

學士學位論文

얼굴인식 기반의 태깅 자동화

Automatic Tagging Scheme
based on Face Recognition

聖潔大學校 工科大學
멀티미디어工學部

이충연

2009年 12月

學士學位論文

얼굴인식 기반의 태깅 자동화

Automatic Tagging Scheme
based on Face Recognition

이 論文을 學士學位 資格要件의 部分으로 提出함

2009年 12月

聖潔大學校 工科大學
멀티미디어工學部

이충연

이 論文을 이충연의 學士學位를 위한 資格要件의
部分으로 認准함

최영미 (印)

주문원 (印)

진성아 (印)

2009年 12月

聖潔大學校 工科大學 멀티미디어工學部

국 문 초 록

최근 웹페이지의 생성 및 웹이 가진 정보량이 기하급수적으로 늘면서 사용자의 검색 목적을 파악하여 효율을 높이기 위한 다양한 방법이 연구되고 있으며, 태깅 시스템을 사용하는 방법이 하나의 대안으로 떠오르고 있다. 태깅 시스템은 인터넷 사용자로 하여금 태그라고 불리는 메타데이터를 글, 사진, 동영상 등에 부여하도록 함으로써 콘텐츠의 검색 및 브라우징을 편리하게 하는 시스템이다.

이처럼 태그는 해당 페이지의 대표 키워드를 의미하므로 콘텐츠 분류의 기준을 마련할 수 있으나, 사용자에게 의해 직접 입력되어야 하는 수고가 필요하고, 또한 무분별한 태깅으로 인해 오히려 분류에 방해가 되는 등의 문제점들이 있다.

본 논문에서는 이러한 태깅의 문제를 해결하기 위한 방법으로 얼굴인식 알고리즘을 활용한 영상콘텐츠 내에서의 인물 태깅 자동화 방법을 제시한다. 이를 위해 먼저 여러 얼굴검출 방법 중 Haar-like feature와 AdaBoost 알고리즘을 이용하여 빠른 속도와 높은 정확도로 영상콘텐츠 내에서 얼굴 영역을 검출해냈다. 이후 PCA와 고유얼굴을 이용하여, 검출해 낸 얼굴을 데이터베이스에 미리 저장해 놓은 프로필 사진과 비교, 인식해냄으로써 해당 인물에 대한 정보를 불러와서 자동으로 태깅하는 시스템을 구현하였다.

이러한 새로운 방식의 태깅 기술은 현존하는 사진공유, 쇼핑, 검색 등의 수많은 웹서비스에 적용이 가능하며, 특히 소셜네트워크서비스에서의 사진관리나 인물검색 등에서 활용할 때 큰 효과를 보일 것으로 기대된다.

목 차

I 서 론	1
1 연구 배경	1
2 연구 동기 및 목적	2
3 연구 목표	2
4 연구 범위 및 구성	4
II 관련 연구 및 기술 동향	5
1 얼굴 인식기술	6
1.1 기하학적 특징 기반 방법	6
1.2 주성분 분석	7
1.3 선형판별 분석	8
1.4 신경망	8
1.5 가버 웨이블릿	9
1.6 서포트벡터머신	9
2 얼굴 검출기술	10
2.1 지식 기반 방법	10
2.2 특징 기반 방법	11
2.3 템플릿 매칭 방법	11
2.4 외형 기반 방법	12
3 소셜네트워크서비스	13
4 태깅 관련 기술	15

III 연구 내용 및 방법	17
1 얼굴 검출	17
1.1 Haar-like 특징	17
1.2 적분 이미지	19
1.3 특징값 계산	21
1.4 분류기 학습	22
1.5 Cascade 구조를 이용한 분류 속도 향상	25
2 얼굴 인식	26
2.1 얼굴인식 과정	26
2.2 얼굴벡터 집합 구성	27
2.3 이미지 정규화	28
2.4 공분산을 이용한 변환행렬 도출	31
2.5 고유얼굴 생성	32
2.6 고유성분벡터 추출	35
2.7 얼굴 인식	35
IV 연구 결과 및 토의	36
1 실험 환경	36
2 시스템 인터페이스	37
3 실험 결과	38
V 결 론	42
참고문헌	43

그림 목차

[1-1] 시스템 개괄도	3
[2-1] PCA를 이용한 데이터 축소 변환	7
[2-2] 세컨드라이프 가입자수 변화	13
[2-3] 대표적인 SNS 제공 사이트와 각각의 특성	14
[2-4] 도요타의 가상 오토쇼	15
[2-5] 태그 클라우드	16
[2-6] 위치 태그	16
[3-1] 얼굴영역 내 Haar-like feature 적용	18
[3-2] Haar-like features	18
[3-3] Lienhart가 확장한 Haar-like features	19
[3-4] 점 P에서의 적분이미지	20
[3-5] 적분이미지를 이용한 계산	20
[3-6] Line 특성 feature에서의 특징값 계산	22
[3-7] 얼굴벡터 구성의 예	27
[3-8] 이미지 정규화 결과	30
[3-9] 평균얼굴 벡터	33
[3-10] 고유얼굴	34

[4-1] 시스템 인터페이스 및 결과 화면	37
[4-2] 얼굴 검출시간	38
[4-3] 얼굴 검출률	39
[4-4] 얼굴 학습시간	40
[4-5] 얼굴 인식시간	41
[4-6] 얼굴 인식률	41

표 목차

[3-1] 얼굴검출을 위한 AdaBoost 알고리즘	23
------------------------------------	----

I 서론

1 연구 배경

전 세계의 인터넷 사용자들은 자신의 미니홈피나 블로그(blog), 또는 카페나 클럽과 같은 참여형 대형 콘텐츠 공간을 통해 자신의 생각을 담은 글이나 직접 제작 또는 촬영한 영상 콘텐츠 등을 올리고 있으며, 다수의 기업들이 이를 활용한 다양한 비즈니스 모델들을 통해 막대한 규모의 수익을 창출해내고 있다. 이처럼 사용자제작콘텐츠(UCC, User Created Contents)는 현재 인터넷에서 가장 각광받는 분야이며, 앞으로도 계속 성장할 잠재성을 지닌 인터넷 비즈니스 아이템이다 [1].

이러한 글이나 사진, 동영상 등의 UCC에 사용자가 직접 키워드를 다는 태그(tag)는 웹 2.0 시대의 새로운 인터넷 정보 유통 채널로 급부상하였으며, 현재 싸이월드, 네이버, 구글(Google), 플리커(Flickr) 등 국내외 주요 인터넷 기업들이 이를 시행하고 있다 [2]. 특히 구글의 이메일 서비스인 Gmail은 검색과 카테고리 분류 서비스가 주로 텍스트나 디렉토리 위주였던 기존의 이메일 관리 시스템과 달리, 이메일 각각에 여러 개의 태그를 사용자가 임의로 붙일 수 있도록 함으로써 보다 가치 있고 효과적인 정보 분류를 가능하게 하였다. 또한 야후(Yahoo)의 사진 공유 서비스인 플리커는 사용자가 사진을 올릴 때 주제에 적합한 태그를 입력할 수 있도록 하고, 이러한 태그로 분류된 정보를 이용하여 관심 있는 사진을 쉽게 검색할 수 있도록 함으로써 폭발적인 인기를 얻을 수 있었다.

하지만 태그의 이렇게 화려한 장점의 이면에는, 스팸 태그(spam tag)나 태그 오남용으로 인한 쓰레기 정보 홍수 등의 다양한 부작용이 있으며, 콘텐츠마다 직접 일일이 태그를 달아야 하는 불편함으로 인하여 일반적으로 사용자들이 잘 사용하지 않아 현재 태그 서비스의 유용성 논란이 발생하고 있는 실정이다 [3].

2 연구 동기 및 목적

사용자가 올린 콘텐츠를 분석하여 이에 맞는 태그를 자동으로 입력해준다면 부적합하거나 불필요한 태그의 입력을 막을 수 있으며, 사용자는 직접 태그를 입력하지 않고도 태그의 정보 분류 효과를 누릴 수 있게 될 것이다. 본 논문에서는 이처럼 태그가 가지고 있는 문제들을 해결하고, 소셜네트워크서비스(Social Network Service, SNS) 사용자들이 부담 없이 태그를 활용할 수 있도록 하기 위해 얼굴인식 기반의 태깅 자동화 시스템을 제안한다.

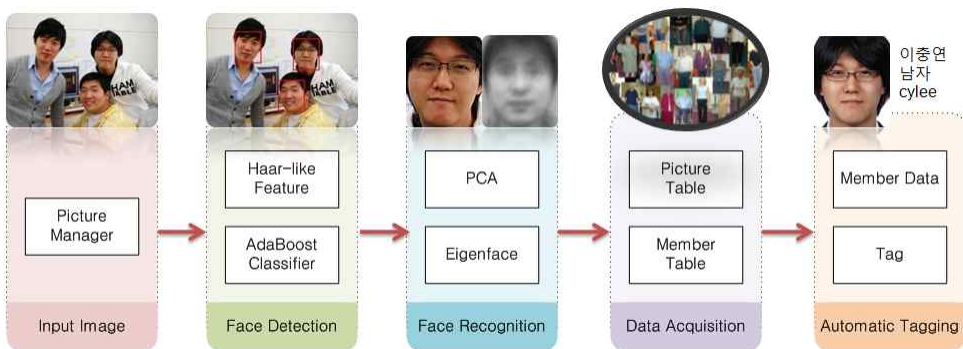
이 시스템은 기존 태그의 단점을 보완할 뿐만 아니라 얼굴인식이라는 재미 요소를 가미함으로써 SNS 이용자들의 적극적인 참여를 유도할 수 있다는 장점이 있다. 또한 추출된 태그 정보를 가공하여 엔터테인먼트나 마케팅 산업에 소비자와의 자연스러운 접점을 제공하는 새로운 비즈니스 모델을 만들어낼 수도 있을 것이다.

3 연구 목표

기존의 태깅 방식은 텍스트 위주이며 검색과 분류 외에 별다른 기능을 제공

하지 못한다. 이를 간파한 국내외의 몇몇 SNS 업체에서는 얼굴 영역과 해당 인물을 매칭시켜 태깅하는 서비스를 선보였다. 하지만 이 방법 또한 사용자가 모든 태그를 직접 입력해야 하기 때문에 많은 호응을 얻지 못하였다. 본 논문에서는 사진 속 다수의 얼굴을 자동으로 검출하고, 인식한 얼굴 영역에 자동으로 태그가 입력되는 시스템을 개발한다.

입력된 이미지로부터 실시간으로 얼굴을 검출하고 이를 인식하여 자동으로 태그를 입력하는 시스템을 [그림 1-1]과 같이 개발한다. 이 시스템은 사진첩 모듈에 얼굴인식 기능을 탑재하여 이미지에 나타난 얼굴 영역을 검출하여 표시해주고, 이를 클릭하면 데이터베이스에 저장된 프로필 사진과 대조하여 인식함으로써 프로필 출력 및 태그 입력이 가능하도록 구현한다.



[그림 1-1] 시스템 개괄도

이를 위해 Haar-like feature와 AdaBoost 알고리즘을 이용하여 얼굴영상을 학습하고, 주성분분석 알고리즘을 통해 추출한 고유얼굴(Eigenface)을 얼굴인식에 이용한다. 시스템은 플랫폼의 제약을 최소화하고, 데이터베이스를 사용하도록 함으로써 이후 웹이나 모바일과의 호환이 용이하도록 한다.

4 연구의 범위 및 구성

본 논문에서는 얼굴인식 알고리즘을 이용한 영상 콘텐츠에서의 얼굴 검출과 인식 방법에 대한 개념 및 구조, 이것을 활용하여 태깅 자동화 시스템을 구현하기 위한 설계 및 구현 방법, 그리고 이와 관련된 배경 지식들을 담고 있다. 전체적인 구성은 다음과 같다.

1장은 서론을 통해, 연구 배경 및 동기, 목적에 대해 밝힌다. 그리고 본 논문에서 제안하는 얼굴인식을 이용한 태깅 자동화 방법에 대해 구체적으로 설명하고 이 방법을 실현화하기 위해 필요한 기술적 과제들에 대해 설명한다.

2장은 제안하는 방법의 기반 기술이 되는 얼굴인식 및 검출 분야에 대해 소개하고, 여기서 사용되고 있는 다양한 방법들의 성능 및 장단점에 대하여 설명한다. 그리고 소셜네트워크서비스와 태깅 관련 기술에 대하여 소개한다.

3장은 연구에 사용된 얼굴 검출 및 인식 방법에 대하여 상세하게 기술한다. 먼저, 얼굴 검출에 사용된 특징추출 기법과 분류기 학습 알고리즘에 대해 설명한 후, 얼굴 인식의 전체적인 과정에 대해 알아본다.

4장에서는 본 논문에서 구현한 시스템의 구성과 각부의 역할에 대해 설명하고, 이를 위해 구축한 실험 환경에 대해 소개한다. 그리고 얼굴 학습 및 검출, 인식에 걸리는 시간 및 성능에 대한 실험과 결과에 대해 설명하고, 이를 바탕으로 향후 연구에서 다루어야 할 인식 방법의 개선과 연구 방향에 대해 토의한다.

마지막으로 5장에서는 결론과 함께 제안한 태깅 자동화 방법의 기대효과에 대해 간략히 기술하고 향후에 수행되어야 할 연구와 실험을 제시하면서 논문을 마무리한다.

II 관련 연구 및 기술 동향

얼굴인식 기술은 사진이나 동영상으로부터 실시간으로 얼굴을 검출하고 인식하는 기술로 사진에서의 얼굴 유무 판별 및 동영상 내에서의 실시간 자동 얼굴 검출 등에 활용된다 [4]. 이 기술을 데이터 관리에 이용할 경우 각 얼굴에 해당하는 인물을 자동으로 판별해냄으로써, 특정 인물을 중심으로 사진을 분류하거나 서로 연관이 있는 인물들을 동적인 그룹으로 설정하여 저장하고 관리하는 등 기존에는 텍스트 데이터에 한정되었던 인공지능적인 작업을 구현해낼 수 있다.

또한 얼굴인식 기술은 지문인식이나 홍채인식 등의 다른 생체인식 기술들과 달리 인간의 신체 일부를 인식장치에 직접 접촉시키지 않아도 되기 때문에 생체 정보 획득 방법에 대한 강제성이 적어 다른 생체인식 기술들이 사용자에게 줄 수 있는 거부감이나 불편함이 존재하지 않는다는 특성이 있다 [5]. 이러한 비접촉성, 비강제성, 편리성의 특징으로 인해 얼굴인식 기술은 신원확인 분야 외에도 신분 확인 대상자가 모르는 사이에 자연스럽게 정보를 획득해야 하는 지능형 무인감시 등의 분야에도 효과적으로 응용될 수 있다 [6].

하지만 인간의 얼굴은 움직임, 표정, 인종, 나이, 성별, 헤어스타일, 타 객체에 의한 일부 영역 가려짐, 조명, 카메라의 기계적인 특성 등에 따라 다양하게 나타나게 되며, 이렇게 복잡하고 많은 변수들을 통제하기 위해 얼굴인식 일련의 과정들이 다양한 방법에 의해 연구되고 있다. 본 장에서는 이러한 얼굴검출 및 얼굴인식 기법에 대한 관련연구들과 기술 동향에 대해 알아본다.

1 얼굴 인식기술 (Face Recognition Methods)

현재 얼굴 인식 및 검증 시스템에 사용되고 있는 방법은 크게 특징 기반 방법과 통계적 방법으로 분류할 수 있다. 특징 기반 방법은 얼굴을 구성하는 요소들의 특징점을 찾아서 각 점들 사이의 위치나 크기, 모양 등을 측정함으로써 얼굴 영상들 사이의 유사도를 비교한다. 주로 사용되는 특징들로는 에지(edge), 그레이 정보(gray information), 색상, 움직임 등이 있다. 통계적인 방법은 개인의 얼굴 정보를 얼굴 공간상에 투영하여 사용하는 방법으로써 대표적인 방법으로는 주성분분석[9], 선형판별분석[12], 신경망[14], 가버웨이블릿과 탄성그래프매칭 방법[16], 그리고 최근에 많이 사용되는 서포트벡터머신[19] 등이 있으며, 검출률을 높이기 위해 여러 가지 방법을 결합하여 사용하기도 한다.

1.1 기하학적 특징 기반 방법

기하학적 특징 기반 방법은 눈, 코, 입과 같은 얼굴의 특징들의 위치나 크기 또는 이들 간의 거리와 같은 기하학적인 인자들만으로도 각 개인의 얼굴을 인식할 수 있다는 사실에 착안한 것으로써, 얼굴의 기하학적 특징점을 추출하여 일치 여부를 판단하여 인식하는 방법이다 [7, 8]. 기하학적 특징들은 상대적으로 명암차가 크기 때문에 해상도를 낮추었을 때 확연히 구분되며, 따라서 이러한 요소들을 각각의 상대적인 위치나, 형태 등의 관계를 비교하여 얼굴을 분류하는데 사용할 수 있다. 이 방법은 얼굴 인식에서 보편적으로 이용하는 방법론 중 하나로, 특징들을 이용하기 때문에 외부환경(조명, 시점)에 대해 강인한 면모를 보인다.

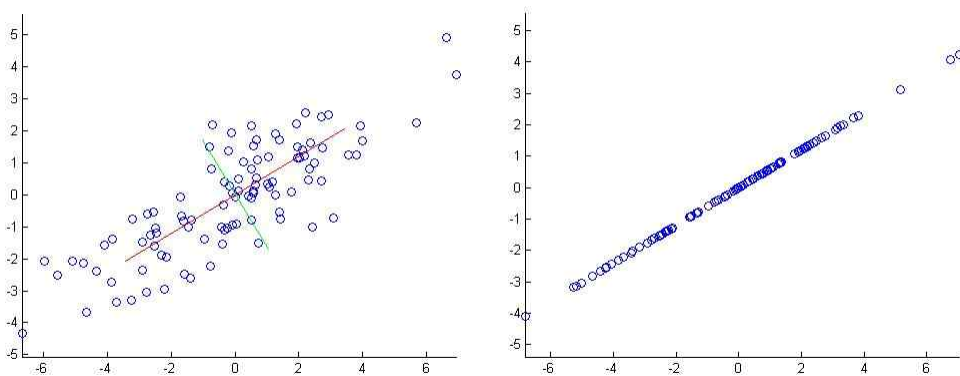
이러한 특성으로 인해 기하학적특징기반 방법은 주로 다른 특징들과 결합하

여 사용되거나, 더 좋은 인식 결과를 얻기 위해 다른 인식 방법의 전처리 단계로 활용된다.

1.2 주성분 분석 (Principal Component Analysis, PCA)

PCA는 다변량 분석 기법 중 하나로 전체 영상의 데이터를 데이터의 분산이 큰 몇 개의 고유방향(eigen-vector)에 대한 축으로 선형 투영(linear projection)시켜 데이터의 차원을 줄여주는 방법이다 [9]. 이는 데이터의 차원을 줄여 간단하게 표현할 수 있는 체계적이고 실용적인 방법으로, Karhunen-Loeve(KL) 변환 [10] 또는 Hotelling 변환[11]이라고도 불리며, 주로 다차원의 특징벡터로 이루어진 데이터에 대하여 높은 차원에서의 정보를 유지하면서 낮은 차원으로 차원을 축소시키는 방법으로 사용된다.

다양한 변형을 표현할 수 있는 국부적 특징을 추출하지 못한다는 문제점을 가지고 있지만, 수학적으로 잘 정의되어 있는 특성으로 인해 PCA의 기본 개념을 조금씩 변형시킨 여러 가지 방법들이 널리 이용되고 있다.



[그림 2-1] PCA를 이용한 데이터 축소 변환

1.3 선형판별 분석 (Linear Discriminant Analysis, LDA)

선형판별 분석은 클래스간 분산(between-class scatter)과 클래스내 분산(within-class scatter)의 비율을 최대화하는 방식을 이용하여 데이터에 대한 특징 벡터의 차원을 축소하는 방법이다 [12].

이 방법은 Fisher Linear Discriminant(FLD) [13]의 개념을 도입한 것으로, 한 사람 내에서의 변화와 다른 사람들 간의 변화의 비를 고려하여 자기 자신들 사이는 최소화하고, 다른 사람들 사이는 최대화시키는 방향으로 선형 변환을 한 후 학습하는 방법이다. 따라서 보다 정확하고 환경 변화에 강인하여 온라인 상태에서 실시간으로 적용 가능한 알고리즘이다.

얼굴 인식에서의 PCA의 가장 큰 문제점은 클래스가 서로 중첩되는 현상이 발생하는 것인데, 이러한 경우 데이터를 분류함에 있어 인식률을 떨어뜨리는 주요한 요인이 된다. 특히 조명이나 표정 등의 요인에 의한 변화를 구분하기 어렵다. 하지만 LDA는 클래스간의 분산뿐 아니라 클래스 내 샘플들 간의 분산도 고려하기 때문에 PCA에 의해서만 인식할 때보다 성능이 향상된다.

1.4 신경망 (Neural Network)

신경망은 성별 분류, 얼굴 인식 그리고 표정들의 분류와 같은 문제들을 다루기 위한 방법으로 얼굴 인식 분야에서 많이 사용되고 있다 [14]. Kohonen의 연상지도(associative map)를 사용하여 신경망을 얼굴 인식에 적용시킨 실험에서는 얼굴 영상들의 작은 집합을 사용함으로써 입력 영상에서 잡음이 매우 심할 때나 영상 일부가 상실되었을 때에도 정확하게 적용됨이 확인되었다. 신경망 방

법으로 얼굴을 검출하는 것은 얼굴 클래스, 비-얼굴 클래스와 같은 이진 클래스 분류 문제로써 고려된다 [15].

1.5 가버 웨이블릿 (Gabor Wavelet)

L. Wiskott 등에 의하여 제안된 가버 웨이블릿과 탄력적 구성 요소 그래프 매칭방법(Elastic Bunch Graph Matching, EBGM)은 생물학적 동적 연결구조 (Dynamic Link Architecture, DLA)에 기반하여 얼굴을 인식하는 방법으로, 가버 웨이블릿을 뇌의 시각 경로를 모형화한 개념(keyhole view)으로 사용하며, 얼굴의 구성 요소들의 그래프(face bunch graph)에 의하여 얼굴을 표현한다 [16].

얼굴 구성 요소 그래프는 눈, 코, 입, 윤곽과 같은 식별이 용이한 부분에 위치한 노드 집합으로 구성되며, 각각의 노드로부터 가버 웨이블릿의 크기로써 표현되는 특징들의 집합을 추출한다. 그리고 탄력 그래프 매칭을 통하여 두 개의 패턴을 비교함으로써 얼굴을 인식하게 된다. 이 방법은 머리 모양을 바꾸거나 수염을 기르거나, 안경 또는 모자를 착용하더라도 인식이 가능하다는 장점이 있다.

1.6 서포트벡터머신 (Support Vector Machines, SVM)

SVM은 AT&T에서 Vapnik 등에 의해 이진패턴 분류 문제를 해결하기 위해 제안된 학습 방법으로 [17], 각 범주의 대표적인 특성을 갖는 벡터를 이용하여 최적의 하이퍼분리면(hyperplane)을 찾아 두 범주를 분류하는 기법이다. 인식 성능과 처리 속도가 뛰어나 최근 얼굴인식, 문자인식, 문서범주화 등 다양한 분야에 적용되고 있다 [18, 19].

2 얼굴 검출기술 (Face Detection Methods)

얼굴인식에 관련된 많은 연구들에서는 얼굴인식 이전에 선행되어야 하는 얼굴검출 기술에 대해 대부분 간과하는 측면이 있으며, 이미지 속에 얼굴이 있다고 가정하거나 이미지 영역에 대하여 명시화 하고 있다. 하지만 얼굴검출 기술은 얼굴인식 시스템 개발 시에 첫 번째로 수행되어야 하는 단계이며, 이것이 충족되지 않을 시에는 응용 프로그램의 사용자 인터페이스가 불편해지거나 감시 시스템, 추적 시스템 등으로의 활용이 불가능해진다. 또한 얼굴은 고유동적인 객체이므로 이것을 찾아낼 수 있다면 다른 다양한 사물의 인식에도 그 기술을 사용할 수 있기 때문에 일반적인 객체에 대한 검출 또한 가능하여 그러한 시스템을 만들기 전에 선행하는 학습 분야이기도 하다.

얼굴검출을 위한 연구들은 크게 지식기반 방법(knowledge-based methods), 특징기반 방법(feature-based methods), 템플릿매칭 방법(template matching methods), 외형기반 방법(appearance-based methods)으로 나눌 수 있다 [20].

2.1 지식 기반 방법 (Knowledge-based Method)

지식 기반 방법은 사람의 얼굴을 구성하는 눈, 코, 입 등의 부위를 얼굴의 특징요소로 인식하고, 각 요소 사이의 거리나 위치와 같은 기하학적인 특성을 분석하여 얼굴을 검출하는 방법이다 [21].

입력 이미지의 히스토그램을 이용하는 방법이 대표적인데, 이는 가로 축의 히스토그램을 이용하여 얼굴의 좌우 특성을 찾고, 세로 축의 히스토그램을 이용하여 눈, 코, 입의 위치 정보를 찾아내는 방법이다. 얼굴 영상은 중앙 영역에 부분

적인 명암도 집중이 있는데, 얼굴 영상과 비교될 영상의 명암도 분포를 보면 집중 과형이 어느 정도 비슷하게 나온다는 것이 이 방법의 핵심이다.

지식 기반 방법은 눈, 코, 입 간의 상관관계를 간단하게 구할 수 있지만 그에 대한 각각의 정의를 내리기 힘들다는 문제가 있다. 정의된 규칙이 지나치게 엄격하면 검출률이 낮아지고, 반대로 규칙을 줄이면 오검출이 많아지게 된다. 또한 포즈 등에 변화에 적용하기 힘들어 가려지지 않은 정면 얼굴의 검출에서만 좋은 성능을 보인다.

2.2 특징 기반 방법 (Feature-based Method)

특징 기반 방법은 얼굴의 부분적인 특징을 찾고 부분 후보영역을 통합하여 얼굴의 위치를 찾는 방법이다. 부분적인 얼굴특징에는 눈, 코, 입과 같은 얼굴요소 [22, 23], 질감(texture)[24], 피부색(skin color)[25, 26], 그리고 외곽선(outline) 등이 있으며 이들의 조합이 적절한 것을 얼굴로 인지한다[27].

이 방법은 빠른 속도의 장점이 있지만, 조명에 따른 피부색의 변화, 기울어짐에 따른 얼굴 모양 변화, 카메라 위치, 잡음 등으로부터 많은 영향을 받는다.

2.3 템플릿 매칭 방법 (Template Matching Method)

템플릿 매칭 방법은 얼굴 영상 데이터의 부분 영역이나 외곽선을 이용하여 미리 정의된 규칙에 의해 표준 얼굴 템플릿을 생성하고, 이를 입력 영상과 비교하여 얼굴을 검출하는 방법이다. 이때 템플릿 정보는 학습 과정을 거치지 않고 수동적으로 생성한다.

이 방법은 얼굴 요소를 찾는 과정이 필요 없고 복잡한 배경에서도 수행이 가능한 장점이 있지만, 개개인의 얼굴마다 차이가 크기 때문에 표준 템플릿의 패턴에 따라 영향을 받게 되며, 지식기반 방법과 같이 각각 다른 포즈에 대한 템플릿을 정의하기 어렵다.

이외에도 미리 정의된 Shape Template을 이용한 검출방법 [28], Deformable Template을 이용한 방법인 Active Shape Model(ASM)이 있다 [29].

2.4 외형 기반 방법(Appearance-based Method)

외형 기반 방법은 학습영상 집합에 의해 학습된 모델을 이용하여 입력영상으로부터 얼굴을 구분하는 분류기(classifier)를 만들어 얼굴을 검출하는 방법이다. 학습 과정에서 얼굴 영상과 얼굴이 아닌 배경 영상을 입력받아 둘의 차이를 잘 나타낼 수 있는 특징들을 찾아내고, 이것을 학습시킨 후, 입력 영상에 대하여 학습된 정보를 이용하여 검출하는 방법이다.

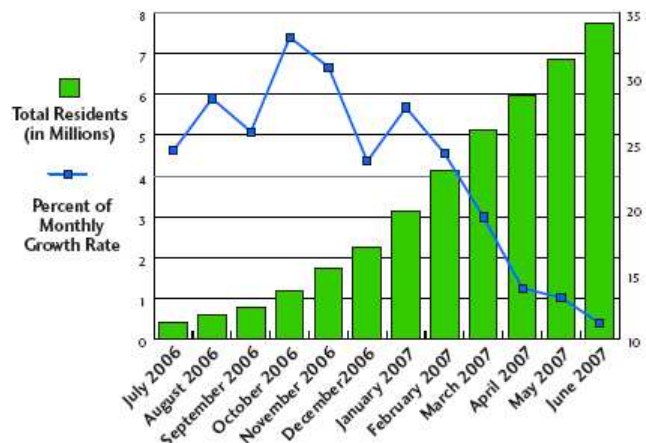
이에 따른 방법으로는 주성분분석과 고유얼굴(Eigenface)을 이용한 방법 [9], 분포 기반(Distribution-based) 방법 [30], Neural Network [14], SVM [19], 순수베이즈분류기(Naive Bayes Classifier)[31], 은닉마르코프모델(Hidden Markov Model, HMM)[32] 등 수많은 패턴인식 기법들이 존재한다.

외형 기반 방법은 현재 얼굴검출 분야에서 가장 많이 사용하는 방법 중 하나이며, 다른 방법에 비해 좋은 검출 성능을 나타내고 있지만, 학습 및 특징추출 과정에 따라 다른 방법에 비해 공간적 비용과 시간이 많이 소비된다.

3 소셜네트워크서비스

최근 국내외를 막론하고 인터넷 상에서 SNS와 관련된 다양한 서비스들이 등장하며 확산이 가속화되고, 새로운 사회 트렌드로 자리 잡고 있다. SNS란 사회적 관계 개념을 인터넷 공간으로 가져온 것으로, 사람 간의 인맥구축 및 네트워크 형성을 지원하는 서비스를 일컬으며, 최근 개인의 일상 및 관심사를 공유·소통시켜주는 서비스 전체를 일컫는 표현으로 발전하면서 블로그, 미니홈피, 카페, 커뮤니티, 가상현실 웹서비스 등을 총칭하는 폭넓은 개념으로도 사용되고 있다[33][34].

특히 미국의 마이스페이스(Myspace), 페이스북(Facebook), 세컨드라이프(Second Life)는 이미 전 세계의 인터넷 사용자가 참여하는 SNS로 자리잡았으며, 우리나라에서도 일찌감치 싸이월드(Cyworld)가 서비스를 시작하여 지속적으로 성장하고 있다. 이들 주요 SNS 제공 사이트들은 최근 큰 폭의 트래픽 증가를 보이는 등 인터넷 사용자들의 관심과 이용이 집중되고 있다. 특히 2006년 초 10만여명에 불과했던 세컨드라이프의 가입자 수는 2007년 6월 700만명, 2009년 1월 1670만명을 돌파하며 폭발적으로 성장하고 있는 추세이다 [그림2-2].



[그림 2-2] 세컨드라이프 가입자수 변화
(자료: Yankee Group)



마이스페이스(Myspace) : 가입자수 2억 5천만명을 보유한 세계 최대 규모의 SNS



페이스북(Facebook) : 최근 급격히 성장하고 있는 SNS로, Open API를 통해 어플리케이션이 운용되는 신개념 플랫폼을 제공



세컨드라이프(Second Life) : 온라인 3D 가상 커뮤니티 서비스로 현실세계와 유사한 다양한 활동이 가능한 SNS



싸이월드(Cyworld) : 국내 최대 규모의 SNS로 미니홈피와 일촌이라는 관계를 기반으로 한국형 소셜 네트워크를 제공

[그림 2-3] 대표적인 SNS 제공 사이트와 각각의 특성

SNS는 사용자 참여·공유·개방 등의 웹 2.0의 핵심 가치가 서비스에 반영되어 있는 웹 2.0 시대의 킬러앱(killer application)이다. 특히 세컨드라이프와 같은 가상현실 서비스는 실제 경제적 이윤을 매개로 사용자의 지속적인 참여를 유도하고, 이러한 사용자들의 관심 증대로 인해 콘텐츠 제공업체의 참여가 가속화되며, 이를 통해 새로운 비즈니스 모델 출현과 서비스 고도화로 이어지는 등



[그림 2-4] 도요타의 가상 오토쇼, 도요타는 시카고 오토쇼에서 선보일 2개의 신차 모델을 세컨드라이프에 동시 공개 하였다.

SNS를 둘러싼 선순환의 구조가 정착되고 있다. 실제로 IBM, 도요타 등 수많은 글로벌 기업들이 이미 세컨드라이프에 진출했으며[그림 2-4], BBC, ABC, MTV 등 주요 미디어들도 이곳에 방송국을 개설하였고, 하버드와 스탠포드 등 미국의 우수한 대학들이 세컨드라이프 내에 캠퍼스를 개설하여 온라인 강의를 진행하고 있다.

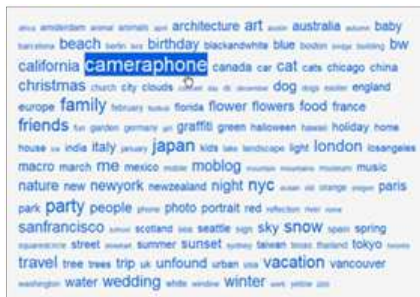
SNS는 이밖에도 다양한 비즈니스 모델을 파생하며 웹서비스의 패러다임 변화를 주도하고 있으며, 향후에도 이러한 시도는 지속될 전망이다. 특히 다양한 가상 서비스에 현실감을 제공하면서 현실세계와 동일한 모양과 경험을 제공하는 방향으로 진화해갈 것으로 예상된다 [35].

4. 태깅 관련 기술

태그(tag)는 어떠한 정보에 대한 메타데이터(metadata)를 의미한다. 웹에서의 검색 텍스트, 이미지, 영상 등과 같이 모든 웹 자원들에 대해 사용자는 태그의 형태로 메타데이터를 기록할 수 있다. 태깅은 많은 사용자들에 의해 블로그나 검색사이트를 통하여 사용되어지고 있으며, 일반적으로 하나의 텍스트박스에 사용자가 일련의 키워드들을 문자열 형태로 입력하는 방식으로 사용된다. 하지만 최근에는 이러한 태깅을 자동으로 추출하는 서비스들이 많이 등장하고 있으며, 이러한 자동화된 태깅 서비스들은 텍스트 기반의 문자들의 패턴을 분석하여 빈도

수가 많은 검색어를 태깅으로 추출하거나 텍스트 자체에 대한 연관된 정보들을 메뉴화 하여 자동으로 연관 있는 정보를 출력하여 주는 등 더욱 지능화되어지고 있는 추세이다 [36]. 또한 태그구름(tag cloud)과 같은 개념을 도입하여 태그들의 집합을 빈도수에 따라 색상이나 굵기, 크기 등과 같은 요소들을 각기 달리 나타냄으로써 보다 시각적인 서비스를 제공하기도 한다 [37].

이밖에도, 웹 자원의 지리적인 메타데이터를 태그형태로 추가하는 기법인 위치태깅(geometric tagging)[38], 공유된 콘텐츠에 다수의 사용자들이 키워드의 형태로 메타데이터를 추가하는 협력태깅(collaborative tagging)[39], 텍스트 분석을 통해 자동으로 태그를 생성하는 자동태깅(automatic tagging)[40] 등 태깅 기술을 응용하고 확장하기 위해 다양한 기법들이 시도되고 있다.



[그림 2-5] 태그 클라우드



[그림 2-6] 위치 태그

Ⅲ 연구 내용 및 방법

1 얼굴 검출

입력된 영상 콘텐츠에는 사람의 얼굴과 얼굴이 아닌 배경 및 객체 영상이 포함되어있다. 따라서 태깅을 위한 얼굴인식 이전에 이미지에서 얼굴 영역만을 검출해내는 작업이 선행되어야 한다. 본 논문에서는 이를 위해 아다부스트(Adaptive Boosting, 이하 AdaBoost) 알고리즘과 Haar-like feature를 이용하여 전체 입력 이미지의 스캔을 빠르고 정확하게 수행할 수 있도록 하고, 시스템에서의 실시간 얼굴 검출을 가능하게 하였다. Viola와 Jones가 제안한 AdaBoost를 이용한 얼굴 검출 방법[42]은 Haar-like feature를 적분이미지를 통해 빠른 속도로 계산하여 얼굴을 검출한다. 학습과정에서 얼굴영상과 비얼굴영상을 입력받아 둘의 차이를 가장 잘 나타내주는 특징들을 선택하고, AdaBoost를 사용하여 이러한 특징들을 선형적으로 결합함으로써 최종적인 분류기를 생성해낸다. 이 방법은 적은 양의 훈련 데이터(training data)로도 학습이 가능하며, 적분이미지를 활용하여 한 번 계산한 위치를 상수시간에 계산해냄으로써 빠른 검출 속도를 가지는 장점이 있다.

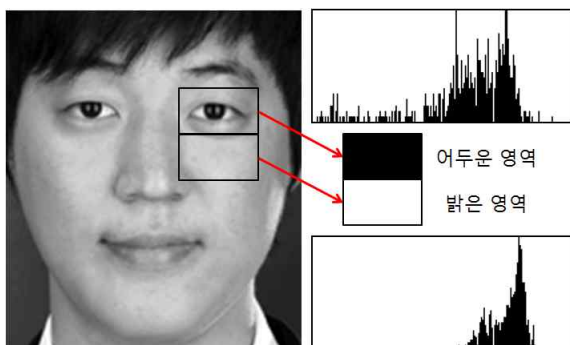
1.1 Haar-like 특징

얼굴 영상은 조명이나 포즈, 배경 등의 다양한 외부환경 요소에 영향을 받기 때문에 이와 같은 환경 요소에 강인하면서도 빠른 연산이 가능한 얼굴 검출 방법이 요구된다. 이러한 문제를 해결하기 위해 [그림 3-1]과 같이 간단하고 객체

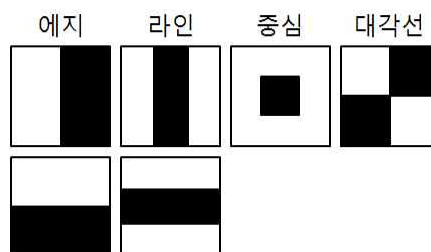
간의 구별 능력이 있는 Haar-like feature를 사용한다.

Haar 웨이블릿(wavelet)은 1910년 Haar가 제안한 간단한 직교 웨이블릿으로, 영상에서의 수평, 수직, 대각 방향에 대한 에지(edge)를 각 스케일별로 분해하는 기법이다 [41]. Papageorgiou는 Haar 웨이블릿을 이용하여 물체를 검출하는 알고리즘을 제안하였으며 [42], Viola와 Jones는 이것을 확장한 Haar-like feature[그림 3-2]를 이용하여 얼굴검출에 활용하였다 [43].

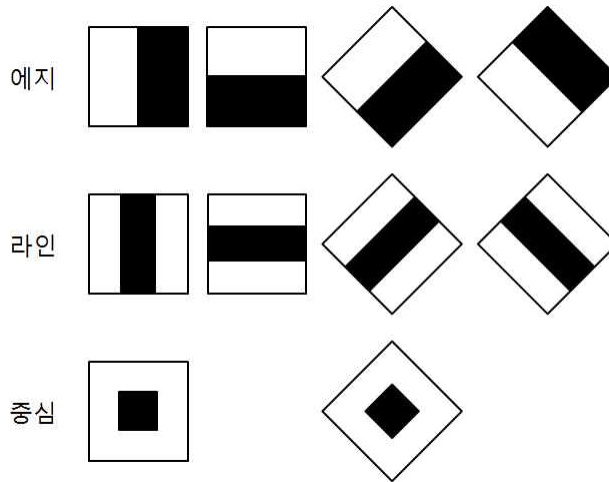
본 논문에서는 기울어진 얼굴에 대한 검출률을 높이기 위해 Viola와 Jones의 Haar-like feature를 확장하여 45도 기울어진 feature들을 추가한 Lienhart의 Haar-like feature를 사용하였다. [그림 3-3]에서 볼 수 있는 것처럼 Lienhart가 제안한 확장된 Haar-like feature들에서는 Viola의 기존 대각선 feature가 제거되고, 대신 에지와 라인, 중심 특성에 각각 대각선 특성이 추가되었다 [44].



[그림 3-1] 얼굴영역 내 Haar-like feature 적용



[그림 3-2] Haar-like features



[그림 3-3] Lienhart가 확장한 Haar-like features

Haar-like feature는 [그림 3-2, 3-3]과 같이 서로 인접한 사각형으로 구성되어 있으며, 검은색과 흰색으로 표시된 각 영역 안에 위치한 픽셀들의 값을 더하여 영역의 합을 구한 뒤, 그 값들에 가중치를 곱한 값의 합을 구하여 특징값을 계산한다. 즉, 흰색 영역과 검은색 영역의 명암 차다. 이때 Haar-like feature는 직사각형 형태를 가지므로, 적분이미지를 이용하여 빠른 속도로 계산이 가능하다.

1.2 적분이미지 (Integral Image)

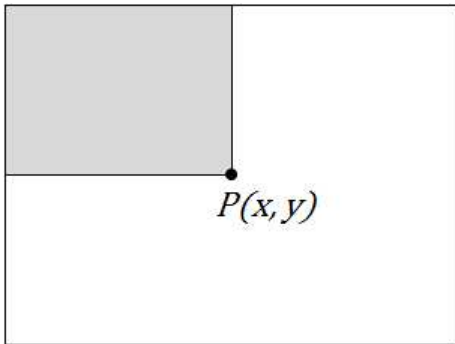
이미지 내에서 Haar-like feature의 픽셀 합계 계산을 빠르게 수행하기 위해 적분이미지를 사용한다. 적분이미지는 처음 영상이 들어왔을 때, 픽셀값을 합한 번의 계산으로 영역합테이블(Summed Area Table, SAT)[45]을 생성하며, 이렇게 미리 구해진 적분이미지를 이용하면 임의의 원하는 위치에 해당하는 픽셀 값들의 합을 구할 때 중복연산을 최소화함으로써 매우 빠른 계산이 가능해진다.

이미지 상의 한 점 $P(x,y)$ 에서의 적분이미지 값은 [그림 3-3]에서 표현된 것과 같이 좌측상단의 시작점에서부터 $P(x,y)$ 까지의 모든 픽셀값들의 합으로, [수식 3-1]에 의해 계산되고, [수식 3-2]에 의해 갱신된다.

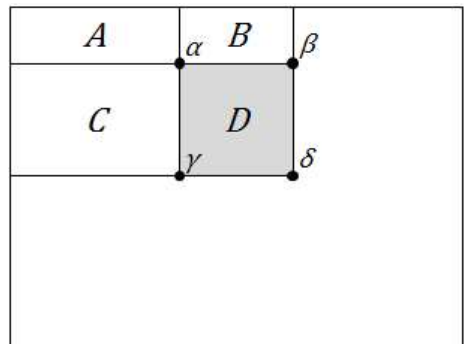
$$i_p(x,y) = \sum_{x' \leq x, y' \leq y} i_i(x',y') \dots\dots\dots [수식 3-1]$$

$$i_p(x,y) = i_i(x,y) + i_p(x-1,y) + i_p(x,y-1) - i_p(x-1,y-1) \dots\dots [수식 3-2]$$

※ i_p 는 P 에서의 적분이미지, i_i 는 입력이미지의 픽셀값



[그림 3-4] 점 P에서의 적분이미지



[그림 3-5] 적분이미지를 이용한 계산

적분이미지를 이용하면 [그림 3-4]의 D영역과 같이 이미지의 중간에 놓여있는 임의의 영역에 포함된 픽셀값의 합을 빠르게 구할 수 있다. 점 α 에서의 적분이미지값을 A영역의 각 픽셀값을 모두 합한 값이라고 하고, 점 β 에서의 적분이미지값은 A영역과 B영역의 합, 마찬가지로 점 γ 에서의 적분값은 A영역과 C영역의 합, 점 δ 에서의 적분값은 A, B, C, D 영역 모두의 픽셀값의 합이라고 할 때,

D영역의 픽셀값의 합을 $\alpha - \beta - \gamma + \delta$ 로 나타낼 수 있기 때문이다.

이와 마찬가지로 Haar-like feature는 이웃한 두 개 이상의 직사각형으로 이미지 상에 놓여지기 때문에, 적분이미지를 이용하면 빠른 속도로 계산이 가능하다. 두 개의 직사각형으로 이루어진 feature는 6개 점의 적분이미지값으로 구할 수 있으며, 세 개의 직사각형으로 이루어진 feature는 8개 점, 네 개의 직사각형으로 이루어진 feature는 9개 점의 적분이미지값을 이용하여 구할 수 있다.

1.3 특징값 계산

특징값은 입력 영상에서 Haar-like feature 마스크(mask)를 이동시키며 구한다. 마스크를 구성하는 사각형 성분들을 크기가 $w \times h$ 인 직사각형 r 이라고 하고, 사각형에 포함된 모든 픽셀값의 합을 $pixelSum(r)$ 이라고 하면, 특징값은 [수식 3-3]으로 구할 수 있다.

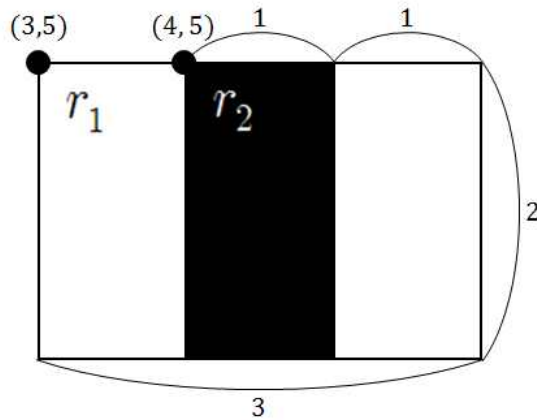
$$feature_i = \sum_{i=1}^N \omega_i \cdot pixelSum(r_i) \dots\dots\dots [수식 3-3]$$

이때, r_i 의 가중치 ω_i 는 [수식 3-4]에서와 같이 서로 반대되는 부호를 갖게 되며, 두 사각형의 넓이비가 다를 경우 이를 보상해준다.

$$-\omega_1 \cdot Area(r_1) = \omega_2 \cdot Area(r_2) \dots\dots\dots [수식 3-4]$$

예를 들면 [그림 3-5]와 같은 라인 특성의 Haar-like feature는 검은색 사각형 r_2 와 검은색 사각형을 포함한 흰색 사각형 r_1 의 차이로 특징값을 계산할 수 있는데, r_1 의 좌상단 좌표가 (3,5), r_2 의 좌상단 좌표가 (4,5)이고, r_1 과 r_2 의 너비는 각각 3과 1이며, 높이는 둘 다 2라고 했을 때, 각 영역의 넓이비는 6:2로 3이므로, [수식 3-5]와 같이 표현할 수 있다.

$$feature = 3 \cdot pixelSum(4,5,1,2) - pixelSum(3,5,3,2) \dots\dots\dots [수식 3-5]$$



[그림 3-6] Line 특성 feature에서의 특징값 계산

1.4 분류기 학습 (Classifier Learning)

Haar-like 특징 추출 단계에서 생성된 얼굴특징계수를 이용하여 얼굴을 분류하기 위해서 본 논문에서는 AdaBoost 알고리즘과 캐스케이드(Cascade) 구조를 이용하였다. AdaBoost 알고리즘은 약분류기(simple, weak classifier)들의 선형

결합을 통하여 최종적으로 높은 검출 성능을 가지는 강분류기(strong classifier)를 생성하여 분류하는 기법으로 Freund와 Schapire에 의해 소개되었으며, 약분류기 수가 증가할수록 에러율이 지수적으로 영점에 근접함이 증명되었다 [46].

이전 단계에서 구한 얼굴특징계수를 이용한 분류 방법은 [수식 3-6]과 같다.

$$h_j = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases} \dots\dots\dots \text{[수식 3-6]}$$

※ f_j 는 feature, θ_j 는 threshold, p_j 는 parity

약분류기는 마스크 내의 Haar-like 특징(feature)들 중 얼굴과 배경을 가장 잘 분류하는 하나의 특징을 선별한다. 이를 위해 적분이미지를 통해 구한 픽셀값의 합을 이용하여 에러가 최소화되는 문턱값(threshold)을 찾는다. 이렇게 구한 약분류기들은 선형적으로 결합하여 최종적인 강분류기를 구성한다. 얼굴 검출을 위한 AdaBoost 학습 알고리즘은 간단하게 [표 3-1]과 같이 요약할 수 있다.

[표 3-1] 얼굴 검출을 위한 AdaBoost 알고리즘 [43]

① 학습이미지 입력: $(x_1, y_1), \dots, (x_N, y_N)$ ※ x_i : 입력영상, y_i : 입력영상이 얼굴일 때 1, 얼굴이 아닐 때 0
② y_i 각각에 부여할 가중치를 다음과 같이 초기화 $w_{t,i} = \left\{ \frac{1}{2m}, \frac{1}{2l} \right\}$ ※ m: 얼굴이 아닌 이미지 수, n: 얼굴 이미지 수
③ 결합할 약분류기의 수만큼 다음 작업을 반복

▶ 가중치 정규화: $w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$

※ $w_{t,i}$ 는 확률분포

▶ 특징점 j 에 대한 약분류기 h_j 를 학습시키며, 에러값을 계산

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$

▶ 가장 작은 에러값(ϵ_t)를 가지는 분류기 h_t 를 선택

▶ 가중치를 다음과 같이 갱신

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad \left(\beta_t = \frac{\epsilon_t}{1-\epsilon_t} \right)$$

※ 학습영상 x_i 가 잘 분류되었으면 $e_i = 0$ 이고, 그 이외에는 $e_i = 1$

④ 최종 결합된 강분류기는 다음과 같다.

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & otherwise \end{cases} \quad \left(\alpha_t = \log \frac{1-\epsilon_t}{\epsilon_t} \right)$$

학습 영상은 얼굴 영상과 얼굴이 아닌 영상으로 구분되어 입력된다. x_i 는 각각의 훈련 영상이며, y_i 는 그 영상이 얼굴 영상인지 아닌지를 나타내주는 인덱스 값이다. 알고리즘 시작 단계에서는 우선 가중치를 초기화한 후, 약분류기의 개수 만큼 반복하면서 에러값이 가장 작은 가중치를 찾는다. 학습 초기에는 큰 특징들을 통해 얼굴 영상의 에러를 줄이고, 뒤로 갈수록 세세한 특징들을 통해 얼굴과 비얼굴을 구별시키는 특징을 판별한다. 따라서 계층의 앞부분은 크기가 크고 특징의 수가 적은 반면, 뒷부분은 크기가 작고 특징의 수는 많게 된다. 이렇게 가중치가 조정된 약분류기들이 선형적으로 결합되면서 최종적으로 강분류기를 생성하게 된다.

1.5 Cascade 구조를 이용한 분류 속도 향상

AdaBoost 알고리즘이 Neural Network나 SVM과 같은 기존의 분류 알고리즘보다 좋은 성능을 보이기 위해서는 많은 수의 훈련 영상과 약분류기들이 필요하며, 이 경우 훈련시간과 검출시간이 늘어나게 된다. Viola는 이러한 계산시간을 줄이기 위해 Cascade 구조를 이용하여 분류기를 구성하였다 [43].

이 방법의 원리는 초기 단계에서 얼굴영상을 유지하면서 비얼굴영상들을 걸러냄으로써 다음 단계로 갈수록 좀 더 복잡하고 구별하기 힘든 특징들을 분류하는데 집중하는 것이다. 이를 위해 단계별로 50% 정도의 오판율(false positive rate, 이하 FPR)을 감수하며, 100%에 가까운 얼굴 검출률(detection rate, 이하 DR)을 갖도록 한다. 이 정도의 높은 FPR 자체는 분류기로서 사용하기에 적합하지 않지만, 많은 수의 후속 단계를 거침으로써 판별력과 효율이 증대되어, 최종적으로는 분류기로 사용하기 적합한 성능을 갖게 된다.

Cascade 분류기의 각 단계에서의 특징수와 총 단계수는 FPR과 DR에 의해 결정된다. K 를 분류기의 수, f_i 와 d_i 를 각각 i 번째 분류기의 FPR과 DR이라고 할 때, 최종 분류기의 전체 FPR과 DR은 다음과 같이 표시할 수 있다.

$$FPR = \prod_{i=1}^K f_i, \quad DR = \prod_{i=1}^K d_i \dots\dots\dots [수식 3-7]$$

따라서, 90%의 DR을 갖는 최종 분류기를 구하기 위해서는 먼저 99%의 DR을 갖는 10단계의 분류기를 구성해야 한다. 그리고 각 단계에서의 FPR을 30%로 만들면, 총 FPR은 $0.30^{10} \approx 6.10^{-6}$ 으로 기하급수적으로 줄어들게 된다.

2 얼굴 인식

본 논문에서는 얼굴 인식을 하기 위한 방법으로 PCA 기법[9]을 이용하였다. PCA는 얼굴에서 구성요소들을 일일이 분리하지 않고 하나의 주성분으로 생각하여 처리하는 방식으로 각 입력영상을 1차원 벡터로 변환한 후 전체 입력영상들의 직교적인 공분산 행렬(orthogonal covariance matrix)을 계산한다. 이 후에 공분산행렬의 고유값(eigen-value)에 따라 고유벡터(eigen-vector)를 구하고, 입력된 얼굴과 기준 얼굴의 고유벡터를 벡터 공간상에서 비교하여 가장 가까운 거리의 벡터를 지닌 얼굴이 인식되도록 하는 방법이다. 이 방법은 인식률을 떨어뜨리지 않으면서도 데이터의 양을 축소하는 효과가 있다.

2.1 얼굴인식 과정

얼굴영상의 다차원의 정보는 얼굴을 표현하기에 적절하지 못하며, 얼굴을 묘사할 수 있는 얼굴 공간(face space)으로 표현 되어야 한다. 따라서 얼굴 인식에서의 PCA 기법은 고차원의 얼굴 정보들을 얼굴 공간의 저차원으로 축소시킴으로써 데이터를 다루기 쉽게 하기 위한 방법이라 할 수 있다.

PCA를 이용한 얼굴 인식은 크게 학습, 입력, 인식의 세 단계로 구분된다. 학습단계는 입력 영상과 비교할 수 있는 학습영상 집합으로부터 평균 얼굴 벡터를 계산한 후 각 학습 얼굴 벡터와의 차를 구하여 공분산 행렬을 통해 고유 값과 고유벡터를 계산한다. 이 값이 구해지면 고유 벡터를 크기순으로 재배열하여 얼굴 모양과 비슷한 고유 얼굴을 생성하며 학습영상을 아주 작은 고유 값에 대응하는 고유 얼굴을 제외한(학습영상의 공통적인 특징을 가장 잘 나타내는 벡터)

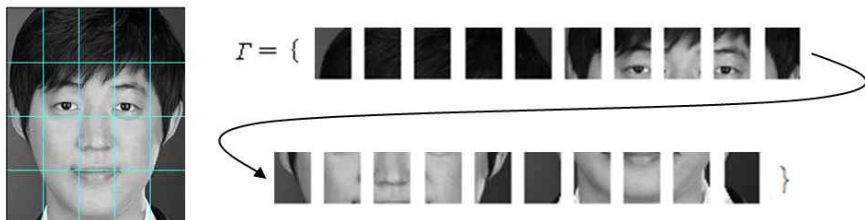
학습영상의 특징을 타나내는 특징 벡터를 생성한다.

입력 단계에서는 학습 단계에서 생성된 평균 얼굴 벡터와 고유 얼굴을 사용한다. 학습 단계와 마찬가지로 평균얼굴의 차영상을 고유 얼굴 공간에 투사시켜 입력 영상의 특징 벡터의 값을 추출한다.

마지막으로 인식 단계에서는 학습 영상의 특징 벡터와의 유클리디안 거리값 (euclidean distance)을 비교하여 그 거리가 최소가 되는 얼굴이 입력 얼굴과 유사한 특징을 가진 얼굴이므로 이 영상을 입력 얼굴과 동일한 인물로 결정한다.

2.2 얼굴벡터 집합 구성

학습 얼굴 영상을 하나의 얼굴 벡터로 표현하기 위해 얼굴 영상의 각 픽셀을 행의 순서로 재구성한다. 각각의 얼굴 영상의 크기가 $W \times H$ 일 때 각 픽셀 정보를 하나의 행으로 연결하여 벡터성분으로 표현하면 [그림3-7]과 같이 $(W \times H) \times 1$ 의 행 벡터인 얼굴 벡터가 생성된다.



[그림3-7] 얼굴벡터 구성의 예

생성된 각 얼굴 벡터들은 하나의 얼굴 벡터 행렬로 구성되어야 하는데, 얼굴 영상의 크기가 $W \times H = N$ 이고 학습 얼굴 영상의 수가 M 개 일 때, 각각의 얼굴벡터 Γ_M 과 각 얼굴 벡터들을 하나의 얼굴 벡터 행렬 S 로 표현한다.

$$\Gamma = \{V_1, V_2, V_3, \dots V_N\} \dots\dots\dots [\text{수식 3-8}]$$

$$S = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots \Gamma_M\} \dots\dots\dots [\text{수식 3-9}]$$

2.3 이미지 정규화

PCA를 이용한 얼굴인식에서는 사진크기, 조명, 배경, 카메라의 각도 및 거리와 같은 외적 요인에 따라 인식률의 변화가 매우 심하게 나타난다. 따라서 얼굴 인식에 영향을 미치는 외적 요인 중 사진크기와 조명에 대한 영향을 줄이기 위해 전처리 과정으로 이미지 리사이징과 히스토그램 평활화를 사용하여 사진크기와 조명에 의하여 발생하는 에러를 최소화 하여야 한다.

이미지 리사이징 과정은 본 연구의 얼굴 검출 과정에서 데이터베이스의 학습 영상과 동일한 얼굴영상을 추출하고자 추출된 이미지를 학습영상의 이미지와 크기를 맞추는 이미지의 변환 과정이 필요하다.

히스토그램 평활화(histogram equalization) 알고리즘은 한쪽에 치우친 명암 분포를 가진 히스토그램을 재분배 과정을 거쳐 일정한 분포를 가지게 하는 알고리즘이다.

히스토그램 평활화의 일반적인 식은 다음과 같다.

















$$HEQ = \frac{h_s}{n} \times I_m \dots\dots\dots [수식 3-10]$$

※ h_s : 히스토그램 누적 합, n : 영상의 픽셀 개수, I_m : 명암값 개수

히스토그램의 평활화 과정은 다음과 같다.

- ① 히스토그램 생성
- ② 히스토그램의 값을 정규화 하여 누적 합을 계산
- ③ 입력 영상의 화소 위치를 파악, 결과 영상 생성

전처리 과정 중 하나로 히스토그램 평활화를 사용하는 이유는 단순히 영상의 히스토그램을 균일하게 해주는 것 뿐 아니라, [그림 3-8]에서 볼 수 있는 것처럼 어두운 영상은 밝게, 너무 밝은 영상은 조금 어둡게 하는 등 적당한 명암 값을 유지하게 함으로써 인식 대상의 특징정보를 균등하게 적용시켜 조명에 대한 인식률을 향상시킬 수 있기 때문이다.

원 이미지		히스토그램 평활화 적용 이미지	
			
			
			
			

[그림 3-2] 이미지 정규화

2.4 공분산을 이용한 변환행렬 도출

얼굴 영상 공간에서의 각각의 픽셀이 다른 픽셀과 연관성이 많기 때문에 얼굴의 자료를 나타낼 때 불필요한 정보의 중복이 나타나게 된다. 실제로 학습단계에서의 얼굴벡터집합의 공분산 행렬은 고차원 비대각화 행렬(high non-diagonal)이라는 것을 알 수 있으며 공분산 행렬은 아래와 같이 나타낼 수 있다.

$$\sum S = S \times S^T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,N \times M} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2,N \times M} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{N \times M, 1} & \sigma_{N \times M, 2} & \cdots & \sigma_{N \times M, N \times M} \end{bmatrix} \dots\dots\dots [\text{수식 3-11}]$$

※ $\sum S$: 공분산 행렬 , σ_{ij} : i번째 얼굴과 j번째 얼굴의 공분산

각각의 얼굴 벡터가 어떠한 벡터와도 연관되지 않는 얼굴 영역을 만드는 것이 목표이므로 공분산 행렬은 대각 행렬이어야 한다.

$$\sum A = A \times A^T = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_{N \times M, N \times M} \end{bmatrix} \dots\dots\dots [\text{수식 3-12}]$$

분산을 최대화하는 방법으로 각 변수의 상관관계가 자기 자신과 최대가 되고 다른 성분들과는 0이 되게 하기 위해 W를 [수식 3-10]을 만족하는 변환 행렬이라고 정의한다.

$$A = P^T \times S \dots\dots\dots [수식 3-13]$$

여기서 어떠한 W 를 구해야만 A 의 공분산 행렬이 대각화 하게 되느냐가 관건이며 [수식 3-13]을 이용하면 다음과 같은 식을 유추할 수 있다.

$$\sum A = A \times A^T = P^T \times S \times S^T \times P = P^T \times \sum S \times P \quad [수식 3-14]$$

A 를 $\sum S$ 의 고유 값을 포함하는 대각화 행렬이라고 할 때 W 를 $\sum S$ 의 고유벡터를 포함하는 다음과 같은 식으로 정의할 수 있다.

$$\sum S \times W = W \times A \dots\dots\dots [수식 3-15]$$

결과적으로 $\sum A$ 는 $\sum S$ 의 고유 값을 포함하는 대각 행렬이 됨을 알 수 있으며 이것을 이용하여 변환행렬 W 를 구할 수 있다.

2.5 고유얼굴 생성

공분산 행렬은 주성분을 분석하고 축약하여 고차원의 자료를 저차원으로 변환함으로써 원 자료의 복잡성을 제거할 수 있게 하는 하나의 절차이며 공분산 행렬 및 주성분 분석을 위한 과정은 다음과 같이 수행된다.

- ① 이미지 정규화 과정을 거친 후 얼굴 벡터의 집합 S 로부터 평균 얼굴 벡터

를 계산한다.

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \dots\dots\dots [\text{수식 3-16}]$$



[그림 3-9] 평균얼굴 벡터

② 각 학습 얼굴 벡터와 평균 얼굴 벡터의 차를 계산한다.

$$\Phi_i = \Gamma_i - \Psi \dots\dots\dots [\text{수식 3-17}]$$

③ 차 벡터 ϕ_i 로부터 공분산 행렬을 계산한다.

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T \dots\dots\dots [\text{수식 3-18}]$$

④ 공분산 행렬로부터 M 개의 고유 값 λ_i 와 이에 대응하는 고유벡터 u_i 를 계산한다.

$$Cu_i = \lambda_i u_i \quad \text{단 } u_i \neq 0 \quad \dots\dots\dots [\text{수식 3-19}]$$

공분산 행렬에서 구한 고유 값은 분산의 정도를 나타내며 이에 대응하는 고유 벡터를 $W \times H$ 로 재배열하면 얼굴의 모양과 닮아 보이므로 이를 “고유 얼굴(Eigenface)” 이라고 한다. 고유 얼굴의 개수는 학습영상 개수와 동일하며 고유 얼굴을 생성하는 식은 다음과 같다.

$$u_l = \frac{1}{M} \sum_{k=1}^M \lambda_{lk} \Phi_k \quad l = 1, 2, \dots, M \quad \dots\dots\dots [\text{수식 3-20}]$$



[그림 3-10] 고유얼굴

고유 얼굴은 얼굴 공간을 구성하는 벡터에서 학습 영상의 공통적인 특징을 가장 잘 나타내는 기저 벡터를 의미하며 각 얼굴 벡터와 평균 얼굴 벡터의 차

벡터로 공분산 행렬의 고유벡터에 해당한다.

2.6 고유성분벡터 추출

공분산 행렬의 고유치는 평균얼굴 영상에 대한 분산의 정도를 나타내기 때문에 가장 큰 고유치에 대응하는 고유벡터로 이루어진 고유 얼굴이 가장 비슷한 얼굴이 되며, 고유치가 작을수록 얼굴의 특징이 작아진다. 본 연구에서는 큰 고유치에 해당하는 M' 의 학습 얼굴영상에 대해서만 고유벡터로 사용하며 입력영상의 성분벡터 비교를 위해 각각의 고유얼굴 성분 벡터[47]를 생성한다.

$$w_k = u_k^T(\Gamma - \Psi) \quad k = 1, 2, \dots, M' \dots\dots\dots [\text{수식 3-21}]$$

$$\Omega^T = [w_1, w_2, \dots, w_{M'}] \dots\dots\dots [\text{수식 3-22}]$$

2.7 얼굴 인식

Ω_k 값이 구해지면 학습 얼굴영상의 고유 얼굴상에서의 성분벡터와 입력영상의 성분벡터와의 유클리디안 거리값을 비교하여 그 거리가 최소가 되는 얼굴이 입력 얼굴과 가장 유사한 얼굴이므로, 최종적으로 이 영상 후보를 인식 결과로 결정한다. 그리고 유클리디안 거리 값의 오차가 일정 임계값보다 큰 경우에는 학습 얼굴영상에 일치하는 얼굴이 없음을 나타내는 결과를 출력한다.

IV 연구 결과 및 토의

1 실험 환경

제안한 얼굴인식 기반의 태깅 자동화 시스템은 AMD Athlon(TM) X2 Dual Core CPU 5200+ @ 2.70GHz, 2G DDR2 RAM, ATI Radeon HD3400 VGA 사양의 환경에서 실험하였다. 2개의 코어로 구성된 CPU지만 하나의 코어만을 사용하도록 프로그램 하였다. 향후 모든 코어를 사용하면 한 방향의 얼굴을 검출하는 시간동안 두 방향의 얼굴 검출이 가능할 것으로 예측된다.

실험에 사용한 코드는 Microsoft Windows XP Professional 운영체제에서 C# 언어와 OpenCV 라이브러리[48]를 사용하여 컴파일(compile)하였다. 입력 영상으로 사용한 인물 사진들은 Nikon D80 카메라를 사용하여 얼굴 기울어짐에 상관없이 1936×1296의 해상도로 실내에서 촬영되었으며, 이후 프로필 사진은 240×200, 그룹 사진은 800×600으로 크기를 조정하였다.

또한, 향후 웹서비스로의 전환이 용이하도록 Microsoft의 닷넷 프레임워크를 플랫폼으로 사용하였으며, SNS 환경과 최대한 유사하게 제작하기 위해 모든 프로필 정보와 영상 콘텐츠들은 각각 독립적으로 구축된 데이터베이스에 저장하여 사용하였다.

2 시스템 인터페이스

제안한 얼굴인식 기반의 자동화 태깅 시스템을 구현하기 위해 20대 남녀 10

명의 프로필 사진과 각각의 이름, ID, 이메일을 기록한 사용자 정보 테이블(member table), 그리고 여러 사람이 모여서 찍은 단체사진과 각 사진의 제목, 태그를 기록한 사진 테이블(picture table)로 구성된 데이터베이스를 구축하였다. 이때 두 테이블은 서로 독립적으로 어떠한 매개체도 가지고 있지 않다.

이후 SNS의 사진 관리 시스템을 모방하여 [그림 4-1]과 같이 시스템 인터페이스를 제작하고, 여기에 구축한 데이터베이스를 연동하여 사진 테이블에 기록된 사진 정보들을 열람할 수 있도록 하였다. 사용자는 시스템을 통해 새로운 프로필이나 사진을 데이터베이스에 등록 또는 수정하거나 삭제하는 것이 가능하며, 아이디를 입력하여 사용자 정보를 검색할 수 있다.



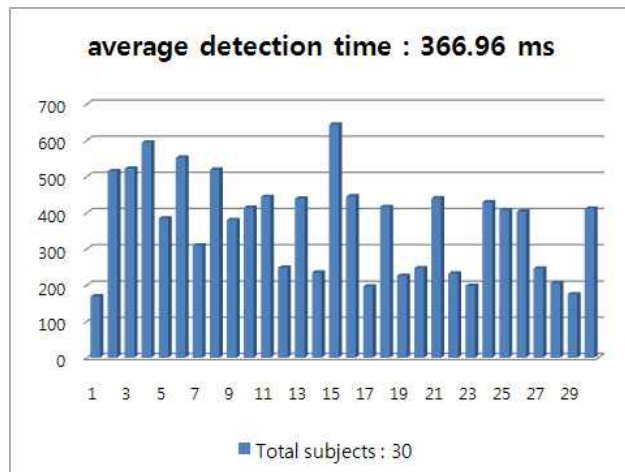
[그림 4-1] 시스템 인터페이스 및 결과 화면

사용자가 활성화버튼(Activate)을 눌러 얼굴 검출 및 인식을 활성화한 후, 마우스 커서를 사진 속 인물의 얼굴에 올리면 검출된 얼굴 영역이 표시된다. 그리고 그 얼굴을 클릭하면 해당 인물의 프로필 사진 및 정보가 출력되고, 태그 목록에는 해당 인물의 이름이 태깅된다. 이때 입력되는 태그는 사용자가 입력하는 것이 아니라, 시스템이 얼굴을 인식하여 자동으로 출력하는 것이다.

3 실험 결과

제작된 시스템의 성능을 평가하기 위해 얼굴 검출시 소요되는 시간과 검출률, 얼굴영상 학습시간, 그리고 얼굴 인식시 소요되는 시간과 인식률을 측정하였다.

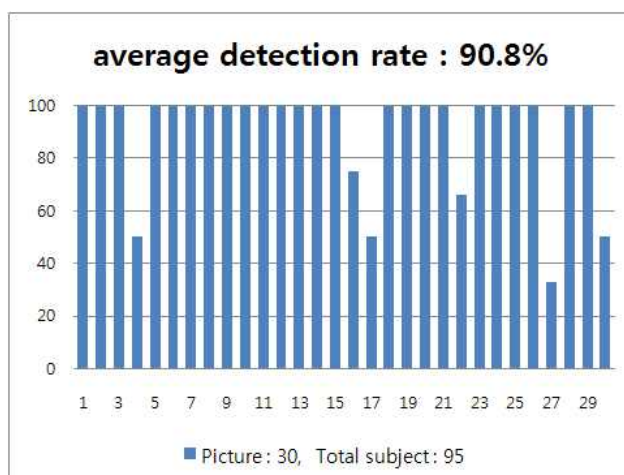
먼저 실험에 사용된 30장의 단체사진을 이용하여 총 95명의 얼굴을 검출한 시간은 [그림 4-2]와 같다.



[그림 4-2] 얼굴 검출 시간

평균 검출시간은 366.69ms 로 매우 양호한 검출 시간을 보였으며 이는 정지 영상이 아닌 동영상에서의 얼굴을 실시간으로도 검출가능한 시간이다. 따라서 정지영상뿐만 아니라 동영상을 다루는 시스템에서도 적극 활용될 수 있다.

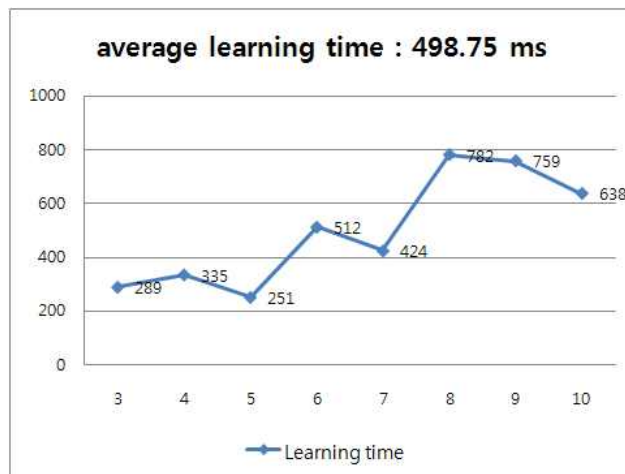
다음으로 30장의 단체사진에 대한 95명의 얼굴 검출률은 [그림 4-3]와 같은 결과를 보였다.



[그림 4-3] 얼굴 검출률

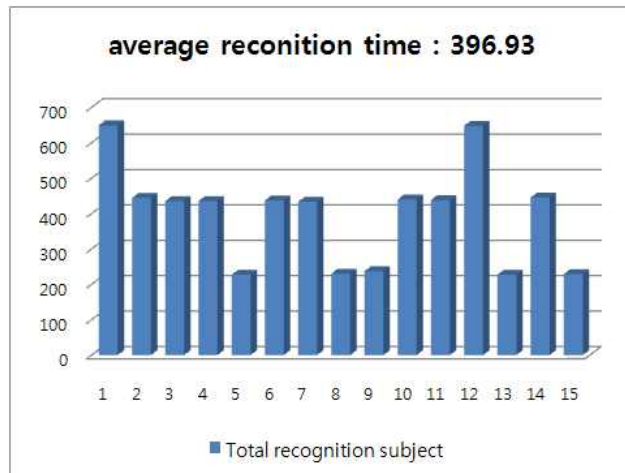
조명변화와 다양한 각도에서의 얼굴 사진에도 강인한 검출률을 보였으며, 얼굴 크기, 피부색, 안경착용, 얼굴가림 등과 같이 환경적인 요인에도 불구하고 평균 90.8%의 높은 검출률을 보였다. 이는 다양한 카메라기기나 모바일 디바이스 기기에 적용되어 얼굴검출을 자동으로 하는 시스템을 적용시킬 수 있다.

얼굴 인식 단계 이전에 입력이미지와 비교를 위한 학습데이터가 필요하므로 10장의 얼굴사진을 트레이닝 하였다. 3~10명에 대한 각각의 트레이닝 시간은 [그림 4-4]와 같으며 학습이미지의 양이 늘어날수록 평균적으로 시간이 증가하는 것을 확인할 수 있다.

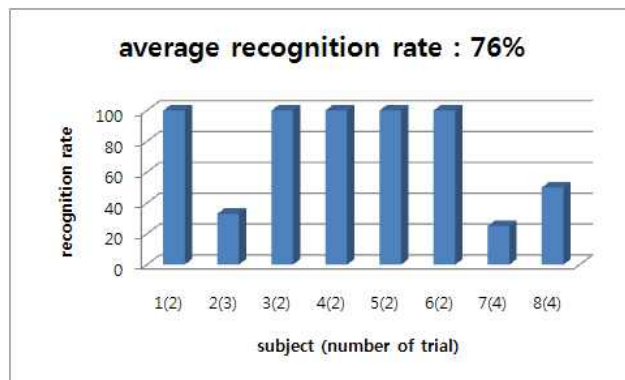


[그림 4-4] 얼굴 학습 시간

학습된 데이터를 대상으로 한 입력영상에 대한 인식시간과 인식률은 각각 [그림 4-5]와 [그림 4-6]에 나타난 바와 같다. 8명의 얼굴 학습 데이터를 이용하여 10장의 단체사진 이미지에 존재하는 21명의 얼굴을 인식 테스트한 결과 평균 76%의 인식률을 보였다. 30~50%의 인식률을 보이는 이미지에 대해서는 조명변화와, 얼굴의 기울어짐의 정도가 심한 것으로 확인되었고, 그 이외의 입력 이미지에 대해서는 높은 인식률을 보였다. 따라서 인식률을 향상시키기 위해서는 이러한 환경적인 변화에 강인한 알고리즘에 대해 지속적인 연구가 필요할 것이다.



[그림 4-5] 얼굴 인식 시간



[그림 4-6] 얼굴 인식률

V 결 론

본 논문에서는 텍스트 기반의 수동적인 기존 태깅 시스템과 차별되는 다양한 패턴인식 알고리즘을 적용한 얼굴인식 기반의 태깅 자동화 방법에 대해 제안하고, 이를 적용한 사진 관리 시스템을 개발하였다. 이를 위해 먼저 얼굴검출 및 인식에 대한 다양한 방법들에 소개하고, 현재 가파르게 성장하고 있는 SNS 시장과 태깅 관련 기술에 대해 알아본 후, 개발한 시스템에서 사용된 방법들에 대해 상세하게 기술하였다. 얼굴검출 과정에서는 Haar-like 특징과 AdaBoost를 이용하여 높은 검출률과 속도를 가진 분류기를 생성했으며, 인식 과정에서는 PCA를 이용하여 고유얼굴을 생성한 후, 여기서 추출된 고유벡터 성분을 비교하여 오차가 가장 적은 얼굴을 선택하여 인식하였다.

웹2.0을 지나 웹3.0으로 도약하는 현 인터넷 시장에 인공지능 알고리즘을 적용하여 새로운 형태의 웹서비스를 제공하고, 자동화된 얼굴인식을 통해 사용자로 하여금 재밌고 편리한 인터넷 콘텐츠 이용을 가능하게 하는데 큰 의의가 있으며, 또한 인맥 네트워크를 중시하는 SNS 에서, 사용자가 더욱 넓은 인맥을 구축하거나, 잃어버린 인맥을 복구할 수 있는 환경을 제공해줄 것으로 기대된다. 2006년 마이스페이스에 이어, 2007년 페이스북, 2008년 트위터(Twitter)와 믹시(Mixi), 그리고 지금도 끊임없이 생겨나고 있는 다양한 SNS 및 웹서비스들에 대해 세계의 관심이 증가하는 가운데, 이러한 사용자 중심의 독창적인 서비스는 기술적으로 정제되어 있는 국내 인터넷 시장에 활기를 불어넣어주고, 유망 기술로 떠오르고 있는 패턴인식 분야의 연구 수준을 향상시킬 것으로 기대된다.

참 고 문 헌

- [1] 박종진, “인터넷 영상 소비 패턴과 업계 주도권 전망: SBSi의 온라인 디지털 콘텐츠 사업전략,” 삼성경제연구소 디지털사례세미나 SBSi, 2006. 5.
- [2] 한국소프트웨어진흥원, “포털사이트를 중심으로 한 국내 웹 2.0 서비스 현황과 전망,” SW산업동향, 2006.
- [3] Scott Bateman, Christopher Brooks and Gord McCalla, “Collaborative Tagging Approaches for Ontological Metadata in Adaptive E-Learning Systems,” Proceedings of 4th International Workshop on Applications of Semantic Web Technologies for e-Learning, pp.3-12, June 2006.
- [4] Ramma Chellappa, Charles L. Wilson, and Saad Sirohey, “Human and Machine Recognition of Faces: A Survey,” Proceedings of the IEEE, Vol.83, No.5, pp.705-741, May 1995.
- [5] Sina Jahanbin, Hyohoon Choi, Alan C. Bovik, and Kenneth R. Castleman, "Three Dimensional Face Recognition using Wavelet Decomposition of Range Images," Proceedings of IEEE International Conference on Image Processing, Vol.1, pp.145-148, September 2007.
- [6] Ji Tao, Yap-Peng Tan, "People Monitoring using Face Recognition with Observation Constraints," Proceedings of IEEE International Conference on Image Processing, Vol.3, pp.1437-1440, October 2004.
- [7] David. D. Zhang, Automated Biometrics: Technologies and System, Kluwer Academic Publishers, Boston, May 2000.
- [8] Athinodoros S. Georgiades, Peter. N. Belhumeur, David. J. Kriegman,

- "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.6, pp.643–660, June 2001.
- [9] Matthew Turk, Alex Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71–86, 1991.
- [10] Chengjun Liu, Harry Wechsler, "Enhanced Fisher Linear Discriminant Models for Face Recognition," Proceedings of 14th International Conference on Pattern Recognition, Vol.2, p.1368, August 1998.
- [11] Y. K. Ham, S. Y. Lee and R. H. Park, "Knowledge-based Face Recognition using Neural Networks," Proceedings of International Joint Conference on Neural Networks, Vol.3, pp.2125–2128, October 1993.
- [12] Dennis Gabor, "Theory of communication," Journal of the Institute of Electrical Engineering, Vol.93, No.26, pp.429–459, November 1946.
- [13] Bernd Heisele, Purdy Ho and Tomaso Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach," Proceedings of 8th International Conference on Computer Vision, Vol.2, pp.688–694, 2001.
- [14] Meng Joo Er, Shiqian Wu, Juwei Lu and Hock Lye Toh, "Face Recognition with Radial Basis Function(RBF) Neural Networks," IEEE Transactions on Neural Networks, Vol.13, No.3, May 2002.
- [15] Su Hongtao, David Dagan Feng, Zhao Rong-Chun, "Face Recognition using Multi-feature and Radial Basis Function Network," Conferences in Research and Practice in Information Technology, Vol.22, pp.51–57, 2003.
- [16] Lakhmi C. Jain, Uqur Halici, Isao Hayashi, S. B. Lee, Shiqeyoshi Tsutsui, Intelligent Biometric Techinques in Fingerprint and Face

- Recognition, CRC Press, Boca Raton, pp.239–245, 1999.
- [17] Vladimir N. Vapnik, "The Nature of Statistical Learning Theory," Springer–Verlag, New York, 1995.
 - [18] Christopher J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, Vol.2, No.2, pp.121–167, 1998.
 - [19] Edgar Osuna, Robert Freund, and Federico Girosi, "Training Support Vector Machines: An Application to Face Detection," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.130–136, 1997.
 - [20] Ming–Hsuan Yang, David J. Kriegman and Narendra Ahuja, "Detecting Faces in Images: A Survey," *Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.1, pp.34–58, January 2002.
 - [21] Guangzheng Yang, Thomas S. Huang, "Human Face Detection in Complex Background," *Pattern Recognition*, Vol.27, No.1, pp.53–63, 1994.
 - [22] Thomas K. Leung, Michael C. Burl, Pietro Perona, "Finding Faces in Cluttered Scenes Using Random Labeled Graph Matching," *Proceedings of 5th IEEE International Conference on Computer Vision*, pp.637–644, June 1995.
 - [23] Kin C. Yow, Roberto Cipolla, "Feature–based Human Face Detection," *Image and Vision Computing*, Vol.15, No.9, pp.713–735, September 1997.
 - [24] Ying Dai, Yasuaki Nakano, "Face–Texture Model based on SGLD and Its Application in Face Detection in a Color Scene," *Pattern Recognition*, Vol.29, No.6, pp.1007–1017, June 1996.
 - [25] Jie Yang, Alex Waibel, "A Real–time Face Tracker," *Proceedings of 3rd IEEE Workshop on Applicatinos of Computer Vision*, pp.142–147, 1996.
 - [26] Stephen J. McKenna, Shaogang Gong and Yogesh Raja, "Modelling

- Facial Colour and Identity with Gaussian Mixtures," *Pattern Recognition*, Vol.31, No.12, pp.1883–1892, December 1998.
- [27] Rick Kjellden, John R. Kender, "Finding Skin in Color Images," *Proceedings of 2nd International Conference on Automatic Face and Gesture Recognition*, pp.312–317, October 1996.
- [28] Ian Craw, David Tock and Alan Bennett, "Finding Face Features," *Proceedings of 2nd European Conference on Computer Vision*, pp.92–96, 1992.
- [29] Andreas Lanitis, Christopher J. Taylor, and Timothy F. Cootes, "Automatic Face Identification System using Flexible Appearance Models," *Image and Vision Computing*, Vol.13, No.5, pp.393–401, June 1995.
- [30] Kah-Kay Sung, Tomaso Poggio, "Example-based Learning for View-based Human Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20, No.1, pp.39–51, January 1998.
- [31] Henry Schneiderman, Takeo Kanade, "Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.45–51, July 1998.
- [32] Ambasamudram N. Rajagopalan, K. Sunil Kumar, Jayashree Karlekar, R. Manivasakan, M. Milind Patil, Uday B. Desai, P. G. Poonacha, and Subhasis Chaudhuri, "Finding Faces in Photographs," *Proceedings of 6th IEEE International Conference on Computer Vision*, pp.640–645, 1998.
- [33] Danah M. Boyd, Nicole B. Ellison, "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Mediated Communication*, Vol.13, No.1, April 2007.

- [34] Su Yun Chae, "Evolution of Social Network Service into Media," ICA 42nd Conference, Seoul, Korea, October 2008.
- [35] John Breslin, Stefan Decker, "The Future of Social Networks on the Internet: The Need for Semantics," IEEE Internet Computing, Vol.11, No.6, pp.86–90, November 2007.
- [36] Tye Rattenbury, Nathan Good, Mor Naaman, "Towards Extracting Flickr Tag Semantics," Proceedings of 16th International Conference of World Wide Web, pp.1287–1288, 2007.
- [37] Martin J. Halvey, Mark T. Keane, "An Assessment of Tag Presentation Techniques," Proceedings of 16th International Conference on World Wide Web, pp.1313–1314, 2007.
- [38] Alexei Pyalling, Michael Maslov, Pavel Branslavski, "Automatic Geotagging of Russian Web Sites," Proceedings of the 15th International Conference on World Wide Web, pp.965–966, 2006.
- [39] Scott A. Golder, Bernardo A. Huberman, "Usage Patterns of Collaborative Tagging Systems," Journal of Information Science, Vol.32, No.2, pp.198–208, 2006.
- [40] Paul–Alexandru Chirita, Stefania Costache, Wolfgang Nejdl, Siegfried Handschuh, "P–TAG: large scale automatic generation of personalized annotation tags for the web," Proceedings of the 16th International Conference on World Wide Web, 845–854, 2007.
- [41] Alfred Haar, "Zur Theorie der Orthogonalen Funktionen–Systeme," Mathematische Annalen, Vol.69, pp.331–371, 1910.
- [42] Constantine P. Papageorgiou, Michael Oren, Tomaso Poggio, "A General Framework for Object Detection," Proceedings of 6th IEEE

- International Conference on Computer Vision, pp.555–562, 1998.
- [43] Paul Viola, Michael Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision, Vol.57, No.2, pp.137–154, May 2004.
- [44] Rainer Lienhart, Alexander Kuranov and Vadim Pisarevsky, "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection," MRL Technical Report, May 2002.
- [45] Franklin C. Crow, "Summed-area Tables for Texture Mapping," Proceedings of 11th Annual Conference on Computer Graphics and Interactive Techniques, Vol.18, No.3, pp.207–212, 1984.
- [46] Yoav Freund, Robert E. Schapire, "A Decision-Theoretic Generalization of On-line Learning and An Application to Boosting," Journal of Computer and System Sciences, Vol.55, No.1, pp.119–139, August 1997.
- [47] Peter N. Belhumeur, João P. Hespanha, David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No.17, pp.711–720, July 1997.
- [48] OpenCV Library, <http://opencvlibrary.sourceforge.net/>