이다. 6가지의 깊이에 대한 등고선 값에 의한 회전된 각도는 기준 축과 최소

$$\theta_z = \frac{1}{6} \sum_{p=15}^{20} \left( \frac{1}{2} \tan^{-1} \left( \frac{2\mu_p^{1,1}}{\mu_p^{2,0} - \mu_p^{0,2}} \right) \right) \quad (6)$$

그림 2는 평면 회전과 깊이 회전에 대한 보상전과 후의 영상이다. 그림 2.(a)는  $\theta_z$ 을  $-32^\circ$  보상한 영상의 코 영역이고, 그림 2.(b)는  $\theta_y$ 을  $-4.8^\circ$ 로 보상한 영상의 횡단면이다.

### 3) 상하회전(X축 회전)

일반적으로 정면을 향하는 얼굴영상은 코밀점이 미간 점보다 높다는 사전지식을 이용하여, 본 논문에서는 모든 영상에 대해 코밀점의 높이 값과 미간점의 높이 값의 차이가 10이 될 때까지  $Y-Z$ 평면을 좌표  $(x_p, y_p)$  중심으로 회전시킴으로써 원영상을 정규화 시키는 회전 보상각  $\theta_x$ 를 구하며 그림 2.(c)는  $2.5^\circ$  회전 보상시킨 결과 영상이다.

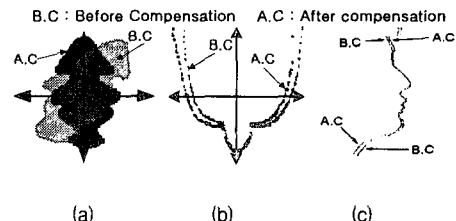


그림 2. 평면, 깊이, 상하회전 보상 전후의 영상

## 3. 특징영역 추출 및 특징값 계산

### (1) 특징영역 추출

정규화된 3차원 얼굴 영상의 최고점에서 일정 깊이에 있는 문턱치의 등고선 값을 취하여 그 내부 영역과 외부 영역을 이진화한다.

본 논문에서는 깊이 20, 30, 40에 대한 등고선 값을 취하고 그 내부와 외부 영역을 이진화한 다음 레이블링 작업을 하여 그림 3과 같이 특징영역을 추출한다. 일반적으로 깊이 20인 영역은 코 전체와 입 일부를 포함하는 부분이고 깊이 30인 영역은 눈을 제외한 영역을 나타내고 깊이 40인 영역은 얼굴전역을 포함한다.

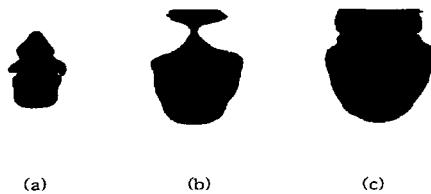


그림 3. (a)깊이 20 (b)깊이 30 (c)깊이 40

## (2) 웨이블릿 변환

웨이블릿 변환은 푸리에 변환과 같이 기저함수들의 집합으로 신호를 분해하는 하나의 방법이다. 그러나 푸리에 변환과 다르게 국부적으로 에너지가 집중된 대역 통과 신호인 웨이블릿을 기저함수로 사용한다. 웨이블릿 변환에서 생성되는 서브밴드를 분석함으로써 영상정보를 얻을 수 있다. 이는 모 웨이블릿이라고 불리는 하나의 원형 웨이블릿을 이동하고 확장함으로써 주파수 영역에 따른 다해상도 영상을 가지게 된다.

웨이블릿 변환의 일반적인 수식은 식 (7)과 같다.

$$WT_{f(a,b)} = \int \psi\left(\frac{t-b}{a}\right) f(t) dt \quad (7)$$

변환의 기본 개념은 임의의 함수  $f(x)$ 를 시간-주파수 공간에서 동시에 지역성을 갖는 웨이블릿 기저함수의 선형 결합으로 표현하는 것이다. 식 (7)에서  $a$ 는 스케일 변수이고  $b$ 는 이동변수이다.

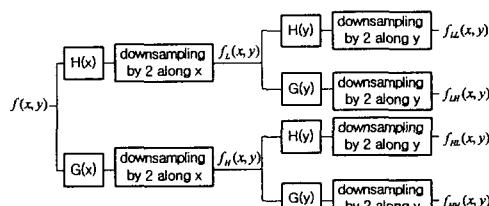


그림 4. 웨이블릿 변환

웨이블릿 변환은 하나의 저주파 서브 밴드와 수평, 수직, 대각 성분의 고주파 서브 밴드를 생성한다. 수평, 수직, 대각 고주파는 각각 수평 방향, 수직 방향 그리고 대각 방향의 에지정보를 제공한다.

본 논문에서는 정규화된 3차원 얼굴 영상 전체를 1 레벨 웨이블릿 변환하고 특징영역의 해당하는 얼굴 영역의 에지 성분을 얻는다.

## (3) 특징영역의 히스토그램

깊이 20, 30, 40의 영역에 대한 각각의 영상 D에서 저주파 LL, 수평고주파 LH, 수직고주파 HL, 대각고주파 HH가 나오는데 저주파 LL을 제외한 나머지 고주파들에 대해서 각각 -1.0에서 1.0 사이로 정규화시킨다. 이 고주파들을  $n$ 개의 값들을 가지고  $c_1, c_2, c_3, \dots, c_n$  양자화 된다고 할 때 영상 D에 각 고주파 대한 히스토그램  $H_{LH}(D), H_{HL}(D), H_{HH}(D)$ 는 식 (8)과 같이 된다.[5]

$$H_{LH}(D) = \langle h_{c_1}(D), h_{c_2}(D), h_{c_3}(D), \dots, h_{c_n}(D) \rangle$$

$$H_{HL}(D) = \langle h_{c_1}(D), h_{c_2}(D), h_{c_3}(D), \dots, h_{c_n}(D) \rangle$$

$$H_{HH}(D) = \langle h_{c_1}(D), h_{c_2}(D), h_{c_3}(D), \dots, h_{c_n}(D) \rangle \quad (8)$$

$h_{c_i}(D)$ 는 표면 변화값  $c_i$ 를 갖는 화소의 수이다.

본 논문에서는 각 고주파 영역에 대해서 9개 구간으로 나누고 각 영역에서 구간에 해당하는 화소를 모아 히스토그램을 만들었다.

## 4. 실험 및 결과

본 실험에 사용된 3차원 얼굴영상은 4D 컬쳐사에서 제작된 스캐너를 사용하여 획득한  $320 \times 320$  영상이다. 인식에 사용된 데이터베이스 영상은 10세이상 60세미만의 남녀로서 정면을 향한 35명에 대해 2장씩 총 70 장의 영상을 사용하였다. 제한조건으로 안경을 착용한 경우와 코 및 턱수염이 있는 경우, 머리카락이 두 눈을 가리는 경우는 배제하였다.

질의영상과 DB영상간의 유사도 측정은  $L_1$  거리함수를 이용하여 식 (9)와 같이 특징값을 비교하였다. 첫째 수평, 수직, 대각고주파 각각에 대해서 특징값을 비교하는 것과 이를 고주파들에서 나온 각각의  $R_{diff}$ 를 모두 더한 값을 가지고 비교하는 것 두 가지를 하였다.

$$\begin{aligned} R_{diff} &= |D - D'|_{L_1}, \\ &= \sum_{i=1}^n |\{dt_i(D)_{query\_img} - dt_i(D')\}_{DB\_img}| \end{aligned} \quad (9)$$

표면 에지정보를 가지고 얼굴의 부분 영역인 눈, 코, 입을 인식할 때 가장 두드러지는 영역은 코이고 나머지 입과 눈은 인식률이 낮다.[6]

표 1. 깊이값 20에 대한 인식률

인식률(%)	수평 고주파	수직 고주파	대각 고주파	합성
3위	30	31	31	50
5위	42	34	40	57
10위	59	59	67	66
15위	69	69	77	82

표 2. 깊이값 30에 대한 인식률

인식률(%)	수평 고주파	수직 고주파	대각 고주파	합성
3위	51	46	60	74
5위	64	62	77	83
10위	80	87	91	93
15위	86	93	99	94

표 3. 깊이값 40에 대한 인식률

인식률(%)	수평 고주파	수직 고주파	대각 고주파	합성
3위	40	40	37	69
5위	51	55	50	76
10위	66	74	74	86
15위	79	81	85	90

본 논문에서 표 2, 깊이값 30의 수평, 수직, 대각고주파가 다른 깊이값 영역에 비해 가장 높은 인식률이 나타났다. 깊이값 30인 영역은 얼굴 영역중 눈이 제외된 영역인데, 인식률이 낮은 눈이 없으므로 인식률이 높았다. 깊이값 20은 눈이 제외된 코와 입일부가 포함된 영역인데도, 인식률은 깊이값 30이 높게 나왔다. 부분 영역인 눈, 코, 입 이외의 얼굴 영역이 사람마다 고유하게 나타나므로 색인 정보로서 가치가 있기 때문이다. 깊이값 40의 경우, 깊이값 30에 비해서 눈, 코, 입 이외의 넓은 얼굴 영역을 가짐에도 깊이값 30이 인식률이 더 높았다. 이것은 깊이값 40 영역에는 인식률이 낮은 눈이 포함되었기 때문이다.

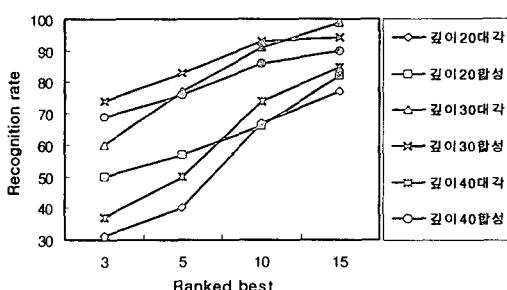


그림 5. 각 영역의 대각고주파와 합성 인식률

수평, 수직, 대각고주파의 유사도가 모두 고려된 합성의 경우, 각각의 깊이값 영역 고주파들 보다 높은 인식률을 나왔다. 얼굴마다 특정고주파에 높은 유사도를 보이거나 반대로 낮은 유사도를 보이는 경우가 있다. 일반적으로 질의영상과 DB영상이 동일인물이라면 유사도가 다른 사람을 비교할 때 보다는 상대적으로 높았다. 그 결과 수평, 수직, 대각고주파 각각의 인식 결과보다 높은 인식률이 나왔다. 깊이값 30 순위 임계값 10위에서 합성이 93% 인식률을 획득하였다.

## 5. 결론

본 논문에서는 표면의 에지특성을 잘 나타내는 웨이블릿을 이용한 3차원 얼굴 인식 방법을 제시하였다. 3차원 얼굴 영상에 1레벨의 웨이블릿 수행하였하여 수평, 수직, 대각고주파에 에지 정보를 획득하였다. 깊이값에 따른 영역에 대하여 고주파별로 특징벡터인 히스토그램을 만들어 실험한 결과, 순위임계값 10위를 기준으로 깊이값 30 대각고주파에서 91%의 인식률을 얻었고 수평, 수직, 대각고주파의 유사도 비교치 합한 합성에서는 93%의 인식률을 획득하였다.

향후 연구 과제로는 웨이블릿의 에지 특성을 더욱 살리기 위하여 에지 직교 방향으로 투영하는 방법을 연구하고자 한다.

## [참고문헌]

- [1] 유명현, 박정선, 이상웅, 최형철, 이성환, “얼굴 기반 생체인식 기술의 현황과 전망,” 정보과학회지 제19권 제7호, 2001.
- [2] P. W. Hallinan, *Two- and three-dimensional patterns of the face*, A k Peters, Ltd., 1999.
- [3] 이영학, 심재창, 이태홍, “코 정보를 이용한 3차원 얼굴 인식,” 신호처리합동학술대회 논문집, 제 14 권, 제1호, p.135-138, 2001.
- [4] Stephane Marchand-Maillet, Yazid M. sharaiha, *Binary Digital Image Processing A Discrete Approach*, pp.194-195, 2000.
- [5] Dimitrios Androutsos, Kostas N. Plataniotis and Anastasios N. Venetsanopoulos “Directional Detail Histogram for Content Based Image Retrieval”, *13th International Conference on*, Vol.1, 1997.
- [6] 배기억, 이영학, 이태홍, “표면 곡률을 이용한 3차원 얼굴인식,” 하계종합학술대회 논문집, 제 26권, 제1호, p.2263-2266, 2003.