

포즈 추이 평면을 이용한 포즈
변화에 강인한 다중영상 얼굴인식
시스템

연세대학교
생체인식협동과정
김 봉 조

포즈 추이 평면을 이용한 포즈
변화에 강인한 다중영상 얼굴인식
시스템

지도교수 손 광 훈

이 논문을 석사 학위논문으로 제출함

2006년 12월 일

연세대학교

생체인식협동과정

김 봉 조

김봉조의 석사 학위 논문을 인준함

심사위원 _____ 인

심사위원 _____ 인

심사위원 _____ 인

연세대학교 대학원

2006년 12월 일

감사의 글

석사과정 동안 많은 지도와 관심을 가져주신 지도교수 손광훈 교수님께 깊이 감사 드립니다. 또한 바쁘신 와중에도 귀중한 시간을 내주시어 논문에 대한 많은 충고를 해주신 김재희 교수님, 이상운 교수님께도 많은 감사 드립니다.

연구실 최고참으로 항상 많은 조언을 해준 용태 선배님, 얼굴인식 팀장으로 고생하는 육일선배님, 논문을 쓰는데 많은 도움을 주고 늘 후배들 일에 많은 관심을 주는 착한 형철이 형, 3차원 비전 팀의 든든한 팀장으로 항상 열심이고 뛰어난 동보 박사, 코텍팀의 카리스마 정동이, 2년간 유일하게 같이 놀아준 귀엽고 능력있는 동현이, 알수 없는 개그를 하는 동안의 재원이, 멋쟁이 상운이, 저음불가의 이쁜 지현이, 4차원의 잠꾸러기 지연이, 코텍 막내 현성이, 2년간 형이라는 이유로 많은 서류작업 및 공무를 도와주고 같이 고생도 하고 좋은 기억도 많은 나의 동기이자 동생인 동주, 너무 착해서 2년간 너무 고생했고 앞으로 몇 년은 계속 막내 할것 같은 우리 팀 막내 민성이, 우리 연구실 살림꾼이자 나의 보급창고 원미씨, 같이 졸업 논문 쓰면서 고생하신 착한 오윤이형, 딱 부러지고 능력 많은 삼성맨 희철이형, 올해 결혼하신 족구계의 이단아 주현이형, 정상으로 보이지만 너무나도 이상한 마인드의 형욱이형 모두에게 감사 드립니다.

마지막으로, 항상 아들을 북돋아 주시고 지켜봐 주신 아버님, 어머님 그리고 누나에게 사랑한다는 말과 함께 깊은 감사의 말씀을 드립니다.

2006년 12월

김봉조 드림

차 례

제 1 장 서 론	1
제 2 장 이론적 배경	12
제 2-1 절 PCA 방법론에 의한 얼굴표현과 인식	12
제 2-2 절 다중영상 얼굴인식 시스템 데이터베이스	19
제 2-3 절 다중영상 얼굴인식 시스템	20
제 3 장 포즈 추이 평면을 이용한 다중영상 얼굴인식	24
제 3-1 절 시스템 개요	24
제 3-2 절 연속적인 포즈의 얼굴영상 획득	25
제 3-3 절 포즈 추이 평면을 이용한 얼굴인식	29
제 4 장 실험 및 결과분석	36
제 4-1 절 실험 환경	36
제 4-2 절 실험 결과	39
제5장 결 론	44
참고문헌	45
영문요약	53

그림 차례

그림 1.1	얼굴인식 응용 분야	2
그림 1.2	얼굴인식 기술	3
그림 1.3	곡률로 표현된 얼굴	6
그림 1.4	곡률지수로 표현된 얼굴	7
그림 1.5	Point signature 정의	8
그림 2.1	얼굴 벡터의 구성	13
그림 2.2	고유얼굴	18
그림 2.3	다중 분류기 기반의 다중영상 얼굴인식	21
그림 2.4	단일 분류기 기반의 다중영상 얼굴인식	21
그림 2.5	상호 부분 공간 기법	22
그림 2.6	다중영상의 궤적을 이용한 매칭	23
그림 3.1	포즈 추이 평면을 이용한 시스템 개략도	25
그림 3.2	정규화된 3차원 얼굴 데이터	27
그림 3.3	연속적으로 변화하는 포즈의 얼굴 영상	28
그림 3.4	연속적으로 포즈가 변하는 다중영상의 궤적	31
그림 3.5	기존 다중영상 얼굴인식 시스템의 오차측정	32
그림 3.6	최소 거리값을 이용한 인식 방법	33
그림 3.7	포즈 추이 평면을 이용한 오차측정	34
그림 4.1	3차원 레이저 스캐너	37
그림 4.2	Flea CCD 카메라	38
그림 4.3	다중 입력 영상	38
그림 5.1	16번째 사람에게 해당하는 입력영상의 실험결과	40
그림 5.2	9번째 사람에게 해당하는 입력영상의 실험결과	41
그림 5.3	포즈 오차가 존재하는 입력 영상	42
그림 5.4	Cumulative Match Score	43

표 차 례

표 2.1 포즈 변화가 존재하는 데이터베이스	19
표 4.1 3차원 레이저 스캐너 사양	37
표 4.2 Flea CCD camera 사양	38
표 5.1 실험 1 인식률 결과	40
표 5.2 포즈 오차가 존재하는 경우의 실험 결과	43

약 기 호 표

2D	2 dimensional
3D	3 dimensional
NN	Nearest Neighbor
PCA	Principal Component Analysis
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
EBGM	Elastic Bunch Graph Matching
AAM	Active Appearance Model
FERET	FacE REcognition Technology
FRVT	Face Recognition Vendor Test
EGI	Extended Gaussian Image
ICP	Iterative Closest Point
SVD	Singular Vector Decomposition
CCD	Charge Coupled Device
KPCA	Kernel Principal Component Analysis
KLDA	Kernel Linear Discriminant Analysis

국문요약

포즈 추이 평면을 이용한 포즈 변화에 강인한 다중영상 얼굴인식 시스템

본 논문에서는 포즈 추이 평면을 이용한 포즈변화에 강인한 다중영상 얼굴인식 시스템을 제안한다. 제안한 시스템은 다중 얼굴 영상을 이용하여 포즈 추이 평면을 정의하고 이를 이용한 새로운 오차측정 방법을 사용한다. 포즈 추이 평면은 포즈 변화에 따른 특징 공간상의 특징점의 변화를 내포하기 때문에 포즈 추정 단계에서 발생하는 오차에 강인한 인식률을 나타낸다.

제안한 시스템의 성능을 검증하기 위하여 2가지 실험을 수행하였다. 첫째 입력 영상의 특징벡터와 데이터베이스의 특징벡터간의 거리를 측정하는 Nearest Neighbor (NN)방법에 의한 인식률과 본 논문에서 제안한 포즈 추이 평면과의 거리를 이용한 인식률을 비교하였다. 또한 제안한 시스템이 포즈 추정 오차에 강인함을 검증하기 위하여 포즈 추정이 정확히 이루어졌을 경우와 $\pm 5^\circ$, $\pm 10^\circ$ 의 포즈 추정 오차가 발생할 경우에 대하여 기존의 다중영상기반 시스템의 인식률과 제안한 다중영상기반 시스템의 인식률을 비교하였다. 기존 다중영상 기반의 얼굴인식 시스템의 인식률은 76%이고 제안한 시스템은 91%로 15%의 인식률의 향상을 확인하였다. 또한 포즈 추정 오차가 발생할 경우 기존 다중영상 기반의 얼굴인식 시스템은 포즈 추정의 오차가 커질수록 인식률이 급격히 감소하나 제안한 시스템은 포즈 추정 오차에 강인한 인식률을 나타내었다.

핵심되는 말 : 포즈 추이 평면, 다중영상

제 1 장 서론

얼굴인식 기술은 광학 장치를 통해 입력되는 정지영상 및 동영상 자료의 얼굴을 분석하여 기존의 데이터베이스 안에서 일치하는 얼굴을 찾아내는 기술이다. 얼굴인식 분야는 기본적인 영상 신호 처리, 영상 분할, 신경회로망 또는 통계적 패턴인식 기술 등을 종합하는 기술로서 선진 외국의 경우, 기업은 물론 정부 심지어 국제협력기구의 주도하에 1970년대부터 학제적 공동연구가 본격적으로 전개되기 시작하였다. 현재 유럽의 경우 여러 국가에서 공동으로 연구단체를 구성하여 얼굴영상 처리에 대한 대규모 프로젝트가 진행되고 있으며, 개발된 기술이 일부 실용화에 근접한 단계에 있다.

그림 1.1과 같이 얼굴인식 기술은 보안시스템이나 법 집행목적, 상업적 목적 등으로 사용되고 있으며 그 응용 범위가 매우 다양하기 때문에 많은 주목을 받고 있다[1]. 현재 선진 외국은 물론 국내에서도 은행 경비 시스템, 건물 무인 경비 시스템 등의 분야에서 자동 얼굴인식 기술을 이용한 지능형 무인 경비/감시 시스템이 개발되고 있다. 그리고 또 다른 중요한 분야로서 국가 안보 분야에 적용될 수 있다. 얼굴인식은 범죄자 색출 작업, 테러나 범죄 용의자 확인 작업, 첩보 활동 등 여러 분야에서도 매우 실용가치가 크다. 또한 ID 카드, 신용 카드, 운전 면허증 등의 사진을 이용한 신분 확인도 법적, 상업적으로 중요한 얼굴인식의 응용분야가 될 것이다. 그밖에도 네트워크상에서 자동적으로 사용자들의 사용 권한 허용 여부를 결정하고, 사용자가 원하는 컴퓨터 환경으로 자동적으로 설정하는 등 보다 고차원의 기계-사람 인터페이스 구현 등에 사용되어 편리한 환경을 제공해 줄 수 있을 것이다.

얼굴인식은 크게 얼굴 검출, 특징 추출, 매칭을 통한 인식과정으로 구분할 수 있다[1]. 이 중 첫째 단계인 얼굴검출은 전체 인식시스템에 처리할 얼굴부분의 데이터만을 제공해 주는 단계이다[2]. 입력 영상은 배경과 얼굴 영역 경계 부분에 존재하는 화소들의 밝기 값의 유사성, 복잡한 배경, 각 개인별 차이, 조명의 영향 등으로 인한 매우 다양한 특징을 나타내므로, 얼굴 영역을 추출하는 문제는 매우

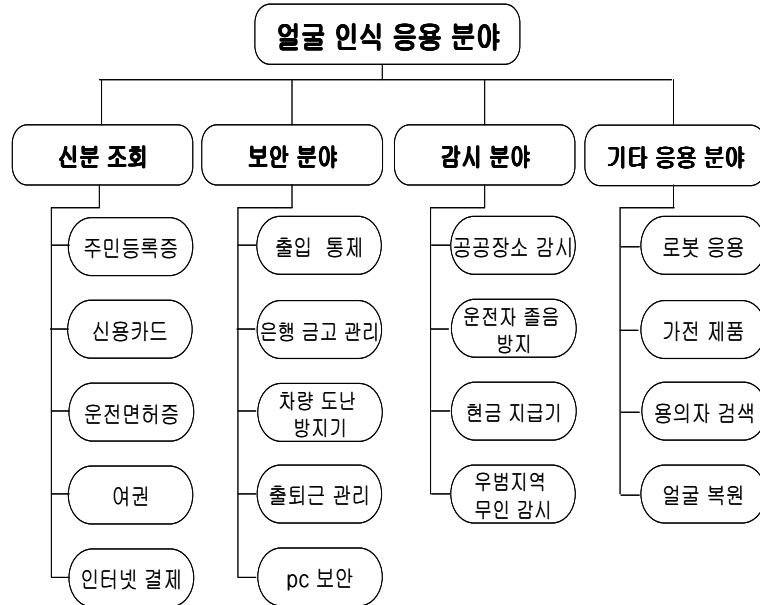


그림 1.1 얼굴인식 응용 분야

어려운 일이다. 얼굴 영역을 추출하는 기술은 크게 밝기 정보를 이용하는 방법, 움직임 정보를 이용하는 방법 [3], 밝기 및 움직임 정보를 동시에 이용하는 방법 [4], 색상정보를 이용하는 방법으로 구분된다[5]. 밝기 정보를 이용한 기법은 배경이 단순하여 얼굴 영역과 배경 영역의 경계 부근 화소들의 밝기 값의 차가 큰 경우에 얼굴 영역을 정확히 추출할 수 있다. 그러나 복잡한 배경에서는 얼굴 영역을 정확히 추출할 수 없는 단점이 있다. 움직임 정보를 이용한 기법은 동영상에서 사람의 얼굴은 움직임이 있는 객체라고 가정하고 인접 프레임간의 차 또는 움직임추정을 통해 유사한 움직임 벡터(motion vector)값을 나타내는 화소들을 그룹화 함으로써 얼굴 영역을 추출한다. 그러나 배경 영역에서의 움직임 추정정보가 얼굴영역에서의 추정 정보와 유사할 경우 움직임 정보만을 이용하여 얼굴 영역을 정확히 추출할 수 없다. 밝기와 움직임 정보를 동시에 고려할 경우 단순한 배경에서 분할된 많은 영역들을 효과적으로 제거하면서 얼굴 영역을 추출 할 수 있는 장점이 있다. 그러나 얼굴 영역 추출시 사용되는 밝기 정보와 움직임 정보의 비에 민감하므로

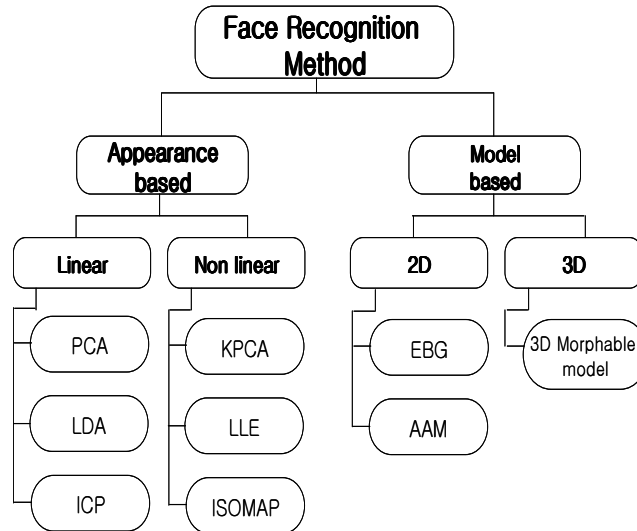


그림 1.2 얼굴인식 기술

이들간의 분할 척도 비를 조절하는 어려움이 있다. 피부색은 얼굴 영역 추출시 유용한 정보가 될 수 있어 널리 이용되어 왔다. 이는 크게 HSI, YIQ, YCbCr 등과 같은 칼라 모델을 기반으로 하여 피부색 구간을 설정하고 피부색 화소들을 추출함으로써 얼굴 영역을 추출하는 기법들과 피부영역의 색분포를 모델링 한 후 입력 화소들에 대한 피부색 및 비피부색 화소 여부를 결정하여 얼굴 영역을 추출하는 기법들로 분류된다. 그러나 배경영역에 피부색과 유사한 색상 정보를 나타내는 화소들이 존재할 경우 얼굴 영역을 정확히 추출할 수 없다. 최근에는 배경이 복잡한 영상[6-8], 여러 명의 얼굴이 있는 영상, 회전한 영상에 대한 연구[9]가 현재 활발히 진행되고 있으며, 그 필요성이 증대되고 있다. 특징 추출 방법과 인식 방법은 인식 방법에 따라 추출되는 특징이 다르기 때문에 함께 얼굴인식 기술이란 범주로 연구되고 있다. 얼굴인식 기술은 그림 1.2와 같이 크게 외형 기반과 모델 기반으로 구분할 수 있다. 외형 기반은 얼굴영상을 고차원의 벡터로 간주하고 벡터 공간에서 다양한 확률적 관점을 바탕으로 얼굴벡터의 분포를 분석하고 인식에 효율적인 벡터 표현을 이끌어낸다. 그리고 벡터 공간상에서 입력 영상의 벡터와 데이

터베이스 벡터간의 유사도를 이용하여 인식을 수행한다. 대표적인 외형 기반의 얼굴인식 기술은 Principal Component Analysis (PCA) [10-12], Independent Component Analysis (ICA) [13-15], Linear Discriminant Analysis (LDA) [16-19] 등이 있다. 각각의 방식은 서로 다른 확률적 관점을 바탕으로 얼굴영상을 벡터로 표현했다. PCA 기법은 Karhunen-Love's Transform 으로부터 유도된 기법으로써, 얼굴 영상을 n 차원의 벡터로 간주할 때 PCA 는 $d(n \gg d)$ 차원으로 차원을 줄여 인식률을 높이면서 동시에 계산의 복잡도를 줄여준다. d 차원의 고유벡터는 원 영상공간에서 분산이 가장 큰 방향을 나타낸다. 이러한 고유벡터로 정의된 새로운 공간은 일반적으로 원래 영상공간보다 낮은 차원으로 표현가능하다. ICA 는 입력 데이터의 2차, 고차의 확률적 종속성을 최소로 하며 동시에 입력데이터간의 확률적으로 가장 독립적인 기저벡터를 찾는다. LDA는 모든 클래스에 속한 데이터를 이용하여 Between-class scatter matrix (S_B)와 Within-class scatter matrix (S_W)를 정의하고 S_B 는 최대로 하며 동시에 S_W 는 최소로 하는 벡터를 찾는다. 이러한 벡터는 특징벡터들의 클래스를 가장 잘 분리한다. PCA, ICA, LDA는 얼굴영상을 선형으로 분류할 수 있다는 가정을 전제한다. 실제로 얼굴영상은 3차원 객체로써 선형으로 완벽하게 분류할 수 없기 때문에 이와 같은 가정은 인식률의 저하를 초래한다. 이러한 이유로 비선형 기법을 적용한 kernel method[20-26] 에 대한 연구가 이루어졌다.

모델기반의 얼굴인식 기술은 선행지식을 이용하여 얼굴 모델을 만들고 이를 인식에 이용한다. 대표적인 모델기반의 인식 기법으로 Elastic Bunch Graph Matching (EBGM)[27]과 Active Appearance Model (AAM)[28, 29], 3D morphable model [30-34]기법 등이 있다. 일반적으로 모델 기반의 얼굴인식 방법은 다음 3단계를 거쳐 인식과정을 수행한다.

- 1) 선행지식을 이용하여 얼굴모델을 생성한다.
- 2) 얼굴모델을 입력된 얼굴 영상과 적합 시킨다.
- 3) 적합된 모델의 매개변수를 특징벡터로 사용하여 데이터베이스의 저장된 특징벡터와의 유사도를 계산하여 인식을 수행한다.

EBGM은 얼굴을 그래프로써 표현한다. 얼굴의 눈, 코, 입과 같은 특징 점을 노드로 연결선은 얼굴의 특징점끼리의 거리를 이용하여 얼굴을 그래프로써 표현하였다. 각각의 노드는 서로 다른 크기와 방향에 따라 40개의 Gabor wavelet 계수를 포함하고 있다. 이와 같은 방법으로 얼굴을 그래프로 표현하고 그래프 매칭기법을 이용하여 얼굴인식을 수행한다. AAM은 얼굴의 구조적 정보와 색상정보를 통합한 얼굴 모델을 정의하고 입력 얼굴 영상과 데이터베이스 얼굴모델과의 차이를 이용하여 얼굴인식을 수행한다. 3D morphable model은 얼굴을 3차원 모델링 기법을 이용하여 표현한다. 얼굴은 3차원 객체이기 때문에 3차원 모델링 기법으로써 표현할 경우 포즈와 조명변화에 무관하게 얼굴을 표현할 수 있다. Blantz [30, 31]은 3차원 평균 얼굴 모델을 정의하고 입력 2D 영상을 이용하여 3차원 평균모델을 변형시키고 이때 사용된 변형 매개변수를 이용하여 얼굴인식을 수행하였다.

위에서 언급한 대부분의 얼굴인식 기법은 입력 영상을 단일 정면 영상으로 가정한다. FRGC [35] 보고서에 의하면 고화질의 단일정면 영상을 입력으로 사용할 경우 단일영상 기반의 얼굴인식 시스템은 99% 이상의 신뢰할만한 인식률을 나타낼 수 있다. 그러나 이러한 단일영상 기반의 얼굴인식 시스템은 입력 영상의 포즈 변화가 있을 경우 인식률의 급격한 저하를 나타낸다. 포즈 변화에 따라서 얼굴의 중요 특징들의 위치와 생김새가 변하기 때문에 단일영상 기반의 얼굴인식 시스템은 인식률의 한계를 가지게 된다. 서로 다른 포즈의 동일한 사람의 두 영상보다는 동일한 포즈의 서로 다른 두영상이 동일인의 영상으로 인식될 가능성이 더 높아지기 때문이다. FacE REcognition Technology (FERET)의 테스트 리포트 [36, 37]과 Face Recognition Vendor Test (FRVT)의 2000년, 2002년 테스트 리포트[38-40]을 보면 포즈 변화 문제의 심각성과 해결의 어려움이 잘 나타나 있다. 이러한 포즈의 변화의 따른 인식률의 저하를 극복하기 위하여 다양한 기법이 소개되었다. 초기에는 포즈의 변화에 무관한 특징점을 찾거나 [27], 정면영상을 합성하는 기법에 초점이 맞춰져 왔다 [41]. 이러한 기법들은 작은 포즈의 변화에는 강인한 인식 결과를 보여주었지만 포즈의 변화가 클 경우에는 인식률의 급격한 저하를 가져왔다. 포즈의 변화가 클 경우에도 강인한 인식률을 획득하기 위한 방법으로 3차원 얼굴인식 기법과 다중영상을 이용한 방법이 주목을 받고 있다.

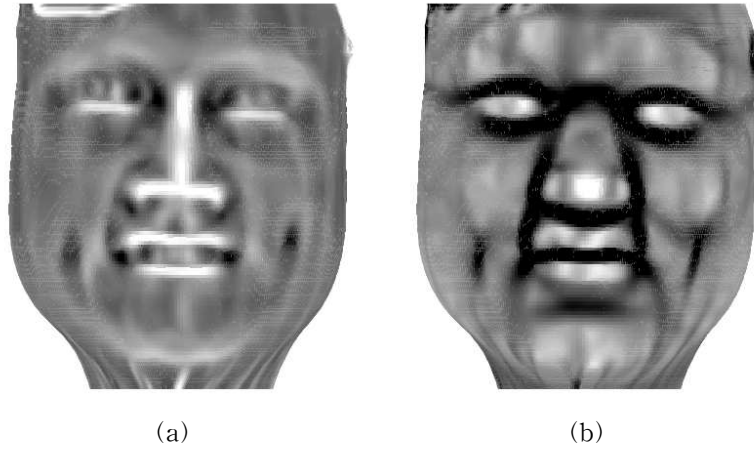


그림 1.3 곡률로 표현된 얼굴 (a) 최대 곡률 (b) 최소 곡률

3차원 얼굴인식 기법은 기하정보, 색상정보, 기하정보와 색상정보를 모두 이용한 방법으로 구분할 수 있다. 3차원 얼굴인식에서 기하정보란 곡률, 곡률지수, Signature와 같이 얼굴의 기하학적 특징을 말한다. 곡률이란 곡선 또는 곡면의 휨 정도를 나타내는 변화율을 말한다. 그림 1.3은 얼굴의 최대, 최소 곡률의 크기를 나타낸 예이다. R^3 공간에 곡선 $C(x,y,z)=0$ 이 존재할 경우 곡선 C 의 곡률 k 는 아래의 식 (1.1)에 의해 정의 된다.

$$k = \left| \nabla \cdot \left(\frac{\nabla C}{\|\nabla C\|} \right) \right| \quad (1.1)$$

R^3 공간에서 2차원 곡면을 생각할 경우, 곡률에는 가우스 곡률(Gaussian curvature)과 평균 곡률(mean curvature)이 있다. 2차원 곡면 S 위의 한 점 P 에서 곡면 S 로의 법선 벡터 \vec{p}_n 을 포함하는 평면을 S_\perp 라 하면, 곡면 S 와 평면 S_\perp 이 만나는 곡선 C_A 가 생성되며, 곡선 C_A 는 점 P 를 항상 지난다. 점 P 에서 C_A 의 곡률을 k 라 할 때, \vec{p}_n 을 중심으로 평면 S_\perp 를 회전시키면 k 는 여러 가지 값을 가지게 된다. 그 값들 가운데 최대값을 k_1 , 최소값을 k_2 라고 하고, 이 두 값들을 곡면 S 위의 점 P 에서의 주요 곡률(principal curvatures)이라 한다. 가우스 곡률은

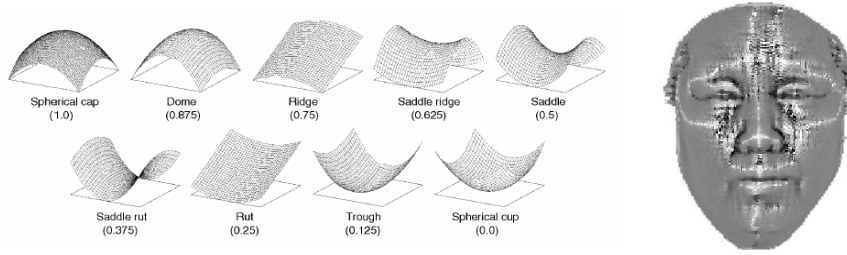


그림 1.4 곡률지수로 표현된 얼굴

주요 곡률 k_1 , k_2 의 곱과 같고, 평균 곡률은 주요 곡률 k_1 , k_2 의 평균과 같다. 평균 곡률은 평면일 경우 0의 값을, 평면이 아닐 경우 0이 아닌 값을 가진다. 가우스 곡률은 위로 볼록(convex) 또는 아래로 볼록(concave)한 지점에서 양의 값을 가지며, saddle한 지점에서 음의 값을 가지며, 평면일 경우 0의 값을 가지게 된다.

곡면 S 위의 한 점 P 에서의 곡률 지수는 아래의 식 (1.2)과 같이 정의 된다.

$$S(p) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (1.2)$$

$k_1(p)$, $k_2(p)$ 는 각각 곡면 S 위의 한 점 P 에서의 최대 곡률과 최소 곡률이다. 곡률 지수는 $[0, 1]$ 범위의 값을 가지게 되며, 평면일 때 0.5의 값을 가지게 된다. 그림 1.4는 곡률지수를 이용하여 표현한 얼굴영상이다.

그림 1.5는 point signature의 정의를 그림으로 나타낸 것이다. 그림 1.5의 (a)의 2차원 곡면 S 위의 한 점 P 에서 곡면 S 로의 법선 벡터 \vec{N} 을 포함하는 평면을 S_p 라고 한다. 평면 S_p 에 수직이고 점 P 를 포함하는 평면을 S_\perp 라 할 때, 점 P 를 중심으로 반지름 r 인 원을 평면 S_\perp 위에 그렸을 때, 그림 1.5의 (b)처럼 원 위의 각 지점으로부터 곡면 S 까지의 수직 거리 d_n 을 point signature라고 한다. 이러한 기하정보를 기반으로 3차원 얼굴인식 기법은 많은 연구가 진행되었고, 많은 결과들이 제시되었다. Tanaka [42]는 깊이 영상으로부터 주요 곡률을 계산하여 얼굴의 특징으로 사용하였다. 얼굴 데이터의 깊이 영상으로부터 곡률을 계산한 다음 Extended Gaussian Image(EGI)를 생성하여 그 영상을 특징으로 사용하고, 테스트

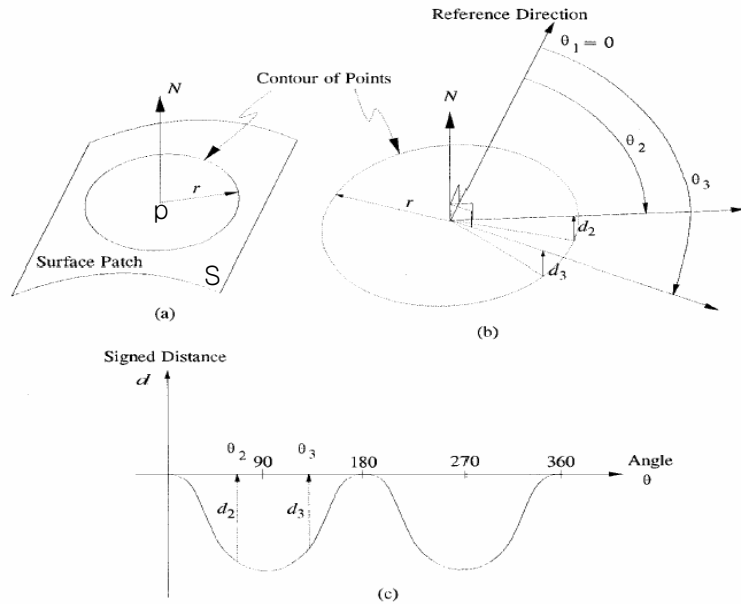


그림 1.5 Point signature 정의 (a) 임의의 곡면 S와 법선벡터 \vec{N} ,
 (b) 접면 S_{\perp} 와 곡면 S와의 거리 d_n , (c) Point signature 특징

데이터와 데이터베이스의 인식은 Fisher's spherical correlation 기법을 사용하였다. 논문에서 37개의 얼굴 데이터를 사용하였고 100%의 인식률을 보인다고 나왔다. 하지만 포즈와 조명 변화에 대한 언급은 제시하지 않았고, 얼굴의 곡률을 계산할 때 한계치(threshold)를 두어 표현을 하는데 한계치가 얼굴 영상마다 다르고 곡률의 위치에 따라 크게 영향을 받는 단점이 있다. Lee와 Milios는 거리 영상(range image)을 이용하여 얼굴 영역의 굴곡 정도를 측정하여 불룩한 영역을 특징으로 사용하는 기법을 제안하였다[43]. 각 불룩한 영역은 EGI로 표현되어지고 불룩한 영역간의 유사도는 EGI의 관계로 측정한다. 입력 영상과 데이터베이스 인식은 EGI로 표현된 불룩한 영역을 그래프 매칭 기법으로 수행한다. 불룩한 영역을 비교하는 것은 얼굴 표정 변화에 오목한 영역보다 영향을 적게 받기 때문에 표정에 강인하지만 EGI를 이용한 매칭기법은 동일한 얼굴이더라도 크기가 다를 경우 인식에 실패하는 단점이 있다. Gordon과 Moreno는 눈, 코, 입 등과 같은 특이점에서의 곡률과 특이점 간의 거리를 특징으로 사용한 기법을 제안하였다.

Gordon[44]의 경우 10개의 서로 다른 특징벡터로 얼굴을 표현하였고, Moreno는 86개의 다른 특징벡터를 사용하였다. 그들은 얼굴 영역을 8개의 구역으로 나누어 각 영역에서 특징을 나타내는 곡률과 특이점 간의 거리, 각도 등으로 특징벡터를 구성하였다. 특징벡터 중에서 중요성 여부에 따라서 가중치를 주도록 하였다. 입력 영상과 데이터베이스 간의 유사도 측정은 가중치를 부여한 특징벡터간의 거리 정합 기법을 사용하였다. Gordon의 경우 25개의 얼굴 데이터로 실험을 수행한 결과 91.7%의 인식률을 보였고, Moreno는 2가지의 포즈 변화와 5가지의 표정 변화를 포함한 60명의 420개의 얼굴 데이터로 실험한 결과 78%의 인식률을 보였다. 특징벡터를 사용할 경우 발생하는 단점은 특이점들의 잡음에 영향을 많이 받게 되고, 곡률을 계산할 경우 2차 미분을 수행해야 되는데 2차 미분의 경우 잡음에 민감하다. 또한 작은 포즈변화가 발생했을 경우 얼굴 영역의 특이점을 가리는 현상이 발생할 수 있다. Lee [45]는 얼굴의 profile을 추출하고 추출된 profile의 곡률을 계산하여 특징으로 사용하는 기법을 제안하였다. 깊이 영상에서 깊이 정보에 따라 얼굴 profile을 추출하여 인식을 수행하였다. 작은 얼굴 데이터들을 이용하여 94%의 성능을 보였다. 그들의 방법은 깊이 값의 변화에 크게 영향을 받고, 다른 3차원 얼굴 데이터 획득 장치를 사용했을 경우 서로 다른 해상도로 인해 비교하기가 힘들어진다. Chua [46, 47]은 point signature의 개념을 얼굴인식에 도입하였다. 얼굴 영상에 4개의 특이점을 선택한 후 특이점 주위 3X3 포인트들에 point signature를 계산하여 9개의 point signature 벡터를 만들고, 입력 영상과 데이터베이스간의 거리 정합을 통해 인식을 수행하였다. 그림 1.5의 (c)가 최종적인 특징이 되며, 그 특징을 서로 비교하여 인식을 수행하였다. 그들의 실험에서 개인별로 3개의 입력 영상과 3개의 데이터베이스를 비교한 결과 85%의 성능을 보였다. Medioni [48]은 stereo-based 시스템으로 획득된 3차원 얼굴의 표면 정보를 이용하여 Iterative Closest Point (ICP)를 수행하였다. ICP는 입력 영상과 데이터베이스를 비교하기 위한 대응점을 추출한 후 주어진 대응점을 기준으로 두 데이터를 가장 비슷하게 만들어주는 선형 변환을 계산한다. 일반적으로 평균자승오차를 측면에서 최적이 되도록 하는 방법을 사용한다. 계산된 선형 변환을 적용하고 계산된 선형 변환이 주어진 임계치보다 작아질 때까지 위의 3단계를 반복한다.

색상정보를 이용한 3차원 얼굴인식 기법은 영상의 픽셀의 값을 이용한 방법으로 2차원 얼굴영상 기법과 동일하다. 대표적인 방법으로 고유 얼굴을 이용한 방법이 있다. 이 방법은 데이터가 가지고 있는 원래의 특징을 잃어버리지 않으면서 효율적으로 데이터의 차원을 줄일 수 있는 선형 변환 방법을 이용하여 얼굴의 특징을 추출하는 방법이다. 2차원 얼굴인식에서는 얼굴 영상을 이용하여 고유 얼굴을 생성한다. 반면에 3차원 얼굴인식의 경우 얼굴 영상이 아닌 얼굴의 구조적 정보를 가지고 있는 깊이 영상 또는 거리 영상(range image)을 이용하여 고유 얼굴을 생성하고, 이것을 특징으로 하여 얼굴인식을 수행하는 방법들이 제안되었다. Chang [49]는 조명 변화에 강인한 얼굴인식을 위하여 2D+3D 얼굴인식 알고리즘을 제안하였다. 3차원 얼굴 데이터를 이용하여 다양한 조명 변화가 존재하는 2D 영상을 생성하고 이를 데이터베이스로 사용하였다. 조명의 변화가 있는 2D 얼굴영상이 입력으로 사용하여 유사도 측정을 통하여 인식을 수행하였다. Heshner [50]은 여러 가지 크기의 깊이 영상을 이용하여 고유 얼굴을 계산하여 인식에 사용하는 방법을 제안하였다. 6가지 표정 변화를 가진 37명의 222개의 얼굴 데이터를 가지고 거리 영상을 이용한 얼굴인식을 수행하였다. 포즈 변화에 대한 얼굴 데이터 거리 영상을 코를 기준으로 정규화 시켜 83%의 인식률을 보였다. Achermann [51]의 경우는 거리 영상을 이용하여 고유 얼굴을 특징으로 추출하고 Markov field를 이용하여 인식하는 방법을 제안하였다. 5가지 포즈 변화를 가진 24명의 120개 얼굴 데이터로 사람당 5개의 훈련 과정을 거친 PCA로 인식한 경우 100%의 인식률을 보였고, Markov field로 인식했을 경우는 89.17%의 인식률을 보였다.

기하 정보와 색상정보를 모두 이용한 방법은 색상정보를 이용하는 방법과 기하정보를 이용하는 방법의 서로의 단점을 보완할 수 있다. Wang [52]는 Gabor filter response를 색상정보로부터 획득하고 Point signature를 기하정보로부터 획득하였다. 두 가지 정보를 하나의 특징벡터로 취합하고 support vector machine을 이용하여 인식을 수행하였다. Godil [53]은 색상정보를 이용하여 PCA 기법을 수행하고 기하정보로부터 획득한 거리영상을 이용하여 PCA를 수행하였다. 두 개의 결과의 합으로써 인식을 수행하였다. Papatheodorou [54]는 색상정보와 기하정보를 취합하여 하나의 특징으로 간주하고 이 특징을 기반으로 ICP를 수행하여 인식을

수행하였다. 62명의 얼굴 데이터베이스를 이용하여 실험한 결과 98%의 인식률을 나타내었다. 하지만 3차원 얼굴 데이터를 사용해야만 하는 3차원 얼굴인식 방법은 막대한 계산량이 요구되어 아직까지는 실용적이지 못하다. 실용적인 측면에서 3차원 얼굴인식 방법보다는 다중영상을 이용한 기법이 더욱 많은 주목을 받고 있다.

일반적으로 다중영상 기반의 얼굴인식은 단일영상기반의 얼굴인식 시스템을 기반으로 한다. 즉 다중영상 시스템은 각각의 얼굴영상을 단일영상 얼굴인식의 사용된 기법을 적용하고 이들의 조합으로 인식을 수행하였다. 조합의 방법으로는 각각의 단일 영상 인식결과의 합을 이용하는 방법과 인식률 향상을 위해 각각의 얼굴영상으로부터 획득된 인식결과를 이용하여 voting 기법을 사용하는 방법이 있다 [55]. 하지만 이러한 방법은 단일영상 기반의 얼굴인식 시스템에서 사용된 방법의 반복일 뿐이다. 이러한 다중영상 기반의 얼굴인식 시스템은 크게 2가지의 문제점을 가지고 있다. 첫째 학습과 데이터베이스에 필요한 다양한 포즈변화가 있는 다수의 얼굴영상을 필요로 한다. 동일한 환경에서 한사람의 다양한 포즈변화가 있는 얼굴영상을 획득하기는 매우 어려운 일이다. 둘째, 다중영상 기반의 인식기법이 기존의 단일영상 기반의 얼굴인식 시스템에서 사용하던 기법의 반복이나 단순한 확장이라는 점이다. 이는 시스템을 단일영상 기반에서 다중영상기반의 시스템으로 확장했지만 다중영상에 존재하는 다양한 특성을 이용하지 못하는 문제점이 있다.

본 논문에서는 위에서 언급된 기존의 다중영상 기반의 얼굴인식 시스템의 2가지 문제점을 해결하기 위한 포즈변화에 강인한 다중영상 얼굴인식 시스템을 제안한다. 제안한 시스템에서 학습과 데이터베이스로 사용될 다중영상은 3D 얼굴 데이터로부터 정확하게 획득한다. 얼굴인식 과정은 다중영상에 존재하는 특성을 고려한 포즈 추이 평면을 이용하여 수행된다.

본 논문의 구성은 다음과 같다. 제 2장에서는 제안하는 다중영상 기반의 얼굴인식 시스템의 배경이론에 대하여 설명한다. 제 3장에서는 포즈 추이 평면을 정의하고 이를 이용한 오차측정 방법을 설명한다. 제 4 장에서는 제안한 방법과 기존 방법과의 비교 실험을 통해 제안한 방법의 우수성을 검증한다. 마지막으로 5장에서 결론을 맺는다.

제 2 장 이론적 배경

본 장에서는 포즈 추이 평면을 이용한 다중영상 기반의 얼굴인식 시스템의 배경이 되는 이론을 기술한다. 제 2.1 절에서는 2차원 얼굴영상을 표현하는 대표적인 방법인 PCA에 대해 기술한다. 제 2.2 절에서는 다중영상 기반의 얼굴인식 시스템의 학습과 데이터베이스에 사용되는 다중영상의 획득 방법에 대하여 기술한다. 제 2.3 절에서는 기존의 다중영상 기반의 얼굴인식 시스템에 대하여 기술한다.

제 2-1 절 PCA 방법론에 의한 얼굴표현과 인식

PCA를 이용한 얼굴인식 기술은 학습 과정과 인식 과정으로 나누어진다. 먼저 학습 과정에서 고유 얼굴을 생성시키고 모든 학습 영상을 이 얼굴 공간에 투사시켜 몇 개의 성분 값을 저장한다. 그 다음 각각의 사람마다 저장된 성분들의 평균을 대표 값으로 가지게 된다. 인식 과정에서는 입력 영상과 고유 얼굴의 내적으로 나온 성분과 저장된 대표 값과 Euclidean 거리로 비교한다. 가장 작은 거리를 갖는 대표값에 해당하는 사람을 입력 영상과 같은 인물로 인식한다.

PCA는 조밀한 코딩(compact coding)을 유도하여 중복 신호를 제거한다. 얼굴 영상 자체에서 불필요한 정보를 조밀하게 바꾸어 차원을 낮추어 주기에 인식과정이 편리해질 뿐만 아니라 높은 인식률을 나타낸다.

2.1.1 얼굴 벡터

얼굴 영상은 벡터로 나타낼 수 있는데, 각각의 픽셀은 하나의 벡터 성분으로 부호화된다. 만약 한 사람의 얼굴 영상의 높이와 너비를 W 와 H 라고 하면 이 벡터 성분의 수는 $W \times H$ 이다. 각각의 픽셀은 하나의 벡터 성분으로 부호화되어진다. 영상으로부터 얼굴 벡터를 만드는 방법은 그림 2.1과 같이 얼굴 영상의 한 행을

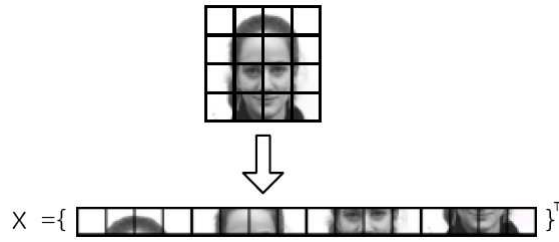


그림 2.1 얼굴 벡터의 구성

다음 행에 연결하는 것이다. 한 사람의 얼굴 벡터는 하나의 열벡터가 된다.

2.1.2 PCA

X 를 p 개의 랜덤 변수를 갖는 벡터라고 하자. PCA 방법은 변수들의 변위와 분산에 의해 주어진 정보의 대부분을 보존하는 적은 수의 변수를 찾는 것이다.

첫 번째 과정은 최대의 변위를 가지도록 선형방정식 $\alpha_1^T x$ 를 찾는 것이다. 여기서 α_1 은 $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$ 의 p 개의 정수를 가지는 벡터이다.

$$\alpha_1^T x = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j \quad (2.1)$$

다음 과정은 $\alpha_1^T x$ 와 상관관계를 가지지 않고 최대의 변위를 가지는 선형방정식 $\alpha_2^T x$ 를 찾는 것이다. 이런 식으로 계속 진행하여 k 번째 과정에서는 $\alpha_1^T x, \alpha_2^T x, \dots, \alpha_{k-1}^T x$ 와 연관되지 않고 최대 변위를 갖는 $\alpha_k^T x$ 를 찾는 것이다. k 번째 유도된 변수 $\alpha_k^T x$ 는 k 번째 주성분이다. 계속 p 개까지의 주성분들을 찾을 수 있다. 그러나 일반적으로 p 보다 훨씬 작은 m 개의 주성분만 가지고 x 에서의 대부분 변화를 설명할 수 있다. 일반적으로 처음 몇 개의 주성분들이 본래의 변수들에서 대부분의 변화를 설명한다. 나머지 주성분들은 매우 작은 차이만을 나타낸다. 그것들은 거의 정적인 것과 일치한다.

주성분들을 정의하기 위하여, 주성분을 찾는 방법을 알아야 한다. 랜덤 변수 x 의 벡터는 공분산 행렬(covariance matrix)인 Σ 를 갖는다고 가정하자. 이 Σ 의 (i, j) 번째 성분은 $i \neq j$ 일 때 x 의 i 번째와 j 번째의 성분 사이에 공분산을 나타낸다. 그리고 $i = j$ 일 때 x 의 j 번째 성분의 분산을 나타낸다. $k = 1, 2, \dots, p$ 에 대하여 α_k 가 Σ 행렬에서 k 번째 큰 고유치(eigenvalue)인 λ_k 에 일치하는 고유벡터(eigen vector)일 때, k 번째 주성분은 $z_k = \alpha_k^T x$ 이다. 그리고 $\alpha_k^T \alpha_k = 1$ 라면 z_k 의 분산은 λ_k 이다.

주성분들을 유도하기 위해서, $\alpha_1^T x$ 를 생각하자. α_1 은 $var[\alpha_1^T x] = \alpha_1^T \Sigma \alpha_1$ 을 최대로 한다. 여기서 var 는 분산을 나타낸다. 그런데 이 경우 무한히 많은 α_1 에서 최대를 가진다. 따라서 정규화된 $\alpha_k^T \alpha_k = 1$ 이 이용된다. $\alpha_k^T \alpha_k = 1$ 이라고 하면 $\alpha_1^T \Sigma \alpha_1$ 을 최대로 하기 위해서 Lagrange multiplier를 사용한다. 그러면 $\alpha_1^T \Sigma \alpha_1 - \lambda(\alpha_1^T \alpha_1 - 1)$ 를 최대로 하는 것을 찾으려면 된다. 여기서 λ 는 Lagrange multiplier이다. 이를 α_1 에 대하여 미분하면

$$\begin{aligned} \Sigma \alpha_1 - \lambda \alpha_1 &= 0 \\ (\Sigma - \lambda I_p) \alpha_1 &= 0 \end{aligned} \tag{2.2}$$

이다. 여기서 I_p 는 $P \times P$ 단위행렬이다. 따라서 λ 는 Σ 의 고유치이고, α_1 은 그것에 해당하는 고유벡터이다. 식 (2.2)를 풀어서 나오는 p 개의 고유벡터 중 어느 것이 최대로 하는 α_1 인지 결정하기 위해

$$\alpha_1^T \Sigma \alpha_1 = \alpha_1^T \lambda \alpha_1 = \lambda \alpha_1^T \alpha_1 = \lambda \tag{2.3}$$

를 최대로 하는 것과 일치하는 α_1 를 찾으려면 된다. 즉 λ 가 제일 큰 것을 찾으려면 된다. 따라서 α_1 은 Σ 의 가장 큰 고유치에 해당하는 고유벡터이고, 가장 큰 고유치는 $var[\alpha_1^T x] = \alpha_1^T \Sigma \alpha_1 = \lambda_1$ 이다.

두 번째 주성분 $\alpha_2^T x$ 는 $\alpha_1^T x$ 과 $cov[\alpha_1^T x, \alpha_2^T x] = 0$ 인 관계를 가진다. 여기서 cov 는 공분산을 나타낸다. 다시 말해서 서로 상관관계를 가지지 않고 $\alpha_2^T \Sigma \alpha_2$ 를 최대로 한다. 그런데

$$cov[\alpha_1^T x, \alpha_2^T x] = \alpha_1^T \Sigma \alpha_2 = \alpha_2^T \Sigma \alpha_1 = \alpha_2^T \lambda_1 \alpha_1 = \lambda_1 \alpha_2^T \alpha_1 = \lambda_1 \alpha_1^T \alpha_2 \quad (2.4)$$

이다. 따라서

$$\begin{aligned} \alpha_1^T \Sigma \alpha_2 &= 0, \quad \alpha_2^T \Sigma \alpha_1 = 0 \\ \alpha_1^T \alpha_2 &= 0, \quad \alpha_2^T \alpha_1 = 0 \end{aligned} \quad (2.5)$$

이 가운데 마지막 방정식을 가지고 구하면

$$\alpha_2^T \Sigma \alpha_2 - \lambda (\alpha_2^T \alpha_2 - 1) - \phi \alpha_2^T \alpha_1 \quad (2.6)$$

를 최대로 하는 α_2 를 찾으려면 된다. 여기서 λ, ϕ 는 Lagrange multiplier이다. 식 (2.6)을 α_2 에 대하여 미분하여 양변의 왼쪽에 α_1^T 를 곱하면

$$\begin{aligned} \Sigma \alpha_2 - \lambda \alpha_2 - \phi \alpha_1 &= 0 \\ \alpha_1^T \Sigma \alpha_2 - \lambda \alpha_1^T \alpha_2 - \phi \alpha_1^T \alpha_1 &= 0 \end{aligned} \quad (2.7)$$

이 된다. 여기서 처음 두개의 식은 0이고, $\alpha_1^T \alpha_1 = 1$ 이다. 따라서 $\phi = 0$ 이 된다. 그러므로 $(\Sigma - \lambda I_p) \alpha_2 = 0$ 이다. 그래서 다시 λ 는 Σ 의 고유치이고, α_2 는 그에 해당하는 고유벡터이다.

여기서 다시 $\lambda = \alpha_2^T \Sigma \alpha_2$ 이고, λ 는 가능한 크게 잡는다. 이때 $\lambda = \lambda_2$ 이고, 두 번째로 가장 큰 고유치이다. 여기서 λ_2 는 λ_1 과 같을 수 없다. 왜냐하면 $\alpha_2 = \alpha_1$ 이면 $\alpha_2^T \alpha_1 = 0$ 이 성립하지 않기 때문이다. 그리고 α_2 는 λ_2 에 해당하는 고유벡터가 된다.

위와 같은 방식으로 세 번째, 네 번째, \dots , p 번째 주성분을 구하면

$\alpha_3, \alpha_4, \dots, \alpha_p$ 의 벡터 계수는 $\lambda_3, \lambda_4, \dots, \lambda_p$ 에 해당하는 고유벡터가 되고, $\lambda_3, \lambda_4, \dots, \lambda_p$ 는 차례로 세 번째, 네 번째, \dots , p 번째로 가장 큰 고유치가 된다. 여기서 z 의 k 번째 원소가 z_k 라 할 때,

$$z = A^T x \quad (2.8)$$

가 된다. 여기서 A 는 k 번째 열이 α_k 인 직교 행렬이다. 따라서 주성분들은 x 를 선형 변환으로 구할 수 있다. 식 (2.8)을 PCA라고 한다. 여기서 A 는 Σ 의 k 번째 고유치인 λ_k 를 k 번째 대각 성분으로 하는 대각행렬이라 하자. 그러면 앞의 유도에 의해

$$\Sigma A = A \Lambda \quad (2.9)$$

임을 알 수 있다. 즉 X 의 주성분은 위의 식에서 A 의 행벡터가 된다.

2.1.3 얼굴인식에서의 PCA

앞에서 설명한 것처럼 얼굴의 영상은 영상 공간에서의 각각의 픽셀은 다른 픽셀들과 매우 연관성이 많다. 그래서 공분산 행렬인 Σ 는 대각 행렬과 전혀 다르게 식 (2.10)의 형태로 나타난다.

$$\Sigma = XX^T = \begin{bmatrix} \sigma_{11}^X & \sigma_{12}^X & \cdots & \sigma_{1,w \times h}^X \\ \sigma_{21}^X & \sigma_{22}^X & \cdots & \sigma_{2,w \times h}^X \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{w \times h, 1}^X & \sigma_{w \times h, 1}^X & \cdots & \sigma_{w \times h, w \times h}^X \end{bmatrix} \quad (2.10)$$

x_i : 입력 얼굴의 i 번째 성분

X : 학습할 얼굴 집합인 $I \times K$ 차원의 행렬

이 행렬의 k 번째 열은 학습할 집합의 k 번째 얼굴인 x_k 이다.

σ_{ij} : i 번째 픽셀과 j 번째 픽셀 사이의 공분산

따라서 서로 다른 성분과 상관관계가 없는 얼굴 공간을 만들어야 한다. 이것은 새

로운 성분들의 공분산 행렬은 대각행렬이 되어야 한다는 것이다.

$$\Sigma_Z = ZZ^T = \begin{bmatrix} \sigma_{11}^Z & 0 & \cdots & 0 \\ 0 & \sigma_{22}^Z & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{w \times h, w \times h}^Z \end{bmatrix} \quad (2.11)$$

공분산 행렬이 대각 행렬이 된다는 것은 서로 다른 변수의 공분산이 영이 되는 반면, 변수 그 자체의 분산은 최대로 됨을 의미한다.

영상 공간을 얼굴 공간으로 변환시키기 위해서 Singular Vector Decomposition (SVD)를 사용하면, X 는

$$X = A\Delta Q^T \quad (2.12)$$

로 분해된다. 여기서 A 는 주성분을 열로 갖는 행렬이다. 여기서 A 를 변환 행렬로 두자. 그러면 $Z = A^T X$, $X = AZ$ 이다. 그런데, $A^T A = I$ 이므로 $A = A^{-1}$ 이다. 이때 Σ_Z 는 아래의 식으로 정리된다.

$$\Sigma_Z = ZZ^T = A^T X X^T A = A^T \Sigma A \quad (2.13)$$

여기서, A 는 Σ 의 고유벡터를 포함하는 행렬이므로

$$\Sigma A = \Lambda A \quad (2.14)$$

이 된다. 따라서

$$\Sigma_Z = A^T \Lambda A = \Lambda A^T A = \Lambda \quad (2.15)$$

가 된다. 여기서 Λ 는 대각 행렬이므로 Σ_Z 는 Σ 의 고유치를 가지는 대각행렬이 된다. Σ_Z 의 대각 성분은 얼굴 공간에서 학습하는 얼굴의 분산을 나타내기 때문에 Σ 의 고유치도 역시 이 분산이다.

PCA를 사용하는 이유는 처리하는 공간상에서 차원을 줄이는 것이다. 주성분의

$$\text{Target Face} - \text{Mean Face} = a * \text{Eigenface 1} + b * \text{Eigenface 2} + c * \text{Eigenface 3} + \dots$$

그림 2.2 고유 얼굴

최대 수는 원래의 공간상의 변수의 수이다. 그러나 차원을 줄이기 위해서 몇몇의 주성분들은 생략되어질 수 있다. 물론 차원을 더 줄이게 되면 원래의 얼굴의 정보 일부분을 잃어버리게 되지만 이것이 인식 문제에서는 중요하지 않다. 즉 고유치를 감소하는 순으로 배열했을 때, 마지막 몇 개의 고유치를 생략하면 큰 오차 없이 복원될 수 있다. 다른 의미로 인식의 큰 오차 없이 차원을 줄일 수 있다는 것이다. 이것이 PCA 방법에서 실질적으로 차원을 줄이는 것이다. 특히 PCA에서 고유치는 지수함수처럼 급속히 감소하므로 마지막 여러 주성분을 생략하더라도 오차가 매우 작다. PCA는 학습하는 벡터들에 의해서 표현되어진 공간의 기저들을 계산하는 것이다. PCA에 의해 계산되어진 기저 벡터는 학습하는 벡터의 변위를 가장 크게 하는 방향이 된다. 이들 기저 벡터가 고유벡터이며 그림 2.2와 같이 영상으로 보일 수 있다. 그리고 이것들은 얼굴과 비슷하며 고유 얼굴이라 한다. 첫 번째는 얼굴의 평균값이다. 그리고 나머지는 이 평균값으로부터 변화를 나타내는 고유 얼굴이다. 학습하는 벡터에서 평균값으로부터 가장 큰 변화는 첫 번째 고유 얼굴에 의해 묘사된다. 두 번째 고유 얼굴은 두 번째로 큰 변화를 나타낸다. 어떤 얼굴 영상이 얼굴 공간에 투사될 때 얼굴 공간에서의 가중치는 그 얼굴에서 고유 벡터의 변화치를 나타낸다. 그림 2.2는 어떤 얼굴을 얼굴공간으로 변환한 것을 나타낸다. 얼굴은 얼굴공간에서 고유 얼굴의 가중치에 의하여 나타낸다. 일단 이 과정을 끝내면 그 다음 인식 과정을 하게 된다. 일단 고유 얼굴들이 계산되면 먼저 데이터베이스에 있는 얼굴을 얼굴 공간으로 변환하고 그것의 성분을 메모리에 저장한다. 이 상태에서 인식할 얼굴 영상을 입력으로 넣으면, 이 시스템은 그것을 얼굴공간에 투사시킨다. 그리고 저장되어진 모든 얼굴의 변환된 값과 거리를 계산한다. 이 때 그 얼굴은 얼굴 공간에서 가장 가까운 거리를 갖는 얼굴과 일치한다고 가정한다.

표 2.1 포즈 변화가 존재하는 데이터베이스

KISA	YALE	CMU
43 illumination combinations	64 light cones	21 flashes
total 7 pose (0°, 0°) (±15°, 0°) (±30°, 0°) (±45°, 0°)	total 9 pose (0°, ±12°) (9°, ±9°) (12°, 0°) (17°, ±17°) (24°, 0°)	total 13 pose (0°, 0°) (±22.5°, 0°) (±45°, 0°) (±67.5°, 0°) (±90°, 0°) (0°, ±22.5°)
15° horizontal gap no vertical variance	five pose at 12° three poses at 24° from camera's axis	22.5° horizontal gap in row.
1000 people (52,000 images)	10 people (4500 images)	68 people (41,368 images)

제 2-2 절 다중영상 얼굴인식 시스템 데이터베이스

다중영상 얼굴인식 시스템의 실험을 위해 연구기관이나 학교를 중심으로 특별히 제작된 환경에서 다양한 포즈의 데이터베이스가 구축되었다. 하지만 아직도 다양한 포즈와 충분한 개수의 데이터베이스의 구축은 어려운 일이다. 가장 큰 이유는 동일한 환경에서 한사람의 다양한 포즈의 데이터베이스를 획득하기가 매우 어려운 작업이기 때문이다. 표 2.1은 포즈 변화가 있는 대표적인 얼굴 데이터베이스의 예이다.

개인의 다양한 포즈의 데이터베이스를 영상획득 장치를 통해 직접 획득하기가 힘든 일이기 때문에 소수의 영상 또는 3D morphable 모델을 이용하여 다중영상을 획득하는 연구가 진행되었다. T. Vetter [56]은 소수의 영상으로부터 다양한 포즈

변화를 포함하고 있는 영상을 합성해 냈다. 그러나 이들이 제시한 기법은 특정얼굴을 기본얼굴들의 선형결합으로 생성할 수 있다는 가정을 전제하기 때문에 얼굴인식에 적합한 고화질의 얼굴 영상 생성에는 문제가 있다. Blanz [30]은 3D morphable 모델을 사용하여 다양한 포즈와 표정의 영상을 만들어냈다. 이 기법은 높은 정확도를 가진 얼굴을 생성해내지만 수행시간, 계산량의 증가 등의 문제로 실용적이지 못하다.

제 2-3 절 다중영상 얼굴인식 시스템

얼굴인식 시스템의 가장 큰 문제점중의 하나로 포즈의 변화에 따른 인식률의 저하를 들 수 있다. 단일영상 기반의 얼굴인식 시스템은 포즈변화에 따른 인식률의 저하를 줄이기 위하여 다양한 방법이 제안되었지만 그 한계를 보이고 있다. 그러나 다중영상을 이용할 경우 포즈변화에 따른 인식률의 저하를 감소시킬 수 있는 다양한 가능성들이 있다.

다중영상을 이용한 얼굴인식 시스템은 크게 단일 분류기 시스템과 다중 분류기 시스템으로 구분할 수 있다. 다중 분류기 기반의 시스템은 특정한 포즈를 잘 나타내는 다수의 분류기를 제작하고 이로부터 다중영상 입력의 특징을 추출하고 이를 이용하여 인식을 수행한다[57]. 그림 2.3은 다중 분류기 기반 시스템의 개략도를 나타낸 것이다. 이에 반해 단일 분류기 기반의 시스템은 하나의 분류기만을 이용하여 다중영상 입력의 특징을 추출하고 이를 이용하여 인식을 수행한다[58]. 그림 2.4는 단일 분류기 기반 시스템의 개략도를 나타낸 것이다. 단일 분류기 기반의 시스템은 다중 분류기 기반의 시스템에 비하여 하나의 분류기만을 이용하기 때문에 시스템의 복잡도가 단순하고 수행시간이 적은 장점이 있지만 단일 분류기만으로 모든 포즈에 대한 특징을 추출하기 때문에 인식률이 낮은 단점이 있다.

기존의 다중영상 기반의 얼굴인식 시스템은 단일영상 얼굴인식 시스템을 기반으로 한다. 즉 다중영상 시스템은 각각의 얼굴영상을 단일영상 얼굴인식의 사용된 인식 기법을 적용하고 이들 결과 값의 조합으로 인식을 수행하였다. 그러나 이러

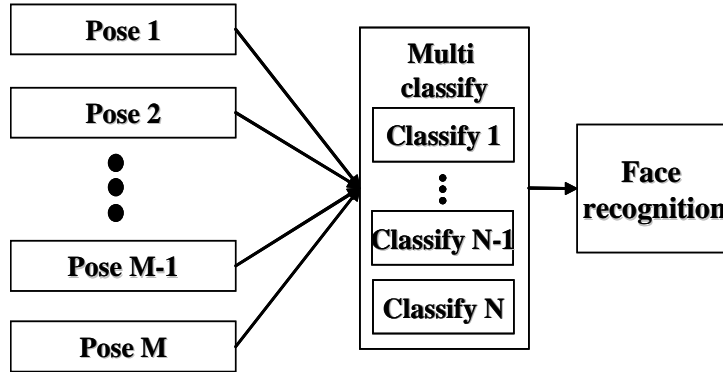


그림 2.3 다중 분류기 기반의 다중영상 얼굴인식

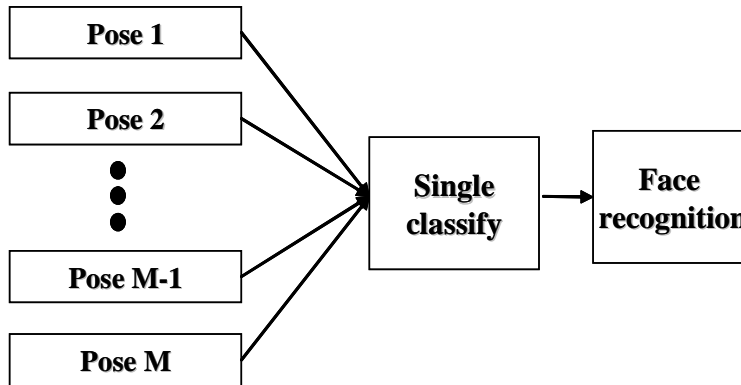


그림 2.4 단일 분류기 기반의 다중영상 얼굴인식

한 방식은 시스템을 다중영상 기반으로 확장했지만 다중영상에 존재하는 특징을 이용하지 못하고 단일영상 기반의 얼굴인식에 사용된 기법의 단순한 반복에 지나지 않는다. 최근의 다중영상 얼굴인식 시스템은 다중영상에 존재하는 특징을 이용하여 단일 분류기만을 사용하더라도 높은 인식률을 나타내고 있다. Yamaguchi [59]는 개개의 다중입력 영상에 대하여 매칭을 수행하고 이를 조합하는 방식 대신 전체 다중입력 영상을 하나의 입력으로 간주하고 이를 이용하여 매칭을 수행하였다 (Mutual subspace method). 그림 2.5에서 보듯이 다중입력 영상을 이용하여 부분 공간을 형성하고 데이터베이스의 부분 공간과의 사잇각을 측정하여 인식을 수

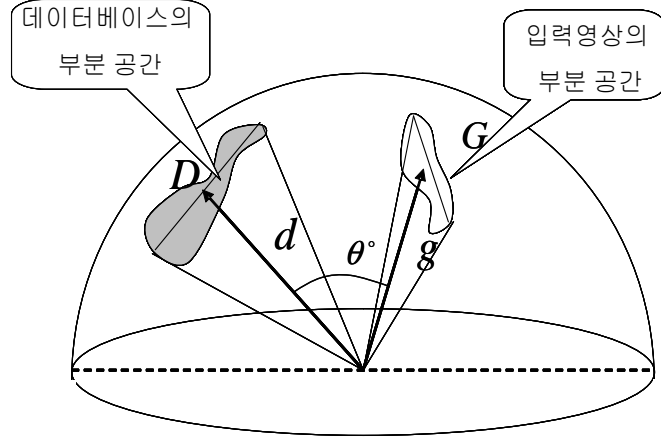


그림 2.5 상호 부분 공간 기법

행하였다. 입력 영상으로 구성된 부분공간 G 와 데이터베이스의 부분공간 D 와의 사잇각은 식 (2.16)과 같이 측정한다. g, d 는 각각 부분공간 G 와 D 의 기저 벡터이다.

$$\cos^2\theta = \text{SUP} \frac{|d \cdot g|}{\|d\|^2 \|g\|^2}, \quad d \in D, g \in G, \|d\| \neq 0, \|g\| \neq 0 \quad (2.16)$$

또한 Yongmin Li [60]은 연속된 다중영상 입력으로부터 “dynamic face”를 모델링하고 이를 이용하여 다중입력 영상을 궤적으로 변화시켜 데이터베이스의 궤적과의 매칭을 통해 인식을 수행하였다. 그림 2.6은 특징공간상에 궤적의 매칭을 나타내고 있다. 입력 영상의 궤적과 m 번째 데이터베이스의 궤적과의 유사도 측정은 식 (2.17)과 같다.

$$D_m = \sum_{l=1}^L w^l d_m^l \quad (2.17)$$

여기서 L 은 입력 영상의 개수이고 d_m^l 은 입력궤적상의 l 번째 포즈영상과 동일한

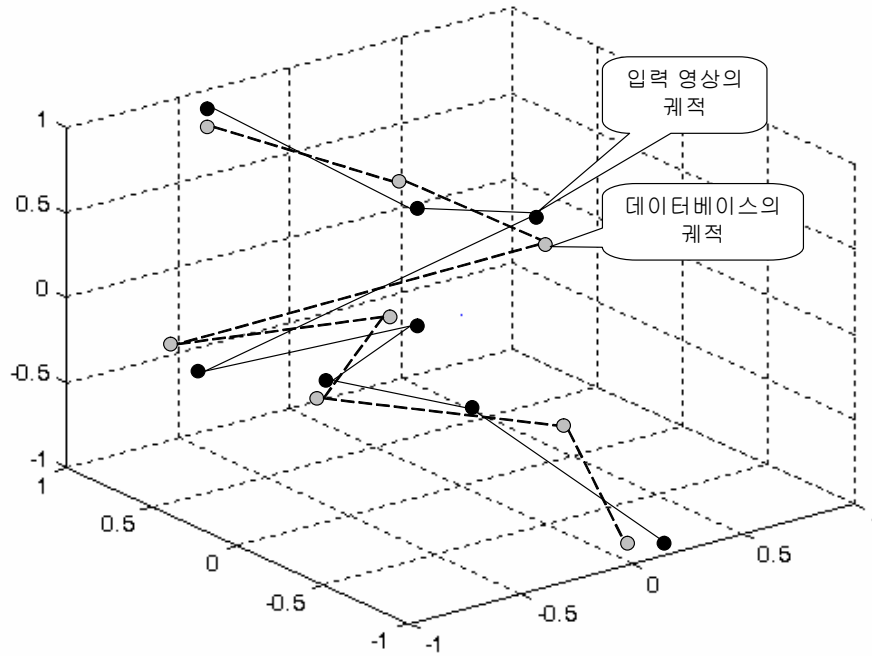


그림 2.6 다중영상의 궤적을 이용한 매칭

포즈의 m 번째 데이터베이스 궤적위의 영상과의 Euclidean 거리이다. w^l 은 궤적의 변화율을 고려하여 결정한다.

제 3 장 포즈 추이 평면을 이용한 다중영상 얼굴인식

본 장에서는 논문에서 제안하는 포즈변화에 강인한 다중영상 기반의 얼굴인식 시스템에 대해서 자세히 서술하였다. 제 3-1 절에서 전체 시스템의 개요에 대해 살펴보고, 제 3-2 절에서는 학습과 데이터베이스에 사용된 연속된 포즈변화가 있는 얼굴영상을 생성하는 과정에 대해 기술한다. 제 3-3 절에서는 포즈 추이 평면을 정의한다. 또한 기존의 다중영상 기반의 얼굴인식 시스템에서 사용된 오차측정 기법과 포즈 추이 평면을 이용한 오차측정 기법에 대하여 비교 서술한다.

제 3-1 절 시스템 개요

본 절에서는 포즈 변화에 강인한 다중영상 기반의 얼굴인식 시스템의 개요에 대하여 기술한다. 이 시스템은 3D 센서로부터 3차원 얼굴데이터를 획득하고 이를 이용하여 연속적으로 포즈가 변화하는 2D 얼굴영상들을 생성한다. 개인당 좌 90° 에서 우 90° 까지 5° 간격으로 포즈가 변화된 37장의 얼굴 영상을 이용하여 PCA 학습과 데이터베이스로 사용되었다. 입력 영상으로는 7대의 일반 CCD 카메라로 구성된 다중영상시스템으로부터 획득된 7장의 포즈가 서로 다른 얼굴영상을 사용하였다. 본 논문에서는 다중영상의 특징을 이용한 다중영상 기반의 얼굴인식에 적합한 인식 방법에 초점을 맞추고 있으므로 입력 영상의 포즈를 추정하는 단계는 수동으로 해결하였다. 정확한 포즈를 알고 있는 7장의 얼굴영상은 미리 학습된 PCA 공간상에 투영되고 본 논문에서 제안한 포즈 추이 평면을 이용한 오차측정 기법을 통하여 인식을 수행한다. 그림 3.1은 제안한 다중영상 기반의 얼굴인식 시스템의 개략도이다.

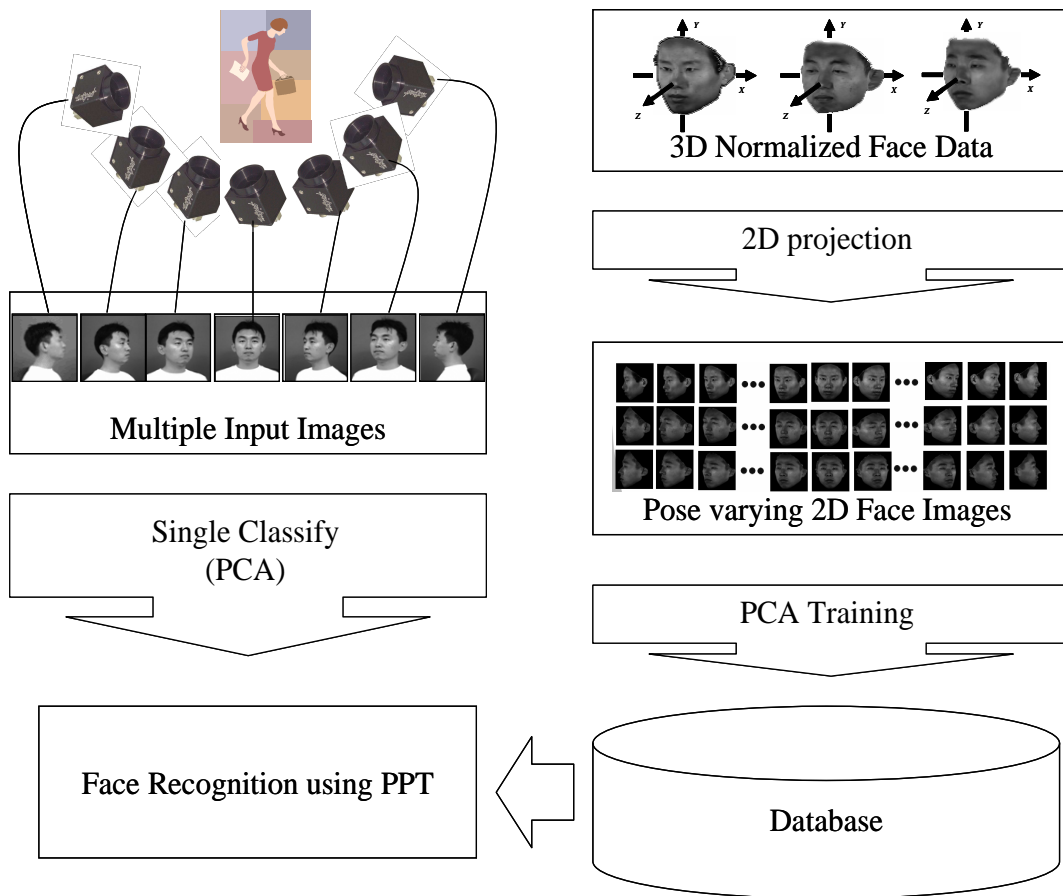


그림 3.1 포즈 추이 평면을 이용한 시스템 개략도

제 3-2 절 연속적인 포즈의 얼굴영상 획득

본 절에서는 다중영상 기반의 얼굴인식 시스템에서 학습과 데이터베이스로 사용된 연속적으로 변하는 포즈의 얼굴 영상획득에 대해서 기술한다. 쉽고도 정확하게 원하는 포즈의 2D 얼굴영상을 획득하기 위하여 본 논문에서는 3차원 센서를 이용하여 획득된 3차원 얼굴 데이터를 사용한다. 잡음 제거와 정규화 과정을 거친 3차원 얼굴 데이터를 개인당 좌 90° 에서 우 90° 까지 5° 간격으로 포즈를 변화시키면서 2차원 영상평면으로 투영시킨다.

3.2.1 3차원 얼굴 획득

3차원 얼굴 데이터는 얼굴인식 시스템에서 입력이 아닌 데이터베이스로써 사용되므로 정확한 데이터가 요구되기 때문에 모델링에 의한 간접적 방법이 아닌 3차원 센서를 이용하여 직접적으로 얼굴 데이터를 획득한다. 본 논문에서 3차원 얼굴 데이터 획득을 위해 레이저를 이용한 3차원 센서를 이용하였다. 레이저를 이용하기 때문에 정확한 값을 측정할 수 있지만, 주변 빛의 세기에 따라서 빛의 간섭 현상에 의한 오차가 발생하기도 한다. 이런 오차가 포함되지 않은 정확한 데이터로 데이터베이스를 생성하기 위해서 주변의 빛을 차단하기 위한 스튜디오를 구성하였으며, 모든 3차원 얼굴 데이터를 이 스튜디오 내에서 획득하였다.

획득한 3차원 얼굴 데이터는 얼굴을 약 10만개의 점으로 표현한다. 각각의 점은 스캐너 자체 기준의 절대 좌표 값과 색상 정보 값을 가지고 있다. 3차원 스캐너에 의해 획득한 얼굴 데이터는 완벽한 3차원 모델이기 때문에 3차원 공간 이동 연산에 의해서 임의의 포즈를 발생시킬 수 있으며, Lambertian reflectance 모델을 적용할 경우 조명 변화까지 임의로 조정할 수 있다.

3.2.2 3차원 얼굴 데이터 정규화

3차원 센서로 획득된 3차원 얼굴 데이터는 기하학적인 위치 정보와 이에 해당하는 색깔 정보로 구성되어있다. 사람에 따라 얼굴의 기하학적인 크기와 위치가 다양하기 때문에 정규화 과정을 거치지 않은 3차원 얼굴 데이터를 이용하여 학습, 데이터베이스로 사용할 경우 인식률의 저하가 발생된다. 강인한 얼굴인식 시스템을 위해서는 3차원 얼굴 데이터는 반드시 정규화 과정이 필요하다. 본 논문에서는 H. Song [61]논문에서 제안한 정규화를 이용하여 정규화 된 3D 얼굴 데이터를 획득하였다. 정규화된 3차원 얼굴 공간을 정의하고 모든 얼굴 입력 영상을 동일 공간으로 정규화 하여, 스케일 변화를 보상시킬 수 있는 과정을 보인다. Cartesian 좌표계의 점들의 집합인 입력 3차원 얼굴 데이터 F 가 주어졌을 때, x -, y -, z -축의 범위는 다양하게 변화되므로, 입력 데이터를 각 축으로 정해진 범위 내로 정규화

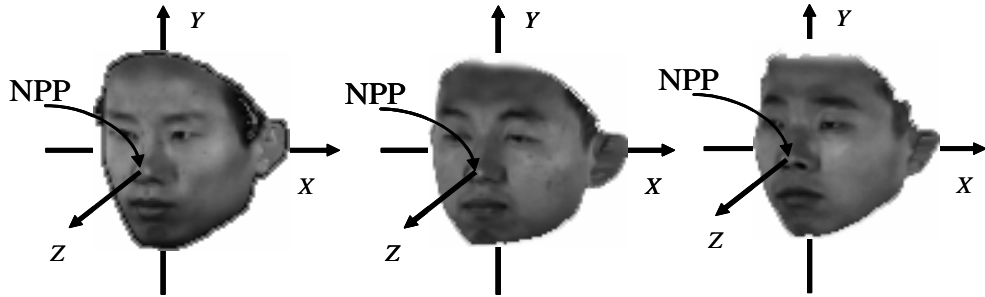


그림 3.2 정규화된 3차원 얼굴 데이터

시키는 과정이 필요하다. 이렇게 설정된 공간을 정규화된 3차원 얼굴 공간이라 정의한다. 먼저 깊이 정보를 정해진 범위 내로 정규화 시킨 후 비례적으로 x 와 y 축에 대해서 정규화 과정을 수행한다. 얼굴 영역에서 다른 구성요소와 가장 구별이 되는 요소는 코이다. 가장 큰 깊이 값을 가지는 부분이 코끝이기 때문에 Z -buffer 알고리즘을 사용하여 코 끝점을 추출하고, 그 점을 $(0, 0, Z)$ 로 정규화 시킨다. x -축과 y -축에 대해서도 동일한 과정을 수행한다. 각 축에 제한된 범위 $[-\sigma, \sigma]$, $[-\epsilon, \epsilon]$, $[0, Z]$ 로서 다음 식 (3.1)로부터 얻을 수 있다.

$$F(x_i, y_i, z_i) = \left(\frac{F_x - F_{\min-x}}{F_{\max-x} - F_{\min-x}} \times \sigma, \frac{F_y - F_{\min-y}}{F_{\max-y} - F_{\min-y}} \times \epsilon, \frac{F_z - F_{\min-z}}{F_{\max-z} - F_{\min-z}} \times Z \right) \quad (3.1)$$

식(3.1)에서 $F = (F_x, F_y, F_z)$ 는 입력 데이터의 공간 좌표이다. $F_{\max-x}, F_{\max-y}$ 와 $F_{\max-z}$ 는 입력 데이터의 각각의 축에 대하여 최대값이고 $F_{\min-x}, F_{\min-y}$ 와 $F_{\min-z}$ 는 입력 데이터의 각각의 축에 대하여 최소값이다. 이와 같은 과정을 마치고 정규화된 3차원 얼굴 데이터에 대한 예를 그림 3.2에 나타내었다.

3.2.3 다양한 포즈의 2차원 얼굴영상 획득

연속적으로 변화하는 포즈의 얼굴 영상을 얻기 위하여 정규화된 3차원 얼굴

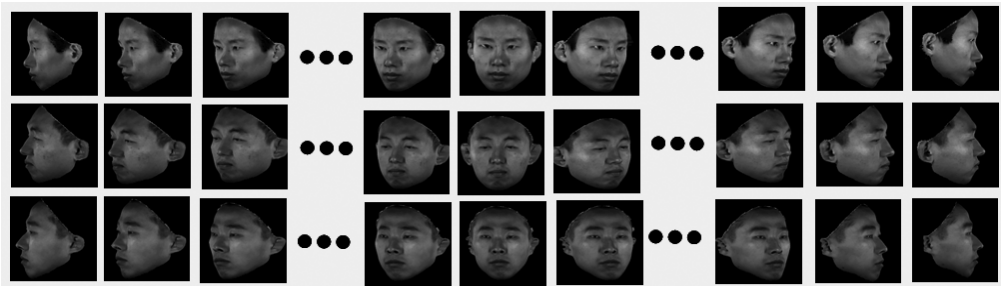


그림 3.3 연속적으로 변화하는 포즈의 얼굴 영상

데이터를 이용하여 개인당 좌 90° 에서 우 90° 까지 5° 간격 회전시키면서 2차원 영상평면으로 투영시킨다. 이와 같은 과정을 거쳐 한 사람당 37장의 연속적으로 포즈가 변하는 얼굴 영상을 획득할 수 있다. 이와 같은 방법은 완벽한 3차원 얼굴 데이터로부터 2차원 얼굴 영상을 획득하기 때문에 사용자가 원하는 포즈를 정확하고 쉽게 획득할 수 있는 커다란 장점이 있다. 그림 3.3은 3차원 얼굴 데이터로부터 획득된 개인당 37장의 얼굴영상의 예이다.

3.2.4 선형기반의 밝기 보정

PCA에 의한 얼굴영상의 표현은 조명의 영향에 따라 많은 영향을 받게 된다. 즉 조명 변화에 의한 화소 값의 변화에 의해 특징벡터의 값이 많은 영향을 받게 된다. 이런 조명 문제를 해결하기 위하여 선형 변환 기반 밝기 보정기법을 사용하였다. 이 방법은 영상의 밝기를 보정하는 방법이다. 기준 영상과 참조 영상의 화소 값의 평균과 분산을 동일하게 해주는 선형 변환을 찾아서 적용하는 방법이다. 현재 영상의 휘도 값을 I_C , 선형 변환된 뒤의 휘도 값을 I'_C 라고 하면 식 (3.2)의 관계를 가지게 된다.

$$I'_C = aI_C + b \quad (3.2)$$

그러므로 I'_C 의 평균 μ'_C 과 분산 σ'^2_C 은 식 (3.3)과 같아진다.

$$\begin{aligned}\mu_C' &= a\mu_C + b \\ \sigma_C'^2 &= a^2\sigma_C^2\end{aligned}\tag{3.3}$$

식 (3.3)을 a 와 b 에 대해서 정리하면 식 (3.4)이 된다.

$$\begin{aligned}a &= \frac{\sigma_C'}{\sigma_C} \\ b &= \mu_C' - a\mu_C\end{aligned}\tag{3.4}$$

여기서 $\mu_C' = \mu_R$, $\sigma_C' = \sigma_R$ 의 조건을 만족해야 하므로 결국 a 와 b 는 식 (3.5)의 값을 가지게 된다.

$$\begin{aligned}a &= \frac{\sigma_R}{\sigma_C} \\ b &= \mu_R - a\mu_C\end{aligned}\tag{3.5}$$

그리고 최종적으로 I_C' 는 식 (3.6)의 값을 가지게 된다.

$$I_C' = \frac{\sigma_R}{\sigma_C}I_C + \left(\mu_R - \frac{\sigma_R}{\sigma_C}\mu_C\right)\tag{3.6}$$

이렇게 전체적인 영상의 밝기를 맞추어 줌으로써 영상 전반에 걸친 조명효과를 보상해 줄 수 있다. 이와 같은 방법은 영상 전체를 a 와 b 로 나타내기 때문에 전체적인 밝기 보상에선 좋으나 부분 조명에 의한 조명 효과를 보상하지는 못한다[62].

제 3-3 절 포즈 추이 평면을 이용한 얼굴인식

본 절에서는 연속적으로 포즈가 변하는 다중영상을 이용하여 포즈 추이 평면을 정의한다. 또한 기존의 다중영상 기반의 얼굴인식 시스템에서 사용된 오차측정 기법과 포즈 추이 평면을 이용한 오차측정 기법에 대해 비교 서술한다.

3.3.1 얼굴 영상의 표현

본 논문에서는 PCA 기법을 적용하여 2차원 얼굴상을 3차원 특징벡터로 표현하였다. 즉 2차원 얼굴영상은 3차원 PCA공간상의 하나의 점으로 표현된다. 2차원 얼굴 영상 I 가 주어졌을 때, 학습된 3차원 PCA 공간상에 투영시키면 식 (3.7)과 같이 3개의 성분을 가진 벡터로 표현된다.

$$F(I) = (\alpha, \beta, \gamma) \quad (3.7)$$

여기서 F 는 PCA공간으로의 투영시키는 함수이다. α, β, γ 는 PCA 공간을 구성하는 3개의 PCA축에 영상을 투영시킨 값이다. 일반적으로 고차원의 얼굴영상을 3차원 특징벡터로 표현할 경우 인식률이 저하된다. 그러나 연속적으로 변화하는 포즈의 다중영상을 이용할 경우 이와 같은 문제점을 해결할 수 있다. 본 논문에서는 제 3-2절 에서 획득한 연속적으로 변화하는 포즈의 다중영상을 3차원 PCA 공간상에 투영시킴으로써 데이터베이스를 구축한다. 개인당 37장의 연속적으로 포즈가 변하는 얼굴 영상은 3차원 PCA 공간상에 하나의 궤적을 형성한다. 그림 3.4는 연속적으로 변화하는 포즈의 다중영상이 3차원 PCA 공간상에 궤적으로 표현됨을 보여준다. 궤적위의 하나의 점은 특정 포즈에 해당하는 얼굴 영상이다.

3.3.2 포즈 추이 평면을 이용한 오차측정

기존의 다중영상 얼굴인식 시스템은 다음 3단계의 과정을 통해 얼굴인식을 수행한다.

- 1) 입력 영상의 얼굴 포즈를 추정한다.
- 2) 입력 영상을 정의된 특징공간으로 투영시킨다.
- 3) 입력 영상의 특징점과 데이터 베이스상의 동일한 포즈의 특징점간의 Euclidean 거리를 계산하고 이중 가장 작은 거리의 데이터베이스를 인식하고자 하는 사람으로 결정한다. 그림 3.5는 기존의 다중영상 기반의 얼굴인식 시스템에서 사용된 오

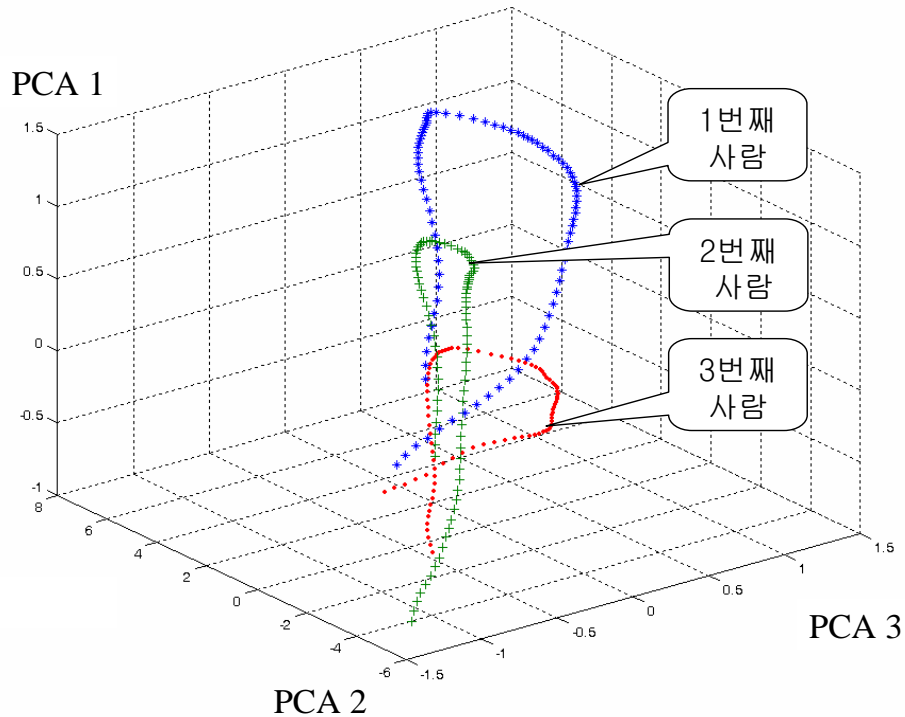


그림 3.4 연속적으로 포즈가 변하는 다중영상의 궤적

차 측정방법이다. 35° 의 포즈로 추정된 입력 영상이 3차원 PCA 공간에 투영되면 동일 공간상에 하나의 점으로 표현된다. 데이터베이스와의 오차측정을 위해 각각의 사람의 동일한 포즈의 점들과 Euclidean 거리를 측정한다. D_1, D_2 는 각각 입력 영상과 1번째 사람과 2번째 사람과의 거리를 나타낸다. 이중 가장 작은 거리를 나타내는 사람을 찾고자 하는 인물로 결정함으로써 인식과정을 수행한다. 이를 수학적으로 나타내면 다음과 같다. L 장의 포즈가 변하는 입력 영상을 가정할 때 입력 영상과 m 번째 사람과의 거리는 식 (3.8)과 같다.

$$D_m = \sum_{l=1}^L d_{point}^l \quad (3.8)$$

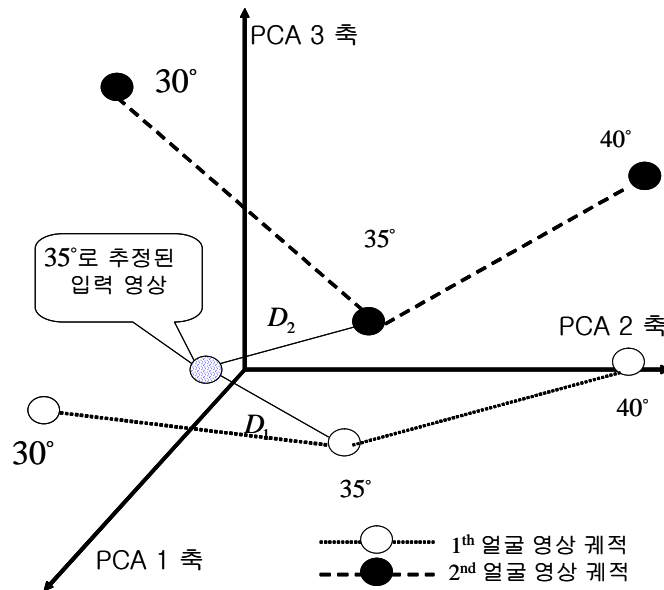


그림 3.5 기존 다중영상 얼굴인식 시스템의 오차측정

위 식에서 d_{point}^l 은 다음과 같이 구한다. θ° 의 포즈로 추정된 l 번째 입력 영상 I 가 PCA 공간에 투영될 때 좌표 값을 $(\alpha^{\theta^\circ}, \beta^{\theta^\circ}, \gamma^{\theta^\circ})$ 라 가정한다. m 번째 사람 궤적상위의 동일한 포즈의 점의 좌표를 $(\alpha_m^{\theta^\circ}, \beta_m^{\theta^\circ}, \gamma_m^{\theta^\circ})$ 라 하면 d_{point}^l 은 식 (3.9)로 구해진다.

$$d_{point}^l = \sqrt{(\alpha^{\theta^\circ} - \alpha_m^{\theta^\circ})^2 + (\beta^{\theta^\circ} - \beta_m^{\theta^\circ})^2 + (\gamma^{\theta^\circ} - \gamma_m^{\theta^\circ})^2} \quad (3.9)$$

그러나 이와 같은 기존의 다중영상 기반의 얼굴인식 시스템은 다중영상에 존재하는 특징을 이용하지 못하기 때문에 인식률 측면에서 신뢰할 만한 수준의 성능을 기대하기 어렵다. 또한 입력 영상의 포즈를 정확히 측정했다는 가정을 전제한다. 실제로 포즈를 정확히 측정하기란 어려운 일이다. 결국 이러한 요소들은 인식률의 저하를 초래한다.

이러한 문제점을 해결할 수 있는 방법으로 35°로 추정된 입력영상에 대하여 데이터베이스의 동일한 포즈에 해당하는 점과 $\pm 5^\circ$ 의 점들중 Euclidean 거리가

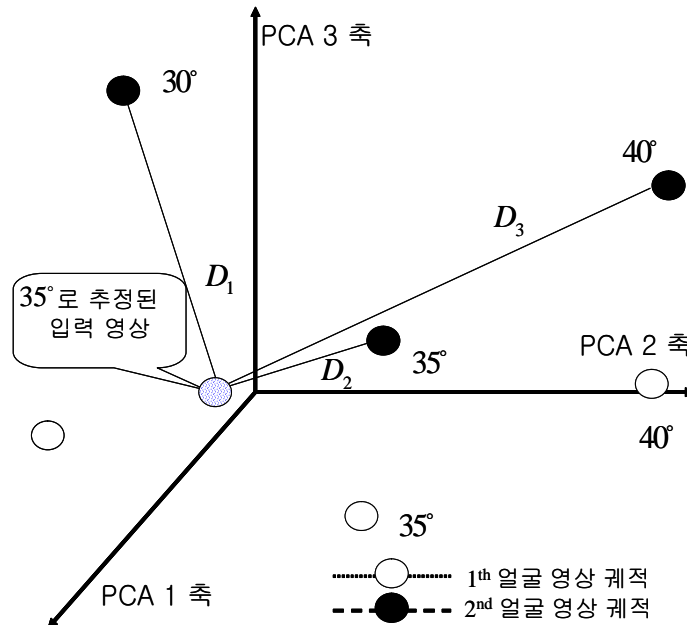


그림 3.6 최소 거리값을 이용한 인식 방법

가장 적은 값을 이용하여 인식을 수행하는 방법이다. 그림 3.6과 같이 35°로 추정된 입력 영상에 대하여 30°, 35°, 40°에 해당하는 거리 D_1, D_2, D_3 을 측정하고 이중 가장 작은 거리값을 입력 영상과 1th 얼굴 영상과의 오차값으로 결정한다. 최종적인 인식을 위하여 모든 데이터베이스의 얼굴 영상과의 오차값을 결정하고 이 들중 가장 작은 오차값을 가지는 얼굴 영상을 인식하고자 하는 사람으로 결정한다. 하지만 이러한 방법은 입력영상의 포즈가 추정된 포즈와 정확히 $\pm 5^\circ$ 오차가 있다는 전제를 가정한다. 하지만 실제로 입력영상의 포즈는 추정된 포즈와 다양한 정도의 오차가 존재하기 때문에 이와 같은 방법도 인식률의 효과적인 개선을 기대할 수 없다.

본 논문에서 연속적으로 포즈가 변하는 영상의 특징을 효율적으로 사용하기 위하여 포즈 추이 평면을 정의하고 이를 이용한 오차측정 기법을 제안한다. 그림 3.7은 포즈 추이 평면을 이용한 오차측정 방법이다. 35°로 추정된 입력 영상에

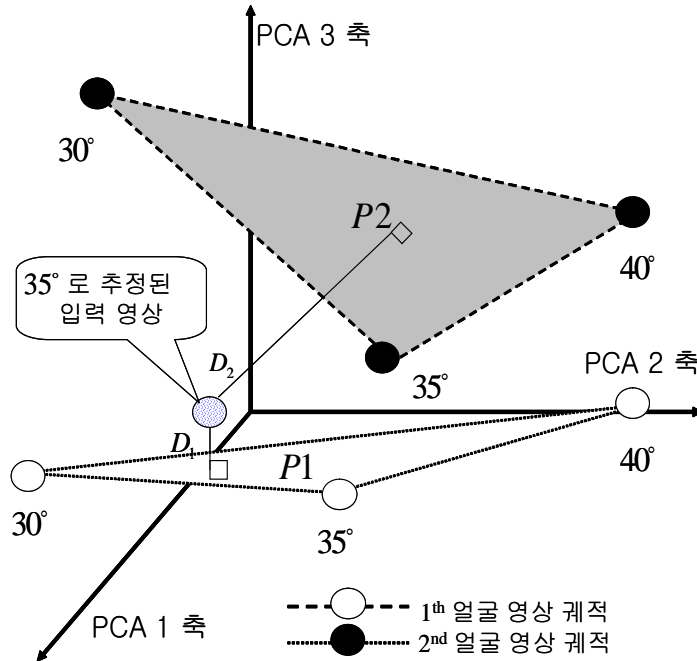


그림 3.7 포즈 추이 평면을 이용한 오차측정

대하여 데이터베이스의 동일한 포즈에 해당하는 점과 $\pm 5^\circ$ 의 점들을 이용하여 하나의 평면을 만든다. 본 논문에서는 이평면을 포즈 추이 평면이라고 정의한다. $P1, P2$ 는 입력 영상과 동일한 포즈의 1번째 2번째 사람의 포즈 추이 평면이다. 3차원 공간상에서 연속적으로 포즈가 변하는 얼굴영상의 위치의 변화를 선형이라 가정할 때 포즈 추이 평면은 이 변화의 추이선을 양변으로 한다. 고확률 얼굴 평면은 포즈가 변하면서 발생하는 추이선을 포함하고 있기 때문에 포즈 추정단계에서 발생한 오차에 매우 강인하다.

포즈 추이 평면을 이용한 인식 과정은 다음과 같다. 입력 영상의 포즈를 추정한 후 이와 동일한 포즈의 데이터베이스의 각각의 사람과의 포즈 추이 평면과의 Euclidean 거리를 측정한다. D_1, D_2 는 입력 영상의 특징점과 각각의 포즈 추이 평면과의 거리이다. 이중 최소의 거리 값을 갖는 사람을 찾고자 하는 인물로 결정하여 인식과정을 마친다. L 장의 포즈가 변하는 입력 영상을 가정할 때 입력 영상과

m 번째 사람의 포즈 추이 평면과의 거리는 식 (3.10)과 같다.

$$D_m = \sum_{l=1}^L d_{plane}^l \quad (3.10)$$

위 식에서 L 은 입력 영상의 개수이고 d_{plane}^l 은 식(3.11)으로부터 구한다.

$$d_{plane}^l = \left| \frac{\overrightarrow{N_m^l} \cdot \overrightarrow{R_m^l X_{input}^l}}{|\overrightarrow{N_m^l}|} \right| \quad (3.11)$$

θ° 의 포즈로 추정된 l 번째 입력 영상 I 가 PCA 공간에 투영될 때 좌표 값을 $(\alpha^{\theta^\circ}, \beta^{\theta^\circ}, \gamma^{\theta^\circ})$ 라 가정한다. m 번째 사람의 궤적위에 동일한 포즈의 점의 좌표와 $+5^\circ$ 포즈의 점 그리고 -5° 인 점의 좌표를 각각 $(\alpha_m^{\theta^\circ}, \beta_m^{\theta^\circ}, \gamma_m^{\theta^\circ})$, $(\alpha_m^{\theta^\circ+5^\circ}, \beta_m^{\theta^\circ+5^\circ}, \gamma_m^{\theta^\circ+5^\circ})$ 그리고 $(\alpha_m^{\theta^\circ-5^\circ}, \beta_m^{\theta^\circ-5^\circ}, \gamma_m^{\theta^\circ-5^\circ})$ 라 하면 이 세 점은 3차원 PCA 공간에서 하나의 평면 P_m^l 을 정의할 수 있다. $X_{input}^l = (\alpha_m^{\theta^\circ}, \beta_m^{\theta^\circ}, \gamma_m^{\theta^\circ})$ 이다. $\overrightarrow{N_m^l}$ 은 평면 P_m^l 의 법선벡터이고 R_m^l 은 동일평면 상의 임의의 점이다.

제 4 장 실험 및 결과분석

본 장에서는 실험 환경에 대하여 기술하고 본 논문에서 제안하는 다중영상 기반의 얼굴인식 시스템의 성능을 다양한 실험을 통하여 검증한다. 제 4-1 절에서는 실험에 사용된 장비에 대하여 기술하고, 제 4-2 절에서는 기존의 다중영상 기반의 얼굴인식 시스템과 제안한 시스템의 인식률을 비교 분석하였다. 또한 포즈변화에 따른 인식률을 비교 분석함으로써 제안한 다중영상 기반의 얼굴인식 시스템의 성능을 검증하였다.

제 4-1 절 실험 환경

본 절에서는 실험에 사용된 장비에 대하여 기술한다. 1024 Mbyte 메모리의 Intel Pentium(R) D CPU 3.20 GHz 환경에서 Visual C++ 6.0을 사용하여 수행하였다. 실험에서 학습과 데이터베이스 구축에 사용된 연속적으로 포즈가 변하는 다중영상은 한국 과학재단의 지원을 받아 연세대학교 생체인식 연구센터에서 생성한 3차원 얼굴 데이터로부터 생성하였다. 3차원 얼굴 데이터는 3차원 레이저 스캐너로 획득되었으며, 획득된 3차원 얼굴 데이터는 정규화 과정을 거친 후 좌 90°에서 우 90°까지 5° 간격으로 포즈를 변화시키면서 37장의 연속적으로 포즈가 변하는 얼굴영상을 생성하였다. 그림 4.1은 3차원 얼굴 데이터를 획득하는데 사용된 3차원 레이저 스캐너이다. 3차원 레이저 스캐너의 자세한 사양은 표 4.1과 같다. 테스트를 위한 입력 영상으로는 7대의 일반 CCD 카메라로 구성된 다중영상 입력 시스템으로부터 획득된 얼굴영상을 사용하였다. 그림 4.2는 다중영상 입력 시스템을 구성하는 Flea CCD camera 이다. 자세한 사양은 표 4.2와 같다. 4.3은 입력에 사용된 얼굴영상의 예이다. 모든 영상은 100×120의 크기로 정규화 하였다.



그림 4.1 3차원 레이저 스캐너

표 4.1 3차원 레이저 스캐너 사양

Spec	3D Laser Scanner (Cyberware)
Resolution	512×512
Field of View	Theta 360, Max 340mm(Y) 300mm
3D format	OBJ, PLY, STL
Acquisition time	90sec
Vertex point	Max 100,00



그림 4.2 Flea CCD 카메라

표 4.2 Flea CCD camera 사양

Spec	2D CCD Camera (Flea)
Resolution	640×480
SNR	50dB
Frame Rates	60 fps
Sensor	Sony 1/3" Type Sony HAD CCD



그림 4.3 다중 입력 영상

제 4-2 절 실험 결과

본 절에서는 제안한 다중영상 기반의 얼굴인식 시스템의 성능을 검증하기 위하여 다음과 같은 2가지 실험을 하였다. 동일한 실험조건에서 기존의 다중영상 기반의 얼굴인식 시스템과 제안한 시스템의 인식률을 비교 분석하였고, 포즈 추정시 오차가 발생할경우의 인식률을 비교 분석함으로써 제안한 다중영상 기반의 얼굴인식 시스템의 성능을 검증하였다.

기존의 다중영상 기반의 얼굴인식 시스템과의 인식률 비교를 위하여 동일한 데이터베이스와, 입력 영상을 사용하는 조건하에서 식 (3.8)을 이용한 인식방법, 최소 거리값을 이용한 인식방법, 그리고 식 (3.10)을 적용할경우의 인식률을 비교하였다. 표 5.1에서 볼 수 있듯이 식 (3.8)을 적용할경우의 인식률은 76% (24명 오인식)이고 최소 거리값을 이용한 인식률은 85%, 그리고 포즈 추이 평면을 이용한 식 (3.10)을 적용할경우의 인식률은 91%(9명 오인식)로써 인식률의 증가를 나타내었다.

그림 5.1은 16번째 사람에 해당하는 영상을 입력으로 사용할 경우 기존의 다중영상 기반의 얼굴인식 시스템의 거리측정 결과와 제안한 다중영상 기반의 얼굴인식 시스템의 거리측정 결과이다. 실험결과에서 정규화된 거리는 최소의 거리값을 1로 정규화한 거리이다. 그림 5.1에서 볼 수 있듯이 기존의 다중영상 기반의 얼굴인식 시스템은 27번째 사람과 입력과의 거리를 최소값으로 측정하여 인식에 실패했다. 그러나 제안한 다중영상 기반의 얼굴인식 시스템은 정확히 16번째 사람과 입력과의 거리를 최소값으로 측정하여 인식에 성공하였다. 그림 5.1에서 더욱 중요한 점은 제안한 다중영상 기반의 얼굴인식 시스템에서 입력과 데이터베이스와의 거리의 분산 값이 기존의 다중영상 기반의 얼굴인식 시스템의 거리의 분산 값에 비하여 매우 크다는 점이다. 이는 포즈 추이 평면을 이용한 오차측정 방법이 입력과 동일한 사람과의 거리는 줄여주고 다른 사람들과의 거리는 증가시킴을 의미한다. 즉 포즈 추이 평면을 이용한 오차측정 기법을 사용할 경우 오인식될 확률을 줄임으로써 인식률을 향상시켰다. 그림 5.2는 9번째 사람에 해당하는 영상을

표 5.1 실험 1 인식률 결과

	NN 방법	최소 거리 방법	PPT 방법
인식률	76% (24 person fail)	85% (15 person fail)	91% (9 person fail)

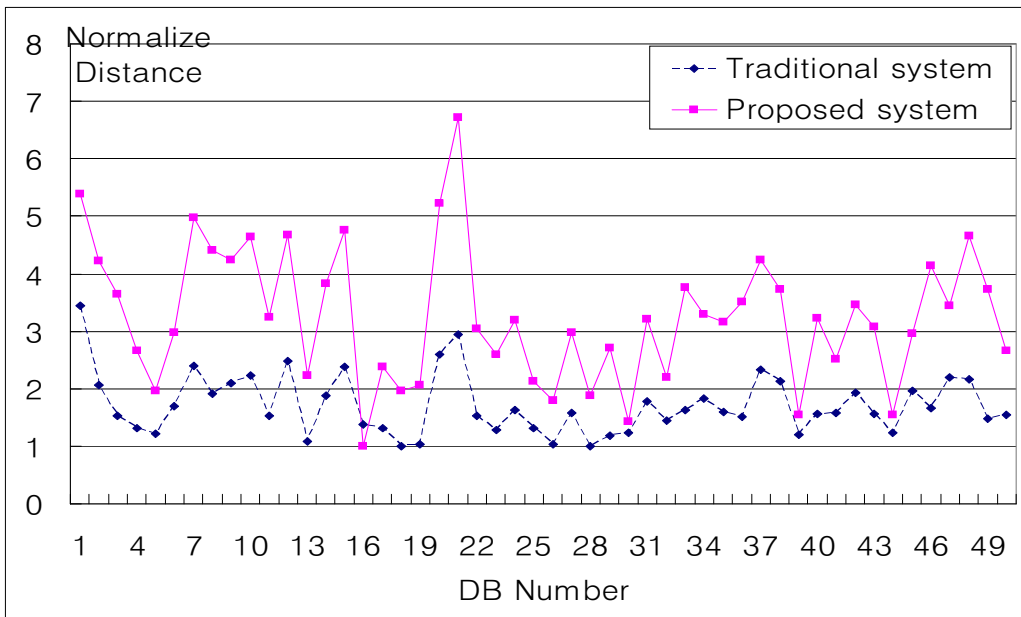


그림 5.1 16번째 사람에 해당하는 입력 영상에 대한 실험결과

입력으로 사용할 경우 기존의 다중영상 기반의 얼굴인식 시스템의 거리측정 결과와 제안한 다중영상 기반의 얼굴인식 시스템의 거리측정 결과이다. 그림 5.2에서 볼 수 있듯이 기존의 다중영상 기반의 얼굴인식 시스템과 제안한 다중영상 기반의 얼굴인식 시스템 모두 41번째 사람과 입력과의 거리를 최소로 측정하여 인식에 실패한 경우이다. 비록 제안한 다중영상 기반의 얼굴인식 시스템이 인식에 실패했지만 기존의 다중영상 얼굴인식 시스템의 실험결과와 비교할 때 제안한 시스템은 9번째 사람과의 입력과의 거리는 줄어들고 다른 사람과의 거리는 증가시켰다. 인식에 실패한 가장 큰 원인으로서는 PCA에 의한 분류성능의 한계라고 생각된다.

다. Kernal Principal Component Analsys (KPCA), Kernal Linear Discriminant Analysis (KLDA)와 같이 PCA에 비해 분류 성능이 좋은 분류기를 사용할 경우 인식률의

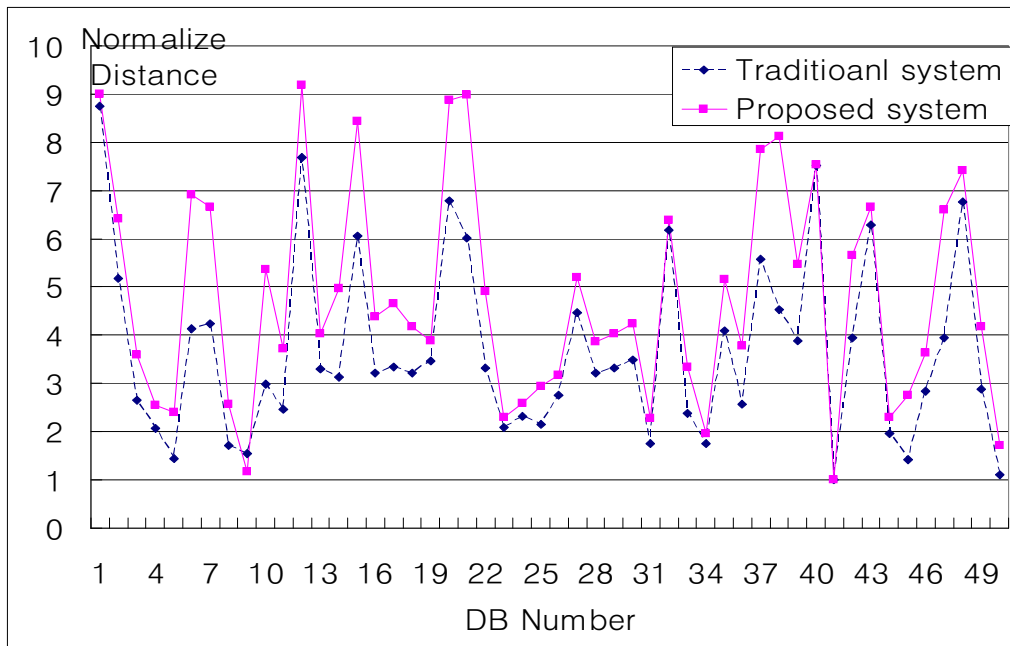


그림 5.2 9번째 사람에 해당하는 입력 영상에 대한 실험결과

향상시킬 수 있으리라 예상된다. 제안한 시스템이 포즈 추정단계에서 발생한 오차에 강인함을 보이기 위하여 다음과 같은 실험을 수행하였다. 실험 방법은 다음과 같다. 입력 영상의 포즈와 동일한 데이터베이스와의 인식률과 $\pm 5^\circ$, $\pm 10^\circ$ 만큼의 포즈 오차가 있는 데이터베이스와의 인식률을 비교하였다. 그림 5.3은 실험에 사용된 입력 영상과 동일한 포즈의 데이터베이스, $\pm 5^\circ$, $\pm 10^\circ$ 만큼의 포즈 오차가 있는 데이터베이스의 예이다. 표 5.2에서 알 수 있듯이 기존의 다중영상 시스템은 동일한 포즈의 데이터베이스와의 인식률은 76%이고 $\pm 5^\circ$, $\pm 10^\circ$ 의 포즈 추정의 오차가 있는 데이터베이스일 경우에는 각각 62%, 54%의 인식률을 나타내고 있다. 그러나 제안한 다중영상 시스템은 동일한 포즈의 데이터베이스와의 인식률은 91%이고 $\pm 5^\circ$, $\pm 10^\circ$ 의 포즈 추정의 오차가 있는 데이터베이스일 경우에는 각각

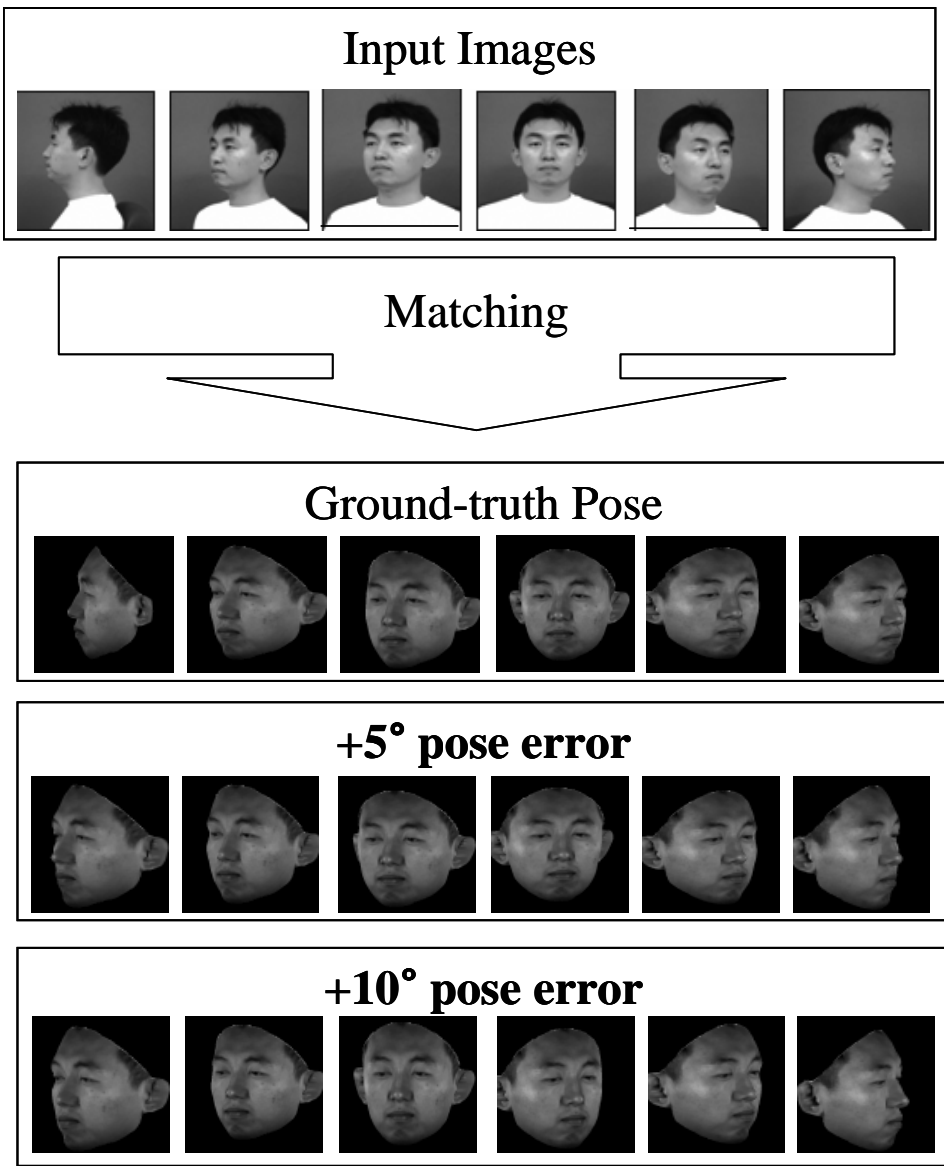


그림 5.3 포즈 오차가 존재하는 입력 영상

89%, 83%의 인식률을 나타내고 있다. 기존의 다중영상 시스템은 포즈 추정단계에서 오차가 발생할 경우 $\pm 5^\circ$, $\pm 10^\circ$ 에서 각각 14%, 22%씩 인식률이 급격히 떨어짐을 보인다. 그러나 포즈 추이 평면을 이용한 오차 측정기법을 사용한 경우 연

표 5.2 포즈 오차가 존재하는 경우의 실험 결과

	Ground-truth pose	$\pm 5^\circ$ pose error	$\pm 10^\circ$ pose error
NN method	76% (24 person fail)	62% (38 person fail)	54% (46 person fail)
PPT method	91% (9 person fail)	89% (11 person fail)	83% (17 person fail)

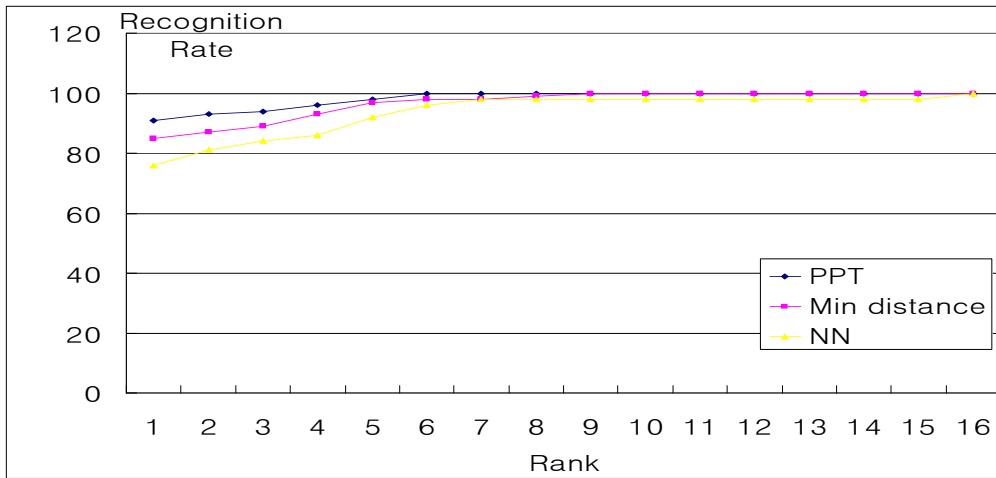


그림 5.4 Cumulative Match Score

속적인 포즈변화의 경향성을 고려하기 때문에 포즈 추정 오차에 강인한 인식률을 나타낸다.

제 5 장 결론

본 논문에서는 포즈 추이 평면을 이용하여 포즈 변화에 강인한 다중영상 얼굴 인식 시스템을 제안하였다. 학습 또는 데이터베이스로 이용할 다중영상 획득의 어려움, 다중영상의 특징을 효율적으로 이용한 매칭방법의 부재는 기존 다중영상 기반의 얼굴인식 시스템의 가장 큰 문제점들이다. 본 논문에서는 정규화된 3차원 얼굴 데이터를 이용하여 학습과 데이터베이스에 이용될 다중영상을 정확하고 쉽게 획득하였다. 또한 다중영상의 특징을 효율적으로 이용하기 위하여 포즈 추이 평면을 정의하고 이를 이용한 새로운 오차추정방법을 제안하였다. 궤적의 자취를 두 번으로 하는 포즈 추이 평면은 포즈 변화에 따른 얼굴영상의 변화를 효율적으로 내포하기 때문에 포즈 추정 오차가 존재하더라도 강인한 인식률을 제공한다. 제안한 시스템의 성능을 검증하기 위하여 다양한 실험을 수행하였다. 첫째 기존의 대표적인 오차 추정 방법인 Nearest Neighbor방법에 의한 인식률과 본 논문에서 제안한 포즈 추이 평면과의 거리를 이용한 인식률을 비교한 결과 각각 91%, 76%로 15%의 인식률의 차이 확인되었다. 둘째 포즈 추정이 정확히 이루어졌을 경우와 $\pm 5^\circ$, $\pm 10^\circ$ 의 포즈 추정 오차가 발생할 경우에 대하여 기존의 다중영상 시스템은 인식률이 급격히 저하되었지만 제안한 다중영상 시스템은 포즈 추정 오차에 매우 강인한 인식률을 나타내었다. 본 논문에서 사용된 PCA 분류기는 선형분류기로서 분류능력이 일반적으로 비선형 분류기인 KPCA, KDA 보다 낮다. 앞으로는 이러한 비선형 분류기를 적용하여 보다 정확한 얼굴인식 시스템을 설계할 것이다. 또한 다중입력시스템을 이용하여 다중 2차원 얼굴영상으로부터 3차원 얼굴영상을 생성한 후 데이터베이스 구축과정과 동일한 과정으로 다중 얼굴영상을 획득하여 이를 입력으로 사용할 것이다. 이와 같은 입력시스템의 확장은 데이터베이스의 모든 정보를 활용할 수 있어 소수의 다중 얼굴영상을 입력으로 사용하는 방법보다 정확한 인식률을 나타낼 것이다.

참 고 문 헌

- [1] R. Chellappa, Charles L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces : A Survey," *Proc. IEEE*, Vol. 83, No. 5, pages 704-740, May 1995.
- [2] Ming-Hsuan Yang, David Kriegman, Narendra Ahuja, "Detecting Faces in Images : A Survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pages 1-38, 2002.
- [3] A. Shio and J. Sklansky, "Segmentation of People in Motion," *Proc. IEEE Workshop on Visual Motion*, pages 325-332, 1991.
- [4] Choong Hwan Lee, Jun Sung Kim and Kyu Ho Park, "Automatic human face location in a complex background using motion and color information," *Pattern Recognition*, Vol.29 (11), pages 1877-1889, 1998.
- [5] Cah, D. and Ngan, K. N., "Face Segmentation Using Skin-Color Map in Videophone Applications," *IEEE Transaction on Circuit and Systems for Video Technology*, Vol. 9, pages 551-564, 1999.
- [6] Liming Zhang, Patrick Lenders, "A new head detection method based on the region shield segmentation in complex background," *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing*, pages 328-331, 2001.
- [7] Kin-Man Lam, "A fast approach for detecting human faces in a complex background", *ISCAS '98, Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, vol. 4, pages 85-88, 1998.
- [8] G. Yang and T. S. Huang, "Human Face Detection in a Complex Background", *Pattern Recognition*, vol. 27, no. 1, pages 53-63, 1994.

- [9] Xiao-guang Lv, Jie Zhou, Chang-shui Zhang, "A Novel Algorithm for Rotated Human Face Detection," *Proceedings. IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pages 760-765, 2000.
- [10] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pages 71-86, Mar. 1991.
- [11] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 586-591, June 1991.
- [12] H. Moon, P.J. Phillips, "Computational and Performance aspects of PCA-based Face Recognition Algorithms," *Perception*, Vol. 30, pages 303-321, 2001.
- [13] M.S. Bartlett, H.M. Lades, and T.J. Sejnowski, "Independent component representations for face recognition," in *Proceedings of SPIE*, vol. 3299, pages 528-539, 1998.
- [14] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, "Face Recognition by Independent Component Analysis," *IEEE Trans. on Neural Networks*, Vol. 13, No. 6, pages 1450-1464, November 2002.
- [15] C. Liu, H. Wechsler, "Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition," *Proc. of the Second International Conference on Audio- and Video-based Biometric Person Authentication, AVBPA'99*, pages 211-216, March 1999.
- [16] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pages 711-720, Jul. 1997.

- [17] K. Etemad, R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images," *Journal of the Optical Society of America*, Vol. 14, No. 8, pages 1724-1733, August 1997.
- [18] W. Zhao, R. Chellappa, A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," *Proc. of the 3rd IEEE International Conference on Face and Gesture Recognition, FG'98*, pages 336, April 1998.
- [19] A.M. Martinez, A.C. Kak, "PCA versus LDA," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pages 228-233, 2001.
- [20] B. Scholkopf, A. Smola, and K. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pages 1299-1319, 1998.
- [21] M.-H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods," *Proc. of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 215-220, May 2002.
- [22] F.R. Bach, M.I. Jordan, "Kernel Independent Component Analysis," *Journal of Machine Learning Research*, Vol. 3, pages 1-48, 2002.
- [23] M.-H. Yang, "Face Recognition Using Kernel Methods," *Advances in Neural Information Processing Systems*, vol. 14, 8 pages, 2002.
- [24] S. Zhou, R. Chellappa, B. Moghaddam, "Intra-personal kernel space for face recognition," *Proc. of the 6th International Conference on Automatic Face and Gesture Recognition, FGR2004*, pages 235-240, May 2004.
- [25] S. Zhou, R. Chellappa, "Multiple-exemplar discriminant analysis for face recognition," *Proc. of the 17th International Conference on Pattern Recognition, ICPR'04*, pages 191-194, August 2004.

- [26] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms," *IEEE Trans. on Neural Networks*, Vol. 14, pages 117-126, January 2003.
- [27] L. Wiskott, J.M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pages 775-779, 1997.
- [28] T.F. Cootes, C.J. Taylor, "Statistical Models of Appearance for Computer Vision," *Technical Report, University of Manchester*, pages 125.
- [29] T.F. Cootes, K. Walker, C.J. Taylor, "View-Based Active Appearance Models," *Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 227-232, March 2000.
- [30] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," *in Proc. ACM SIGGRAPH*, pages 187-194, Mar. 1999.
- [31] V. Blanz, T. Vetter, "Face Recognition Based on Fitting a 3D Morphable Model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9, pages 1063-1074, September 2003.
- [32] J. Huang, B. Heisele, V. Blanz, "Component-based Face Recognition with 3D Morphable Models," *Proc. of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA 2003*, pages 27-34, June 2003.
- [33] B. Moghaddam, J.H. Lee, H. Pfister, R. Machiraju, "Model-Based 3D Face Capture with Shape-from-Silhouettes," *Proc. of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, AMFG*, pages 20-27, October 2003.

- [34] J. Lee, B. Moghaddam, H. Pfister, R. Machiraju, "Finding Optimal Views for 3D Face Shape Modeling," *Proc. of the International Conference on Automatic Face and Gesture Recognition, FGR2004*, pages 31-36, May 2004.
- [35] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," *Proc. IEEE Conf Computer Vision and Pattern Recognition*, 2005.
- [36] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET Database and Evaluation Procedure for Face Recognition Algorithms," *Image and Vision Computing Journal*, Vol. 16, No. 5, pages 295-306, 1998.
- [37] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face-recognition Algorithms," *IEEE trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pages 1090-1104, Oct. 2000.
- [38] D.M. Blackburn, J.M. Bone, and P.J. Phillips, "FRVT 2000 Evaluation Report," February 2001.
- [39] P.J. Grother, R.J. Micheals and P. J. Phillips, "Face Recognition Vendor Test 2002 Performance Metrics," *Proceedings 4th International Conference on Audio Visual Based Person Authentication*, 2003.
- [40] P.J. Phillips, P. Grother, R.J Micheals, D.M. Blackburn, E Tabassi, and J.M. Bone, "FRVT 2002: Evaluation Report," March 2003.
- [41] Lanitis, A., Taylor, C. J., and Cootes, T. F. 1995. "Automatic face identification system using flexible appearance models." *Image Vis. Comput.* 13, pages 393-401, 1995.

- [42] H.T. Tanaka, M. Ikeda, H. Chiaki, "Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition," *Third International Conference on Automated Face and Gesture Recognition*, pages 372-377, 1998.
- [43] J. C. Lee and E. Milios. "Matching range images of human faces." *Int'l Conf on Comp. Vision*, pages 722-726, 1990.
- [44] G. Gordon. "Face recognition based on depth maps and surface curvature. Geometric Methods in Computer Vision," *SPIE*, July 1991.
- [45] Y. Lee, K. Park, J. Shim, and T. Yi. "3D face recognition using statistical multiple features for the local depth information." *16th International Conference on Vision Interface*, June 2003.
- [46] Chua, C.S., Jarvis, R.: Point Signatures, "A New Representation for 3D Object Recognition." *International Journal on Cumputer Vision*. 25, pages 63-85, 1997.
- [47] Chua, C.S., Han, F., Ho, Y.K., "3D human face recognition using point signature." *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*. pages 233-238, 2000.
- [48] G. Medioni and R. Waupotitsch. "Face recognition and modeling in 3D." *IEEE Int'l Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2003)*, pages 232.233, October 2003.
- [49] S. Chang, M. Rioux, and J. Domey. "Face recognition with range images and intensity images." *Optical Engineering*, pages 1106-1112, April 1997.
- [50] C. Heshner, A. Srivastava, and G. Erlebacher. "A novel technique for face recognition using range images." *Seventh Int'l Symp. on Signal Processing and Its Applications*, 2003.

- [51] B. Achermann, X. Jiang, and H. Bunke. "Face recognition using range images." *International Conference on Virtual Systems and MultiMedia*, pages 129–136, 1997.
- [52] Y. Wang, C. Chua, and Y. Ho. "Facial feature detection and face recognition from 2D and 3D images." *Pattern Recognition Letters*, pages 1191–1202, 2002.
- [53] A. Godil, S. Ressler, P. Grother, "Face recognition using 3D facial shape and color map information: comparison and combination," *Biometric Technology for Human Identification, SPIE*, vol. 5404, pages 351–361, April 2005.
- [54] T. Papatheodorou, D. Reuckert, "Evaluation of automatic 4D face recognition using surface and texture registration," *Sixth International Conference on Automated Face and Gesture Recognition*, pages 321–326, May 2004.
- [55] Gong, S., Mckenna, S., and Psarrou, A. "Dynamic Vision: From Images to Face Recognition." *World Scientific*, Singapore. 2000.
- [56] T. Vetter et. al "Linear object classes and image synthesis from a single example image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 733–742, 1997.
- [57] D.J. Beymer, "Face recognition under varying pose," *MIT AI Lab, Technical Report*, 1993.
- [58] Praveen Sankaran, Vijayan Asari, "A Multi-View Approach on Modular PCA for Illumination and Pose Invariant Face Recognition," *Applied Imagery Pattern Recognition Workshop (AIPR'04)*, pages 165–170, 2004.
- [59] O. Yamaguchi, K. Fukui, and K. Maeda, "Face recognition using temporal image sequence," *IEEE International Conference on Automatic Face & Gesture Recognition*, pages 318–323, 1998.

- [60] Li, Y., Gong, S. and Liddell, H., "Recognising Trajectories of Facial Identities Using Kernel Discriminant Analysis," *Image and Vision Computing*, pages 1077-1086, 2003
- [61] H. Song, S. Lee, and K. Sohn, "3D Sensor Based Face Recognition," *Applied Optics*, vol. 44, no.5, pages 677-687, Feb. 2005
- [62] Cox I.J., Roy S., Hingorani S.L., "Dynamic histogram warping of images pairs for constant image brightness," *Images Processing*, Vol. 2, pp 366-396, 1995.

Abstract

A POSE ROBUST MULTI-VIEW FACE RECOGNITION SYSTEM USING PLANE OF POSE TENDENCY

Bongjoe kim

Biometrics Engineering Research Center

The Graduate School

Yonsei University

In this paper, we propose a pose robust multi-view face recognition system using Plane of Pose Tendency (PPT). The proposed system defines the PPT which represents the tendency of pose variation well presented in multiple images. Recognition is done by matching points of the input images with PPT. Since PPT contains the tendency of pose variation, this plane might have the tolerance about pose estimation error.

To demonstrate the effectiveness of the proposed system, the recognition rate based on method of Nearest Neighbor is compared with the recognition rate obtained from proposed method. From experimental results, the recognition rate for the proposed system increases by 15%. Additionally, we analyze the pose robustness of the proposed system based on the sensitivity to the error of pose estimation. The recognition rate of the traditional system steeply decreases under the existence of pose error. However, the proposed system is more robust than the traditional system because the proposed system utilizes the tendency of pose variation based on the concept of the proposed PPT.

Key words : Plane of Pose Tendency, multi-view