

KL 변환과 템플릿매칭을 이용한 얼굴 인식 방법

강환일*, 송영기**, 이세영***, 정요원****

Comparative Study on the Recognition of Face Image Using the KL transform and the Template Matching

Hwan IL Kang*, Young GI Song**, Se Young Lee***, Yowon Jung****

요약

얼굴인식의 방법 중 하나인 전체얼굴에 대한 인식 방법으로, 고유벡터를 이용한 인식 방법과 템플릿 매칭을 이용한 방법의 차이점을 비교 연구한다. 고유벡터를 이용한 방법은 얼굴 영상에 대한 벡터공간을 얻은 후 각 얼굴 영상을 구별할 수 있는 공간에 대한 투영을 통하여 인식에 이용한다. 템플릿 매칭에 기반한 방법은 몇가지 유사도 정의를 이용한 것이다. 또한 얼굴 영상에 대한 전처리 과정이 인식에 끼치는 영향도 분석한다. 본 논문은 두가지의 얼굴 영상 인식기술의 비교를 통하여얼굴 영상의 인식에 대한 유용한 도구로서 에지영상을 이용한 KL변환 방법이 더 우수함을 보인다.

key words: 얼굴인식, KL(Karhunen-Loeve) 변환, 패턴인식, 템플릿 매칭

1. 서론

컴퓨터를 이용한 물체 인식 분야는 현재 대부분 이차원의 데이터를 통하여, 데이터 내부에 존재하는 물체를 적절하게 표현해서 인식하게 된다. 이러한 이차원 데이터의 인식분야 중에서 출입통제, 보안 시스템, 시스템에 의한 자동화 및 네트워크를 통한 인증분야에 활용되고 있는 기술중의 하나가 얼굴인식이다[1]. 인간의 얼굴은 나이, 시간 및 다른 주변의 여건 및 개인적인 감정의 변화등에 의해 그 특성이 달라질 수 있기 때문에 그 이용이 다른 생체 인식(Biometric identification)방법에 비해 제한되기도 하지만 직관적으로 인식할 수 있는 가장 손쉬운 방법이기 때문에 많은 연구가 진행되어 오고 있다. 얼굴의 자동인식 방법은 크게 다음과 같은 두가지 방법으로 이루어진다.

첫째는, 얼굴 내부에 존재하는 눈, 코, 입 및 다른 구성 요소로부터 얻어진 수치적인 데이터를 이용하는 것이다.[2] 이 방법은 수치적인 데이터가 정확히 얻어지고 이 데이터들이 다른 얼굴 사이에 분별력을 지닌다면 아주 효과적인 인식 방법이다. 그러나 실제로 수치적인 특징을 얻기 위해서는 전처리 단계를 거쳐야 하며, 이 전처리에 의해 특징들이 상호 연관이 되어 변별력을 가진 특징을 얻기 어렵다. 또한 궁극적으로 이 수치적인 데이터는 입력 영상의 조건에 의해 많이 좌우가 된다. B. Brunelli[2]는 35개의 기하학적인 특징을 추출하여 얼굴인식에 이용하였다. 두 번째 방법은 얼굴 영상 자체의 강도(Intensity)를 이용하는 방법이다. 이 방법은 템플릿에 기초한 방법으로 얼굴 영상사이의 직접적인 유사도를 이용하는 방법이다.

본 연구는 정보통신부의 정보통신 우수 대학 시범학교 지원에 의해 수행되었습니다. * 명지대학교 정보 제어 공학부 부교수, ** 기아정보시스템 시스템사업본부 연구팀, *** 국립 경상대학교 제어계측공학과 대학원, **** 명지대학교 정보 제어 공학부 대학원 석사과정.

이 템플릿은 얼굴 구성 성분을 이용할 수도 있고 혹은 전체적인 얼굴을 이용할 수 있다. 이 방법은 직관적으로 인간의 인식 방법과 유사하며, 매우 직관적이다. 물론 크기 및 조명등의 영향을 고려하기 위하여 정규화를 한 다음 인식에 이용한다. [3] M. Turk 와 A. Pentland는 얼굴 영상을 인식하기 위해, 이 영상의 저차의 벡터를 이용하여 기술한 뒤, 이들의 선형 결합을 통하여 인식하는 방법을 제안하였다[4]. 실제 입력 얼굴 영상이 손실 없이 입력 되었을 때의 해상도를 유지하기 위해서는 매우 큰 계산량이 필요하기 때문에 이 고유얼굴 (eigenface) 인식방법은 이것을 줄이기에 아주 효과적인 방법이다.

본 논문에서는 얼굴 영상 사이의 직관적인 인식 방법인 템플릿매칭에 기초한 인식방법과 KL변환에 기초한 고유얼굴 인식 방법을 비교함으로써, 얼굴 인식에 유용한 방법을 제시하고자 하는 것이다. 제 2장에서는 고유얼굴 인식 방법에 관해서 기술하고, 3장에서는 템플릿 인식 방법에 대해 설명한다. 제 4장에서는 두가지 방법을 이용한 인식 결과를 제시하고, 제 5장에서는 앞으로 얼굴 인식에 대한 새로운 방향을 제시한다.

2. KL변환에 의한 얼굴 영상 기술

KL(Karhunen-Loeve)변환을 이용한 얼굴 영상의 인식은 전 얼굴 영상내에 존재하는 각 영상을 가장 잘 표현하는 벡터를 찾아내는 것이다. 실제 이 벡터는 전 얼굴 영상내에서 얻어진 각 영상을 기저로 하는 벡터이다. 이 방법을 이용한 얼굴 영상 방법은 학습을 통한 기저 벡터의 생성과 이 기저 벡터에 의해서 투영된 미지 영상과 학습 영상 사이의 유사도를 결정하는 인식의 두 단계로 나뉘어진다.

고유 얼굴의 학습

M 개의 학습 영상 ($N \times N$), $I_k (i=1, \dots, M)$ 를 일차원벡터 ($N^2 \times 1$)로 표시하고 난 후, 이들의 평균 영상을 Ψ 라고 하면,

$$\Psi = \frac{1}{M} \sum_{k=1}^M I_k$$

가 되고, 각 영상은 평균으로부터 다음 값만큼 차이를 지니게 된다.

$$\Phi_k = I_k - \Psi,$$

여기서, $k=1, \dots, M$ 이다. 실제 이 벡터의 크기는 대체로 학습할 데이터의 크기보다 크며, 여기서 우리는 이 데이터의 분포를 가장 잘 표현할 수 있는 M 개의 정규 직교 벡터 u_n 을 찾는 것이다.

k 번째 벡터 u_k 는 다음과 같은 식을 최대화시킬 수 있도록 선택된다.

즉,

$$\lambda_k = \frac{1}{M} \sum_{k=1}^M (u_k^t \Phi_k)^2$$

여기서, t 는 전치관계를 나타내고, 벡터 u_k 와 스칼라 λ_k 는 각각 다음과 같은 공분산 행렬의 고유벡터와 고유치가 된다.

$$C = \frac{1}{M} \sum_{k=1}^M \Phi_k \Phi_k^t \\ = AA^t$$

여기서, 행렬 $A = [\Phi_1, \dots, \Phi_M]$ 이고, 행렬 A 는 각 학습 영상을 열에 포함하고 있다. 그러나 행렬 C 는 $N^2 \times N^2$ 로서 이것의 고유치와 고유벡터를 구하는 것은 계산량이 많이들게 된다.

그러므로, AA^T 대신에 차원이 학습 영상의 갯수와 일치하는 $A^T A$ 를 구함으로써 계산량을 줄이고 저차의 고유벡터를 구해 낼 수 있다.

$A^T A$ 의 고유 벡터를 v_i 라 하면,

$$A^T A v_i = \mu_i v_i$$

가 된다. 여기에 앞의 위치에 A 를 곱하면

$$AA^T A v_i = \mu_i A v_i$$

가 되고, 따라서 $A v_i$ 는 앞의 식으로부터 AA^T 의 고유 벡터가 되므로,

$$u_k = A v_k = [\Phi_1, \dots, \Phi_M] v_k$$

가 된다. 이렇게 학습 단계에서 구해진 고유치에 대한 고유벡터들은 각 클래스사이에 존재하는 여러 가지 변형들을 가장 잘 표현할 수 있도록 정렬된다.

고유 얼굴의 인식

학습에서 얻어진 M 개의 고유치에 해당하는 고유 벡터에 의해서 표현된 고유얼굴을 가지고 인식에 이용한다. 실제 인식에서는 구해진 고유치중 그 값이 큰 것들로 정렬했을 때, 그 중 상위 몇 개의 값만 선택해서, 인식에 이용한다. Kirby와 Sirovich는 115개의 학습 영상으로부터 40개의 고유 얼굴만 이용했으며 M. Turk와 A. Pentland는 16개의 학습 영상에서 7개의 고유얼굴만 이용하였다. D. Swets와 J. Weng[4]은 사용하는 고유치와 사용하지 않은 고유치를 다음과 같은 식처럼 고정된 백분율을 만족하도록 $n < M$ 을 선정하였다. 즉, 사용하지 않은

고유치의 합을 $\sum_{k=n+1}^M \lambda_k$ 라고 했을 때, 다음 식을 만족하도록 n 의 값을 선정하였다.

$$\frac{\sum_{k=n+1}^m \lambda_k}{\sum_{k=1}^m \lambda_k} < P$$

D. Swets와 J. Weng은 $P=5\%$ 가 되도록 선정 하였다. 그러면 선정된 L 개의 고유치에 해당하는 고유 얼굴을 통하여 얼굴 영상은 다음과 같이 이루어진다. 고유값이 큰 순서대로 L 개의 u_k 를 선택해서 학습에 이용한 각 클래스의 얼굴 영상을 다음과 같은 변환을 통하여 소위, 얼굴 영역으로 투영시킨다.

$$\omega_k = u_k^t(I - \Psi)$$

여기서, $k=1, \dots, L$. 투영된 값은 얼굴 영역에서 각 클래스를 규정하는 특징 벡터가 된다. 즉,

$$\Omega^t = [\omega_1, \dots, \omega_L]$$

인식은 만일 미지의 얼굴 X 가 입력되었을 때, 먼저,

$$\hat{X} = X - \Psi$$

를 구한 다음, 이 값을 고유 얼굴 u 를 통하여 얼굴 영역으로 투영시킨다.

$$\Omega_x = [u_1 \hat{X}, \dots, u_L \hat{X}]^t$$

이 투영된 가중치 벡터 Ω_x 를 학습단계에서 구해진 가중치 벡터와의 유사도 즉, 유클리디안 거리를 이용하여 인식에 이용한다.

3. 템플릿매칭을 이용한 얼굴 인식

템플릿 매칭은 관심있는 대상(object)에 대하여 그것에 대한 템프릿(원형)을 작은 부영상이나 전체적으로 표현하여, 이를 주어진 영상 내에서 이동시키면서, 겹치는 부분의 유사도를 조사하여 그 유사도 값으로 대상체의 유, 무 및 위치를 구하는 방법이다. 이 방법은 직관적으로 두영상 사이의 유사도를 측정하기 때문에, 영상 내에서 발생하는 특이적인 변이들(크기, 방향, 조명)에 민감한 성질을 지니고 있다. 얼굴 인식은 얼굴 영역의 탐색 후 분리를 통하여 이루어지므로 초기단계에서 이들의 영향을 최소화하는 작업이 요구된다. 템플릿 매칭기법은 템플릿 영상을 $T(x, y)$ 라고 하고 압력 영상을 $f(x, y)$ 라고 했을 때, 입력 영상내의 점 (i, j) 에 그 중심이 겹치게 하고, $T(x, y)$ 와 겹치는 영상 광의 겹치는 부분패턴(sub image block)과의 유사도를 측정해서 미리 정한 임계값을 만족시키는가를

판단해서 (i, j) 의 위치에 대상이 존재하게 되는 것을 보여주게 된다. 이 유사도 측정은 대부분 상관 관계값에 근거하며, 동일 영상에서 계조치의 차이에 의해 발생하는 오인식 경우를 제거하기 위해 다음과 같은 정규화된 상관 관계를 이용한다.

$$M(u, v) = \frac{\sum [T - \hat{T}] * [f - \hat{f}]}{\sqrt{\sum [T - \hat{T}]^2 \sum [f - \hat{f}]^2}}$$

이 된다. 여기서 \sum 은 이중 합을 나타내고 ‘ $\hat{\quad}$ ’은 평균 값을 나타낸다.

얼굴 인식을 위한 템플릿 매칭 각 구성요소의 템플릿을 먼저 자동으로 만들고 난 뒤 다음과 같은 절차를 통하여 이루어진다.

투영에 의한 얼굴 구성 요소의 위치 탐색

먼저 sobel연산자를 이용하여 에지 영상을 만든다. 이 에지 영상을 얻은 후 임계값 설정은 다음과 같은 LAT(Locally Adaptive Thresholding)을 이용한다. 즉,

$$LAT = \frac{\max\{|M_k|, k=0, 1, 2, 3\}}{F_0(I(x, y))}$$

여기서, M_k 는 4방향 에지 연산자를 각각 적용하여 얻어지며, 저주파 필터 F_0 는 영상 $I(x, y)$ 를 필터링한 값이다.

이렇게 한 후 끊어진 에지의 연결을 위하여 모폴로지 연산을 이용하였고 수직 수평 방향의 투영을 통하여 먼저 코와 입의 위치를 결정하게 된다. 이렇게 해서 템플릿 매칭에 사용할 템플릿은 두 눈과 코 입을 이용하였다. 그림 1은 사용한 이미지의 템플릿을 보여주고 있다.

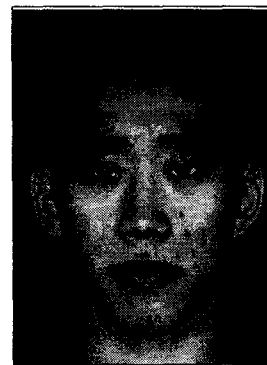


그림 1: 원 얼굴 영상

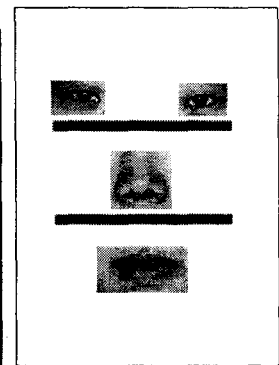


그림 2: 템플릿 영상

이렇게 해서 구해진 템플릿을 이용하여 최종 인식은 다음과 같이 수행된다.

- 방법1. 4개의 템플릿의 유사도를 합하는 방법
- 방법2. 가중 합을 이용하는 방법
- 방법3. 최대값을 나타내는 템플릿으로 결정
- 방법4. 하이브리드 방법

4. 얼굴 영상 인식 실험 결과

각 26명의 사람에 대하여 4개의 영상을 256×256로 만들어 각각 실험에 이용하였다. 실험에 이용한 얼굴은 정면 얼굴이면, 영상에 따라 약간의 표정 변화와 조명의 밝기가 다르다.

실험 1.

실험 1에서는 고유얼굴을 통한 인식 결과를 제시한다. 그림 3은 전체 얼굴에 대한 평균 영상을 나타낸다. 이 그림은 입력 영상에 아무런 처리없이 얻어진다.

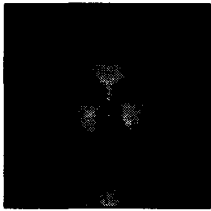


그림 3: 평균 영상

인식을 위하여 본 논문에서 2장에서 기술 한 대로 $M=10$ 이 되고 여기서 L 값의 선택은 각 고유치의 영향에 따라 선정되었다. 그림 4에서는 가장 큰 고유치에 해당하는 고유 얼굴을 나타내었다.



그림 4: 고유 얼굴

본 논문에서는 원 영상 및 에지 영상에 대하여 인식을 수행하였다. 인식시 $L=3$ 으로 선정했을 때, 각 영상은 실제 입력 영상의 차수가 큰 것에서 3개의 얼굴 영역으로 투영이 된다. 원영상을 이용했을 경우는 상위 3개의 고유치가 전체의 99% 이상을 차지하고 있다. 인식은 거리에 기반을 둔 유클리디안 거리를 유사도로 이용하였다. 가장 지배적인 특성을 지닌 고유값으로 인식 했을 때, 90% 즉 9개의 영상이 정확하게 인식되었으며, 3개의 고유치를 이용했을 경우에도 같은 인식 특성을 보였다.



그림5:
에지평균영상

에지 영상에 대한 인식은 먼저 입력 영상을 sobel 연산자를 통하여 에지를 추출한 다음 위와 같은 절

차를 반복적으로 적용하여 인식한다. 에지 영상은 상위 9개가 99% 이상의 값을 차지하고 있었다. 이것은 곧 에지 영상이 더 변별력을 지니는 것으로 판단되며, 인식 또한 9차의 가중치로 표현되며, 인식률은 100%를 나타내었다.

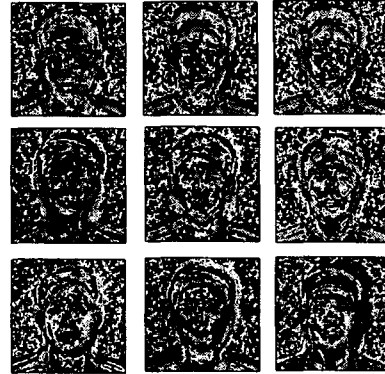


그림6: 에지 영상에서의 고유 얼굴

실험 2.

실험 2에서는 템플릿 매칭에 대한 결과를 제시한다. 학습 단계에서는 각 얼굴 영상마다 4개의 템플릿(두 눈, 코, 입)을 준비한다. 템플릿의 크기는 눈의 경우 대략적으로 30×20이고 코는 40×50이고 입은 50×30이하의 크기를 사용하였다. 이렇게 각 구성요소를 만들어 템플릿을 이용했을 경우 전체 영상을 이용했을 경우보다 10배정도 시간이 단축되었다. 얼굴에서 구성 요소는 전술한 방법에 의해 수행했을 때 영상에서 얼굴 구성 요소를 정확하게 찾지 못하는 경우도 발생하였다. 템플릿 매칭을 이용한 얼굴 인식에서는 총 템플릿이 있는 20개의 영상과 그렇지 못한 6개를 통하여 인식하였다. 각 템플릿에 의해 구해진 정합 결과중 최고의 값으로 인식 하였을 경우, 인식율은 97.78%였으며, 오인식률은 2.22%였다. 또한 템플릿이 존재하지 않는 6개의 영상에 대해서는 거부를 나타내었다. 여기서 사용한 거부 기준은

$$RC = \frac{O_1 - O_2}{O_1 + O_2}$$

를 거부기준으로 선정하였다. 여기서 O_1, O_2 는 Correlation의 결과로써, 얻어지는 출력 값들 중에서 가장 큰 출력 값과 두 번째 큰 출력값을 나타낸다. 본 논문에서는 이 값을 0.8로 정하였다.

5. 결론

본 논문에서 얼굴 영상의 기하학적인 특징벡터를 추출하지 않고 얼굴영상내의 데이터 분포를 이용하여 직관적으로 인식하는 방법인 고유 얼굴 방법과 템플릿 매칭에 기반한 얼굴 인식 방법에 대해서 기술하였다. 두 가지 방법은 복잡한 전처리 과정이 필요 없고 직관적이며, 계산이 간단하다. 또한 대체로 인식을 또한 좋은 편이다. 하지만 얼굴 영상

에서 존재하는 얼굴의 특징적인 변이(크기, 기울어짐, 조명)등의 영향에는 민감한 특성을 보인다. 그러므로 본 논문에서는 에지영상을 고유 얼굴로 표현하여 인식에 이용하여 좋은 인식률을 보였다. 앞으로 더 많은 데이터를 이용하여, 제한이 덜해진 환경에서 얻어진 얼굴영상을 인식할 수 있는 방법을 연구할 계획이다. 본 논문에서 이용한 방법들을 기준으로 앞으로 장면사진 내에서 얼굴 영역의 탐색과 인식을 위한 새로운 방법을 개발할 예정이며, 각 얼굴 내에 존재하는 구성 요소들을 이용한 가변 구조 모델에 대한 연구를 진행할 계획이다.

참고 문헌

- [1] 강영미, 특징 영역 기반의 얼굴 인식, 창원대학교 전자 계산 학과
- [2] R. Brunelli and T. Poggio, "face recognition : Feature Vs Template", IEEE Trans. on PAMI, Vol.15, No.10, pp. 1042-1052, 1993
- [3] M. Turk and A. Pentland, "eigen face For recognition", J. of Cognitive NeuroScience, vol.3, No.1, pp. 71-86, 1991
- [4] D.L. Swets and J. Weng, "SHOSLIF object recognition and Image retrieval", TR-cps-95-39, Michigan State University.
- [5] A.K. Jain, Fundamentals of Digital Image processing, Prentice-hall, 1989
- [6] A. Crouzil et al., " A new correlation criterion based on the gradient fields similarity", Proceedings of ICPR 96, pp. 632-635.