

Haar Wavelet Transform 전처리 기법을 적용한 PCA, LDA기법의 얼굴 인식

이동훈⁰
일전방사선엔지니어링
dhlee@ijrad.co.kr

Haar Wavelet Transform Preprocessing Technique to Face Recognition of PCA, LDA

Dong-Hun Lee⁰
Iljin Radiation Engineering

요 약

얼굴 인식을 위한 주요 기법인 PCA, LDA에 의한 mapping기법은 조명조건의 미세한 변화에 민감한 특성을 가진다. 얼굴 인식 연구에 있어서 인식률의 향상뿐만 아니라 실용적인 얼굴 인식 시스템을 구현하기 위해서는 조명 변화를 최소화 시키는 전처리 과정이 중요한 고려사항이다. 따라서 본 논문에서는 조명의 변화를 최소화 할 수 있는 전처리 방법으로 Haar 웨이블릿 변환으로 얻어진 웨이블릿 계수공간의 조정 후 역변환을 통한 영상향상을 제안한다. 실험 결과 제안한 방법은 기존의 전처리 방법으로 널리 쓰이는 히스토그램 평활화 방법에 비해 우수한 성능을 나타내었을 뿐만 아니라 메모리 절감효과에 따른 처리속도 증가를 보였다.

1. 서 론

널리 알려진 얼굴 인식 기법 중 PCA기법은 클래스 정보를 사용하지 않고 특징 공간에서의 데이터 분포의 특성을 유지하면서 영상의 차원을 줄이는데 목적[1]이 있으므로, 인식을 위한 분류의 목적으로는 사용하기 적합하지 않다. LDA방법은 클래스 정보를 사용하여 동일 클래스의 데이터들은 서로 모이고, 다른 클래스 간의 데이터들은 서로 분리하는 목적[2]이 있기 때문에 입력 데이터들을 비교적 정확하게 분류한다. 대표적인 예로 FLD(Fisher linear discriminant)[3]기법이 있다. 하지만, 조명등의 변화에 따라 인식특성이 민감하여 eigenface나 fisherface의 적용이 어렵다. 따라서 신뢰도가 높은 얼굴 인식의 구현을 위해 해선 조명의 변화를 최소화하는 전처리 과정이 필수적으로 요구된다.

기존의 얼굴 인식에 관련된 연구에서는 조명의 밝기 변화를 보정하기 위한 간단한 전처리로 히스토그램 평활화를 수행하여 밝기를 정규화하는 방법이 널리 사용되어 왔다 [4]. 활발히 연구되고 있는 웨이블릿 변환(wavelet transform)은 신호처리 분야의 신호코딩, 영상압축, 컴퓨터비전에서 사용되고 있다[5]. 웨이블릿(wavelet)을 이용한 변환은 영상을 서로 다른 해상도로 표현하는 다해상도 표현을 결합시킨 형태로서 주파수영역과 공간영역 모두에서 분할 처리가 가능하다[6].

본 논문에서는 정규화된 입력 얼굴 영상을 Haar 웨이블릿변환하여 영상의 웨이블릿 계수공간의 조정후 역변환을 이용하여 조명변화를 최소화 할 수 있는 전처리 방법을 제안한다.

2. 본론

2.1 주요성분 분석법(PCA)

PCA 방법은 패턴인식의 여러가지 기법 중 확률적인 접근 방식을 이용한 것으로, 서로 상관하는 영상벡터들을 공분산 행렬의 고유벡터들에 대한공간으로 mapping하여 서로 상관하지 않게 하는 방법이다. m차원을 가진 영상벡터에 대한 공분산 행렬은 다음과 같이 주어진다.

$$C_x = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})(x_n - \bar{x})^T = \Omega \Omega^T \quad (1)$$

$\Omega^T \Omega$ 의 고유벡터 E를 구한 후 Ω 을 곱하면 원하는 고유 벡터는 ΩE 가 되는 것을 아래의 식에서 보인다.

$$(\Omega \Omega^T) \Omega E = \Omega E \lambda \quad (2)$$

이 고유벡터를 고유얼굴이라고도 부른다. 특징 벡터는 다음의 식으로 구할 수 있다.

$$y = (\Omega E)^T (x - \bar{x}) \quad (3)$$

공분산 행렬은 대칭행렬이므로 생성되는 고유벡터들은 서로 직교한다, 따라서 특징벡터들은 서로 상관하지 않게 된다.

2.2 선형판별 분석법(LDA)

FLD방법은 위의 PCA방법을 개선한 것이다. 이 방법은 한클래스 내에서의 변화와 다른 클래스들간의 변화의 비를 고려한 것이다. 즉, 다른 클래스들 간의 scatter 행렬(S_b)의 행렬식의 값을 최대화하고, 동시에 같은 클래스간의 scatter행렬(S_w)의 행렬식 값을 최소화한다. S_w 와 S_b 는 다음과 같이 정의된다.

$$S_w = \sum_{i=1}^N \Pr(C_i) E[(V_i - \mu_i)(V_i - \mu_i)^T] \quad (4)$$

$$S_b = \sum_{i=1}^N \Pr(C_i) (\mu - \mu_i)(\mu - \mu_i)^T \quad (5)$$

μ 는 전체 평균벡터를, μ_i 는 i번째 클래스의 평균벡터를 나타낸다. FLD 기법에서는 $|W^T S_b W| / |W^T S_w W|$ 를 최대로 하는 프로젝션 matrix W 를 사용한다. 이는 eigenvalue 문제로서 다항식에 의해 구할수있다.

$$S_b u_i = \lambda_i S_w u_i \quad (6)$$

eigenvalue의 크기가 큰 순으로 배열했을 때 q 개의

eigenvector로 구성된 matrix를 다음과 같이 구할 수 있다.

$$W_{LDA} = [u_1, u_1, \Lambda, u_q] \quad (7)$$

$m \cdot n$ 크기의 영상벡터 x_i 는 q 차원의 새로운 벡터 z_i 로 다음 식에 의해 재 구성된다.

$$z_i = W_{LDA}^T x_i \quad (8)$$

생성된 고유벡터를 PCA에서의 고유얼굴에 구분하기 위해서 fisher얼굴이라고 부른다. PCA에서 얻어진 특징벡터를 fisher얼굴에 투영하여 특징벡터를 얻을 수 있다. 인식은 학습과정에서 만들어진 특징벡터들과 입력 얼굴 영상의 특징벡터들간의 특별한 분류기의 설계 없이 단지 최소 유클리디안 거리(Euclidean distance)를 사용하여 분류할 수 있다.

2.3 Haar Wavelet Transform

웨이블렛 변환의 기본개념은 주어진 얼굴 영상을 웨이블렛 평면에 투영하여 여러 개의 해상도로 분해하는 것이다. 웨이블렛 평면은 서로 다른 해상도를 갖는 기본 함수들로 구성되어 있어서, 얼굴 영상은 다른 해상도를 갖는 부합수들로 변환된다. 웨이블렛 변환은 신호에 대해 시간과 주파수의 지역성을 갖는다. 예로, L단계 분해를 수행한 후에는 3L 개의 고주파 분해 영상과 L개의 저주파 분해 영상을 얻을 수 있다. 2단계 분해는 1회 분해한 저주파 서브 이미지에 대하여 재귀적으로 분해를 반복하여 얻을 수 있다.

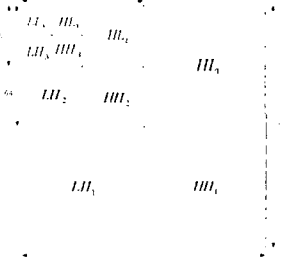
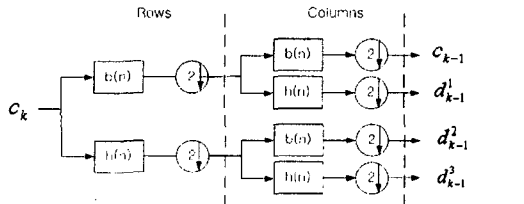
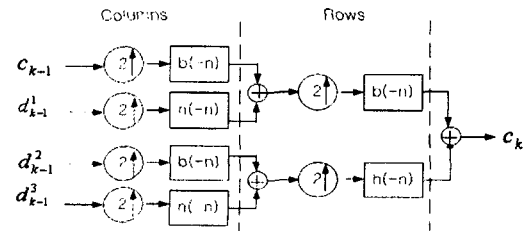


그림 1. 2단계 웨이블렛 분해

그림 1에서 부영상 LL, LH, HL, HH 는 2단계 웨이블렛 변환된 결과들이고, 이는 저주파 분해 영상 LL , 수평 방향 고주파 분해 영상 HL , 수직 방향 고주파 분해 영상 LH , 대각선 방향 고주파 분해 영상 HH 로 표기 된다.



(a) 다해상도 분해 방법을 이용한 이미지 사영도



(b) 다해상도 분해를 원래 이미지로 복원

그림 2. 웨이블렛 변환 필터뱅크

그림 2에서 이미지 영상은 확장계수 $b(n)$ 과 웨이블렛 계수 $h(n)$ 을 이용하여 이미지를 주파수가 다른 여러 개의 이미지로 분할하거나 다시 분할되어진 이미지로부터 원래의 이미지를 복원한다. 2차원 웨이블렛 계수 공간은 2차원 유사계수 (approximation coefficient) c_{k-1} 와 2차원 상세 계수(detail coefficient) $d_{k-1}^1, d_{k-1}^2, d_{k-1}^3$ 와 같이 정의 되어진다. 이들 계수들은 입력 이미지를 주파수 대역이 상이한 사영 이미지로 나타내며 c_{k-1} 는 저주파 영역, $d_{k-1}^1, d_{k-1}^2, d_{k-1}^3$ 는 고주파 영역으로 주파수 대역이 세분화 되어진다.

Haar 웨이블렛 함수는 시간적으로 에너지가 집중하며 진동하는 특성을 가지므로, 웨이블렛의 기본성질인 허용조건과 진동조건을 만족하며, 서로 상호직교하므로 웨이블렛 합수로 사용할 수 있고 시간, 주파수 축상에서 영상을 저주파 영역의 평균영상과 고주파 영역의 차분영상으로 분리할 수 있다[7]. Haar 웨이블렛 변환은 기저함수가 단순하므로 분해와 합성이 고속이다. 본 논문에서 이용한 Haar filter는 유사계수를 구하기 위한 필터 계수로 $\{1/\sqrt{2}, 1/\sqrt{2}\}$ 를 상세 계수를 구하기 위한 필터 계수로 $\{1/\sqrt{2}, -1/\sqrt{2}\}$ 를 사용하였다. Haar의 유사 계수 c 와 상세 계수 d 의 변환 원리는 다음과 같다.

$$\text{유사계수를 얻기 위한 이동평균(Moving Average)} \quad c_k^m = \frac{1}{\sqrt{2}}(c_{k-1}^m + c_{k-1}^m)$$

$$\text{상세계수를 얻기 위한 이동차(Moving Difference)} \quad d_{k-1}^m = \frac{1}{\sqrt{2}}(c_{k-1}^m - c_{k-1}^m) \quad (9)$$

단, $1 \leq k \leq L/2$ 이다.

이 과정에서 자동적으로 저주파 대역은 평균 분해 필터로 인해 평활화가 이루어지고, 고주파 대역은 차분 필터에 의해 영상 에지화가 이루어진다. 평활화된 저주파 대역의 유사계수는 그대로 두고 고주파 대역의 상세계수에 실수가 중치를 곱하여 변화된 웨이블렛 계수 공간을 역변환 함으로서 윤곽이 뚜렷하고 평활화된 복원 영상을 얻게된다. 여기서 c_m^l, d_m^l 은 분해레벨 l , 왼쪽 위의 이미지를 기준으로 한 m 번째의 화소값, L 은 저주파 성분의 높이(height), 또는 너비(width)를 나타낸다. 위의 식이 나타내는 바와 같이 계산은 매우 단순하여 고속으로 수행되고 그림 3에 나타낸 바와 같이 웨이블렛 변환은 1회 분해할 때마다 웨이블렛 변환은 2:1의 다운 샘플링을 수행한다. 이미지는 2차원이므로 화소 수는 1/4이다. 이미지의 축소로 특징추출시의 계산량이 적어지므로 계산이 고속으로 수행된다[8]. 아래의 그림은 제안한 시스템을 나타낸다.

이 과정에서 자동적으로 저주파 대역은 평균 분해 필터로 인해 평활화가 이루어지고, 고주파 대역은 차분 필터에 의해 영상 에지화가 이루어진다. 평활화된 저주파 대역의 유사계수는 그대로 두고 고주파 대역의 상세계수에 실수가 중치를 곱하여 변화된 웨이블렛 계수 공간을 역변환 함으로서 윤곽이 뚜렷하고 평활화된 복원 영상을 얻게된다. 여기서 c_m^l, d_m^l 은 분해레벨 l , 왼쪽 위의 이미지를 기준으로 한 m 번째의 화소값, L 은 저주파 성분의 높이(height), 또는 너비(width)를 나타낸다. 위의 식이 나타내는 바와 같이 계산은 매우 단순하여 고속으로 수행되고 그림 3에 나타낸 바와 같이 웨이블렛 변환은 1회 분해할 때마다 웨이블렛 변환은 2:1의 다운 샘플링을 수행한다. 이미지는 2차원이므로 화소 수는 1/4이다. 이미지의 축소로 특징추출시의 계산량이 적어지므로 계산이 고속으로 수행된다[8]. 아래의 그림은 제안한 시스템을 나타낸다.

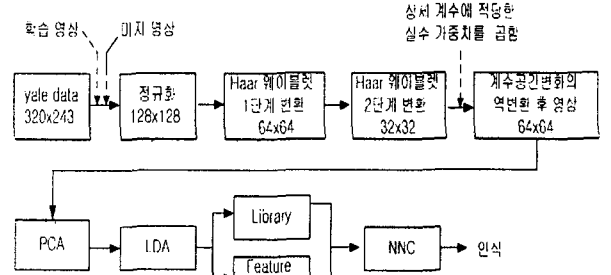
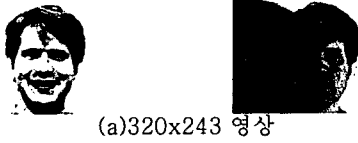


그림 3. 제안된 시스템 블록도

3. 실험 결과 및 고찰

실험에 사용되는 얼굴 영상은 YALE 데이터베이스이다. 이 데이터베이스는 15명의 클래스에 대해 개인별 11장의 서로 다른 정면 얼굴 영상을 가지고 있다. 이 영상들은 조명의 변화(중앙조명, 왼쪽조명, 오른쪽조명), 얼굴표정(행복, 무표정, 슬픔, 놀람, 씩씩), 안경착용의 변화요소를 가지고 있다. 원영상의 크기는 320x243이었으나 실험에

서는 128x128로 정규화하였다. 그 것을 아래의 그림에 보인다.



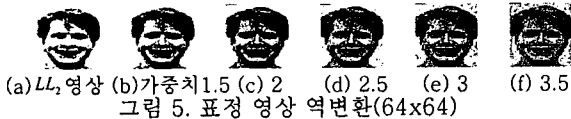
(a)320x243 영상



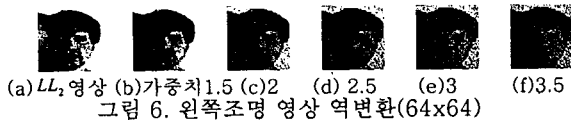
(b) 128x128로 정규화된 영상

그림 4. 실험에 사용된 YALE 데이터베이스의 정규화

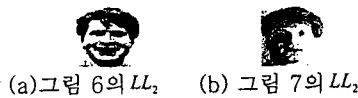
각각의 정규화된 얼굴 영상(128x128)에 Haar 웨이블릿 변환을 2단계까지 수행한 후 2단계 변환된 웨이블릿 계수 공간에서 상세계수에 1보다 큰 실수 가중치를 곱한후 역변환한 결과(64x64)와 1단계 수행한 후의 LL_2 저주파 영상(64x64)을 그림 5, 그림 6, 그림 7에 보인다.



(a) LL_2 영상 (b)가중치 1.5 (c) 2 (d) 2.5 (e) 3 (f) 3.5
그림 5. 표정 영상 역변환(64x64)



(a) LL_2 영상 (b)가중치 1.5 (c) 2 (d) 2.5 (e) 3 (f) 3.5
그림 6. 왼쪽조명 영상 역변환(64x64)



(a)그림 6의 LL_2 (b) 그림 7의 LL_2

그림 7. LL_2 를 히스토그램 평활화한 영상(64x64)

본 논문에서는 2단계 웨이블릿 분해한 주파수 대역에서 LL_2 유사계수는 그대로 두고 LH_3, HL_3, HH_3 상세계수에 1보다 큰 여러 실수 가중치값을 곱하여, Haar 웨이블릿 역변환을 한다음 각각의 역변환된 64x64 영상을 3종류의 영상과 성능 비교 하였다. 3종류의 영상은 첫째 1단계 분해된 후의 LL_2 (64x64), 두번째 LL_2 를 히스토그램 평활화한 영상(64x64) 그리고 3번째 128x128영상을 아무런 조작없이 64x64로 축소한 영상이다.

제안된 방법의 성능을 검증하기 위해 정규화된 Yale 데이터베이스로부터 세가지 방식을 선정하여 실험하였는데, 첫 번째 방식은 Yale 데이터베이스에서 각 클래스당 11장의 모든 영상을 데이터로 사용하였고, 두 번째 방식은 각 클래스당 조명 변화가 있는 3장의 영상을 제외한 8장의 영상을 사용하였다. 그리고 세 번째 방식은 각 클래스당 조명 변화 있는 영상 3장을 사용하였다. 실험 데이터 영상들을 그림 5와 그림 6과 같이 제안한 방법으로 각각의 가중치에 대해 역변환된 영상으로 구성하였다. 실험 데이터 영상들에 대해 각각 PCA, LDA를 적용한 후 얼굴특징 벡터들과의 최소 유클리드 거리를 구하는 nearest neighbor classifier (NNC)를 적용하여 얼굴을 인식한다.

얼굴 인식 성능 검증하기 위해 각각 Leave-one-out 방식을 적용하였다. 클래스가 8명일때의 결과와 클래스가 15명 일때의 결과를 표 1과 표 2에 보인다. 표 1과 표 2에서 보면, 가중치가 1.5일 때 클래스 수에 영향없이 제안한 전처리방법이 정규화된 영상을 1/2한 영상과 1단계 Haar 웨이블릿 변환의 LL_2 대역 영상을 그대로 사용한 얼굴 인식 보다 낮은 에러율을 가진다. 그리고 가중치가 2일 때 전

처리로 잘 알려진 히스토그램 평활화보다 낮은 에러율을 보인다. 전체적으로 가중치가 1.5일때 어떤 비교영상 보다도 가장 적은 에러율을 가진다. 부가적으로 웨이블릿 변환을 함으로써 영상의 차원이 16384(128x128)에서 4096(64x64)로 4배의 감소를 가지게 된다. 이로 인해 기존의 전처리 방법에 비해 약 4배의 인식 속도 향상과 메모리 절감의 효과를 보였다.

표 1. 얼굴 클래스가 8명일때 인식 error rate (PCA+LDA)

영상 대역 영상 종류	비교 영상 (64x64)		제안된 가중치 역변환 영상 (64x64)					
	최소한 영상	히스토그램 평활화	LL_2	1.5	2	2.5	3	3.5
조명변화 영상(3)	50%	30%	50%	40%	30%	40%	60%	60%
조명+아래 조명(4)	5.56%	2.22%	6.67%	1.11%	3.61%	3.61%	5.61%	5.61%
전체 영상(7)	2.22%	1.67%	2.22%	1.11%	2.22%	2.22%	2.22%	4.48%

표 2. 얼굴 클래스가 15명일때 인식 error rate (PCA+LDA)

영상 대역 영상 종류	비교 영상 (64x64)		제안된 가중치 역변환 영상 (64x64)					
	최소한 영상	히스토그램 평활화	LL_2	1.5	2	2.5	3	3.5
조명변화 영상(3)	46.67%	33.33%	46.67%	26.67%	26.67%	43.33%	53.33%	53.33%
조명+아래 조명(4)	2.22%	1.48%	2.98%	1.48%	2.22%	2.22%	2.98%	3.48%
전체 영상(7)	1.48%	1.06%	2.14%	0.74%	1.48%	1.06%	1.06%	2.48%

4. 결론

본 연구에서는 시간-주파수축상에서 영상을 저주파 영역의 평균영상과 고주파영역의 차분영상으로 분리가 가능한 Haar 웨이블릿 변환을 적용하였다. 얼굴 인식기법에서 널리 사용되어 온 PCA, LDA 등의 mapping 기법은 학습 영상의 조명 변화에 민감하게 반응하여 높은 인식율을 유지하기 어려운 단점을 가진다. 따라서, 신뢰도가 높은 얼굴 인식 기법의 구현을 위해서 조명변화를 최소화 시키는 전처리 과정이 필수적으로 요구된다. 본 논문에서 제안한 방법은 Haar 웨이블릿 역변환 특성으로 조명의 변화에 대해 우수한 한 얼굴인식 성능을 보여주었을 뿐만아니라 메모리 절감 효과에 따른 인식속도 증가를 가져왔다.

5. 참고문헌

- [1] M.A. Turk and A. Pentland, "Face recognition using eigenfaces," Proc. IEEE Conf. on CVPR, pp586-591, 1991
- [2] R. Brunelli and T. Poggio, "Face recognition: feature vs templates," IEEE Trans. On PAMI, Vol.15 No.2, pp 1042-1052, 1993
- [3] A. Yuille, P. Hallinan, and D. Cohen, "Feature extraction from faces using deformable templates," Int. J. Computer Vision, Vol.8, No.2, pp. 99-111, 1992
- [4] A.S. Georghades, D. J. Kriegman, P. N. Belbumeur, "Illumination Cones for Recognition Under Variable Lighting: Faces," Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, PP.52-58, 1998
- [5] M. A. Turk, A. P. Pentland, Face recognition using eigenfaces, Int. Conf. on Pattern Recognition, (1991) 586-591.
- [6] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd Edition, John Wiley & Sons, (2001)
- [7] D. Gubbins, I. scollar, and Wisskirchen, "Two Dimensional Digital Filtering with HAAR and WALSH Transforms", Ann. Geophys, Tome 27, Numero 2, pp.85-104, 1971.
- [8] G C Feng, P C Yuen and D Q Dai, Human face recognition using PCA on Wavelet subband, Journal of Electronic Imaging, Vol. 9, No.2 pp.226-233, 2000