

얼굴 영역 정보를 이용한 동적 링크 구조상에서의 얼굴인식

김동훈*, 신대규*, 양민영*, 김현술**, 박상희*

*연세대학교 전기전자공학과, **삼성전자

Face Recognition using Face Region Information in the Dynamic Link Architecture

Dong-Hoon Kim*, Dae-Kyu Shin*, Min-Young Eum*, Hyun-Sool Kim**, Sang-Hui Park*
 *Dept. of Electrical and Electronic Engineering, Yonsei University, **Samsung Electronics

Abstract - 본 논문에서는 다양한 크기와 왜곡을 함유한 절의 얼굴 영상이 들어 왔을 때에도 적용 가능한 동적 링크 구조상에서 얼굴 인식에 관한 것이다. 기존의 동적 링크 구조상에서의 얼굴 인식 기술이 얼굴의 크기와 상관없이 일정한 크기의 그래프가 적용이 되기 때문에 얼굴의 크기가 다른 경우에는 인식을 하지 못하였다. 이러한 단점을 보완하고자 절의 영상의 얼굴 타원 정보를 이용하여 고정된 그래프에 맞게 얼굴 크기를 보정함으로써 기존의 우수한 성능을 유지함과 동시에 동적 링크상에서 다양한 크기와 왜곡을 가진 절의 얼굴 영상을 인식할 수 있도록 한 것이다.

1. 서 론

인간이 사람을 알아볼 때 일반적으로 지문이나 홍채가 아닌 그 사람의 얼굴을 보고 사람을 판단한다. 즉 가장 쉽게 확인 가능할 것이 바로 얼굴 인식이다. 이러한 측면에서 보안을 위한 신뢰성과 사용의 편리성을 동시에 충족시키는 기술로 얼굴인식이 부각되고 있다.

얼굴 인식 기술은 일반적으로 주어진 정지 영상이나 동영상에 대하여 입력 영상내의 한 사람이나 그 이상의 사람을 기준의 데이터베이스 안에서 찾아내는 것으로 정의할 수 있으며, 일반적으로 얼굴 인식 문제는 다음의 세 단계로 나뉘어진다[1].

- 1) 영상으로부터 얼굴 영역의 분리
- 2) 얼굴의 특징 추출
- 3) 판단

위의 기준을 바탕으로 1990년대 이후의 얼굴 인식 분야는 크게 "EigenFace"를 이용한 PCA 방법론[2], "동적 링크 구조를 이용한 방법론"[2], "신경회로망을 이용한 방법론"[2] 등으로 나뉘어서 독립적으로 연구되어지고 있다[1].

"동적 링크 구조에 의한 방법론"은 나머지 두 방법론에 비해서 처리 속도 면이나 인식률의 면에서 우수함과 더불어 여러 가지 조명이나 왜곡에도 인식이 가능함이 이미 입증되었다.

동적 링크 구조에 의한 얼굴인식 시스템은 전처리 과정을 거친 영상을 토대로 얼굴 검출 과정과 얼굴 인식 과정 부분을 탄력적인 그래프 정합 알고리즘에 의하여 수행된다. 이 알고리즘은 입력 영상에 씌워진 직사각형의 그래프에 의한 동적 링크 구조에 의해서 가능하게 된다. 동적 링크 구조의 교차점의 특징 벡터들간의 비교 통하여 얼굴간의 유사성을 판단하게 되며, 에지들의 거리를 비교함으로서 왜곡 정도를 고려할 수 있다[3]. 그러나 이 방법론은 얼굴의 크기가 다른 경우에는 고정된 구조에 의한 세약 때문에 인식이 가능하지 않다[4].

본 논문에서는 얼굴인식 과정에서 기본적으로 요구되는 얼굴 검출 방법으로 기존의 타원 검출법을 이용하여 인식에 있어 문제시되는 크기가 다른 영상들간의 인식 문제를 해결하고자 한다.

본론에서는 얼굴검출 방법, Gabor Jet. 그리고 동적 링크 구조에 대해서 알아본다. 또한 기존의 동적 링크

구조에서의 인식 결과와 인터플레이션을 통해 보완된 영상에서의 실험 결과를 보고자 한다.

2. 본 론

2.1 얼굴 검출 방법

얼굴이 타원이라는 정보를 이용한 Shiroey의 검출 방법이 사용되었다[4]. 타원형 에지를 찾기 위한 첫 단계로서 우선 데이터베이스 내의 얼굴영상에서 Canny의 에지 검출기[5]를 이용하여 에지를 찾아낸다. 이렇게 구한 에지 영상에서는 수많은 에지간의 교차점(break-points)이 존재한다. 얼굴부분과 배경부분의 에지를 구분하기 위해서, 두 개의 교차하는 에지가 서로 다른 물체에서 나온 것이라면 이러한 교차점을 없애주는 과정 후 각각에 대해 단위벡터의 내적을 구해서 최소값을 가지는 것들이 물체의 경계라고 할 수 있다. 이런 에지들에 대해 동일한 번호를 부여한 후 선형화된 타원 방정식을 이용하여 중심과 축의 길이를 알아낸 후 찾아낸 파라미터들(P_i) 중 일정한 범위를 만족하는 파라미터들을 가지는 타원 후보를 결정한다. 이 후보 타원들 중에서 j 의 번호가 붙은 모든 에지 픽셀 (x_i, y_i) 에 대해 그 타원에 대한 오차를 구하여 만약 문턱값보다 작으면 (x_i, y_i) 는 파라미터 P_j 를 갖는 타원에 적합하다고 할 수 있다. 각각의 번호가 붙은 에지에 대해서 이런 픽셀이 전체 픽셀 수의 반 이상이면 이 번호의 에지는 그 타원에 속한다고 판단한다.

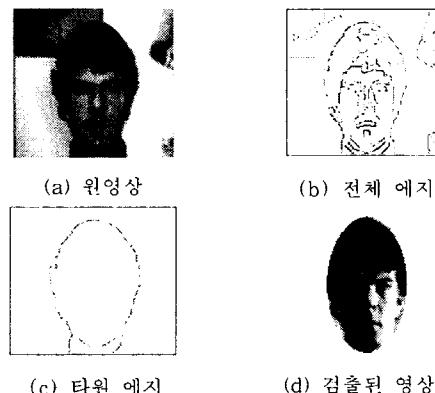


그림 1. 얼굴 영역 검출

2.2 Gabor Jet

Gabor 웨이브렛 변환[6]을 통한 영상의 전처리 과정으로 얼어진 값들은 얼굴인식 시스템에서의 영상의 특징을 규정짓는 특징 파라미터로서 쓰이게 된다. 즉, Gabor 웨이브렛의 실수부에 의한 응답과 허수부에 의한 응답을

이용한 크기를 얼굴의 특징 파라미터로서 사용하였다. 다시 말하면 복소 응답의 절대값을 사용하였다.

중심 주파수 벡터를 크기와 방향에 따라 표현하면 식(1)과 같다.

$$\vec{k}_{\nu\mu} = k_o (\cos(\varphi_\mu) + i \sin(\varphi_\mu)) \quad (1)$$

여기서 $k_o = k_{\max}/f^o$, $\varphi_\mu = \frac{\pi\mu}{4}$ 으로 표현될 수 있다.

f 은 주파수 영역에서의 간격을 조정하는 인자로서 대부분의 영역을 포함하기 위해서 2 또는 $\sqrt{2}$ 로 대수적으로 간격을 선택하였다.

Gabor 필터는 식(2)와 같이 표현되며 이에 대한 전처리 과정의 결과는 식(3)와 같다.

$$\psi_{\nu\mu}(\vec{x}) = \frac{k_{\nu\mu}^2}{2\sigma^2} \exp\left(-\frac{k_{\nu\mu}^2 x^2}{2\sigma^2}\right) [\exp(i\vec{k}_{\nu\mu}\vec{x}) - \exp(-\frac{\sigma^2}{2})] \quad (2)$$

$$(W\mathcal{I})(\vec{k}, \vec{x}_0) := \int \psi_{\nu\mu}(\vec{x}_0 - \vec{x}) I(\vec{x}) d_2x = (\psi_{\nu\mu} * I)(\vec{x}_0) \quad (3)$$

$(W\mathcal{I})(\vec{k}_{\nu\mu}, \vec{x}_0)$ 의 크기는 위치 \vec{x}_0 에서의 특징 벡터를 나타낸다. 이하 이 특징 벡터를 "Jet" 라 한다.

$$J_{\nu\mu}(\vec{x}_0) := |(W\mathcal{I})(\vec{k}_{\nu\mu}, \vec{x}_0)| \quad (4)$$

2.3 동적 링크 구조

2.3.1 얼굴의 특징 추출을 위한 그래프

영상에서 얼굴에 대해 필요한 정보를 추출하고, 인식 과정에서 왜곡을 고려하기 위해서 그래프 구조가 요구된다. 128×128 영상에 대해서 대부분의 영상에서 얼굴 부분을 전부 포함한 격자를 구성하기 위해서 얼굴의 크기에 따라 각 교차점 사이의 거리가 일정한 거리가 되게 세로 부분을 10개의 점으로, 그리고 가로 부분을 7개의 점으로, 다시 말하면 7×10 의 부호화된 그래프가 되도록 설계한다. 그림1은 부호화된 그래프의 일반적인 모형을 나타낸다.

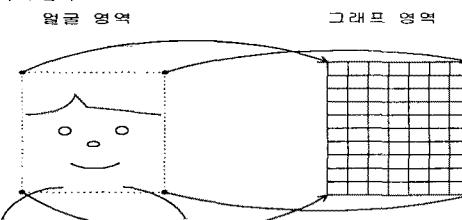


그림 2. 부호화된 그래프

2.3.2 탄력적인 그래프 정합법

탄력적인 그래프 정합법은 동적 링크 구조를 이용한 많은 얼굴 인식 시스템에서 정합 알고리즘으로 사용되고 있으며, 이는 크게 대략적인 그래프의 위치를 선정하는 "전체적인 이동" 과정과 선정된 위치를 토대로 왜곡을 고려하는 "지역적인 이동" 과정으로 나뉜다. 동적 링크 구조상에서의 일반적인 알고리즘은 식(5)와 같이 표현 할 수 있다.

$$C_{total} = \lambda \sum_{(i,j) \in E} S_e(\vec{A}_{ij}^T, \vec{A}_{ij}^M) - \sum_{i \in V} S_v(J'(i), J_i^M) \quad (5)$$

여기서, C_{total} : 총 비용함수

S_e : I 영역과 M 영역의 교차점간의 유사도.

S_v : I 영역과 M 영역의 애지간의 유사도.

λ : 왜곡의 정도를 나타내는 상수

λ 는 왜곡의 정도를 나타내는 상수로서 왜곡을 전혀 고려하지 않은 경우, 즉 그래프의 직사각형의 구조를 유지할 때의 값을 ∞ 로 하였다. 이 경우는 식(5)의 두 번째 항은 고려하지 않아도 된다.

먼저 입력 영상에 대해 데이터베이스의 그래프가 구성되어야 하는 대략적인 위치를 선정하기 위해서 왜곡을 고려하지 않고 직사각형의 모양을 유지하면서 S_e 만을 고려해서 최소 비용함수 값을 갖는 위치를 구한다. 이 과정은 그래프가 전체적으로 이동하는 것이며 이를 "전체적인 이동" 과정이라 표현한다.

최종적으로 "전체적인 이동" 과정을 토대로 선정된 위치를 이용하여 그래프의 각 교차점의 위치의 왜곡 정도를 고려해서 최소 비용함수 값을 구한다. 이 과정은 각 교차점의 위치가 변하는 것이기 때문에 이를 "지역적인 이동"이라 한다. 본 과정을 행할 때는 왜곡 가능한 정도에 따라서 λ 의 값을 설정해야 하며 이 때의 총 비용함수 값은 식(5)과 같다.

후보 얼굴 영상은 데이터베이스의 각 얼굴 영상들 중에 총 비용함수인 C_{total} 의 최소값을 갖는 얼굴 영상이 선택된다.

그럼 3의 (a) 영상은 다른 데이터베이스 안의 영상과의 정합 과정을 통하여 얻은 그래프의 대략적인 위치를 나타내며 (b) 영상은 (a)에서 대략적인 위치를 토대로 왜곡을 고려하여 최종적으로 최소 비용함수 값을 갖는 위치에서 구성된 그래프의 모양을 나타낸다.

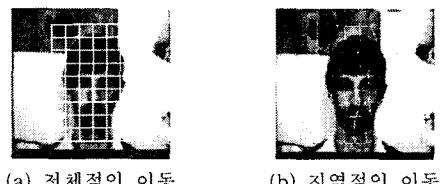


그림 3. 전체적인 이동과 지역적인 이동에 의해 구성된 그래프

3. 타원 정보를 이용한 방법

기존의 동적 링크 구조에서의 인식 방법에 있어서는 얼굴의 크기가 다른 경우에는 고정된 구조에 의한 제약 때문에 인식 성능이 저하되었다. 이러한 단점을 보완하기 위해서 제안된 얼굴 인식 방법은 얼굴 영역을 검출시 얻은 타원 정보를 이용하였다.

질의 영상의 얼굴 크기가 실험적으로 얻어진 임계값보다 작거나 클 경우 데이터베이스 영상들의 얼굴의 평균 크기만큼 질의 영상의 얼굴 영역 부분을 B-Spline 인터플레이션을 통하여 보정한다[7]. 그리고 보정된 질의 영상을 가지고 동적 링크 구조를 이용하면 인식 성능을 개선할 수 있다.



그림 4. 크기 보정된 얼굴 영상

그림 4에서는 (a)는 얼굴 검출시 얻은 타원의 정보가 임계값보다 작은 영상이고 (b)는 인터플레이션을 통하여

보정한 영상, 그리고 (c)는 데이터베이스에 있는 영상이다.

4. 실험 및 결과

실험 영상은 복잡한 배경을 가진 MIT 데이터베이스의 영상과 균일한 배경을 가진 Yale 데이터베이스를 사용하였다. MIT 데이터베이스의 영상들은 128×120 인 레이 영상으로 구성되며, 총 16인에 대하여 세 가지 크기에 세 가지 조명 변화를 가진 1인당 9장씩 총 144장으로 구성된다. Yale 데이터베이스의 얼굴 영상은 128×128 크기의 그레이 영상으로 구성되며, 총 15인에 대하여 조명변화 및 표정변화 그리고 안경의 착용 유무 등 각 개인당 11장씩 총 165장의 영상으로 구성되어 있다.

실험 장비로는 영상의 저장과 시스템 제어를 위하여 IBM 호환 펜티엄 700MHz 퍼스널 컴퓨터를 사용하였다.

데이터베이스는 MIT의 경우 중간크기를, Yale의 경우는 무표정한 영상을 데이터베이스로 선정하였다.

인식 결과는 타원이 검출된 후 인식 여부를 가지고 산출하였다.

그림 5는 Yale(50%)과 MIT 질의 영상에 대해서 보상 전과 보상 후의 각각의 실험 결과를 보여주고 있다.

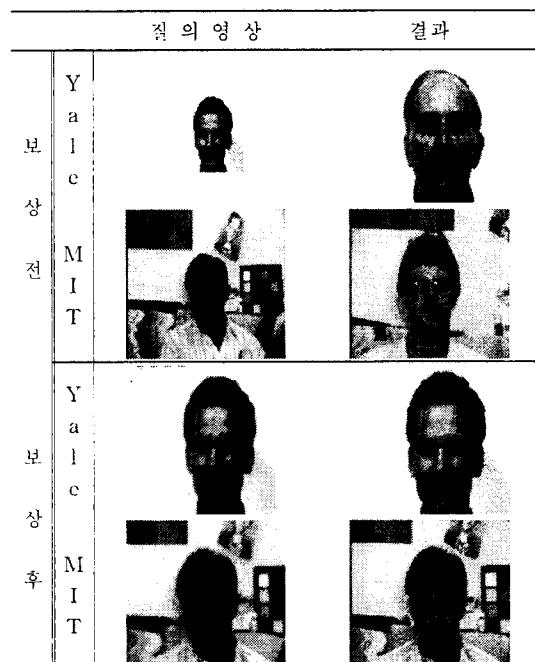


그림 5. 실험 결과

질의 영상의 크기와 데이터베이스 영상의 크기의 차이가 큰 경우들에 대해서는 표 1에서 알 수 있듯이 Yale이나 MIT의 데이터 베이스 모두에서 좋지 않은 성능을 보인다. 그러나, 검출된 얼굴 영역의 타원 정보를 가지고 B-spline 인터플레이션을 통하여 크기를 보정하였다. 이때 블러링 현상이 나타나지만 강인한 Gabor 특성벡터의 성능으로 상쇄된다. 또한, 인식률은 보상 전에 비해 개선된 높은 인식률을 나타내었다. 이는 비슷한 크기의 영상에 대한 결과에는 미치지 못하지만 거의 근접한 인식률이다.

DB	크기	성공률	
		보상 전	보상 후
Y	90%	73.3% (121/165)	94.4% (153/162)
Y	80%	24.2% (40/165)	94.4% (153/162)
A	70%	8.5% (40/165)	93.2% (151/162)
L	60%	10.9% (18/165)	91.9% (149/162)
E	50%	6.7% (11/165)	94.4% (153/162)
M	크기 1	62.5% (30/48)	81.8% (36/44)
I	크기 2	96.7% (29/30)	보상 불필요
T	크기 3	66.7% (32/48)	94.7% (36/38)

*크기 2 : 기준 크기, *크기 1 >> 크기 2 >> 크기3

표 1. 크기가 다른 경우의 동적 링크 방법의 얼굴 인식 결과

3. 결론

본 논문에서는 기존에 강력한 성능을 가진 동적 링크 구조를 이용한 얼굴 인식 방법의 단점인 크기가 다른 질의 영상의 인식 문제를 해결할 수 있는 방법을 제시하였다.

동적 링크를 이용한 얼굴 인식 실험에 있어서, 질의 영상이 데이터베이스의 영상과 크기가 다른 경우에는 고정된 그래프 구조에 의한 제약 때문에 성능이 저하되었으나, 본 논문에서는 추출된 영상의 얼굴 영역 정보를 이용하여 질의 영상의 크기를 데이터베이스의 얼굴 영상의 크기로 인터플레이션을 통하여 보정 후 인식하면, 보정 전에 인식이 불가능했던 크기의 질의 영상에 대해서도 높은 인식률을 얻을 수 있었다.

(참고문헌)

- [1] Sami Romdhani, "Face Recognition using Principal Components Analysis"
- [2] J. Zhang, Y. Yan, M. Lades, "Face Recognition Eigenface, Elastic Matching, and Neural Nets" Proc. Vol. 85, No. 9, 1997.
- [3] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, Malsburg, and R. Wurtz, "Distortion invariant recognition in the dynamic link architecture," IEEE Computers, Vol. 42, pp. 300-311, 1993.
- [4] R. Chellappa, Charles L. Wilson, and S. Sirohey, and Machine Recognition of Faces : A Survey," Proc. Vol. 83, No. 5, pp. 704-740, May 1995.
- [5] J. Canny, "A Computational Approach to edge Detection," IEEE Trans. on PAMI, Vol. 8, pp. 679-689, Nov. 1986.
- [6] J.P. Jones and L.A. Palmer, An evaluation of two-dimensional gabor filter model of simple receptive fields in cat striate cortex, J. Neurophysiol., 58(6):1233-1249, 1987.
- [7] R. Crane, "A Simplified approach to image processing," Prentice Hall, 1997