

효율적인 다 해상도 얼굴 인식을 위한 가보 웨이블릿 연구

정원구¹, 이필규

인하대학교 지능기술연구실

jnlss¹@im.inha.ac.kr, pkrhee@inha.ac.kr

Multi-Resolution using Gabor Wavelet for Efficiency Face Recognition

Won Gu Jung¹, pil kyu Rhee

Inha University, Intelligent Technology lab.

요약

본 논문에서는 여러 해상도로 입력되어지는 얼굴 이미지를 효율적으로 인식시키는 작업을 수행하는 방법에 대한 내용을 소개하고 있다. 정해지지 않은 예측이 불가능한 사람들이 드나드는 공공장소인 공항이나 항만 같은 곳에서의 얼굴인식은 고정된 크기가 아닌 다양한 크기와 조명을 갖는 등, 매우 많은 가지 수의 환경 변수를 가지고 있다. 이러한 환경에서의 얼굴인식은 그만큼 다양한 변수와 그 변수의 조건에 대한 대응을 요구하게 된다. 여기서 제안하는 방법은 다양한 해상도를 갖는 입력 얼굴 이미지에 대하여 최적의 가보 커널과 그에 따르는 적절한 파라미터를 찾는 것으로 효과적인 얼굴인식을 수행하는 방법을 제안한다.

1. 서론

최근 정보 보안 수단의 하나로 여러 가지의 생체인식 기술들이 많은 각광을 받으며 개발되고, 사용되어 지고 있다. 생체인식 기술은 사람 개개인마다 확연히 구별되어지는 생체정보를 이용하기 때문에 높은 신뢰성을 보이며 편리성 면에서도 다른 보안 기술의 한계점이나 불편함을 여러 부분에서 보완해줄 것으로 기대된다. 이중 얼굴 인식 기술은 인식하고자 하는 대상이 직접 센서에 접촉하지 않아도 된다는 점과 인증 시 고정된 자세와 위치를 요구하는 다른 인식 기술과는 달리 얼굴인식기술은 영상을 획득하기 손쉬우며, 공항이나 항만 등의 출입구에 관련된 보안시설과 같은 곳에서 감시용으로도 사용할 수 있다는 점에서 그 쓰임새가 아주 높다고 할 수 있겠다.

하지만 인식하고자 하는 대상을 인증하려 할 때 입력 받는 얼굴 이미지의 해상도와 그에 따른 전처리 방법 및 인식 방법을 어떻게 선택하여 적용하느냐에 따라 결과가 다르게 나타나는 경우가 발생할 수 있다. 예를 들어 입력 되어지는 얼굴 이미지의 크기가 지나치게 작다면 등록 당시 크게 입력 받았던 비교 대상이 되는 얼굴

이미지에 비해 특징이 많이 감소 된 것처럼 되어 인식률에 영향을 미칠 수 있다. 더구나 공항과 같은 곳의 감시 시스템에 적용을 했을 경우, 입력 받는 얼굴 이미지는 그 해상도 카메라에 포착 되어지는 거리의 차이와 위치된 지점의 조명에 따라 큰 영향을 받게 되는데, 거리의 차이로 인한 이미지의 크기는 기존 실험의 방식을 택했을 경우에는 많은 오류를 내포한다. 본 논문에서는 이렇게 각기 다른 해상도로 입력 되어지는 얼굴 이미지를 효과적으로 인식 작업을 수행하기 위해 가보 웨이블릿을 이용하여 각각의 해상도에 대하여 가장 좋은 인식 결과를 나타내는 최적의 파라미터를 추출해 내었다.

2. 관련 연구

신경망을 이용한 방법은 학습이 가능한 몇몇 소수의 클래스를 분류하는 방법으로써, 추상적인 표현을 함축하는 기능을 가지고 있다[1]. 복잡한 얼굴이 추상화되어 압축 저장된다는 장점을 가지고 있지만, 새로운 패턴이 들어왔을 때 전체적인 학습을 다시 수행해야 하기 때문에 매우 느리다는 단점을 가지고 있다. 이러한 신경망 대신 좀 더 실용적인 방법으로써 주성분 분석이

쓰이기도 하는데, 다변량 통계 분석 방법중 하나로서 데이터 차원의 감소라는 큰 장점이 있다. Turk 와 Pentland 는 이 주성분분석을 이용하여 고유얼굴이라는 것을 만들어 얼굴 인식에 적용하여 좋은 결과를 얻었다. 하지만 얼굴의 전역적인 특징을 이용하기 때문에 눈, 코, 입등 세부적인 표현을 하지 못한다. 이러한 점은 얼굴에서 나타나는 작지만 다양한 세부 변형을 표현해 내지 못한다는 단점이 있다[2].

국소 특징 분석은 고유얼굴을 기반으로 한 방법인데 통계적으로 계산된 국소 특징과 위치로 객체를 표현하는 것으로 얼굴의 표정 변화, 조명 등의 변화 때문에 나타날 수 있는 문제들을 해결할 수 있다고 하지만 대부분의 특징들이 얼굴의 외곽에서 추출되기 때문에, 왜곡으로 인한 인식을 감소를 가져올 수 있다[3].

3. 가보 웨이블릿

가보 웨이블릿은 방향성분, 공간 주파수, 그리고 공간 지역성을 효율적으로 추출한다. 가보 웨이블릿은 인간의 시각 신경 계통 특성에 대응하는 실험적 필터에 대한 접근으로 볼 수 있다. 가보 웨이블릿은 생물학적 연관성과 계산적 특성 때문에 이미지 인식에 널리 사용되고 있다. 가보는 인간의 기본적인 시각 피질내에 있는 일반적인 세포의 수용계와 비슷한 모형을 가지고 있다. Receptive function 은 인간의 기본적인 시각 피질에서 뉴런의 수용계는 독자적인 주파수와 방향을 갖는다[4]. 이러한 것들은 2-D 가보 필터로 모델링된다. 가보 필터는 이미지에서의 노이즈와 중복성을 감소시키는데 효율적이다. 가보 웨이블릿은 가보 커널에 의해 한정된 평면파의 형상에서 회선 커널의 생물학적인 모티브가 되었다. 가보 웨이블릿은 이미지의 압축과 재표현에 특히 적합하다는 것을 보여주며, 서로 다른 주파수와 방향의 커널을 위한 회선 계수는 어떤 특정한 얼굴의 특징점에서 시작하는 것으로 계산되어 진다. Malsberg 가 제안한 가보 커널은 아래 수식 (1)과 같다.

$$w(x, y) = \frac{k_f^2}{\sigma^2} \exp\left(-\frac{k_f^2(x^2 + y^2)}{2\sigma^2}\right) \exp\left(i k_f(x^2 + y^2)\right) - \exp\left(-\frac{\sigma^2}{2}\right)$$

$$\vec{k}_f = (k_x \cos \theta_f, k_y \sin \theta_f)^T, k_f = 2 \frac{\pi}{\lambda} \mu, \theta_f = \mu \frac{\pi}{n}$$

$$j = \mu + 8v, v = 0, \dots, m, \mu = 0, \dots, n$$

수식 (1)

수식 1 에서 첫번째 괄호 안의 항목은 커널의 주파수를 결정하며 두번째 괄호는 DC 값을 상쇄하며, DC-free 한 커널을 만들게 된다. σ 는 조절가능한 파라미터로서 wavelength 에 대한 가우시안 윈도우 길이의 비율을 나타낸다. 가보 웨이블릿은 보통 5 개의 주파수, $v=0, \dots, 4$ 와 8 개의 방향, $\mu=0, \dots, 7$ 을 사용한다[5].

4. 커널 파라미터 산출을 위한 실험 및 결과

첫번째 실험은 윗 절에서 열거한 세 개의 가보 웨이블릿 중 어떤 커널이 가장 최적의 성능을 산출하는지, 몇몇 커널 고유의 파라미터를 제외한 나머지 변수들은 모두 동등한 조건으로 실험하였다. 실험에 사용한 얼굴 데이터는 자체 데이터를 사용하여 인식을 측정하였으며, 실험 과정 중 인식을 측정에서 유사도 측정은 거리계산 공식인 코사인 거리 계산을 이용하여 인식을 측정하였다. 아래표 1 , 2, 3 과 같이 다양한 해상도로 들어오는 이미지를 각각의 커널 크기에 맞게 적용하기 위해서 주파수와 방향수치를 달리 실험하여 결과를 내었다.

F \ O	2	3	4	5	6	7	8
3	235	83	42	21	14	8	10
4	274	107	55	27	22	17	17
5	143	52	25	14	9	9	8
6	146	65	28	19	12	11	12
7	110	44	22	13	13	11	11
8	118	44	19	13	12	10	8
12	x	41	23	12	12	10	9
16	x	42	23	13	13	10	9
20	x	41	22	13	13	10	9

(표 1) 이미지 크기 : 128x128, 커널 크기 : 33x33, 시그마 : 2PI(F:주파수 , O:방향)

F \ O		3	4	5	6	7	8
3	80	22	17	14	15	15	16
4	90	40	32	23	20	19	19
5	34	19	14	14	13	13	13
6	48	29	21	17	18	17	18
7	29	25	14	12	12	13	13
8	28	22	18	16	14	15	15
12	x	25	14	15	13	13	13
16	x	26	17	15	13	13	13
20	x	25	16	15	13	13	13

(표 2) 이미지 크기 : 64x64, 커널 크기 : 17x17, 시그마 : 2PI (F:주파수 , O:방향)

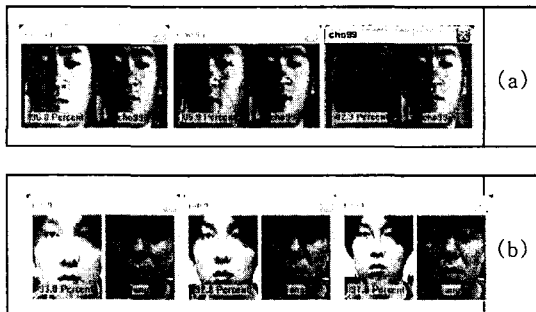
F \ O	2	3	4	5	6	7	8
3	31	18	20	19	21	20	20
4	30	17	19	19	19	18	18
5	25	14	15	17	17	19	19
6	22	17	17	18	18	18	18
7	22	15	18	17	16	16	16
8	20	15	18	18	16	17	17
12	21	15	16	18	17	18	18
16	23	15	16	18	17	18	18

(표 3) 이미지 크기 : 32x32, 커널 크기 : 9x9, 시그마 : 0.75PI (F:주파수 , O:방향)

앞서 구한 가보 커널수치를 이용하여 온라인 서버일런스 시스템에 확장, 적용하여 실험을 해 보았다. 기존의 서버일런스 시스템의 얼굴인식은 다 해상도로 입력되는 얼굴 이미지의 경우 보간법을 이용, 확대/축소하여 128x128 크기로 통일시킨 후 동일한 커널과 거기에 따르는 내부수치를 적용하여 인식하였다. 본 논문에서 제안하는 시스템은 실험에서 얻어낸 다해상도의 얼굴 이미지 각각을 위한 커널을 이미지 사이즈에 따라 여러 부류로 나누어서 각각의 크기에 합당한 커널을 해당 해상도로 입력된 이미지에 적용하여 그 결과를 보기 위한 파라미터를 아래 표 4 와 같이 산출하였고, 그 실험 결과가 그림 1 에 나타나 있다.

커널	이미지 사이즈	커널크기	시그마	주파수	방 향
128	128x128 ~80x80	33x33	2.0	5	7
64	79x79~3 3x33	17x17	1.0	5	7
32	32x32~..	9x9	0.75	3	5

(표 4) 이미지 해상도에 따른 커널 크기 및 파라미터



(그림 1) 서버일런스 시스템 적용 결과
(a): accept, (b) reject

기존의 시스템의 경우, 조명변화에 크게 변화가 없는 안정된 인식률을 보였다. 하지만 현재 개선된 시스템에 적용하여 실험을 했을 경우, 조명이 별로 없는 상태에서는 좀 더 좋은 인식률을 보여주지만, 조명이 강렬한 얼굴 영상에서는 심한 오인식을 나타내었다. 이것은 기존 시스템의 커널 크기가 33x33 으로서 고주파수 성분을 잘 잡아내는 특성이 있었기 때문에

조명의 영향에도 관계없는 좋은 인식률을 보였지만, 현재 개선된 시스템에 적용한 17x17, 9x9 커널의 경우 고주파수 성분보다는 저주파수 성분쪽에 강하기 때문에 조명을 밝게 주어서

고주파수 성분이 강조된 이미지의 경우 특징을 잘 검출하지 못하여 오인식이 나오지 않았나 생각한다. 현재 개선된 시스템에서 조명쪽에 대한 대비를 좀 더 보강한다면 좀 더 안정적인 인식을 가져오지 않을까 기대한다

5. 결론

본 논문에서는 얼굴 인식을 위해 얼굴 이미지가 다 해상도로 입력 되었을 때 각각의 해상도에 따라 어떠한 커널 사이즈와 내부 파라미터에서 가장 좋은 결과를 보이는지 실험, 분석하였다. 가보 웨이블릿은 자동얼굴 인식을 위한 특징을 추출하는데에 매우 효과적인 방법으로 알려진 바와 같이 우수한 성능을 보였으며, 그 커널의 크기와 내부 파라미터를 어떻게 선택하느냐에 따라 결과에 많은 영향을 미친다는 것도 알 수 있었다. 따라서 입력되는 얼굴 이미지를 실험한 결과를 바탕으로 크기에 따라 적합한 파라미터의 범위를 설정, 서버일런스 시스템에 적용하여 성능을 검증하였다. 그리고 가보 벡터를 추출하여 유사도를 측정 함으로써 다해상도로 입력되어지는 얼굴 인식에 적합하며, 안정된 인식결과를 기대할 수 있게 되었다. 하지만 작은 이미지인 경우 조명과 같은 환경에 매우 민감하기 때문에 항상 안정된 결과를 가져온다고 볼 수는 없었다. 향후 연구 방향은 조명, 회전, 크기에 안정 적이라고 알려진 가보 웨이블릿의 특성을 살려 주변 환경과 크기에 좀 더 안정적인 방법을 연구하는 것이다.

6. 참고문헌

[1] H.A.Rowley, S.Baluja, T.Kanade, "Neural network-based face detection" Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on, 18-20 June 1996

[2] A.Pentland, B.Moghaddam, T.Starner, "View-based and modular eigenspaces for face recognition" Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on, 21-23 June 1994

[3] Penio S. Penev, Jeseoph J. Atick, "A General Statistical Theory for Object Representation " Network: Computation In Neural Systems, Vol.7, pp.477-500, 1996

[4] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," J. Neurophysiol., 58(6):1233-1258, 1987.

[5] L.Wiskott, J.K.Fellous, N.Kruger, von der Malsburg, C, "Face recognition by elastic bunch graph matching" Image Processing, 1997. Proceedings., International Conference on, Volume: 1, 26-29 Oct. 1997