

GWT 계수 에너지와 원영상 결합을 이용한 얼굴 인식

한정훈⁰, 홍소범, 김우생
 광운대학교 컴퓨터 과학과
 {sopure⁰, hongallen, kwsrain}@cs.kw.ac.kr

Face recognition in conjunction between GWT coefficients' energy and original image

Jeonghoon Han⁰, Xiaofan Hong, Woosaeng Kim
 Dept. of Computer Science, Kwangwoon University

요 약

본 논문에서는 GWT(Gabor Wavelet Transform) 계수 에너지와 원 영상간의 영상 결합을 수행한 영상을 주성분 분석법(Principal Component Analysis)에 적용하여 얼굴 인식을 하는 방법을 제안한다. GWT는 가버 함수의 크기 변화와 방향 변화에 의해 생성된다. 따라서 GWT는 다양한 크기 변화와 방향 변화를 가지는 변환으로 특정 주파수 성분과 방향성을 가지는 영상 구조가 어디에 있는지의 지역적 정보를 효과적으로 표현할 수 있는 변환으로 알려져 있다. GWT를 통해 나온 계수 에너지를 추출하고 원 영상에 더하여 지역적 특성을 크게 만든 후에 통계적 방법 중 가장 많이 사용되고 어지고 검증들 받은 PCA를 사용하여 인식한다. GWT 계수의 에너지는 얼굴 윤곽선, 눈과 입, 얼굴과 머리의 경계 등 색감의 급격한 변화를 나타내는 곳의 정보를 표현을 해주기 때문에 특징점 추출에 사용되고 있지만 이를 전역적으로 이용하여 인식하는 방법에 관한 연구가 이루어지지 않고 있다. 본 논문에서는 에너지 값만으로 전체 얼굴 영상의 세부적 표현을 할 수 없기 때문에 원 영상과의 1:1 비율의 영상 결합을 한 후 얼굴 인식 처리에 사용한다. 이 영상을 얼굴 인식에 사용하였을 때 원본 영상을 사용하였을 때보다 오인식이 줄었다.

1. 서 론

얼굴 인식은 오래 전부터 연구되어진 컴퓨터 응용 분야로서 얼굴은 인식과 감정 전달에 있어서 큰 역할을 하기에 관심의 초점이 되어 왔고, 최근 인터넷의 발달로 인한 인터넷, 모바일 사용자의 급증으로 보안, 인증뿐만 아니라 다양한 멀티미디어 콘텐츠 서비스를 요구하게 되면서 캐리커처 서비스, 아바타 서비스 등의 분야에서 활성화 되고 있는 추세이다.

얼굴 인식은 다른 생체 인식 기술에 비해 강압적이지 않고 다른 생체 기술이 요구하는 별도의 센서가 필요하지 않고 가정용 구비된 카메라나 PC 카메라 등을 이용할 수 있다는 장점이 있다. 뿐만 아니라 영상 획득의 용이성과 은행 ATM등의 보안 시설과 같은 곳에서 감시용으로 사용될 수 있다는 점에서 응용 분야가 넓다[1]. 얼굴 인식은 크게 특징 기반 방법과 영상 기반 방법으로 나누어진다[2]. 특징 기반 방법은 얼굴의 특징을 추출하여 인식하는 방법이고, 영상 기반 방법은 전체 얼굴 영상을 통계적 방법을 사용하여 인식하는 방법이다.

본 논문에서는 사람의 시각계와 비슷한 특징을 가지고 있으며, 좋은 특징 추출법으로 알려진 가버 웨이블릿 계수의 에너지와 원본 영상을 1:1 비율로 결합하여 나온 영상을 이용하여 인식하는 방법을 제안한다. 이렇게 나온 영상은 원 영상을 사용하였을 때보다 개인 간 특성을 더욱 잘 살려주는 영상이 된다. 인식 알고리즘은 현재 얼굴 인식 방법에서 많이 사용되는 주성분 분석법(PCA)을 이용하였으며, 본 논문의 타당성은 원 영상을 사용하여 인식 하였을 경우의 인식율과 제안하는 방법을 사용하였을 때의 인식율과의 비교를 통하여 검증한다.

본 논문의 구성은 2장에서 가버 웨이블릿 변환과 주성분 분석법에 대해서 소개를 하고 3장에서는 얼굴 인식 시스템 구조 및 설계에 대해 설명한다. 4장에서는 실험 및 결과를 보이며 5장에서는 결론을 맺는다.

2. GWT 계수 에너지와 주성분 분석법

이차원 가버 웨이블릿 변환을 위한 2-D함수와 그것의 푸리에 변환은 다음과 같이 정의된다[3].

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + j2\pi u x \right]$$

$$G(u, v) = \exp \left[-\frac{1}{2} \left(\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right) \right]$$

여기서 σ_x 와 σ_y 는 x 와 y 축의 표준 편차이다. σ_u 와 σ_v 는 u 와 v 축의 표준 편차이다. 가버 웨이블릿은 가버 함수의 크기 변화와 방향변화로 생성된다.

$$g_m(x, y) = a^{-m} g(x, y),$$

$$a > 1, m = 1:M, n = 1:N.$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = a^{-m} \begin{bmatrix} \cos\theta_n & \sin\theta_n \\ -\sin\theta_n & \cos\theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

$$\theta_n = n\pi/N$$

a 는 크기 파라미터이고 m 은 주파수 변화 파라미터이고, n 은 방향 변화 파라미터이다.

가버 웨이블릿 변환은 다음처럼 원본 영상과 가버 웨이블릿의 컨볼루션으로 정의된다.

$$\mathcal{X}(x, y, m, n) = \int \mathcal{X}(x', y') g_m(x - x', y - y') dx' dy'$$

GWT 계수의 에너지는 [4]에서 다음처럼 정의해 놓았다.

$$E(x, y) = \sum_{m=1}^M \sum_{n=1}^N \mathcal{I}(x, y, m, n)$$

주성분 분석법(principal component analysis)은 화상과 같이 큰 데이터의 분석을 작은 크기의 데이터로 변환하여 위해 사용되는 기술이다. 관찰 벡터 $[X_1, X_2, \dots, X_N]$ 를 각각 $N \times N$ 의 크기를 가지는 영상 X 이라 할 때 영상 X 에 대한 공분산 행렬 X_c 는 다음과 같다.

$$X_c = E[(X - m_x)(X - m_x)^T] = \frac{1}{M} \sum_{i=1}^M (X_i - m_x)(X_i - m_x)^T$$

영상 X 의 공분산 행렬 X_c 는 대칭 행렬이고 $A \Lambda A^T$ 로 분해가 가능하다. 여기서 Λ 는 대각행렬이고 이것의 대각 원소는 X 의 내림차순 고유치와 같고, A 는 X 의 고유벡터가 된다. 주성분 분석 변환식은 다음과 같다.

$$Y = A(X - m_x)$$

이를 이용하되 큰 고유치에 해당하는 k 개의 고유벡터만을 계산하여 차원을 줄인다. 줄여진 k 개의 Y_k 벡터를 사용하여 원본 X 를 재구성할 수 있다.

$$X = A^{-1}Y + m_x = A^T Y + m_x$$

줄여진 k 개의 Y_k 벡터로부터 X 의 약근치 \hat{X} 는 아래와 같이 구할 수 있다.

$$\hat{X} = A_k^T Y_k + m_x$$

3. 제안하는 얼굴 인식 시스템

이 장에서는 본 논문에서 설계한 얼굴 인식 시스템 구조와 사용된 인식 기술들에 대해서 설명한다. 학습 및 테스트 원 영상들을 가버 웨이블릿 변환(GWT)을 시키고 계수 에너지를 추출한다. GWT 계수 에너지 영상은 시각적으로 중요한 얼굴 윤곽 예를 들어 눈, 코, 입, 얼굴 윤곽선 등은 보여 주지만 세부적인 정보를 표현하지 못하기 때문에 원 영상과의 결합을 통하여 세부적인 형태를 잃지 않도록 한다. 따라서 아래의 식과 같이 1:1 비율로 더해서 나온 결과 영상을 원 영상 대신 사용하게 된다.

$$\mathcal{I}(x, y) = \frac{E(x, y) + \mathcal{I}(x, y)}{2}$$

여기서 $E(x, y)$ 는 영상의 (x, y) 좌표에서의 GWT 계수 에너지 값이고, $\mathcal{I}(x, y)$ 는 원본 영상의 (x, y) 에서의 계수값, $\mathcal{I}(x, y)$ 는 두 영상이 영상 결합된 영상의 값을 의미한다. 그림 1은 원 영상(왼쪽 그림)과 원 영상의 GWT 계수 에너지 영상(중간 그림),

원 영상과 GWT 계수 에너지를 1:1 비율로 결합된 영상(오른쪽 그림)을 보여준다.



그림 1. 원본 영상(왼쪽)와 GWT 계수의 에너지를 표현한 영상(중간), 두 영상이 영상 결합된 영상(오른쪽)

그림 1에서 보는 바와 같이 각기 사람의 GWT 계수 에너지는 서로 다르기 때문에 1:1 비율로 GWT 계수 에너지가 추가된 영상은 각 사람마다 제각기 다른 GWT 계수 에너지를 원 영상에 추가한 것이기에 원 영상에 비해 각 개인의 특징을 더욱 잘 살려주고 개인 간 차이가 더 커진 영상이 된다.

이 영상을 이용하여 주성분 분석을 통해 인식을 한다. 학습 데이터베이스의 모든 영상을 고유 벡터들의 선형 결합으로 나타낼 수 있으므로 아래 식을 이용하여 k 개로 축소된 성분값을 구할 수 있다.

$$Y_k = A_k(X - m_x)$$

따라서 각 영상들은 k 개의 원소를 가지는 성분 벡터를 가지게 되고 인식은 k 개의 원소의 거리차를 비교하여 인식할 수 있다. 인식되어 나오는 결과는 같은 사람인지 여부만을 판단하여 인식율을 결정하도록 한다. 인식은 학습 데이터의 고유 얼굴(Eigenface)을 이용하여 입력 영상 간의 유사도를 구하되 방식은 유클리드 거리가 최소가 되는 것을 인식 결과로 나타내기로 한다. 유클리드 거리 d 는 아래와 같다.

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n |x_k^{(i)} - x_k^{(j)}|^2} = (x_i - x_j)^T (x_i - x_j)$$

이때 유클리드 거리 d 값이 최소인 집합을 유사한 패턴으로 결정하게 된다.

4. 구현 및 결과

결과를 구하기 위해 사용된 컴퓨터는 Pentium-4 2.8GHz의 성능을 가지고 메인 메모리의 크기는 1GB이고 운영체제는 Windows XP Professional을 사용하였다. GWT 계수 에너지를 구하기 위해 사용된 언어는 Visual Studio 6.0의 C++을 사용하

였으며 PCA 변환 및 인식을 하기 위해서 Matlab 7.0.4를 사용하였다. 계산의 부하를 고려하여 원본 영상의 크기를 Bicubic interpolation을 이용하여 64×64로 만들었고, 축소된 차원수는 30차원으로 할 경우 원 영상의 97%를 표현할 수 있음을 그림 2와 같이 확인하여 30을 사용하였다.

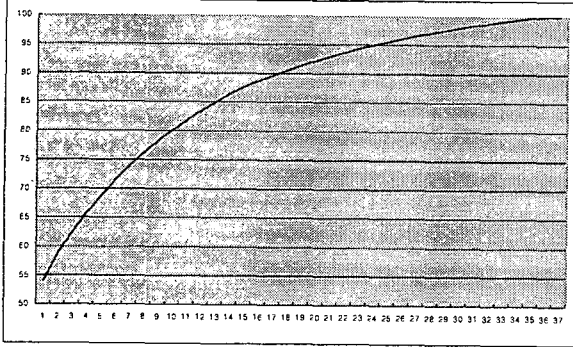


그림 2. 차원수에 따른 고유값의 누적 비율

영상 데이터베이스에 각 사람당 1장의 사진을 학습 영상으로 입력 받는다. 이를 제외한 영상을 테스트 영상으로 사용한다. 가버 웨이블릿의 비직교성(nonorthogonality) 때문에 가버 웨이블릿 변환을 거친 정보는 중복된다. 이 중복성을 최소화하기 위해서 B.S. Manjunath 등은 다음 식을 사용했다[3]. 본 논문에서는 이 중에서 인식식 가장 높은 확률을 보인다고 알려진 $W=U_h$, $m=5$, $n=8$ 을 선택하였다.

$$a = (U_h/U_d)^{-\frac{1}{s-1}}, \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}$$

$$\sigma_v = \tan\left(\frac{-\pi}{2k}\right) \left[U_h - 2\ln\left(\frac{\sigma_u^2}{U_h}\right) \right] \left[2\ln 2 - \frac{(2\ln 2)^2 \sigma_u^2}{U_h^2} \right]^{-\frac{1}{2}}$$

인식에 사용된 데이터는 얼굴 표정 변화와 좌우 빛의 변화를 갖는 영상을 포함한 IMM 영상 데이터[5]와 밝기 변화만으로 이루어진 영상을 포함하는 YaleB 영상 데이터[6]와 작은 밝기 변화와 포즈의 변화를 갖는 영상으로 이루어진 ORL 영상 데이터[7]를 사용하였다. 인식을 결과는 표1과 같다.

	원영상 사용시 인식률	제한한 방법의 인식률
IMM 영상	83.33%	85.00%
YaleB 영상	65.00%	72.50%
ORL 영상	53.33%	65.93%

표 1. 각 영상 데이터의 인식률

표1에서 보는 바와 같이 얼굴 표정 변화와 포즈의 변화 그리고 빛의 변화에는 원 영상에 비하여 제안하는 방법이 더욱 강건함을 알 수 있었다. 빛의 변화에 따른 오인식률의 감소는 가

버 웨이블릿 변환의 특성인 빛에 의한 영향이 적은 성질[3]을 통해 알 수 있고, 표정 변화나 포즈의 변화시의 오인식률의 감소는 원 영상에 비해 각 개인의 특징을 더욱 잘 살려주고 개인 간 차이가 큰 영상을 가공하여 사용함으로써 주성분 분석의 결과인 성분값을 가장 가까운 유클리드 거리로 인식할 때 각 개인의 거리 차이를 넓혀주어 작은 변화를 가진 영상의 경우에 정인식 영상의 거리에 더욱 가깝게 함으로 오인식을 줄여주는 역할을 했음을 알 수 있다.

5. 결론

본 논문에서는 원 영상과 얼굴의 중요도를 나타내는 GWT 계수 에너지 영상을 결합하여 얻은 영상을 이용하여 PCA를 사용한 얼굴 인식을 하는 방법을 제시하였다. GWT 계수의 에너지 분포는 사람의 얼굴 중에서 시각적으로 중요한 눈, 코, 입, 윤곽 등을 높은 에너지 분포로 나타낸다. 따라서 본 논문에서는 시각적으로 중요한 부분을 원 영상보다 더욱 큰 차이를 보이게 하여 각 개인의 차이를 크게 하기 위하여 원영상에 GWT 계수 에너지 영상을 1:1 비율로 결합하여 이를 가지고 PCA를 사용하여 인식을 하였다. 결과는 원영상을 이용한 PCA를 사용하는 경우에 비해서 높은 인식율을 보였다. 추후 과제로는 다른 인식 알고리즘에 적용하여 제안한 특징 추출 방법이 적합한지 여부를 밝혀내고자 한다.

6. 참고문헌

- [1] R. Chellappa, C.L. Wilson and S. Sirohey "Human and machine Recognition of faces", a survey. Proc. of the IEEE, 83, 705-750, 1995
- [2] R. Brunelli, T. Poggio, " Face recognition : Feature Versus Templates", IEEE. Trans. on Pattern Analysis and Machin Intelligence, Vol.15, No.10, Jan. 1993
- [3] B.S. Manjunath and W.Y. Ma, "Texture Features for Browsing and Retrieval of Image Data" IEEE Trans. Pattern Anal. Mach. Intelligence, Vol. 18, No. 8, Aug. 1996
- [4] Chao He, Yuan F.Zheng, "Object Tracking Using the Gabor Wavelet Transform and the Golden Section Algorithm" IEEE TRANSACTION ON MULTIMEDIA, VOL.4, NO.4, Dec. 2002
- [5] M. B. Stegmann, "Analysis and segmentation of face images using point annotations and linear subspace techniques", Technical Report IMM_REP_2002_22, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Aug 2002.
- [6] Georghiadis, A.S. and Belhumeur, P.N. and Kriegman, D.J. "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose" IEEE Trans. Pattern Anal. Mach. intelligence, Vol. 23, No. 6, pp. 643-660, Aug 2001.
- [7] Olivetti Research Laboratory "www.uk.research.att.com/facedatabase.html"