
SIFT-Grid를 사용한 향상된 얼굴 인식 방법

김성훈*, 김형호**, 이현수***

An Improved Face Recognition Method Using SIFT-Grid

Sung Hoon Kim*, Hyung Ho Kim**, Hyon Soo Lee***

요약 본 논문은 SIFT-Grid 기반의 얼굴 인식 시스템에서 식별 능력의 향상과 계산량 감소를 목적으로 한다. 첫 번째는 한 얼굴 클래스의 다양한 훈련 이미지로부터 비슷한 SIFT 특징점들은 제거하고, 상이한 특징점들은 병합하는 통합템플릿의 구성 방법을 제안한다. 통합템플릿은 SIFT-Grid를 통해 나누어진 훈련 이미지들의 동일 부분영역 내의 특징점들에 대한 유사도 행렬의 계산과 임계치 기반의 히스토그램의 계산을 통해 구성하였다. 두 번째는 구성된 통합 템플릿들로부터 테스트 이미지의 효과적인 식별을 위한 유사도 계산 방법을 제안한다. 유사도의 계산은 테스트 이미지와 각 클래스의 통합템플릿간의 일대일 비교로 수행된다. 이때 동일 부분영역 별로 유사도 점수와 임계치 기반의 보팅 점수가 계산된다. 얼굴 인식 작업에 대한 실험 결과 제안된 방법이 SIFT-Grid 기반의 다른 두 방법보다 정확한 것으로 확인 되었고, 또한 계산량도 감소하였다.

주제어 : 얼굴 인식, 패턴 인식, 템플릿 매칭

Abstract The aim of this paper is the improvement of identification performance and the reduction of computational quantities in the face recognition system based on SIFT-Grid. Firstly, we propose a composition method of integrated template by removing similar SIFT keypoints and blending different keypoints in variety training images of one face class. The integrated template is made up of computation of similarity matrix and threshold-based histogram from keypoints in a same sub-region which divided by applying SIFT-Grid of training images. Secondly, we propose a computation method of similarity for identify of test image from composed integrated templates efficiently. The computation of similarity is performed that a test image to compare one-on-one with the integrated template of each face class. Then, a similarity score and a threshold-voting score calculates according to each sub-region. In the experimental results of face recognition tasks, the proposed methods is founded to be more accurate than both two other methods based on SIFT-Grid, also the computational quantities are reduce.

Key Words : Face Recognition, Pattern Recognition, Template Matching

1. 서론

얼굴 인식 시스템은 IT 기술의 발전으로 정보 보안, 사람과 컴퓨터간의 상호작용(Human Computer Interaction) 그리고 상황 인식 컴퓨팅 (Context Aware Computing)등의 다양한 응용 분야에서 활용되고 있다. 하지만 얼굴은 조명, 포즈, 표정의 변화, 장신구의 착용

그리고 시간의 변화에 따른 노화에 의해 다양한 형태로 변하기 때문에 얼굴인식을 통한 개인의 식별은 어려운 문제이다. 그러므로 얼굴 인식 시스템의 성능 향상을 위해서는 개인 얼굴의 변화에 강인한 특징을 추출하여야 한다.

대표적인 특징 추출 방법은 주성분분석법(PCA : Principal Component Analysis)과 선형판별분석법(LDA

*경희대학교 컴퓨터공학과 박사

**세한대학교 컴퓨터공학과 교수

***경희대학교 컴퓨터공학과 교수(교신저자)

논문접수: 2013년 1월 25일, 1차 수정을 거쳐, 심사완료: 2013년 2월 20일, 확정일: 2013년 2월 20일

: Linear Discriminant Analysis)으로 선형변환행렬을 사용하여 조건에 만족하는 특징들을 추출한다. PCA는 학습 얼굴 이미지들로부터 얼굴공간을 구성하는 기저벡터인 고유얼굴(eigenface)을 계산하여 얼굴 인식에 사용된다[1][10]. LDA는 학습 얼굴 이미지들로부터 클래스 내부의 분산은 줄이고 클래스 간의 분산은 증가시키는 방향으로 입력 특징 값을 변화시키는 행렬을 계산하여 얼굴 인식에 적용한다[5][11]. 또한 PCA나 LDA의 변형된 알고리즘에 대한 연구를 통해 얼굴 인식 문제를 해결하고 있다[15][14]. 하지만 PCA또는 LDA로부터 검출된 특징들은 빛의 변화와 얼굴의 표정 변화에 민감하기 때문에 얼굴의 인식 성능이 낮다. 또한 시스템의 입력은 훈련 이미지와 동일하게 얼굴 크기 및 방향 등의 조건이 일치하여야 정확한 인식 결과를 획득할 수 있다.

SIFT(Scale Invariant Feature Transform) 알고리즘은 객체 인식, 중복성 검출 및 객체 추적 등의 다양한 응용분야에서 우수한 성능을 보이고 있다[3][7][12]. Lowe에 의해 제안된 SIFT 기반의 객체인식 방법은 특징 추출과 특징점(keypoint) 매칭의 두 단계로 구성된다[4]. 먼저 특징 추출 단계는 이미지의 다양한 해상도 상에서 독특하고 안정적인 지역적 특징점들을 검출한 후, 특징점을 크기와 회전에 불변하는 기술자(descriptor)로 표현한다. 매칭 단계에서는 각 훈련 이미지로부터 추출된 특징점의 데이터베이스 상에서 테스트 이미지의 각 특징점과 가장 일치되는 특징점 쌍을 찾음으로써 수행된다. 이 매칭 방법은 특징점의 기술자에 대한 일치만을 고려하고, 지역적인 위치에 대한 일치는 고려하지 않는다. 따라서 일반적인 객체와 비교해 비강체이며, 매끈하고 대칭적인 얼굴 영역을 대상으로 Lowe의 매칭 방법을 적용할 경우 지역적인 불일치가 발생되며, 이는 인식 성능이 낮아지는 원인이 된다.

최근 얼굴 인식에서 특징점의 지역적인 불일치 문제를 해결하기 위한 연구들이 진행되고 있다[2][6][8][9][13]. [2][6][13]의 논문에서는 두 눈과 입을 기준으로 얼굴 영역을 분할한 후, 각 분할영역내의 특징점들을 대상으로 매칭 과정을 수행함으로써 지역적인 불일치 문제를 해결하였다. 하지만 이 방법들은 눈과 입의 검출 작업이 추가적으로 수행되기 때문에 얼굴 인식 과정에서 계산량이 증가되며, 눈과 얼굴 검출기의 정밀도가 얼굴 인식의 성능에 많은 영향을 준다. 반면 SIFT-Grid 방법을 사용한 [8]와 [9]의 논문에서는 얼굴 영역의 가로와 세로를 특정 비율로

나누어 분할영역들을 결정하고 해당 영역에 대해 매칭을 수행함으로써 지역적인 불일치 문제를 해결하였다.

SIFT 기반의 얼굴 인식 시스템에 있어 또 다른 문제는 매칭 과정에서 테스트 이미지의 각 특징점은 훈련 데이터베이스의 모든 특징점을 대상으로 유사성의 비교 연산이 수행되기 때문에 많은 계산량이 필요하다. 따라서 동일 객체의 다양한 얼굴 이미지로부터 유사한 특징점들은 제거하고 상이한 특징점들의 병합을 통해 템플릿을 구성함으로써 매칭 과정의 계산량은 감소시키고, 인식율은 향상시킬 수 있다.

본 논문은 훈련 얼굴 이미지들로부터 특징점들에 대한 통합템플릿의 구성 방법을 SIFT-Grid 기반의 얼굴 인식 시스템에 추가하여 인식 성능의 향상을 목적으로 한다. 제안된 방법은 먼저 SIFT-Grid 방법을 통해 훈련 이미지들의 얼굴영역을 몇 개의 분할영역으로 나누고, 각 분할영역을 구성하는 특징점들 간의 유사성을 비교 분석하여 통합템플릿을 생성하였다. 또한 통합템플릿으로부터 테스트 이미지의 효과적인 식별을 위한 유사도 계산 방법을 제안한다. SIFT-Grid 기반의 제안된 얼굴 인식 시스템의 유효성을 확인하기 위해 다양한 실험을 수행한 결과 성능 향상을 확인할 수 있었다.

2. 관련 연구

2.1 SIFT 알고리즘

Lowe에 의해 제안된 SIFT는 크기, 회전 그리고 빛의 변화에 강인한 지역적인 특징점을 검출하는 알고리즘으로 최근 영상 처리 및 패턴 인식 분야에서 많이 적용되고 있다. SIFT는 4단계를 통해 특징점들의 특징을 정의한다 [4]. 먼저 크기와 회전 변화에 강인한 특징점의 위치를 검출하기위해 입력 이미지의 다양한 해상도를 대상으로 DoG (Difference-of- Gaussian) 이미지들을 생성한다. 이때 DoG 이미지는 식 (1)과 같이 각 해상도에 해당하는 가우시안 함수를 이미지와 컨볼루션한 것의 차로 표현되며, 가우시안 함수는 식 (2)와 같다.

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \quad (1)$$

$$= L(x,y,k\sigma) - L(x,y,\sigma)$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2)$$

다음 단계에서는 각 해상도의 DoG 이미지에서 모든 픽셀의 값을 이웃하는 8 픽셀 그리고 이전 및 이후 해상도에서 이웃하는 각 9 픽셀과 비교하여 지역적으로 최소 또는 최대값인 경우 특징점의 후보로 선택한다. 세 번째 단계에서는 명암 대비가 낮은 후보 특징점과 모서리가 아닌 에지(edge) 위치의 특징점을 제거하기 위해 각각 DoG 값의 임계치와 Hessian 행렬 기반의 비율을 사용한다. 마지막 단계에서는 gradient 벡터의 크기(magnitude)와 방향(orientation)을 계산한 후 특징점의 기술자를 정의한다. 크기 $m(x,y)$ 와 방향 $\theta(x,y)$ 은 특징점 픽셀과 인접한 픽셀들로부터 식 (3)과 식 (4)와 같이 계산된다.

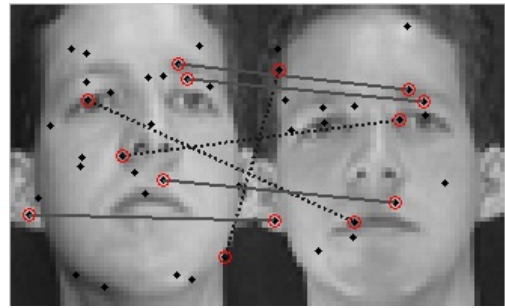
$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (3)$$

$$\theta(x,y) = \tan^{-1}((L(x+1,y) - L(x-1,y)) / (L(x,y+1) - L(x,y-1))) \quad (4)$$

특징점의 기술자는 16×16 픽셀의 인접 영역에 대해 가장 빈도수가 높은 각도를 기준으로 방향을 정규화한 후 인접 영역을 4×4 의 영역으로 나누어 각 영역의 방향들을 8의 빈을 가지는 히스토그램으로 표현한다. 따라서 SIFT 특징점은 위치, 크기, 방향 그리고 128 차원의 기술자로 구성된다.

2.2 SIFT 기반의 이미지 매칭 방법

Lowe가 제안한 이미지 매칭 방법[4]은 두 이미지의 특징점들의 기술자간의 유클리드 거리를 계산하고, 한 특징점에 대응되는 첫 번째와 두 번째로 가까운 특징점들의 거리에 대한 비율을 확인하여 매칭된 특징점 쌍을 결정한다. 그리고 특징점 쌍들 간의 기하학적 일치성을 확인한 후 두 이미지간의 유사도를 계산한다. 그림 1은 두 개의 얼굴 이미지를 대상으로 Lowe의 매칭 방법을 수행한 결과를 나타내고 있다. 그림 1에서 점은 각 특징점의 위치를 표시한 것이고, 실선과 점선은 매칭된 특징점 쌍을 의미한다. 이중 점선으로 연결된 쌍은 지역적인 불일치를 나타낸다. 이러한 지역적 불일치 쌍들은 기하학적 일치성의 확인 과정에서 제거되며, 동일 얼굴의 이미지에도 불구하고 낮은 유사도 점수가 결정되어 오인식의 원인이 될 수 있다.



[그림 1] Lowe의 이미지 매칭 결과

Bicego는 이와 같은 지역적 불일치 문제를 해결하기 위해 이미지를 특정 비율의 중첩된 부분영역들로 나누는 SIFT-Grid 방법을 제안하였다[8]. 그리고 두 이미지간의 유사도는 먼저 각 부분영역에 대응되는 특징점의 기술자간의 유클리드 거리를 계산하여 최소값을 부분영역의 거리로 결정하고, 모든 부분영역 거리의 합에 대한 평균을 두 이미지의 유사도 점수로 결정한다. 얼굴 인식은 테스트 이미지와 훈련 데이터의 모든 이미지간의 유사도 점수를 계산한 후 가장 작은 점수를 획득한 훈련 이미지의 클래스로 결정된다.

Lowe와 Bicego는 테스트 이미지와 훈련 이미지 각각에 대해 유사도 점수를 계산하여 테스트 이미지의 얼굴을 식별하기 때문에 유사도 점수에 얼굴 클래스 전체에 대한 특징정보가 반영되지 않는다. [9]의 논문은 이미지의 각 분할영역의 특징점과 전체 훈련이미지의 동일 분할영역에 포함된 특징점들 중 가장 비슷한 특징점을 가진 얼굴 클래스를 찾고, 해당 클래스의 유사도 점수를 증가시키는 특징점 보팅(Keypoints Voting)방법을 제안하여 인식율을 향상시켰다.

본 절에서 살펴본 매칭 방법들은 테스트 이미지의 모든 훈련 이미지에 포함된 특징점들 간의 유사연산이 반복적으로 수행되기 때문에 많은 계산량이 요구된다. 그러므로 훈련 데이터에 포함된 특징점의 수는 감소시키면서 얼굴 인식율은 향상될 수 있는 방법이 필요하다.

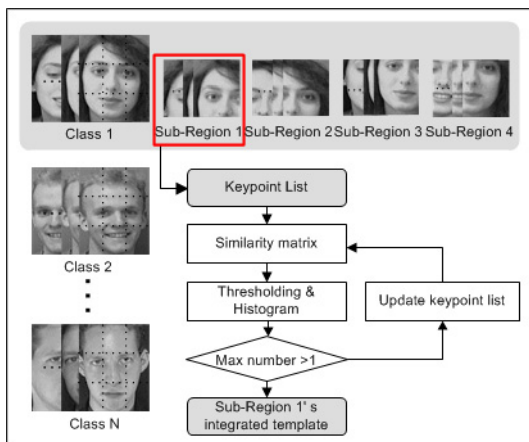
3. 제안하는 얼굴 인식 시스템

본 장에서는 얼굴 인식의 매칭 과정에서 발생하는 계산량의 감소를 위해 제안된 통합템플릿의 구성 방법과 통합템플릿으로부터 테스트 이미지의 효과적인 식별을

위해 제안된 유사도 계산 방법에 대해 기술하였다.

3.1 통합템플릿 구성

제안된 방법은 먼저 인식을 위한 N개의 얼굴 클래스를 대상으로 동일한 해상도를 갖는 얼굴 이미지들을 수집하여 훈련 데이터를 생성하고 모든 이미지의 SIFT 특징점을 검출한다. 그리고 각 훈련 이미지는 가로와 세로를 각각 동일한 비율로 분할한 후 K개의 중첩된 부분영역들을 결정한다. 한 얼굴 클래스의 통합템플릿은 동일 부분영역의 특징점을 대상으로 비슷한 특징점은 제거하고, 상이한 특징점들을 병합함으로써 획득된다. 따라서 하나의 통합템플릿은 부분영역의 수와 동일한 K개의 병합된 특징점 리스트로 구성된다. 그림 2는 N개로 구성된 훈련 집합 중 1번 얼굴 클래스의 통합템플릿을 획득하기 위한 부분영역의 특징점 병합 과정의 예를 나타내고 있다. 이때 부분영역을 결정하기 위한 가로와 세로의 분할 비율은 1/3을 사용하였다. 부분영역 특징점 병합 과정은 유사도 행렬 계산, 임계치를 통한 히스토그램 계산 그리고 특징점 리스트 갱신인 3단계로 구성된다.



[그림 2] 제안된 통합템플릿의 구성을 위한 부분영역 특징점 병합 알고리즘

n번째 클래스의 부분영역 k에 포함된 M개의 현재 ($t=0$) 특징점 리스트는 식 (5)와 같이 정의되며, f_m 은 m번째 특징점의 기술자이다.

$$L_n^k(t) = (f_1, f_2, f_3, \dots, f_m) \quad (5)$$

첫 번째 단계는 특징점 리스트 $L_n^k(t)$ 의 i 와 j 번째 기

술자간의 거리를 계산하여 유사도 행렬을 획득한다. 유사도 행렬 s 는 식 (6)과 같으며, $d(f_i, f_j)$ 는 두 특징점의 기술자에 대한 유클리드 거리를 의미한다.

$$S_{i,j} = d(f_i, f_j), i = 1, 2, \dots, m \quad j = 1, 2, \dots, m \quad (6)$$

다음 단계에서는 빈도수가 높은 비슷한 특징점들을 찾기 위해 식 (7)과 같이 유사도 행렬 $S_{i,j}$ 에 비슷한 정도에 대한 임계치 T_d 를 이용하여 이진 유사도 행렬 $BS_{i,j}$ 을 획득한다. 그리고 이진 유사도 행렬의 각 열에 대해 식 (8)을 적용하여 히스토그램 H 을 계산하고, H 중에서 최대값을 H_{\max} 에 기록한다. 또한 H_{\max} 에 속하는 i 열의 요소 값에 대해 $BS_{i,j} = 1$ 인 j 인덱스를 H_{\max}^{list} 에 기록한다.

$$BS_{i,j} = \begin{cases} 1 & \text{if } S_{i,j} \leq T_d \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$H_i = \sum_{j=1}^M BS_{i,j} \quad (8)$$

이때 H_{\max} 가 1보다 큰 경우 특징점 리스트 갱신 단계를 수행한다. 갱신 단계는 H_{\max}^{list} 에 포함된 비슷한 특징점들을 $L_n^k(t)$ 로부터 제거하여 새로운 특징점 리스트 $L_n^k(t+1)$ 을 생성한다. 그리고 $L_n^k(t)$ 로부터 H_{\max}^{list} 에 포함된 특징점의 기술자에 대한 평균을 계산하여 $L_n^k(t+1)$ 에 추가한다. 이와 반대로 H_{\max} 가 1보다 작은 경우 현재 특징점 리스트 $L_n^k(t)$ 을 n번째 클래스의 부분영역 k의 병합된 특징점 리스트로 결정한다.

모든 부분영역에 대한 특징점 병합 과정의 수행을 통해 n번째 얼굴 클래스의 통합템플릿 T_n 은 식 (9)와 같이 정의되며, 식에서 K 는 부분영역의 수를 그리고 $M_{n,k}$ 는 k번째 부분영역의 병합된 특징점의 수를 의미한다.

$$T_n = \{f_{n,m}^k | k = 1, 2, \dots, K, m = 1, 2, \dots, M_{n,k}\} \quad (9)$$

3.2 유사도 계산 방법

얼굴은 동일한 지역적 특징들로 구성되기 때문에 다른 객체간의 동일한 부분영역에서 비슷한 SIFT 특징점들이 검출될 수 있으며, 같은 객체의 다른 이미지 상에서 SIFT 특징점들이 표정, 빛 등의 영향으로 다른 특성을 가지고 나타날 수 있다. 따라서 테스트 이미지의 한 특징점에 대한 유사한 얼굴 클래스를 결정하기 위해 모든 훈

런 이미지의 동일 부분영역의 특징점들의 거리비교를 수행하는 [9]의 특징점 보팅의 결과가 동일 객체보다 다른 객체의 클래스로 판정될 수 있다.

본 논문에서는 각 클래스의 통합된 템플릿과 테스트 이미지를 일대일로 비교 분석하여 유사도 점수와 보팅 점수를 계산한다. 테스트 이미지의 k 부분영역에 존재하는 p 번째 특징점 $f_{test,p}^k$ 와 n 번째 클래스의 통합템플릿 T_n 의 k 부분영역간의 유사도 점수 $s_{n,p}^k$ 은 식 (10)과 같이 유클리드 거리의 계산 결과들로부터 최소값을 선택하여 결정한다.

$$s_{n,p}^k = \min(d(f_{test,p}^k, f_{n,m}^k)), m = 1, 2, \dots, M_{n,k} \quad (10)$$

테스트 이미지와 n 번째 클래스의 유사도 점수 S_n 은 식 (11)과 같이 모든 부분영역의 특징점들의 유사도 점수 $s_{n,p}^k$ 가 유사도 임계치 T_s 보다 작거나 같은 경우 누적하여 계산한다. 이때 유사도 임계치를 만족하는 특징점의 수를 카운팅하는 유사도 보팅 점수 VS_n 는 동시에 계산된다.

$$S_n = \sum_{k=1}^K \sum_{m=1}^{M_{n,k}} s_{n,p}^k, \text{ if } s_{n,p}^k \leq T_s \quad (11)$$

테스트 이미지에 대한 식별은 가장 큰 유사도 보팅 점수 VS_n 을 획득한 얼굴 클래스로 결정하며, 만약 최대 유사도 보팅 점수를 가지는 클래스가 둘이상인 경우 이들 중에서 유사도 점수 S_n 중 가장 작은 값을 가진 얼굴 클래스로 결정한다.

4. 실험 및 분석

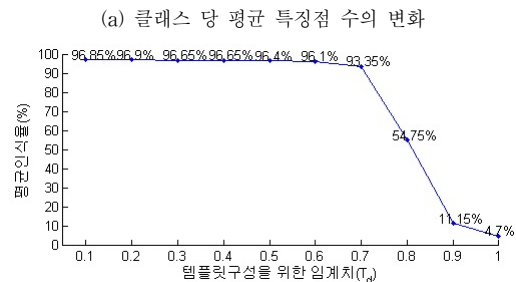
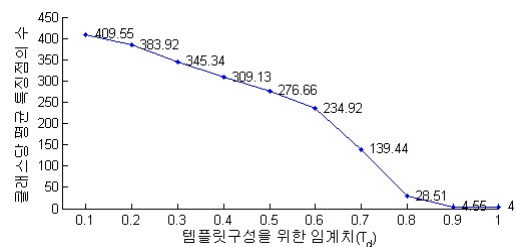
본 연구의 첫 번째 실험은 제안된 얼굴 인식 시스템에서 최적의 성능을 획득하기 위한 것으로 통합템플릿의 구성을 위한 임계치 T_d 와 테스트 이미지의 식별을 위한 임계치 T_s 의 값을 결정하기위해 수행하였다. 다음은 통합템플릿의 구성을 위한 훈련 이미지의 수에 따른 성능을 평가하였다. 그리고 기존의 다른 방법들과의 얼굴 인식 성능 비교 결과를 기술하였고, 마지막으로 다양한 해상도의 얼굴 이미지를 대상으로 실험 결과를 확인하였다. 실험에서 모든 SIFT-Grid 기반의 얼굴 인식 시스템은

가로와 세로의 분할비율을 1/3로하고, 4개의 부분영역을 사용하였다.

실험은 40명의 사람을 대상으로 총 400개의 이미지로 구성된 ORL 얼굴 데이터베이스[16]을 사용하였다. 그리고 각 사람 클래스는 1부터 10까지의 인덱스를 가진 92x112 픽셀의 다른 이미지들로 구성되어있다.

4.1 최적의 임계치 결정 실험

제안된 얼굴 인식 시스템에 영향을 주는 두 개의 임계치 T_d 와 T_s 의 최적화된 값을 결정하기 위해 다음과 같은 실험을 수행하였다. 먼저 식별 과정에서 사용되는 임계치 T_s 을 특정 값으로 고정하고, 통합템플릿 구성에 사용되는 임계치 T_d 을 0.1부터 1까지 0.1씩 증가시키면서 템플릿의 구성에 사용된 특징점의 수와 인식율의 변화를 확인하여 최적의 T_d 값을 결정하였다. 이후 최적의 T_d 값을 기준으로 임계치 T_s 에 다양한 값을 적용하여 인식율의 변화를 확인하였다. 실험은 ORL 데이터베이스의 얼굴 이미지를 60x74 픽셀로 축소하여 사용하였다. 훈련 과정은 10개의 인덱스 중 5개의 번호를 랜덤하게 선택하여 각 얼굴 클래스의 훈련 이미지로 사용하여 통합템플릿들을 구성하였다. 그리고 훈련 이미지로 포함되지 않은 200개의 얼굴 이미지들을 대상으로 구성된 통합템플릿에 대한 얼굴 인식의 성능을 평가하였다. 이와 같은 통합템플릿의 구성과 얼굴 인식의 성능 평가를 각기 다른 랜덤 번호를 대상으로 10회 수행하였다.

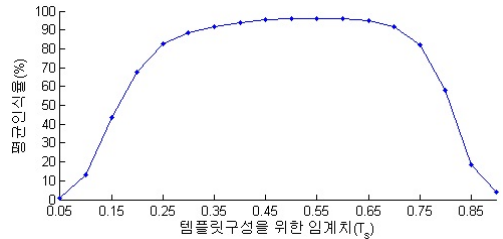


[그림 3] 임계치 $T_s = 0.5$ 일 때 T_d 의 변화에 따른 성능 평가

그림 3은 식별 과정에서 사용되는 임계치 T_s 을 0.5로 고정하고, 임계치 T_d 값의 변화에 따른 제안된 시스템의 성능 평가 결과를 보이고 있다. 그림 3 (a)는 한 클래스의 통합템플릿 구성에 필요한 평균 특징점의 수를 나타내고 있으며, 그림 3 (b)는 10회 반복 실험에 따른 평균 인식율의 변화를 나타내고 있다. 그림 3 (a)을 살펴보면 T_d 가 0.1에서 0.6으로 변경되는 동안 특징점의 수는 평균 35개씩 감소된 반면 0.6에서 0.8사이에서는 평균 103개씩 급격하게 감소됨을 볼 수 있다. 평균 인식율의 변화인 그림 3 (b)을 보면, 특징점 수의 변화가 적은 T_d 의 0.1과 0.6의 범위에서 96.9%와 96.1% 사이의 인식율이 획득되었으며, 이후의 T_d 에서는 인식율의 낮아지는 폭이 큰 것을 볼 수 있다. 그러므로 임계치 T_d 의 값을 0.1과 0.6의 범위에서 결정할 때 제안된 시스템으로부터 최적의 성능을 획득할 수 있다. 이때 단지 인식율만을 고려할 경우 96.9%가 획득 되는 0.2을 T_d 의 값으로 결정하는 것이 합당하다. 그러나 특징점의 수는 인식을 위한 매칭 과정에서 테스트 이미지의 특징점과의 비교 횟수를 의미하기 때문에 클래스당 특징점의 수가 많을수록 계산량은 증가된다. 그러므로 본 논문에서는 제안된 실험으로부터 인식율과 계산량을 고려하여 T_d 의 임계값을 0.6으로 결정하였다. 이때 인식율은 평균 96.1%가 획득되며, 클래스당 특징점의 수는 가장 높은 인식율을 획득 할 때 보다 149개 감소된 평균 234.92개의 특징점이 사용된다.

다음 실험은 식별 과정에서 사용되는 임계치 T_s 에 대한 최적의 값을 결정하기 위해 수행하였다. 이 실험에서는 임계치 T_d 을 최적의 값인 0.6으로 고정하고 임계치 T_s 의 값을 [0.05, 0.9]의 범위에서 0.05씩 변경하여 실험을 수행하였으며 실험 결과는 그림 4와 같다. 결과를 살펴보면 T_s 의 임계값이 0.35에서 0.7의 범위에 존재할 때 평균 91.95%와 96.1%의 인식율을 획득할 수 있음을 알 수 있다. 따라서 제안된 실험 결과를 바탕으로 최고의 인식율이 획득된 0.5를 T_s 의 임계값으로 결정한다.

수행된 임계치 결정 실험에서는 제안된 얼굴 인식 시스템의 인식율과 계산량을 고려하여 두 임계치 T_d 와 T_s 의 값을 각각 0.6과 0.5로 결정하였으며, 이후의 모든 실험들은 결정된 임계치를 바탕으로 수행된다.



[그림 4] 임계치 $T_d = 0.6$ 일 때 T_s 의 변화에 따른 성능 평가

4.2 훈련 이미지의 수에 따른 성능 평가

본 실험에서는 제안된 얼굴 인식 시스템의 통합템플릿 구성에 사용된 훈련 얼굴 이미지의 수에 따른 성능을 평가하기 위해 수행하였다. 실험은 60×74 픽셀의 ORL 데이터베이스를 대상으로 각 얼굴 클래스에 사용되는 훈련 이미지의 수를 2개부터 5개까지 1씩 증가시키면서 수행하였다. 각 얼굴 클래스의 통합템플릿은 실험의 각 훈련 이미지 수에 대응되는 수만큼의 랜덤 번호를 10개 중에서 선택하여 구성하고 나머지 이미지들로부터 인식율을 계산한다. 이와 같은 실험 과정은 각 훈련 이미지의 수에 대해 10회씩 반복 수행하였다. 그리고 SIFT-Grid 기반의 다른 얼굴 인식 시스템인 Bicego[8]와 Cho[9]의 방법들에 대해 동일한 실험 과정을 수행하여 비교 분석 하였다. 이때 Cho의 방법에서 보팅 과정에서 사용되는 임계값은 [9]의 논문을 참조하여 0.3으로 결정하였다.

실험 결과는 표 1과 같이 2, 3, 4 그리고 5개의 훈련 이미지 수에 대해 10회의 반복 실험에 따른 평균 인식율, 훈련 집합에 따른 인식율의 표준편차 그리고 하나의 얼굴 클래스 당 사용된 특징점 수의 평균으로 평가하였다. 실험 결과 제안된 방법이 비교 방법들보다 모든 훈련 이

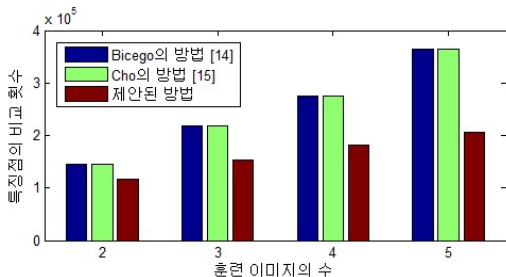
<표 1> 훈련 이미지의 수에 따른 성능 평가

	Bicego의 방법 [8]				Cho의 방법 [9]				제안된 방법			
	2	3	4	5	2	3	4	5	2	3	4	5
훈련 이미지 수	2	3	4	5	2	3	4	5	2	3	4	5
평균 인식율	73.59%	81.50%	87.54%	92.10%	72.03%	82.36%	88.25	93.40%	81.09%	89.18%	93.42%	96.10%
표준 편차	2.74	2.38	2.95	1.90	3.51	1.74	3.05	1.63	2.88	1.54	1.90	1.26
평균 특징점 수	165.94	248.91	311.67	414.81	165.94	248.91	311.67	414.81	132.16	174.90	207.32	234.92

미지 수에 대해 높은 인식율을 보이고 있으며, 훈련 이미지 수가 5인 경우에 대해 살펴보면 Bicego와 Cho의 방법보다 각각 4%와 1.7%로 향상된 96.10%를 획득하였다. 그리고 표준 편차의 평가에서 제안된 방법이 다른 방법보다 낮음을 알 수 있다. 이는 제안된 템플릿 구성 방법이 훈련 데이터의 구성을 위한 이미지 종류의 차이가 인식율의 변화에 미치는 영향이 기존 방법보다 적음을 의미한다. 마지막으로 클래스 당 구성된 특징점의 평균수들을 살펴보면 제안된 방법이 기존의 방법보다 적은수가 사용됨을 알 수 있으며, 이를 기반으로 SIFT-Grid 기반의 얼굴 인식 시스템에서 식별을 위해 수행되는 비교 횟수 계산하였다. 비교 횟수는 식 (12)과 같으며, $f_{test_number}^k$ 와 $f_{n_number}^k$ 는 각각 테스트 이미지와 n 번째 얼굴 클래스의 k 번째 부분영역의 특징점 수를 의미한다.

$$C = \sum_{n=1}^N \sum_{k=1}^K (f_{test_number}^k \times f_{n_number}^k) \quad (12)$$

그림 5는 표 1의 평균 특징점 수를 바탕으로 60×74 픽셀의 한 테스트 이미지가 식별되는데 필요한 비교 수행 횟수의 계산 결과를 나타내고 있다. 이때 테스트 이미지의 부분영역에 대한 특징점 수는 실험에 사용된 모든 얼굴 이미지의 부분 영역들의 평균인 22를 사용하였으며, 부분 영역은 4개이고, 전체 얼굴 클래스의 수는 40이다. 실험 결과 훈련 이미지의 수가 5인 경우 제안된 방법이 기존의 방법보다 비교횟수가 43.37% 적음을 볼 수 있고, 그 결과로써 제안된 템플릿 구성 방법이 인식 과정에서 계산량을 효과적으로 감소시킬 수 있음을 알 수 있다.



[그림 5] 훈련 이미지의 수에 따른 특징점 비교 횟수

4.3 다른 방법들의 얼굴 인식 성능 비교

본 절에서는 SIFT-Grid 기반의 특징, PCA, LDA 또는 PCA/LDA의 변형된 알고리즘을 통해 추출된 특징들

이 얼굴 인식의 정확도에 미치는 영향을 비교하였다. 비교 방법은 얼굴 인식의 실험에서 ORL 데이터베이스를 대상으로 각 클래스 당 5개의 이미지가 훈련 데이터로 사용된 [14]과 [15] 논문의 실험 결과를 참조 표 2에 정리하였다. 논문 [15]는 전통적인 PCA, 2D PCA의 다른 두 가지 방법 그리고 2개의 2D PCA의 특징을 혼합한 fusion approach의 특징 정보들을 적용한 얼굴 인식 실험 결과의 최대 인식율을 표 2의 첫 행부터 순서대로 4행에 기록하였다. 그리고 Baseline, EigenFaces, Fisherfaces 및 Intrinsicfaces의 특징 정보를 적용한 논문 [14]의 얼굴 인식 실험 결과의 최대 인식율은 표 2의 5번부터 8번 행에 기술하였다. 표 2의 마지막 3행은 본 장 2절의 SIFT-Grid 기반의 Bicego, Cho 그리고 제안된 방법에 대한 10회 반복 실험 결과 중 가장 높은 인식율을 기술하였다.

<표 2> 다양한 방법에 대한 최대 인식율의 비교

방법	인식율	이미지 크기
PCA [15]	90.50%	92×112
2D PCA1 [15]	85.50%	92×112
2D PCA2 [15]	56.00%	92×112
fusion approach [15]	92.00%	92×112
Baseline [14]	91.00%	32×32
Eigenfaces [14]	91.00%	32×32
Fisherfaces [14]	94.00%	32×32
Intrinsicfaces [14]	97.00%	32×32
SIFT-Grid Bicego의 방법	95.00%	60×74
SIFT-Grid Cho의 방법	96.00%	60×74
SIFT-Grid 제안된 방법	98.00%	60×74

표 2의 결과를 살펴보면 SIFT-Grid 기반의 특징을 이용한 얼굴 인식 시스템이 PCA, LDA 또는 PCA/LDA의 변형된 알고리즘에 기반을 둔 특징을 이용한 것보다 높은 인식율이 획득됨을 볼 수 있다. 그리고 제안된 SIFT-Grid 기반의 얼굴 인식 시스템이 다른 방법들보다 우수한 인식 성능을 가짐을 알 수 있다.

4.4 다양한 해상도의 얼굴 인식 성능 비교

다음 실험에서는 얼굴 이미지의 해상도에 따른 제안된 방법의 성능을 평가하기 위해 20×25부터 80×98픽셀까지 1.5%씩 증가된 7가지의 해상도와 ORL 데이터베이스

스의 원 해상도인 92×112픽셀에 대한 성능 비교를 수행하였다. 실험 결과는 랜덤하게 선택한 5개의 이미지를 훈련 데이터로 사용하고, 해상도별로 10회 반복 수행한 결과로부터 평균 특징점수, 평균 인식율, 최대 인식율 그리고 최소 인식율을 획득하여 표 3에 정리하였다. 실험 결과로부터 해상도가 증가됨에 따른 평균, 최대 그리고 최소 인식율의 변화를 살펴보면, 20×25부터 60×74픽셀의 해상도까지 인식율들이 증가됨을 볼 수 있으며, 60×74픽셀 이후의 해상도의 증가에 따른 인식율은 거의 변화가 없음을 볼 수 있다. 반면 클래스 당 특징점 수는 해상도가 커짐에 따라 최소 36개에서 최대 58개씩 증가되는 것을 볼 수 있다. 이 결과로부터 제안된 시스템은 얼굴 이미지의 해상도가 60×74픽셀일 때 인식율과 계산량 측면에서 최적의 성능이 획득됨을 알 수 있다.

〈표 3〉 이미지 해상도에 따른 성능 평가

이미지 해상도	평균 특징점수	평균 인식율	최대 인식율	최소 인식율
20×25	42.84	83.05%	86.50%	78.00%
30×37	78.76	91.45%	95.00%	90.00%
40×49	124.36	93.45%	96.00%	91.00%
50×61	178.12	95.25%	96.00%	94.00%
60×74	234.92	96.10%	98.00%	94.50%
70×86	291.32	96.65%	98.00%	95.00%
80×98	349.47	96.90%	99.00%	94.00%
92×112	396.40	96.80%	98.00%	95.00%

5. 결론

SIFT 기반의 얼굴 인식 시스템들은 PCA 또는 LDA 변환에 기반을 둔 전체론적인 방법(Holistic method)보다 조명, 포즈, 표정의 변화에 강인한 특성을 가지고 있다. 하지만 얼굴 인식 과정에서 테스트 이미지와 훈련 데이터에 포함된 SIFT 특징점들 간의 반복적인 비교 연산이 수행되기 때문에 많은 계산이 필요하다. 본 논문은 SIFT-Grid에 기반을 둔 얼굴 인식 시스템에서 계산량 감소와 인식율 향상을 목적으로 두 가지를 제시하였다. 첫 번째 제시 사항은 한 얼굴 클래스의 다양한 훈련 이미지로부터 비슷한 SIFT 특징점들은 제거하고, 상이한 특징점들은 병합하는 통합템플릿의 구성 방법이고, 두 번째는 구성된 통합템플릿들로부터 테스트 이미지의 효과

적인 식별을 위한 유사도 계산 방법이다.

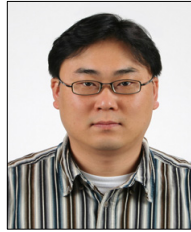
제안된 얼굴 인식 시스템에서 최대의 성능을 획득하기 위한 다양한 실험들의 수행 결과 통합템플릿 구성을 위한 임계치 T_d 는 0.6, 유사도 계산 과정의 임계치 T_s 는 0.5, 그리고 얼굴 이미지의 해상도는 60×74픽셀일 때 인식율과 계산량 두 측면에서 가장 효과적임을 확인하였다. 그리고 기존의 SIFT-Grid 기반의 얼굴 인식 방법과의 비교 실험 결과, 5개의 훈련 이미지가 사용된 경우 제안된 방법이 기존의 방법보다 43.37% 적은 특징점의 비교 연산이 수행되며, 최대 인식율은 98%로 2%향상됨을 확인하였다.

참 고 문 헌

- [1] A. M. Martinex (2000), "Recognition of Partially Occluded and/or Imprecisely localized faces using a probabilistic approach", EEE International Conference on Computer Vision and Pattern Recognition, 1, 712-717.
- [2] C. Cruz, L. E. Sucar and E. F. Morales (2008), "Real-time face recognition for human-robot interaction", EEE International Conference on Automatic Face Gesture Recognition, 1 - 6.
- [3] C. W. Ngo, W. L. Zhao and Y. G. Jiang (2006), "Fast Tracking of Near-Duplicate Keyframes in Broadcast Domain with Transitivity Propagation", ACM Multimedia, 845-854.
- [4] D. G. Lowe (2004), "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60(2), 91-110.
- [5] D. L. Swets and J. Weng (1996), "Using discriminant eigenfeatures for image retrieval", IEEE Transactions On Pattern Analysis And Machine Intelligence, 18(8), 831-836.
- [6] D. R. Kisku, M. Tistarelli, J. K. Sing and P. Gupta (2009), "Face recognition by fusion of local and global matching scores using ds theory: an evaluation with uni-classifier and multi-classifier paradigm", IEEE Workshop on Computer Vision and Pattern Recognition, 60 - 65.
- [7] H. Zhou, Y. Yuan and C. Shi (2009), "Object

- tracking using SIFT features and mean shift”, *Computer Vision and Image Understanding*, 113(3), 345-352.
- [8] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli (2006), “On the use of sift features for face authentication”, *Workshop on Computer Vision and Pattern Recognition*, 35 - 40.
- [9] M. Cho, H. Park (2009), “A Robust Keypoints Matching Strategy for SIFT: An Application to Face Recognition”, *Computer Science Neural Information Processing, Lecture Notes in Computer Science*, 5863/2009, 716-723.
- [10] M. Turk and A. Pentland (1991), “Eigenfaces for recognition”, *Journal of Cognitive Neuroscience*, 3(1), 71-86.
- [11] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman (1996), “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 19(7), 711-720.
- [12] W. -T. Chu and C. -H. Lin (2010), “Consumer photo management and browsing facilitated by near-duplicate detection with feature filtering”, *Journal of Visual Communication and Image Representation*, 21(3), 256-268.
- [13] Y. B. Han, J. Q. Yin and J. P. Li (2008), “Human face feature extraction and recognition base on sift”, *International Symposium on Computer Science and Computational Technology*, 1, 719 - 722.
- [14] Y. Wang and Y. Wu (2010), “Face recognition using Intrinsicfaces”, *Pattern Recognition*, 43, 3580-3590.
- [15] Y. Xu, D. Zhang, J. Yang and J. Y. Yang (2008), “An approach for directly extracting features from matrix data and its application in face recognition”, *Neurocomputing*, 71, 1857-1865.
- [16] <http://www.cl.cam.ac.uk/research/dtg/attarchive/face-database.html>

김 성 훈



(주) 연구소장

- 2002년 2월 : 대불대학교 컴퓨터공학과(공학사)
 - 2004년 2월 : 경희대학교 전자계산공학과(공학석사)
 - 2011년 8월 : 경희대학교 컴퓨터공학과(공학박사)
 - 2011년 9월~현재 : 대양전자통신
- 관심분야 : 신경망, 패턴인식, 얼굴 검출, 이미지 처리, 음성 처리, SoC
- E-Mail : shoonkim@khu.ac.kr

김 형 호



교육과 교수

- 1989년 2월 : 경희대학교 전자계산공학과 졸업(공학사)
 - 1992년 8월 : 경희대학교 전자계산공학과 졸업(공학석사)
 - 1998년 3월 : 일본 게이오대학 계산기공학과(박사수료)
 - 1998년~현재 : 세한대학교 컴퓨터교육과 교수
- 관심분야 : 컴퓨터구조, 신경회로망, ICT교육, 물류정보
- E-Mail : hhkim@sehan.ac.kr

이 현 수



정보대학 학장 및 정보통신대학원 원장

- 1979년 2월 : 경희대학교 전자공학과(공학사)
- 1982년 4월 : 일본 게이오대학원 전기공학과(공학석사)
- 1985년 4월 : 일본 게이오대학원 전기공학과(공학박사)
- 2005년~2008년 : 경희대학교 전자정보대학 학장 및 정보통신대학원 원장
- 1985년~현재 : 경희대학교 컴퓨터공학과 교수
- 관심분야 : 컴퓨터구조 및 VLSI, 병렬처리, SoC, 패턴 인식, 신경망, 음성처리
- E-Mail : leehs@khu.ac.kr