creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:24011-200000264609

조선대학교
CHOSUN UNIVERSITY

2015年 2月

博士學位論文

거리별 얼굴 영상을 이용한 실시간 원거리 얼굴 인식에 대한 연구

朝鮮大學校 大學院

情報通信工學科

文 解 旻

조선대학교
CHOSUN UNIVERSITY

거리별 얼굴 영상을 이용한 실시간 원거리 얼굴 인식에 대한 연구

A Study on the Real-time Long Distance Face Recognition using Face Images by Distance

2015年 2月 25日

朝鮮大學校 大學院

情報通信工學科

文 解 旻

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

거리별 얼굴 영상을 이용한 실시간 원거리 얼굴 인식에 대한 연구

指導敎授 潘 聲 範

이 論文을 工學 博士學位申請 論文으로 提出함

2014年 10月

朝鮮大學校 大學院

情報通信工學科

文 解 旻

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

文解旻의 博士學位 論文을 認准함

委員長	公州大學校	教授	徐 彰 浩	印
委　員	한국인터넷진흥원	室長	朴 海 龍	印
委　員	朝鮮大學校	教授	郭 根 昌	印
委　員	朝鮮大學校	教授	金 潤 泰	印
委　員	朝鮮大學校	教授	潘 聲 範	印

2014年　12月

朝鮮大學校 大學院

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

목 차

표목차 ····· iii

도목차 ····· iv

ABSTRACT ····· vi

제1장 서 론 ····· 1

제1절 연구 배경 ····· 1

제2절 연구 목적 ····· 4

제3절 연구 내용 및 방법 ····· 6

제2장 영상 감시 시스템에서의 얼굴 인식 ····· 9

제1절 영상 감시 시스템 ····· 9

제2절 대표적 얼굴 인식 기술 ····· 13

1. 얼굴 밝기 정규화 ····· 14

2. 얼굴 크기 정규화 방법 ····· 18

3. 특징 추출 방법 ····· 23

제3절 기존 원거리 얼굴 인식 방법 ····· 26

1. 초해상도 복원 방법을 이용한 저해상도 얼굴 인식 ····· 26

2. 구조적 특징을 이용한 저해상도 얼굴 인식 ····· 28

3. PTZ 카메라를 이용한 원거리 얼굴 인식 ····· 30

4. 기존 원거리 얼굴 인식 방법 분석 ····· 33

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3장 제안하는 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 ····· 35
제1절 기존 단일 거리 얼굴 영상을 이용한 얼굴 인식 ····· 37
1. 단일 영역기반 얼굴 영상 ····· 37
2. 다중 영역기반 얼굴 영상 ····· 40
제2절 제안하는 거리별 얼굴 영상을 이용한 얼굴 인식 ····· 42
1. 실제 거리별 얼굴 영상 ····· 42
2. 줌 카메라를 이용한 가상의 거리별 얼굴 영상 ····· 48
3. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상 ····· 50
제4장 실험 조건별 원거리 얼굴 인식 성능 분석 ····· 56
제1절 실험 방법 ····· 56
1. 얼굴 DB ····· 56
2. 성능 평가방법 ····· 62
제2절 거리별 얼굴 인식률 분석 ····· 63
1. 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 ····· 63
2. 저역 통과 필터를 이용한 근거리 얼굴 영상의 화질 열화 ····· 67
3. 양선형 보간법을 이용한 원거리 얼굴 크기 정규화 ····· 69
4. 히스토그램 평활화를 이용한 원거리 얼굴 밝기 정규화 ····· 72
제3절 기준 영상 크기에 따른 얼굴 인식률과 수행 시간 관계 ····· 74
제4절 제안하는 원거리 얼굴 인식 기술 분석 ····· 77
제5장 결 론 ····· 80
참고문헌 ····· 83

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 목 차

표 1. 지능형 CCTV 시스템의 시장 규모 ····· 2
표 2. 기존 원거리 얼굴 인식 방법 비교 ····· 33
표 3. ETRI 얼굴 DB ····· 58
표 4. IPES 얼굴 DB ····· 59
표 5. 화질 열화에 따른 얼굴 인식 실험 조건 ····· 68
표 6. 얼굴 크기 정규화 방법에 따른 얼굴 인식 실험 조건 ····· 70
표 7. 얼굴 밝기 정규화 방법에 따른 얼굴 인식 실험 조건 ····· 72
표 8. 기준 영상 크기에 따른 얼굴 인식 실험 조건 ····· 75
표 9. 원거리 얼굴 인식 방법에 따른 전체 거리 평균 인식률(%) ····· 78
표 10. 원거리 얼굴 인식 방법 비교 ····· 79

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 목차

그림 1. 시간 경과에 따른 감시자의 감시 효율 변화 ····· 3
그림 2. CCTV 기반 영상 감시 시스템 ····· 10
그림 3. 영상 감시 시스템의 진화 ····· 11
그림 4. 영상의 해상도 종류 ····· 12
그림 5. HD급 CCTV와 SD급 CCTV의 화질 비교 ····· 12
그림 6. 얼굴 밝기 정규화 방법에 따른 얼굴 영상 ····· 17
그림 7. 이중 선형-3차회선 보간법의 참조화소 ····· 19
그림 8. 수직 선형-3차회선 보간법의 동작 ····· 20
그림 9. 수평 선형-3차회선 보간법의 동작 ····· 22
그림 10. 예제 기반 초해상도 복원의 예 ····· 27
그림 11. CLPMs을 이용한 저해상도 얼굴 인식 방법의 개념도 ····· 30
그림 12. 고정 카메라와 PTZ 카메라를 이용한 얼굴 인식 시스템 ····· 31
그림 13. 제안하는 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 알고리즘 36
그림 14. 근거리와 원거리 얼굴 영상의 비교 ····· 37
그림 15. 1m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법 ····· 38
그림 16. 2m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법 ····· 39
그림 17. m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법 ····· 39
그림 18. 얼굴 영역에 따른 후보 간 유사도 특징 ····· 40
그림 19. 근거리와 원거리에서 검출된 얼굴 영역 비교 ····· 41
그림 20. 1m 얼굴의 다중 영역을 이용한 학습 데이터 구성 방법 ····· 41
그림 21. 실제 거리별 얼굴 영상 획득 방법 ····· 43

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 22. 실제 거리별 얼굴 영상을 이용한 학습 데이터 구성 방법 ·····43
그림 23. 보간법을 통해 정규화된 얼굴 영상 ·····45
그림 24. 가우시안 저역 통과 필터링이 적용된 거리별 얼굴 영상 ·····47
그림 25. 광학식 줌을 이용한 거리별 얼굴 영상 획득 방법 ·····49
그림 26. 줌 카메라를 이용한 가상의 거리별 얼굴 영상 구성 방법 ·····49
그림 27. 실제 거리별 얼굴 영상과 가상의 거리별 얼굴 영상의 비교 ·····50
그림 28. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상 생성 과정 ···51
그림 29. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상 생성 예 ·····52
그림 30. 가상의 거리별 얼굴 영상과 실제 거리별 얼굴 영상의 비교 ·····54
그림 31. ETRI DB의 예 ·····58
그림 32. IPES-640 DB의 예 ·····60
그림 33. IPES-1280 DB의 예 ·····61
그림 34. QVGA급 영상에서의 원거리 얼굴 인식률 비교 ·····64
그림 35. VGA급 영상에서의 원거리 얼굴 인식률 비교 ·····65
그림 36. SXGA급 영상에서의 원거리 얼굴 인식률 비교 ·····66
그림 37. 화질 열화에 따른 전체 거리 평균 인식률 ·····69
그림 38. 얼굴 크기 정규화 방법에 따른 PCA 기반 원거리 얼굴 인식률 ··71
그림 39. 얼굴 크기 정규화 방법에 따른 LDA 기반 원거리 얼굴 인식률 ···71
그림 40. 얼굴 밝기 정규화 방법에 따른 PCA 기반 원거리 얼굴 인식률 ··73
그림 41. 얼굴 밝기 정규화 방법에 따른 LDA 기반 원거리 얼굴 인식률 ·73
그림 42. 기준 영상 크기에 따른 PCA 기반 원거리 얼굴 인식률 ·····76
그림 43. 기준 영상 크기에 따른 얼굴 인식 알고리즘 수행 시간 ·····76

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

A Study on the Real-time Long Distance Face Recognition using Face Images by Distance

Moon, Hae-Min
Advisor : Prof. Pan, Sung Bum, Ph.D.
Department of Information & Communications,
Graduate School of Chosun University

The recent surveillance system is evolving into an intelligent system where it judges on its own and acts based upon image analysis technology. To satisfy the needs of high level intelligent surveillance system, it shall not only be able to extract objects and classify but also to identify precise information on the object.

The representative method to identify one's identity in the field of surveillance is face recognition. Since recently face recognition has been applied to wide distance of applications including criminal investigation, searching for a missing child, and Human-Robot Interaction(HRI), the needs for long distance face recognition system has been dramatically increased. Since the conventional face recognition system is applied to short distance face recognition such as access control, the need for long distance face recognition has not been considered. In other words, if conventional face recognition system is to be applied to surveillance, it may be hard to expect satisfying quality on long distance face recognition

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

with low image resolution by changes in camera resolution and distance.

Recently researches on long distance face recognition technology based on low resolution face information extracted from low resolution images or changes in distance haven been in progress. As representative methods of the foregoing, there are super resolution image reconstruction method, method using structural characteristics of images, face recognition using Pan Tilt Zoom(PTZ) camera. However, these conventional methods also have problems in which low recognition rate under changes in lighting and distance together with low processing speed for reconstruction of face images with low resolution, and dependability of learned data. In addition, the method using PTZ camera improves the rate of recognition, but does not solve fundamental issues in long distance face recognition such as additional costs. Thus, long distance face recognition technology, where it can be operable under environment with conventional low resolution and high resolution Closed Circuit TeleVision(CCTV), and possess strong real-time processing capability with changes in lighting and distance, is needed without additional costs.

This paper propose a method that face images by actual distances shall be used in the stage of registering user' s face images for the purpose of real-time long distance face recognition. In order to remedy low recognition rate incurred by changes in distance, we would like to propose learned data composition method using face images by distances. This method using face images by distances can be classified into the followings depending upon the manner to acquire face images, for instance, face images by actual distances, face images by virtual distances using zoom camera, and face images by virtual distances using automatic creation

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

method. The proposed method in comparison with the conventional method, which only acquires short distance face information, has merits in acquiring precise characteristics of individual faces in both short and long distance.

The conventional method of long distance face recognition uses reconstruction method turning face images of long distance with low resolution into higher resolution whereas the proposed method is to make long distance face images with low resolution using short distance face images with high resolution through deterioration of image quality. In addition, the problems of nonconformity in face size and brightness of face by changes in distance can be solved by bilinear interpolation and histogram equalization respectively. The proposed method, that is learned data composition method using face images by distances, in this paper has 98.3% face recognition rate on average for entire distance in QVGA resolution achieving 58.7% higher than the conventional method using structural characteristics whereas 93.7% face recognition rate on average for entire distance in VGA resolution achieving 40.6% higher than the conventional method. Finally, the proposed method achieved 36.2% higher than the conventional method in SXGA resolution as well.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제1장 서 론

제1절 연구 배경

정보사회가 급진전됨에 따라 공공기관에서는 행정편의를 도모하고, 국민생활의 안정과 사회질서를 유지하기 위해 영상 감시 시스템의 사용을 늘리고 있다[1][2]. 영상 감시 시스템은 범죄수사의 증거확보 또는 범죄예방활동에 있어서 중요한 역할을 하고 있고, 과속 및 신호위반 등과 같은 교통단속 목적이나 지하철, 공공장소 등에서 안전사고 예방을 위해 설치 및 운용되고 있다. 최근에는 주택가의 쓰레기 불법투기 단속이나 경범죄 단속 등 그 활용 범위가 점차 확대되고 있고, 네트워크 카메라를 통해 놀이방, 유치원, 학원 등에서 아동의 생활모습을 부모가 실시간으로 관찰할 수 있는 서비스도 등장하고 있다.

국내 CCTV의 운영 현황을 분석하면 정부부처와 공공기관, 전국 지자체에서 설치 운영 중인 CCTV만 35만여 대이며, 개인이 설치한 CCTV를 포함한 전체 설치 현황은 274만 대 이상으로 추정하고 있다. 표 1의 지능형 CCTV 시스템의 서계 시장 규모에 따르면 전 세계 지능형 영상 분석기기 시장의 경우 지난 2010년에는 총 2억 1,840만 달러로 집계되었다[2]. 이 중에서 수송 관리 분야를 제외한 보안 분야는 9,210만 달러가 된다. 지금 시장에서 차지하는 규모는 작은 편이지만, IMS Research에서는 지능형 영상 분석 분야의 2015년 전망치를 5억 9,020만 달러로 추정하고 있으며, 이는 연평균 22%의 성장률을 나타내고 있다.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

표 1. 지능형 CCTV 시스템의 시장 규모

[단위 : US$ Millions]

구분	2010	2011	2012	2013	2014	2015	CAGR (10~15)
Security& Safety	82.5	101.4	130.9	167.9	210.4	263.0	26.1%
Business intelligence	9.6	18.1	31.6	45.2	61.8	78.4	52.2%
Traffic Management	126.3	135.0	155.4	180.8	212.1	248.8	14.5%
Total	218.4	254.5	317.9	393.9	484.3	590.2	22.0%

전통적인 영상 감시 시스템은 카메라를 통하여 획득된 영상을 모니터를 이용하여 사람이 직접 감시하거나 DVR 등을 이용해 사후에 저장된 영상을 검색한 후 대응하는 수동적인 시스템이었다. 이러한 전통적인 영상 감시 시스템의 구조는 CCTV 카메라 보급의 증가에 따라 사람 감시를 통한 한계를 나타내고 있다. CCTV와 같은 카메라를 기반으로 하는 기존 영상 감시 시스템은 감시 인력이 모니터를 통해 감시하기 때문에 능동적인 감시 및 효율성이 낮은 문제점을 가진다. 최근 연구 결과에 따르면 한 사람이 두 대의 모니터를 감시할 때, 10분이 지나면 사건의 45%를 인식하지 못하고 22분 후에는 95%를 인식하지 못한다고 한다. 그림 1은 시간 경과에 따른 사람 감시자의 감시 효율의 변화를 나타낸다[3].

최근 감시 시스템은 영상 분석 기술을 기반으로 스스로 판단하고 능동적 대응이 가능한 지능형 시스템으로 진화하고 있다. 지능형 영상 감시 시스템은 영상 카메라의 정보를 디지털화하여 사전에 정의한 패턴과 규칙에 따라 객체의 위치, 형태, 모양 컬러, 움직임, 질감 등 다양한 특징을 분석한 후 그 결과를 바탕으로 사전 예방, 경보 및 차단 시스템을 작동하는 기술을 말한다[4][5]. 지능형 영상 감시의 핵심

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

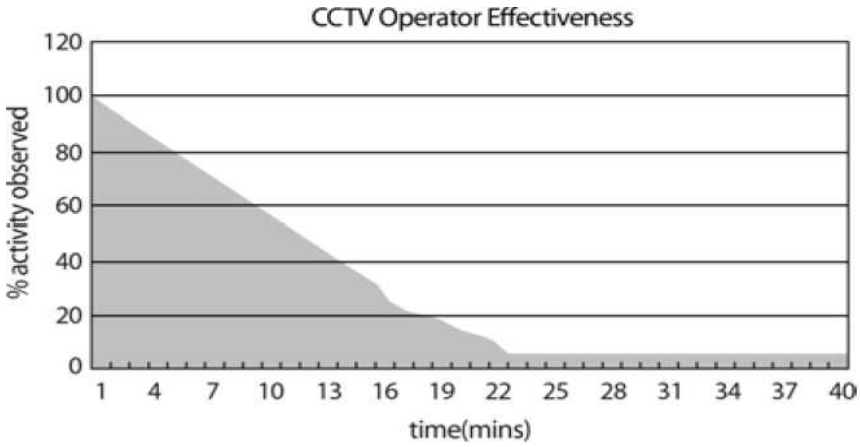


그림 1. 시간 경과에 따른 감시자의 감시 효율 변화

기술로는 영상에서 사람을 검출, 분류하거나 검출된 사람 정보를 기반으로 얼굴 및 사람 등을 인식 및 검색하는 기능 등이 있다[6-9]. 그밖에도 객체를 분석해 비정상 행동 및 위험 상황을 인식하는 기술[10-13]과 영상에 존재하는 사생활 정보를 보호할 수 있는 프라이버시 보호 기술들도 연구되고 있다[14-18].

높은 수준의 지능형 영상 감시 시스템을 만족하기 위해서는 단순히 객체를 검출해서 분류하는 것뿐만 아니라 대상에 대한 정확한 신원 정보까지 확인할 수 있어야 한다. 지능형 영상 감시 시스템에서 대상의 신원을 확인할 수 있는 방법으로는 사람의 얼굴, 홍채, 지문, 정맥, 걸음걸이 등 사람의 바이오 정보를 이용한 방법이 대표적이다. 기존 홍채, 지문, 정맥을 이용한 바이오 인식 기술은 비교적 근거리에서 이루어지면서 인식 시스템에 대한 물리적인 접촉이 필요하기 때문에 거부감이 높은 단점이 있다. 걸음걸이 인식의 경우도 현재 연구 초기 단계로써 실제 감시 시스템 환경에 적용하기에는 한계가 있다. 반면, 얼굴을 이용한 방법은 비접촉 및 원거리 인식이 가능한 장점으로 최근 영상 감시 시스템에서 사용자 인식 방법으로 가장 많이 연구되고 있다[19-23].

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제2절 연구 목적

얼굴 인식 기술의 확산과 더불어 사용자 편의성에 대한 요구가 증가함에 따라 얼굴 인식 기술을 활용한 보안 시스템의 수요가 증가하고 있다. 과거 얼굴 인식 시스템은 근거리 얼굴 인식을 기반으로 출입통제 분야에서 주로 이용되었기 때문에 카메라의 해상도나 외란에 의해 발생할 수 있는 인식률 저하의 문제가 발생하지 않았다. 최근 지능형 영상 감시 시스템은 범죄수사, 미아 찾기, HRI 등 폭 넓은 응용 분야에서 이용되고 있기 때문에 근거리뿐만 아니라 원거리 얼굴 인식에 대한 상황도 고려되어야 한다. 특히, 하드웨어 기술의 발달에 의해 CCTV를 통해 획득된 영상은 낮게는 30만 화소부터 높게는 500만 화소까지 다양한 해상도의 영상을 제공하고 있다. 즉, 동일한 얼굴 인식 방법이라도 영상의 해상도에 따라 처리 속도 및 인식 성능이 달라질 수 있다. 기존 얼굴 인식 기술을 지능형 영상 감시 시스템에 적용 후 얼굴 인식을 수행할 경우 카메라 해상도 및 거리 변화에 대해 만족할 만한 성능을 기대할 수 없다.

최근 거리 변화 혹은 저해상도 영상에서 추출된 저해상도 얼굴 정보를 이용한 원거리 얼굴 인식 기술에 대한 연구가 진행되고 있다. 대표적인 방법으로 초해상도 복원 방법을 이용한 방법, 영상간 구조적 특징을 이용한 방법, PTZ 카메라를 이용한 얼굴 인식 방법 등이 있다. 초해상도 복원을 이용한 원거리 얼굴 인식 방법은 원거리에서 획득되는 저해상도 영상을 고해상도로 복원하여 얼굴 인식에 이용하는 방법이다[24-29].

초해상도 복원 알고리즘은 영상 복원을 통해 저해상도 영상에 포함된 잡음이나 흐릿함을 제거한다. 또한, 영상의 크기가 확대됨에 따라 기존에 존재하지 않던 영상

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

정보를 효과적으로 추정하여 보다 선명한 고해상도 영상을 복원해 원거리 얼굴 인식을 하는 방법이다. 초해상도 복원 방법은 시각적인 측면에서 우수한 고해상도 복원 성능을 보이지만, 얼굴 인식 측면에서는 만족할 만한 성능을 나타내지 못하고 있다. 또한, 학습 데이터에 의존적이기 때문에 학습 데이터의 수가 매우 많이 필요하고 그에 따라 처리 시간이 증가하는 단점이 있다.

일반적으로 저해상도 얼굴 영상은 고해상도 얼굴 영상에 비해 데이터의 양이 적기 때문에 얼굴 인식에 주요한 특징이 되는 질감, 모양, 색상과 같은 정보를 충분히 추출하기 어렵다. 동일한 얼굴로부터 생성된 저해상도 영상과 고해상도 영상을 쌍으로 매핑시켜 학습하는 방법이 연구되고 있다[30-33]. 구조적 특징을 이용한 방법은 사전에 저해상도 영상과 고해상도의 영상을 쌍으로 매핑시키고 추후 입력되는 저해상도 영상의 특징이 가장 유사한 공간을 찾아 그에 상응하는 고해상도 영상을 출력하는 방법이다. 이 방법은 초해상도 복원 방법에 비해 계산 복잡도가 적은 장점이 있지만 정의되지 않은 조명 변화, 거리 변화, 표정 변화 등 외부 변수 및 매핑 관계 가중치 값에 따라 인식률의 변화가 심하게 발생한다.

얼굴 인식률에 큰 영향을 미치는 요인으로는 획득된 얼굴 영상의 해상도, 선명도, 포즈, 조명 등을 들 수 있다. 특히, 원거리 얼굴 인식의 경우 해상도에 따른 선명도에 의해 얼굴 인식률이 결정된다. PTZ 카메라를 이용한 원거리 얼굴 인식은 화질 열화에 대한 근본적인 문제를 해결할 수 있기 때문에 원거리 얼굴 인식에서 매우 높은 인식률을 자랑한다[34-37]. PTZ 카메라를 이용한 방법은 카메라 자체의 가격이 매우 비싸기 때문에 대중적으로 사용되기에는 한계가 있다.

위와 같이 다양한 원거리 얼굴 인식 방법이 연구되고 있지만 저해상도 얼굴 영상의 복원 과정의 처리 속도와 학습 데이터 의존도 및 실제 원거리 얼굴 인식 상황에서 발생할 수 있는 조명 및 거리 변화에 의해 인식률이 저하되는 문제점을 가지고 있다. 또한, PTZ 카메라를 이용한 방법은 인식률은 향상되지만 추가적인 비용의

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

문제로 원거리 얼굴 인식의 근본적인 문제는 해결하지 못하고 있다. 즉, 추가적인 장비 없이 기존 저해상도 및 고해상도 CCTV 환경에서 동작 가능하고 조명 및 거리 변화에 강인한 실시간 처리가 가능한 원거리 얼굴 인식 기술이 필요하다.

제3절 연구 내용 및 방법

얼굴 인식 시스템은 크게 얼굴 영역 추출과 얼굴 인식 부분으로 나눌 수 있다. 얼굴 영역 추출은 영상에서 얼굴의 위치, 크기 등을 고려하여 얼굴 인식을 위한 얼굴 영상을 만들어 내는 과정이다[38-39]. 정확한 얼굴 영역의 추출은 얼굴 인식 단계에서 인식률을 높이는 매우 중요한 단계로 초기 얼굴 인식 연구에서 단순히 전처리 단계로 취급하던 얼굴 검출 단계가 최근에는 하나의 독립적인 연구 분야로 다루어지고 있다. 얼굴 인식 단계는 크게 얼굴의 기하학적 특징을 이용한 특징 기반 방법(Feature-based methods)과 얼굴의 전영역 특징을 이용한 전영역 방법(Holistic methods)으로 분류될 수 있다[40]. 기하학적 특징을 이용한 방법은 얼굴 영상으로부터 눈, 코, 입 등의 얼굴 요소들을 추출하여 이들 요소의 특징들 간의 지형적인 특성의 차이를 이용해 인식하는 방법이다. 전영역 특징을 이용한 방법은 얼굴 영상의 전역적인 특징을 고려하는 방법으로 학습 영상 집합에 의해 생성된 모델을 이용해서 얼굴 인식을 하는 방법이다. 본 논문에서는 기존 CCTV 환경에 적용가능한 원거리 얼굴 인식을 위해 전영역 특징을 이용한 PCA(Principal Component Analysis)와 LDA(Linear Discriminant Analysis) 기반 얼굴 인식 방법을 사용한다.

전영역 특징을 이용한 얼굴 인식은 처리 시간이 빠르고 비교적 구현이 쉽기 때문에 얼굴 인식 시스템에서 가장 많이 이용되고 있지만 원거리 얼굴 인식 상황에 적

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

용할 경우 다음과 같은 문제점이 발생된다. 첫째, 원거리 얼굴 인식의 경우 거리별로 추출되는 얼굴 영상의 크기가 모두 다르기 때문에 학습과 검증 영상간 얼굴 특징 차원의 부적합 문제가 발생된다. 둘째, 얼굴 영상 하나의 데이터는 크지만 클래스당 훈련 영상의 개수가 매우 한정되어 있을 때 소표본 문제가 발생된다. 셋째, 원거리 얼굴 인식의 경우 카메라와 대상의 거리가 매우 다양하기 때문에 조명의 위치가 고정되어 있더라도 자연스럽게 조명의 세기와 방향에 대한 얼굴의 밝기 차이가 발생한다. 넷째, 고해상도 카메라를 통한 얼굴 인식은 저해상도 영상보다 원본 영상에서 추출되는 얼굴의 크기가 매우 크다. 즉, 카메라의 해상도가 높아질수록 전처리 및 얼굴 인식 과정 등에서 처리 시간이 길어지는 문제가 발생한다.

본 논문에서는 위에서 언급한 네 가지 문제점을 해결하기 위해 다음과 같은 방법을 이용한다. 첫째, 학습과 검증 영상간의 얼굴 특징 차원의 부적합 문제를 해결하기 위해 실제 CCTV 환경에서 추출되는 거리별 평균 얼굴 영상 크기를 분석한 후 다양한 보간법을 통해 얼굴 크기 정규화를 수행한다. 둘째, 소표본 문제를 해결하기 위해 실제 거리별 얼굴 영상을 학습으로 사용하는 방법, 줌 카메라를 통해 가상의 거리별 얼굴 영상을 획득하는 방법, 거리별 얼굴 영상을 자동으로 생성하는 방법들을 사용한다. 제안하는 방법을 통해 해상도별 원거리 얼굴 인식 상황에서 최소한의 학습 영상으로 최대의 인식률을 얻을 수 있는 조건을 분석한다. 셋째, 조명에 의한 얼굴의 밝기 차이 문제를 해결하기 위해 히스토그램 평활화(Histogram Equalization), MCT(Modified Census Transform) 변환과 LBP(Local Binary Pattern) 등을 통해 최적의 얼굴 밝기 정규화 방법을 분석한다. 넷째, 고해상도 영상 기반 얼굴 인식에서 처리 속도 개선을 위해 학습 영상의 기준 영상의 크기를 변화시키는 방법을 사용한다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어 2장에서는 영상 감시 시스템 환경과 원거리 얼굴 인식 기술에 대해 소개한다. 3장에서는 본 논문에서 제안하는

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

거리별 얼굴 영상을 이용한 원거리 얼굴 인식 방법에 대해 세부적으로 설명한다. 4장에서는 실제 CCTV를 통해 획득된 다양한 해상도에 따른 원거리 얼굴 인식 실험과 결과를 분석하고 5장에서 논문의 결론과 향후 연구 방향을 제시한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제2장 영상 감시 시스템에서의 얼굴 인식

바이오 인식은 살아있는 사람의 신원을 생리학적 또는 행동 특징을 기반으로 인증하거나 인식하는 시스템으로 대표적인 기술로는 얼굴 인식 방법이 있다. 영상 감시 시스템에서 얼굴 인식은 카메라를 통해 획득되는 영상의 품질에 따라 얼굴 인식의 성능이 달라지기 때문에 카메라의 성능에 매우 의존적인 특징이 있다. 즉, 저해상도 기반 CCTV를 통해 획득된 영상에서 추출한 얼굴 정보는 화질 열화로 인해 고해상도 영상 기반 얼굴 인식보다 인식률이 감소하는 문제가 발생한다. 특히, 이러한 문제는 저해상도 영상이나 CCTV로부터 대상이 멀리 존재하는 원거리 상황에서 심하게 발생된다. 최근 CCTV 기반 얼굴 인식 기술의 한계를 극복하기 위해 초해상도 복원을 이용한 얼굴 인식, 구조적 특징을 이용한 얼굴 인식, PTZ 카메라를 이용한 얼굴 인식 방법 등이 연구되고 있다.

제1절 영상 감시 시스템

영상 감시 시스템은 카메라에서 촬영된 영상 정보를 수신자에게 전송하고 전송된 영상 정보를 기반으로 사용자가 원하는 지역을 관찰하는 방식을 말한다. 이러한 영상 감시 시스템은 보안용, 산업용, 차량감시용 및 교통관리용 등의 다양하고 광범위한 용도로 사용되고 있다.

그림 2는 CCTV 기반 영상 감시 시스템을 나타낸 것으로 특정 공간에 설치된 CCTV 카메라를 통하여 획득된 영상 정보를 폐쇄적인 유선 전송로를 이용하여 관제 센터로 전송하고 관리자가 모니터를 통하여 직접 감시하던 구조였다. 최근에는

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

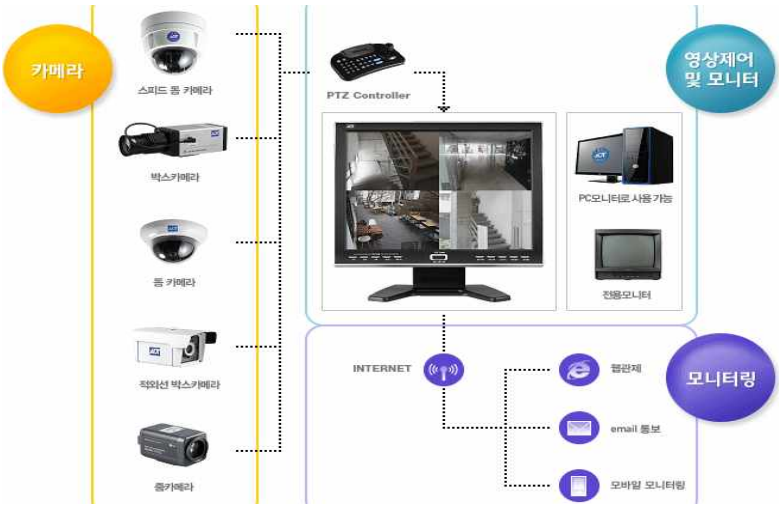


그림 2. CCTV 기반 영상 감시 시스템

IT 기술의 진보로 VCR(Video Cassette Recorder) 저장 장치 기반의 아날로그 시대에서 DVR 기반의 영상 압축 및 디지털 전송기술을 사용하는 디지털 시대로 발전하였으며, 광대역 통신망 및 개방 프로토콜을 사용하는 IP 기반의 네트워크와 영상 자동분석 및 인식 기술이 결합된 지능형 영상 보안 기술로 진화하고 있다. 그림 3은 영상 감시 시스템이 진화하는 형태를 나타내고 있다. 지능형 영상 감시 시스템의 기본적인 목표는 감시 카메라를 통하여 획득된 영상 정보를 실시간으로 분석하여 자동으로 목표물 또는 이동물체를 탐지, 추적, 식별, 행위분석 및 검색하여 관찰된 객체의 행위나 상호작용을 해석하는 것이다.

CCTV의 기본 구조는 피사체를 촬영하여 전기신호로 변환하는 촬영계, 변화된 전기신호를 원격체로 전송하는 전송계, 그리고 전송된 얼굴신호를 재생하여 표시하는 수상계 등 크게 3부분으로 구성된다. 영상 감시 시스템은 이들 3가지 요소 중 하나라도 빠지면 시스템이 구성되지 않으며, 시스템의 크기가 커질수록 이들 각 영역들을 통괄 제지하기 위한 제어계가 포함 된다. 이외에도 상호간에 얼굴의 기록과 재생이 요구되며 정보처리장치에 의한 얼굴의 해석 및 처리를 요구하는 경우도 생

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

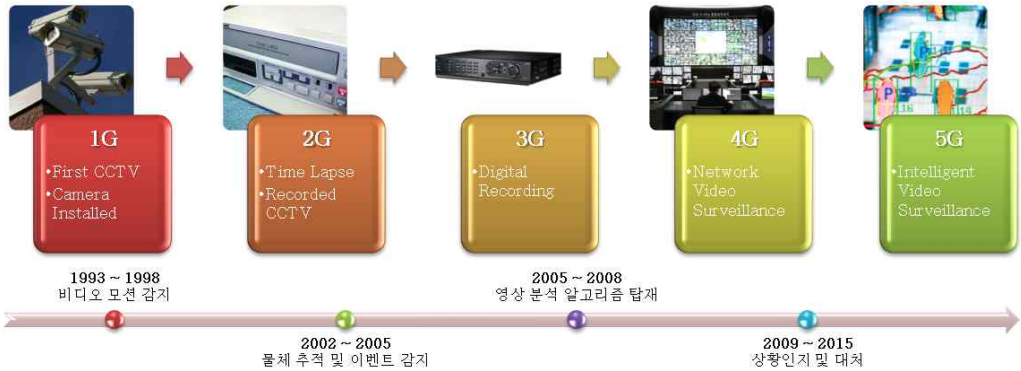


그림 3. 영상 감시 시스템의 진화

긴다. 이로 인하여 얼굴의 기록, 재생, 가공, 처리를 포함하는 얼굴 처리계를 결합하여 CCTV 시스템이 구성된다[41]. CCTV 카메라는 사용 용도와 형태에 따라 크게 돔, 박스, IR(Infrared Ray), PTZ, IP(Internet Protocol) 카메라 등으로 구분된다.

영상 감시 시스템에서 적절한 카메라를 선택하는 것만큼 중요한 것이 카메라 성능의 고려이다. 즉, 카메라를 통해 획득되는 영상의 품질에 따라 감시 시스템의 용도가 결정된다고 볼 수 있다. 고해상도 및 고품질의 영상은 많은 정보를 포함하고 있기 때문에 영상처리를 기반으로 하는 지능형 감시 시스템에서는 매우 중요한 요소이다. 그림 4는 카메라를 통해 획득되는 영상의 해상도를 나타내고, 그림 5는 HD급 영상과 SD급 영상을 비교한 결과이다. 그림에서와 같이 동일한 장소에서 촬영된 영상이라도 HD급 영상의 경우 추가적인 영상처리 없이 직관적으로 차량의 번호판을 확인할 수 있지만 일반 화질의 경우 사람이 직접 번호판의 내용을 판단하기는 어렵다. 즉, 이미지 해상도가 높을수록 더 선명하고 자세한 영상을 볼 수 있지만 카메라 성능에 따라 가격도 비싸지고 영상 데이터 크기가 커져서 처리 시간이 추가로 요구되는 단점이 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

포맷	해상도	픽셀 수
QVGA (Quarter VGA)	320x240	0.077Mega
VGA (Video Graphics Array)	640x480	0.3Mega
SVGA (Super VGA)	800x600	0.5Mega
XGA (eXtended Graphics Array)	1024x768	0.8Mega
SXGA (Super eXtended Graphics Array) - An aspect ratio of 5:4 instead of 4:3	1280x1024	1.3Mega
UXGA (Ultra eXtended Graphics Array) - 4times resolution of SVGA (800X600)	1600x1200	1.9Mega
QXGA (Quad eXtended Graphics Array) - 4times resolution of XGA	2048x1536	3.1Mega

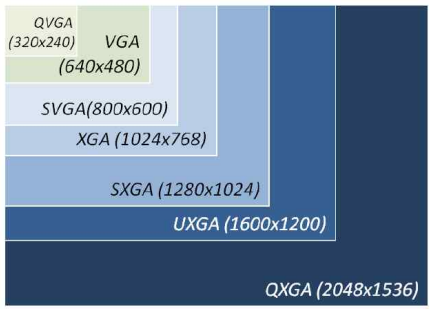


그림 4. 영상의 해상도 종류

그림 5. HD급 CCTV와 SD급 CCTV의 화질 비교

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

제2절 대표적 얼굴 인식 기술

본 절에서는 대표적인 얼굴 인식 방법들에 대해 소개하고, 제안하는 원거리 얼굴 인식 방법의 바탕이 되는 전영역 방법을 이용한 얼굴 인식 방법에 대해 상세히 설명한다. 얼굴 인식 방법은 얼굴을 처리하는 방법에 따라서 전영역 방법, 특징 기반 방법, 혼합 방법(Hybrid methods) 등으로 나눌 수 있다[42][43]. 전영역 방법은 얼굴 영상 전체를 하나의 특징벡터로 사용해 인식 알고리즘을 적용하는 방식이다. 대표적인 전영역 방법에는 PCA[44], LDA[45], ICA(Independent Component Analysis)[46]등이 있다. 전영역 방법은 높은 차원에서의 정보를 유지하면서 차원을 축소한 후 축소된 데이터를 이용해 얼굴 인식 알고리즘을 수행한다. 특징 기반 방법은 국소적인 데이터를 입력 데이터로 사용하는 방법으로 얼굴에서 눈, 코, 입 등과 같은 영역을 입력으로 사용해 인식 알고리즘을 적용하는 방식이다. 특징 기반 방법에는 PG(Pure Geometry) 방법[47], EBGM(Elastic Bunch Graph Matching) 방법[48], HMM(Hidden Markov Model) 방법[49]등이 있다. 특징 기반 방법은 원하는 국소적인 부위를 정확히 추출할 수 있다면, 전영역 방법에 비해 높은 인식률을 얻을 수 있는 장점이 있다. 반대로 정확한 국소부위 추출이 어렵다면 현저히 낮은 인식률을 나타낸다. 마지막으로, 혼합 방법은 인간의 눈이 얼굴의 전체적인 부분과 부분적인 부분을 모두 보면서 얼굴을 인식한다는 점을 이용하여 전영역 방법과 특징 기반 방법을 함께 사용하는 방식이다. 대표적인 혼합 방법에는 Modular Eigenspaces[50], Hybrid Local Feature[51], Shape-normalized[52], Component-based 방법[53] 등이 있다.

본 논문에서는 전영역 방법을 이용한 얼굴 인식 방법을 기반으로 구성되고, 얼굴 인식 시스템은 다음과 같은 순서로 진행된다. 첫째, 원본 영상이 입력되면 영상에서

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

얼굴을 검출한다. 둘째, 검출된 얼굴의 밝기 및 크기 등을 정규화 해주는 전처리 과정을 수행한다. 셋째, 정규화된 얼굴에서 특징을 추출하여 템플릿을 발생시킨다. 마지막으로, 기존의 DB에 존재하는 템플릿과 현재 템플릿을 비교하여 누구인지를 판단한다. 다음은 전영역 방법을 이용한 얼굴 인식에서 중요한 역할을 하는 얼굴 밝기 정규화, 얼굴 크기 정규화 및 특징 추출방법에 대해 설명한다.

1. 얼굴 밝기 정규화

얼굴 인식에서 대표적인 얼굴 밝기 정규화 방법에는 히스토그램 평활화, MCT 변환과 LBP 변환 방법이 있다. 히스토그램은 영상의 중요한 정보인 명암 값의 분포 상태를 파악할 수 있는 방법이다[54-57]. 히스토그램의 구성은 0부터 255까지의 명암 값을 인덱스로 하고 영상을 구성하는 각 화소의 명암 값에 따라 해당 인덱스 테이블에 빈도수를 누적하는 방법이다. 히스토그램 평활화는 영상의 히스토그램을 그레이 스케일 전 구간에서 골고루 나타나도록 변경하는 방법이다. 단순히 영상의 히스토그램을 넓게 펼쳐주는 방식이 아니라, 원본 히스토그램의 특성에 근거하여 히스토그램의 분포를 변경시키는 방법이다. 다음 식 (1)과 같이 영상의 화소 값을 변환하는 함수를 정의 했을 때, r은 입력 그레이 스케일이고, s는 변환 함수 T에 의한 출력 그레이 스케일 값이다. 일반적으로 변환 함수 T는 단조 증가 함수로 가정하며, 히스토그램 평활화도 단조 증가 변환 함수의 형태로 나타낼 수 있다.

$$s = T(r) \tag{1}$$

위의 내용을 바탕으로, 입력 영상의 히스토그램과 출력 영상의 히스토그램을 확률 밀도 함수(Probability Density Function)의 형태로 각각 $P_r(r)$과 $P_s(s)$로 나

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

타낼 수 있다[58][59]. 히스토그램 평활화의 변환 함수는 $P_r(r)$을 누적시킨 값에 대응하는 형태로 정의되며, 다음과 같은 식 (2)로 표현할 수 있다. 여기서, τ는 적분을 위한 임시 변수이다. 확률 밀도 함수를 적분을 사용하여 더해준 형태의 함수를 누적 분포 함수(Cumulative Distribution Function)라고 부른다.

$$s = T(r) = \int_0^r P_r(\tau)d\tau \tag{2}$$

영상에서 얼굴의 구조적 특징 정보만을 추출한 후 이를 이용함으로써 얼굴 인식 또는 검출, 손동작 인식 등에 이용되는 대표적인 방법으로 MCT 변환이 있다[60-65]. MCT 변환은 영상 내의 구조적인 특징을 추출하는 기법으로 영상 내의 밝기 변화나 조명에 의한 영향을 최소화하면서 영상 내의 정보를 추출하는 영상 처리 기법이다. 변환이라는 용어를 쓰기 위해서는 원 영상에 대한 MCT 변환 후 다시 원본 영상이 복원이 가능해야 하는데 이 기법의 경우 원 영상 복원이 가능하지 않다. MCT 변환을 위해서는 보통 3×3 크기의 윈도우를 이용하는데 이 크기의 윈도우 내에서는 조명의 변화가 거의 없는 것으로 가정할 수 있고, 이에 의해서 조명의 변화에 무관한 영상의 구조적인 특징을 추출할 수 있다. MCT 변환은 다음 식 (3)으로 정의할 수 있다.

$$\Gamma(X) = Y \overset{\otimes}{\in} N' \zeta(\bar{I}(X), I(Y)) \tag{3}$$

여기서, X는 영상 내의 화소의 위치를 나타내고 X를 중심으로 3×3 윈도우를 $W(X)$라 한다면, N'는 $W(X)$내의 화소들의 집합이고, Y는 이 윈도우 내의 화소들

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

의 각각의 밝기 값이다. $\zeta()$는 비교 함수로써 만약 $\bar{I}(X) < I(Y)$일 경우 $\zeta(\bar{I}(X), I(Y))$는 1이 되고 그 이외의 경우 0이 된다. $\otimes$는 접합 연산자로 $\zeta()$ 함수의 이진 형식들을 연결해 주는 역할을 하며 $\otimes$ 연산을 통해 이진 형식이 연결되게 된다. 이론적으로 9개의 화소 값이 모두 1이 되는 경우는 발생하지 않기 때문에 총 511개의 구조가 표현될 수 있다. 이렇게 연결된 이진 형식은 10진수로 변환된 후 MCT 변환 영상의 화소 값이 된다.

LBP 변환은 영상 데이터에서 텍스쳐의 높은 식별력, 조명의 변화에 대한 내구성, 간단한 연산 등의 장점으로 영상의 복구 및 분류, 바이오 영상 분석, 얼굴 영상 인식 등 다양한 분야에 이용되고 있다[66-70]. LBP는 이웃하는 8개의 값을 측정하는 방식으로 얼굴 이미지를 작은 블록으로 나누어 LBP 히스토그램으로 만들어 히스토그램 간이 거리 차이를 이용하여 얼굴 이미지를 구별하는 방식이다. 얼굴 영상은 미세 질감의 구성으로 볼 수 있기 때문에 LBP를 사용하면 그레이 스케일의 얼굴 영상을 질감의 형태로 표현이 가능하고, 화소 값의 오프셋을 제거하기 때문에 영상의 조명 변화와 회전 및 스케일링 등에 강인한 성능을 나타낸다. LBP를 구하는 방법은 다음 식 (4)와 같다. 중심 값은 (x_c, y_c)으로 표시되며, (x_c, y_c)주변의 화소간의 농도를 비교하여 이진 집합으로 추출한 후 8비트의 결과 값을 10진수로 변환한다. 여기서, l_n은 8개의 주변 이웃 화소의 그레이 스케일의 명도 값을 나타내고 l_c는 중심 화소의 값을 나타낸다. 함수 $s(x)$는 식 (5)와 같이 정의된다.

$$LBP(x_c, y_c) = \sum_{n=0}^{p-1} s(l_n - l_c)2^p \tag{4}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{5}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

(a) 원본 얼굴 영상

(b) 히스토그램 평활화된 얼굴 영상

(c) MCT 변환된 얼굴 영상

(d) LBP 변환된 얼굴 영상

그림 6. 얼굴 밝기 정규화 방법에 따른 얼굴 영상

그림 6은 정규화 방법에 따라 얼굴 밝기가 정규화된 결과 영상을 나타낸 것으로 그림 6(a)는 동일한 환경에서 조명의 변화에 따른 원본 얼굴 영상이고, 그림 6(b)는 히스토그램 평활화를 통해 얼굴의 밝기가 정규화된 얼굴 영상이다. 그림 6(c)와 그림 6(d)는 각각 MCT 변환과 LBP 변환 방법을 통해 얼굴 밝기가 정규화된 얼굴 영상으로 얼굴의 구조적인 특징을 중심으로 변환된 결과 영상이다. MCT와 LBP의

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

경우 히스토그램 평활화에 비해 정규화된 얼굴 영상의 밝기 변화가 적음을 확인할 수 있다.

2. 얼굴 크기 정규화 방법

PCA 및 LDA와 같이 얼굴의 전영역 특징을 이용한 얼굴 인식은 학습 영상 샘플의 차원이 검증 영상 샘플의 차원보다 크거나 작다면 클래스내 분산 행렬의 특이점에 의해 소표본 문제가 발생할 수 있다. 즉, 다양한 거리에서 추출된 얼굴 영상은 거리에 따라 추출되는 얼굴 영상의 크기가 다르게 추출되기 때문에 소표본 문제에 따른 영상의 차원을 정규화 해주어야 한다. 영상 크기를 정규화하는 대표적인 방법으로는 보간법이 있다. 영상 보간법은 디지털 영상 처리에서 영상을 확대하거나 축소하기 위해 사용된다. 영상의 확대는 크게 영상 확장과 영상 보간의 단계로 나누어 설명할 수 있다. 영상의 확장이란 원 영상의 화소를 확대하고자 하는 영상의 크기의 간격에 맞게 일정 간격으로 재배치하는 것을 말한다. 영상 보간은 확대된 영상의 비어있는 공간에 적절한 화소 값으로 채우는 과정을 말한다. 일반적으로 사용되는 영상 보간법으로는 최근접 이웃 보간법, 양선형 보간법, 양3차회선 보간법 등이 있다[71-73].

최근접 이웃 보간법은 값을 할당받지 못한 목적 영상의 화소에서 가장 가까운 원시 화소 값을 할당받은 목적 영상의 화소 값을 복사해서 사용하는 방법이다. 이 방법은 단순히 이웃 화소를 복사하여 사용하므로 처리 속도가 빠르다는 장점이 있다. 그러나 새로운 화소 값을 계산하지 않고 입력 화소 내에서만 찾기 때문에 결과 영상에서 대부분 계단 현상 또는 블록 현상이 나타나는 단점이 있다.

양선형 보간법은 가장 일반적인 보간 알고리즘으로 새롭게 생성될 화소에 대해 맵핑되는 원본 영상의 인접한 4개의 화소와 각각의 화소에 대한 가중치를 곱한 값

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

의 합을 이용하는 방법이다. 이때, 가중치들은 인접한 참조 화소로부터의 거리에 반비례하고, 확대되는 영상의 배율에 따라 선형적으로 결정된다. 가중치로 사용되는 값은 0부터 1사이의 실수 값이다. 양선형 보간법은 계산이 비교적 간단하지만 확대된 영상의 에지부분이 매끄럽지 못하는 단점이 있다.

고차 다항식을 이용한 보간법은 가중치 함수를 정의하고, 원본 영상의 주변 화소 값에 가중치를 곱한 값을 모두 합하여 화소 값을 계산하는 방식이다. 고차 다항식을 이용한 대표적인 방법인 양3차회선 보간법은 원본 영상의 16개 화소를 이용하여 새로운 보간 화소를 생성하는 방법이다. 양3차회선 보간법은 4개의 화소를 참조하는 양선형 보간법보다 많은 16개의 화소를 참조하므로 화질이 우수하다. 그러나 더 많은 화소를 사용하기 때문에 각 화소에 대한 가중치를 계산하는데 더 많은 연산량이 요구된다.

기존의 양선형 보간법은 연산량이 적고 연산식이 간단해 구현이 쉽지만 보간된 영상의 화질이 우수하지 못하다. 양3차회선 보간법은 양선형 보간법에 결과 영상의

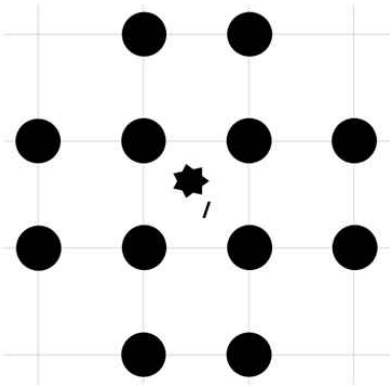


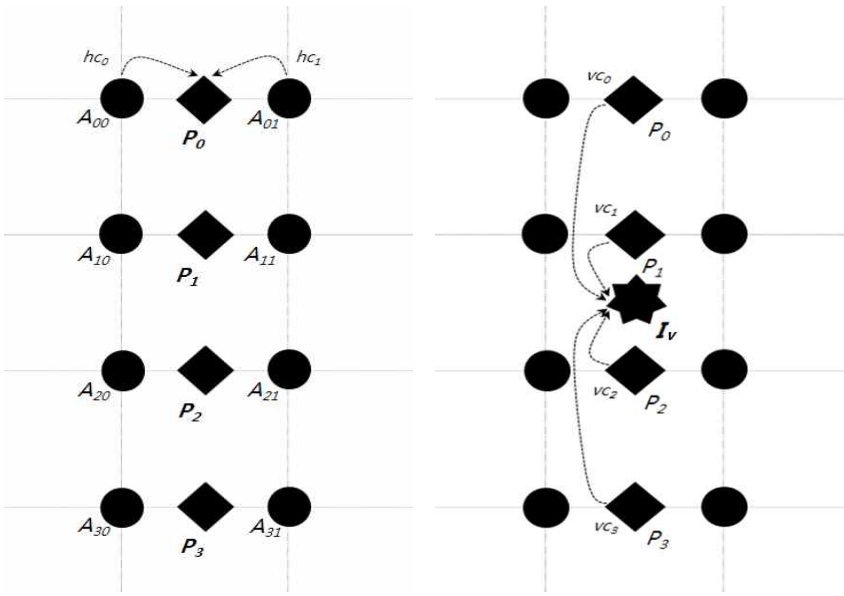
그림 7. 이중 선형-3차회선 보간법의 참조화소

화질은 우수하지만 연산량이 많은 단점이 있다. 본 논문에서는 기존 보간법의 특징을 고려해 선형 보간법과 3차회선 보간법을 이용한 이중 선형-3차회선 보간법을

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제안한다[76]. 그림 7은 제안하는 보간법의 보간 화소 I를 계산하기 위해 참조되는 12개의 화소를 나타내고 제안하는 방법은 수직 선형-3차회선 보간법과 수평 선형-3차회선 보간법을 이용한다. 그림 8은 수직 선형-3차회선 보간법으로 4개의 1차원 선형 보간법과 1개의 1차원 3차회선 보간법으로 분해된 2차원 보간법이다. 그림에서와 같이 하나의 보간된 화소 I_v를 계산하기 위해서 2×4의 이웃하는 참조 화소를 이용한다. 수직 선형-3차회선 보간법은 4번의 선형 보간법을 수행한다. 식(6)을 통해 원본 화소와 각각의 가중치로부터 4개의 보간 화소(P_0, P_1, P_2, and P_3)를 생성한다. 여기서, P_i는 선형 보간법에 의해 보간된 화소이고, A_{ij}는 2×4의 이웃하는 참조 화소이다. hc_j는 화소 A_{ij}의 가중치로 0부터 1사이의 실수 값이다.



(a) 수평 선형 보간법 (b) 수직 3차회선 보간법

그림 8. 수직 선형-3차회선 보간법의 동작

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

$$P_i = \sum_{j=0}^{1} A_{ij} \times hc_j \qquad i = 0,1,2,3 \tag{6}$$

수직 선형-3차회선 보간법의 최종 보간 화소 I_v는 1차원 3차회선 보간법에 의해 생성된다. 식 (7)에서와 같이 I_v는 선형 보간법에 의해 생성된 화소(P_0, P_1, P_2, and P_3)와 가중치(vc_0, vc_1, vc_2 and vc_3)에 의해 계산된다. 여기서, vc_i는 보간화소 P_i의 가중치이다.

$$I_v = \sum_{i=0}^{3} P_i \times vc_j \tag{7}$$

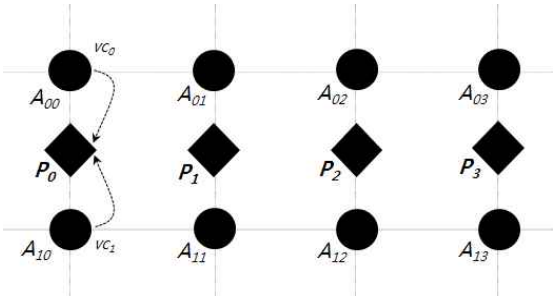
그림 9는 수평 선형-3차회선 보간법으로 4개의 1차원 선형 보간법과 1개의 1차원 3차회선 보간법으로 분해된 2차원 보간법이다. 그림에서와 같이 하나의 보간된 화소 I_h를 계산하기 위해서 4×2의 이웃하는 참조 화소를 이용한다. 수직 선형-3차회선 보간법은 4번의 선형 보간법을 수행한다. 식 (8)을 통해 원본 화소와 각각의 가중치로부터 4개의 보간 화소(P_0, P_1, P_2, and P_3)을 생성한다. 여기서, P_i 는 선형 보간법에 의해 보간된 화소이고, A_{ij}는 4×2의 이웃하는 참조 화소이다. vc_j는 화소 A_{ij} 의 가중치이다.

$$P_i = \sum_{j=0}^{1} A_{ij} \times vc_j \qquad i = 0,1,2,3 \tag{8}$$

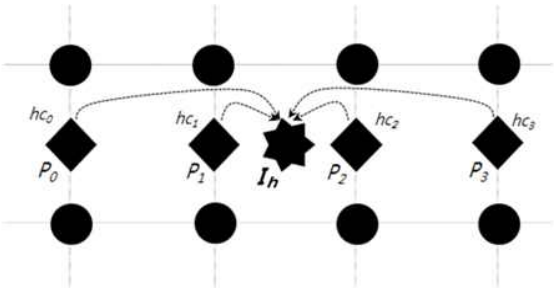
수직 선형-3차회선 보간법의 최종 보간 화소 I_h는 1차원 3차회선 보간법에 의해 생성된다. 식 (9)에서와 같이 I_h는 선형 보간법에 의해 생성된 화소(P_0, P_1, P_2,

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY



(a) 수직 선형 보간법



(b) 수평 3차회선 보간법

그림 9. 수평 선형-3차회선 보간법의 동작

and P_3)와 가중치(hc_0, hc_1, hc_2 *and* hc_3)에 의해 계산된다. 여기서, hc_i는 보간 화소 P_i의 가중치로써 가중치 함수를 이용해 계산된다[73]. 이중 선형-3차회선 보간법의 최종 화소 I는 다음 식 (10)에 의해 결정된다.

$$I_h = \sum_{i=0}^{3} P_i \times hc_j \tag{9}$$

$$I = \frac{I_v + I_h}{2} \tag{10}$$

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

3. 특징 추출 방법

얼굴 인식에 사용되는 입력 얼굴 영상은 고차원의 데이터로 구성되어 있으므로 직접 해석은 많은 시간적, 공간적 복잡도를 요구한다. 이를 해결하기 위해 고차원의 데이터를 효과적으로 저차원의 데이터로 변환하는 방법이 연구되었다. PCA는 대표적인 차원 축소 방법으로 데이터들 사이에 가능한 최대 분산을 갖도록 하는 상호 직교 벡터들의 선형 결합으로 데이터를 표현하는 방법이다[44][75]. PCA를 이용해 데이터를 축소하는 방법은 다음과 같다.

N(height×width)크기의 M개의 학습 영상을 $x = [x_1, x_2, x_3, \ldots x_m]$라 하고 평균 영상을 μ라고 하면 차이 영상의 공분산 행렬은 다음 식 (11)과 같이 정의할 수 있다.

$$\sum = \frac{1}{M}\sum_{m=1}^{M}[x_m - \mu][x_m - \mu]^T \tag{11}$$

여기서, $[x_m - \mu]$를 X라고 하면, $\sum = XX^T$라 할 수 있고, $\sum$의 고유벡터와 고유값을 다음 식 (12)와 같이 구할 수 있다. 여기서, u_k는 고유벡터이고, λ_k는 고유값을 나타낸다.

$$\sum u_k = \lambda_k u_k \tag{12}$$

X의 크기는 $N \times M$이고, X^T의 크기는 $M \times N$이 되기 때문에 $\sum$의 크기는 $N \times N$이 된다. 만약 학습영상 x_i의 크기가 50×50이라면, N은 2,500이 되고 $\sum$의

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

크기는 6,250,000이 되기 때문에 고유벡터의 크기가 너무 커지게 된다. $\sum$의 크기는 다음 과정을 통해 줄일 수 있다.

$$X^T XV = VA \tag{13}$$

다음 식 (13)에서 V를 고유벡터라 하고, A를 고유값이라고 할 때, 양변에 X를 곱해 다음 식 (14)를 구한다.

$$XX^T XV = XVA \tag{14}$$

여기서, XV는 XX^T의 고유벡터가 되고 A는 고유값이 된다. 따라서 X에 식 (14)의 $X^T X$의 고유벡터를 구해서 곱하면 공분산 행렬 $\sum$의 고유벡터를 구할 수 있고, 고유값은 A로써 변함이 없음을 알 수 있다. $X^T X$의 연산은 $(M \times N)(N \times M)$이므로 크기는 $M \times M$이 된다. M은 학습 얼굴 영상의 개수이므로 영상의 크기인 N보다도 훨씬 작은 값이 된다. 즉, $X^T X$의 고유벡터는 $M \times M$이 되고, X에 고유벡터를 곱하면 X의 크기가 $N \times M$이기 때문에 최종 고유벡터의 결과는 $N \times M$이 된다. 일반적으로 고유벡터는 M개 모두를 사용하지 않고 $M-C$개를 사용하고 PCA의 성능 향상을 위해 다음 식 (15)와 같은 정규화 과정을 사용한다. 여기서, u_k는 k번째 고유벡터 $(1 \le k \le M-C)$이고, N은 영상의 크기, M은 학습 영상의 수, C는 클래스의 수를 나타낸다.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

$$
\begin{aligned}
u_k &= \{v_1, v_2, \cdots, v_n\} \\
\overline{u_k} &= \frac{u_k}{\| u_k \|} \\
\| u_k \| &= \sqrt{v_1^2 + v_2^2 + \cdots + v_N^2} \\
u_{\neq w} &= \{\overline{u_1}, \overline{u_2}, \cdots, \overline{u}_{M-C}\}
\end{aligned} \tag{15}
$$

얼굴 인식 분야에서는 고차원의 데이터를 저차원으로 줄이는 방법으로 PCA와 함께 LDA가 가장 많이 사용되고 있다. LDA는 흔히 FLD(Fisher' s Linear Discriminant)라고도 불린다[45]. PCA가 데이터들 사이에 가능한 최대 분산을 갖도록 하는 상호 직교벡터들을 검출하는데 반해, LDA는 클래스 내의 분산은 작게 하고 클래스 간의 분산은 크게 하는 방향의 벡터들을 검출하는 방법이다. 기본적인 구성은 PCA와 비슷하며 실제 구현에서는 연산량을 줄이기 위해 PCA의 고유 얼굴을 사용한다. LDA를 이용해 데이터를 축소하는 방법은 다음과 같다. i번째 클래스의 학습 영상을 $x_{i=}\left[x_1, x_2, x_3, \cdots, x_m\right]$이라고 하고 평균 영상을 μ_i라 하면 클래스내의 분산을 나타내는 행렬 S_W는 다음 식 (16)와 같이 정의된다.

$$
S_W = \sum_{i=1}^{C} \sum_{x_i \in X_i} U^T (x_k - \mu_i)(x_k - \mu_i)^T U \tag{16}
$$

여기서, U는 PCA를 통해 구해진 고유벡터이다. U의 크기는 $N \times (M-C)$이고 U^T의 크기는 $(M-C) \times N$이 된다. $\sum_{i=1}^{C}(x_k - \mu_i)^T$의 크기는 $C \times N$이고, $\sum_{i=1}^{C}(x_k - \mu_i)$의 크기는 $N \times C$이므로 S_W의 크기는 $(M-C) \times (M-C)$가 된다. 클래스간 분산을 나타내는 행렬 S_B는 다음 식 (17)와 같이 정의된다.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

$$S_B = \sum_{i=1}^{C} N_i U^T (\mu_i - \mu)(\mu_i - \mu)^T U \tag{17}$$

여기서, μ_i는 전체 평균 영상을 나타내고 U의 크기는 $N \times (M-C)$이다. U^T의 크기는 $(M-C) \times N$이 되고, $\sum_{i=1}^{C}(\mu_i - \mu)^T$의 크기는 $C \times N$이 된다. $\sum_{i=1}^{C}(\mu_i - \mu)$의 크기는 $N \times C$이므로 S_B의 크기는 S_W와 동일하게 $(M-C) \times (M-C)$가 된다. S_B, S_W로부터 다음 식 (18)을 유도할 수 있다. 여기서, U는 LDA의 축이 되고 식 (18)로 부터 식 (19)를 유도할 수 있다. U는 $S_W^{-1} S_B$의 고유벡터가 된다.

$$S_B U = \lambda S_W U \tag{18}$$

$$S_W^{-1} S_B U = \lambda_i U \tag{19}$$

제3절 기존 원거리 얼굴 인식 방법

1. 초해상도 복원 방법을 이용한 저해상도 얼굴 인식

본 절에서는 현재 연구되고 있는 원거리 얼굴 인식 방법에 대해 소개한다. 원거리에서 촬영된 얼굴 영상은 근거리에서 촬영된 영상에 비해 포함하는 정보의 수가 매우 적기 때문에 원거리에서 촬영된 얼굴 영상을 얼굴 인식에 이용할 경우 인식률이 떨어지는 문제점이 있다. 이를 해결하기 위해 원거리의 저해상도 얼굴 영상을

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

복원하는 초해상도 영상 복원에 관한 연구가 진행되고 있다. 초해상도 복원 알고리즘은 저해상도 영상에 포함된 잡음이나 흐릿함의 제거뿐만 아니라 기존에 존재하지 않던 영상 정보를 효과적으로 추정하여 선명한 고해상도 영상을 복원하는 것을 목표로 한다. 초해상도 복원 방법은 크게 예제 기반 방법과 여러 장의 영상을 이용한 방법으로 구분된다.

그림 10은 예제 기반 초해상도 복원 방법을 이용해 저해상도 얼굴 영상을 고해상도 영상으로 재구성한 예를 나타낸다[26]. 예제 기반 초해상도 영상복원은 크게 훈련과 합성의 과정으로 구분되며, 훈련 단계에서 다양한 크기와 많은 조각 쌍을 확보한 후 이를 참고하여 고해상도 영상으로 복원시키는 방법이다[24-26]. 훈련 단계에서는 저해상도 영상을 업 샘플링한 후, 고해상도 영상과의 차를 이용하여 조각이 될 수 있는 부분을 DB에 저장한다. 그 후 입력된 저해상도 영상을 업 샘플링한 후 원하는 패치와 비슷한 영상이 있다면 비슷한 저해상도 영상에 대응하는 고해상도 영상을 출력해서 영상을 복원하는 방법이다.

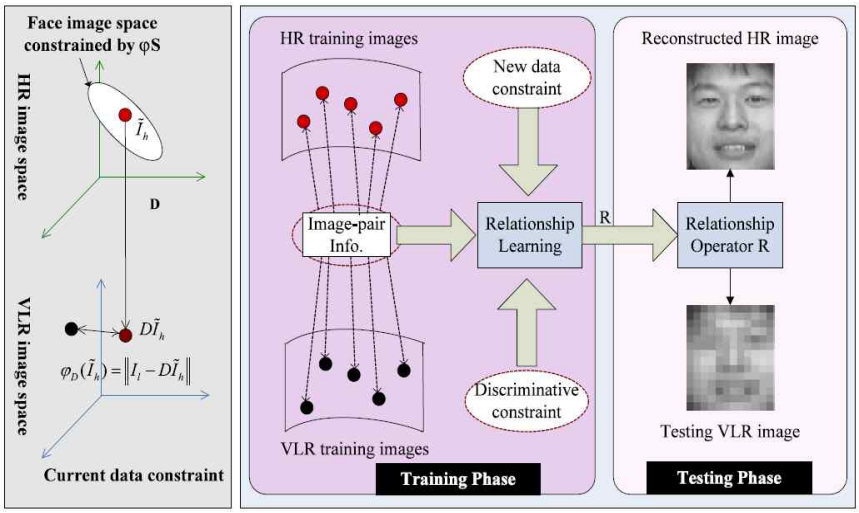


그림 10. 예제 기반 초해상도 복원의 예

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

그 밖에도 가우시안 이론에 근거하여 얼굴을 복원하고 Markov Random Field에 근거하여 국부적으로 얼굴 영상을 복원하는 방법[27], PCA를 이용하여 얼굴 영상을 복원하는 방법[28], ICA를 이용해 얼굴 영상을 복원하는 방법들이 연구되고 있다[29]. 예제 학습 기반의 초해상도 복원 방법은 영상의 대응점이 미리 계산된 학습 예제가 구축되어야하고 미리 구축된 학습 예제에 크게 의존하여 고해상도 영상이 복원되는 단점이 있다.

여러 장의 영상을 이용하여 초해상도 영상 복원을 하는 방법은 공간 영역에서의 초해상도 영상 복원[82], 주파수 영역에서의 영상 복원[83][84], 베이시안 이론을 이용한 초해상도 영상 복원 방법 등이 있다[85]. 여러 장의 영상을 이용한 방법들은 먼저 초해상도 영상을 생성하고 초해상도 영상을 다시 다운 샘플링해 입력 영상과의 유사도를 비교한다. 비교한 유사도를 바탕으로 영상 취득 모델에 대한 매개변수를 갱신하고 그에 대한 초해상도 영상을 반복적으로 복원하는 방식이다. 여러 장의 영상을 이용한 방법들은 초해상도 영상을 복원하는 방식이 다를 뿐 근본적으로는 영상 취득 모델을 이용하는 점에서 유사하다. 즉, 영상간 뒤틀림 정보, 센서의 흐림 정보, 다운샘플링 정보 및 잡음 등을 모델링하여 저해상도 영상으로부터 고해상도 영상을 복원하는 방식이다. 따라서 입력 영상의 개수가 많을수록 성능이 좋아지지만 계산 복잡도 및 처리시간도 증가한다. 초해상도 복원 방법의 경우 학습 단계에서 많은 양의 데이터가 필요하고 얼굴을 구별하는 목적 보다는 시각적으로 화질을 개선하는 목적이 더 크기 때문에 원거리 얼굴 인식에서 만족할 만한 성능을 제공하지 못하고 있다.

2. 구조적 특징을 이용한 저해상도 얼굴 인식

저해상도 기반의 얼굴 인식은 얼굴 인식에 사용되는 특징을 저해상도 얼굴 영상

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

에서 충분히 추출하기 어렵다는 문제점이 있다. 고해상도 얼굴 인식에서 전형적으로 사용되는 특징인 질감, 모양, 색상 같은 경우가 이에 해당한다. 그러므로 저해상도 영상에서 존재하는 잠재적인 특징을 추출하거나 저해상도 영상과 고해상도 영상의 특징 차원의 크기의 오류에 대한 문제를 해결하기 위해 영상의 구조적인 특징을 기반으로 하는 저해상도 얼굴 인식 방법이 연구되고 있다[30] [31].

구조적인 특징을 이용한 방법은 분류 관점에서 저해상도의 DB와 고해상도의 검증 DB들의 직접적인 비교를 효과적으로 하기 위해 저해상도 영상과 고해상도 영상의 특징 공간상의 관계를 구축하는 것에 중심을 두고 있다. 특히, 저해상도와 고해상도 영상간의 특징 차수에 따른 문제를 해결할 수 있다. 특징 차수의 부조화 문제를 해결하는 대표적인 방법으로 CMs(Coupled Mappings) 방법이 있다[32]. CMs 모델은 다음 식 (20)과 같이 표현할 수 있다. 여기서, l_i와 h_i를 각각 저해상도 영상과 고해상도 영상이라고 할 때, A_L와 A_H를 두 개의 쌍으로 매핑된 매트릭스라고 정의할 수 있다.

$$J(A_L, A_H) = \sum_{i=1}^{N_t} \| A_L^T l_i - A_H^T h_i \|^2 \tag{20}$$

$$J(A_L, A_H) = \sum_{i=1}^{N_t} \| A_L^T l_i - A_H^T h_i \|^2 W_{ij} \tag{21}$$

저해상도 기반 얼굴 인식의 경우 저해상도 영상이 입력되면 이 영상과 가장 유사한 특징 공간을 찾고, 그에 상응하는 고해상도 영상을 출력하는 방법이다. 이 경우, 단순히 저해상도 영상을 학습해서 사용하는 방법으로 분류 과정에서 만족할 만한 성능을 기대할 수 없다. 그래서 식 (21)과 같이 관계 가중치 W_{ij}를 이용한

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

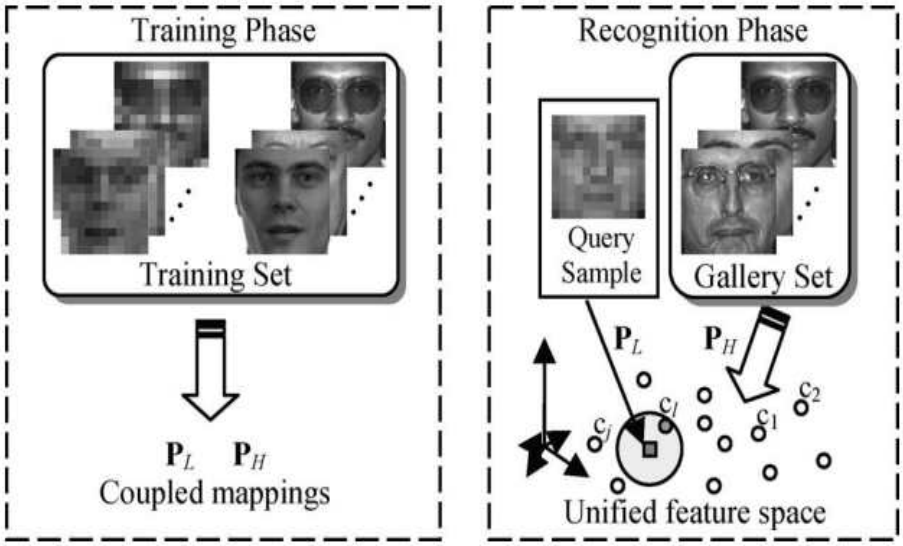


그림 11. CLPMs을 이용한 저해상도 얼굴 인식 방법의 개념도

CLPMs(Coupled locality preserving mappings)라는 방법이 제안되었다[33]. 그림 11은 CLPMs를 이용한 저해상도 얼굴 인식의 전체적인 개념도를 나타낸다. 구조적 특징을 이용한 저해상도 얼굴 인식 방법의 경우 영상의 질보다는 얼굴 인식에 초점을 둔 방법으로 초해상도 복원 방법에 비해 비교적 계산 복잡도가 적은 장점이 있다. 구조적 특징을 이용한 방법은 정의되지 않은 외부 변수 및 관계 가중치 값에 따라 인식률의 변화가 심하게 발생하는 특징이 있다.

3. PTZ 카메라를 이용한 원거리 얼굴 인식

CCTV에서 획득된 얼굴 영상의 품질은 크게 해상도, 선명도, 포즈, 조명 등의 기준으로 측정될 수 있다. 이러한 기준 중 시스템 구성에 가장 의존적인 요소는 해상도 측면을 들 수 있다. 즉, 얼굴 인식에서 카메라와 사람의 거리가 멀어지게 되면 얼굴 인식에 필요한 최소한의 얼굴 정보를 획득할 수 없게 되는데, 이와 같은 문제

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

를 PTZ 카메라를 통해 해결할 수 있다[34-36]. PTZ 카메라를 이용한 원거리 얼굴 인식은 화질 열화에 대한 근본적인 문제를 하드웨어로 해결하기 때문에 PTZ 카메라가 동작하는 환경에 초점을 두어야 한다.

PTZ 카메라를 이용한 얼굴 인식 기술은 PTZ 카메라와 고정 카메라의 수에 따라 크게 단일 PTZ 카메라를 이용하는 방법[86], 한 대의 고정형 카메라와 PTZ 카메라를 함께 이용하는 방법[37][87] 그리고 다수의 고정형 카메라와 PTZ 카메라를 이용하는 방법 등으로 나뉠 수 있다[88].

한 대의 PTZ 카메라를 이용한 방법은 줌 아웃된 화면에서 얼굴을 추정 및 검출하고 추정된 얼굴 위치를 줌 인시켜 고해상도 얼굴 영상을 획득하는 방식이다. 이 방법은 고정형 카메라를 사용하지 않기 때문에 추가적인 비용이 발생하지 않지만 고정형 카메라를 통한 깊이 정보를 획득할 수 없는 단점이 있다. 또한, 지속적으로 줌인과 아웃을 계속해서 반복해야하기 때문에 영상에서 빠르게 움직이는 객체를 잃어버리기 쉬운 문제가 있다.

한 대의 고정형 카메라와 단일 혹은 다수의 PTZ 카메라를 이용한 방법은 고정형 카메라를 통해 얼굴 위치를 추정하고 PTZ 카메라를 제어하여 고해상도 얼굴 영상

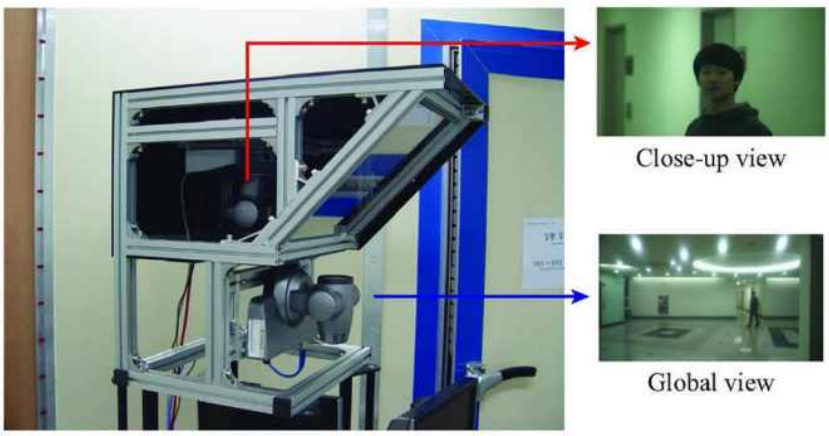


그림 12. 고정 카메라와 PTZ 카메라를 이용한 얼굴 인식 시스템

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

을 획득하는 방식이다. 한 대의 고정형 카메라를 이용하는 경우도 영상에서 객체의 깊이 정보를 파악하기 어렵기 때문에 고정형 카메라의 제한된 범위 내에서만 사용 가능하다. PTZ를 활용한 얼굴 인식 시스템의 경우 한 대의 고정 카메라와 한 대의 PTZ 카메라로 구성되는 것이 일반적이다. 고정 카메라를 통해 입력되는 영상에서 움직이는 물체를 추적하여 추적된 물체에 대해 PTZ 카메라로 영상의 중심으로 이동 시키고 얼굴을 검출한다. 이때, 검출되는 얼굴 영상은 이미 고해상도 영상이기 때문에 바로 얼굴 인식 시스템에 적용할 수 있는 장점이 있다. 그림 12는 1대의 고정 카메라와 1대의 PTZ 카메라를 이용한 얼굴 인식 시스템의 구성을 나타낸다 [37]. 그림에서와 같이 고정형 카메라를 통해 원거리 얼굴을 추정하고 PTZ 카메라를 통해 고해상도의 얼굴 영상을 획득하게 된다.

마지막으로 다수의 고정형 카메라와 한 대의 PTZ 카메라를 이용한 방법은 다수의 고정형 카메라를 통해 3D 좌표값을 추정할 수 있고 입력 영상을 입체적으로 복원할 수 있는 시스템이다. 이러한 구조는 입체 복원에 필요한 연산이 복잡하고 실제 적용 범위가 제한적이기 때문에 많이 사용되고 있는 형태는 아니다. 일반적으로 다수의 고정 카메라는 감시 범위를 증가시키기 위해 사용되고 PTZ 카메라는 다수의 객체를 추적하기 위해 사용된다.

PTZ 카메라를 이용하면 화질 저하로 인한 인식률 저하의 근본적인 문제를 해결할 수 있기 때문에 매우 우수한 성능을 나타내고 있지만, PTZ 카메라를 이용해야 한다는 이유로 비용의 부담이 발생하게 된다. 또한, PTZ 카메라의 특성상 PTZ 카메라 한 대만을 이용하는 경우, 하나의 물체를 추적하는 동안 다른 영역을 감시할 수 없기 때문에 감시를 위해서는 반드시 고정형 카메라가 동반되어야 하는 단점이 있다.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

4. 기존 원거리 얼굴 인식 방법 분석

대표적인 원거리 얼굴 인식 방법으로는 초해상도 복원 방법을 이용한 방법, 영상 간 구조적 특징을 이용한 방법, PTZ 카메라를 이용한 얼굴 인식 방법 등이 있다. 표 2는 기존 원거리 얼굴 인식 방법별 특징을 나타낸다[30].

초해상도 복원을 이용한 원거리 얼굴 인식 방법의 경우 시각적 화질 개선이 목적이기 때문에 실제 원거리 얼굴 인식률의 성능은 우수하지 못하고 영상 복원을 위해서는 사전에 많은 양의 데이터를 학습시켜야하는 단점이 있다. 복원에 필요한 연산이 매우 복잡해 처리시간이 많이 요구되지만 입력 영상의 크기가 매우 작아도 영상을 복원할 수 있는 장점이 있다.

구조적 특징을 이용한 저해상도 얼굴 인식 방법은 초해상도 복원 방법과 달리 얼굴 인식률 향상에 초점을 둔 기술로 계산 복잡도가 작고 학습에 사용되는 얼굴 영상이 적어도 우수한 인식률을 나타내는 장점이 있다. 학습에 사용되는 얼굴 영상이 적은만큼 외부 변수에 민감하게 작용하고 학습에 사용된 고정된 입력 영상의 크기

표 2. 기존 원거리 얼굴 인식 방법 비교

비교 사항	초해상도 복원 기반	구조적 특징 기반	PTZ 카메라 기반
주목적	**시각적 화질 개선**	얼굴 인식률 향상	얼굴 인식률 향상
해상도 허용 범위	매우 작음	작음	매우 작음
계산 복잡도	**높음**	낮음	낮음
필요 학습 영상 수	**매우 많음**	적음	적음
외부 변수	민감	**매우 민감**	민감
비용	적음	적음	**높음**
실험 환경	**근거리 제한적인 환경**	**근거리 제한적인 환경**	원거리 실제 환경

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

에 의존적이기 때문에 다양한 거리에서 추출되는 얼굴 영상 크기에 대해서는 직접 적용이 어려운 단점이 있다.

PTZ 카메라를 이용한 원거리 얼굴 인식 방법은 화질 열화에 의해 발생되는 인식률 저하 문제를 PTZ 카메라를 통해 해결함으로써 매우 우수한 인식률을 나타낸다. 또한, 초해상도 복원과 구조적 특징을 이용한 원거리 얼굴 인식 방법이 근거리 제한적인 환경에서 촬영된 얼굴 DB를 통해 실험이 되었다면 PTZ 카메라를 통한 방법은 실제 원거리 얼굴 인식 환경에서 우수한 성능을 나타냈다. 이 방법은 카메라 자체의 가격이 매우 비싸기 때문에 대중적으로 사용되기에는 한계가 있다.

이와 같이 다양한 얼굴 인식 기술들이 연구되고 있지만 잘 제어된 환경에서 얼굴 인식률 향상에만 초점을 두거나 실제 CCTV 환경이나 원거리에서 획득된 얼굴 영상을 기반으로 진행된 연구는 부족한 상태이다. 또한, 현재 연구되고 있는 기술들이 새로운 CCTV 환경에는 가능할 수 있으나, 처리 속도 및 비용의 문제 등으로 기존 CCTV 기반 감시 시스템 환경에는 적용하기 어렵다. 그러므로 본 논문에서는 기존 얼굴 인식 기술의 문제점을 보완하고 실제 CCTV 환경에서 동작이 가능한 원거리 얼굴 인식 방법을 제안한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3장 제안하는 거리별 얼굴 영상을 이용한 원거리 얼굴 인식

본 장에서는 제안하는 원거리 얼굴 인식 알고리즘의 전체적인 흐름과 학습 데이터 구성에 이용되는 거리별 얼굴 영상 취득 방법에 대해 설명한다. 얼굴 인식 시스템에서 학습에 사용되는 얼굴 영상의 선택은 인식률과 직접 연관이 있는 매우 중요한 과정이다. 원거리 얼굴 인식은 얼굴의 포즈 및 조명 변화뿐만 아니라 원거리에서 획득되는 얼굴 영상의 크기와 영상의 질에 민감한 변화를 나타낸다. 본 논문에서는 거리 변화에 의해 발생되는 화질저하에 따른 인식률 저하 문제를 보완하기 위해 거리별 얼굴 영상을 이용한 학습 데이터 구성 방법을 제안한다.

거리별 얼굴 영상을 이용한 방법은 얼굴 영상을 취득하는 방법에 따라 실제 거리별 얼굴 영상, 줌 카메라를 이용한 가상의 거리별 얼굴 영상, 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상으로 구분된다. 실제 거리별 얼굴 영상을 이용한 방법은 사용자 등록 단계에서 추가적인 사용자 협조를 통해 근거리와 원거리에 해당하는 정확한 개인의 얼굴 특징을 획득할 수 있는 장점이 있다. 줌 카메라를 이용한 가상의 거리별 얼굴 영상을 이용한 방법은 추가적인 장비를 통해 정확한 원거리 얼굴 특징과 비교적 정확한 근거리 얼굴 특징을 획득한다. 자동 생성 방법을 통한 가상의 거리별 얼굴 영상을 이용한 방법은 사용자 협조 및 추가적인 장비 없이 근거리 및 원거리 얼굴 특징을 획득할 수 있다. 기존 방법이 사용자 등록 단계에 필요한 개인의 얼굴 특징을 근거리에서만 정확히 획득하는 반면 제안하는 방법은 근거리 및 원거리에서 정확한 개인의 고유 얼굴 특징을 획득해 사용자 등록에 사용하기 때문에 근거리 및 원거리에서 우수한 인식 성능을 제공한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 13은 제안하는 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 알고리즘의 순서도를 나타낸다. 얼굴 인식 알고리즘은 크게 학습 영역과 검증 영역으로 나뉜다. 거리별 얼굴 영상을 이용한 학습 과정은 다음과 같다. 학습으로 사용되는 얼굴 영상으로 실제 거리별 얼굴 영상 혹은 가상의 거리별 얼굴 영상이 입력되면 양선형 보간법에 의해 얼굴 영상의 크기는 정규화 된다. 정규화된 얼굴 영상은 히스토그램 평활화에 의해 밝기가 정규화 되고, PCA 및 LDA를 통해 특징을 추출한다. 학습 영상별 특징들은 DB에 저장된다. 검증 과정은 다음과 같다. 다양한 거리에서 추출된 실제 거리별 얼굴 영상들이 입력되면 양선형 보간법과 히스토그램 평활화를 이용해 얼굴 크기와 밝기를 정규화 한다. 정규화된 얼굴 영상을 PCA와 LDA를 통해 특징을 추출한다. 최종적으로 검증 과정의 특징과 DB의 특징을 비교해서 가장 유사한 값을 가지는 얼굴 영상을 찾아낸다. 이때, 유사도 측정 방법으로는 유클리디언 디스턴스(Euclidean Distance) 방법을 사용한다.

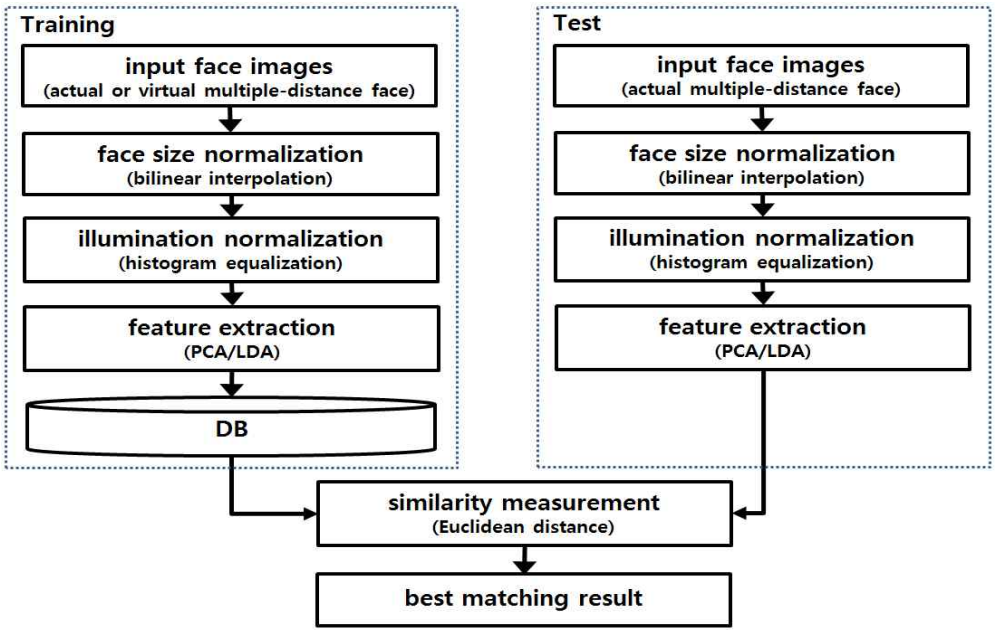


그림 13. 제안하는 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 알고리즘

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제1절 기존 단일 거리 얼굴 영상을 이용한 얼굴 인식

단일 거리 얼굴 영상을 이용한 학습 데이터 구성 방법은 동일한 거리에서 추출한 영상의 단일 영역기반 얼굴 영상을 사용하는 방법과 프레임별 다중 영역기반 얼굴 영상을 사용하는 방법이 있다.

1. 단일 영역기반 얼굴 영상

단일 거리 얼굴 영상을 학습 데이터로 구성하는 방법은 전영역 방법을 이용한 얼굴 인식 기술에서 주로 사용했던 방법으로 얼굴 인식 환경에 관계없이 동일한 거리에서 추출된 얼굴만을 학습 데이터로 사용하는 방식이다. 그림 14는 근거리와 원거리에서 추출된 실제 얼굴 영상을 나타낸 것으로 영상 정규화를 하지 않은 원본 얼굴 영상이다. 그림에서와 같이 원거리 얼굴 인식에서 얼굴 영상은 카메라와 대상의 거리에 따라 추출되는 얼굴 영상의 크기 및 화질의 차이가 발생한다. 그리고 근거리와 원거리 얼굴 영상의 차이는 거리가 멀어질수록 크게 발생한다.

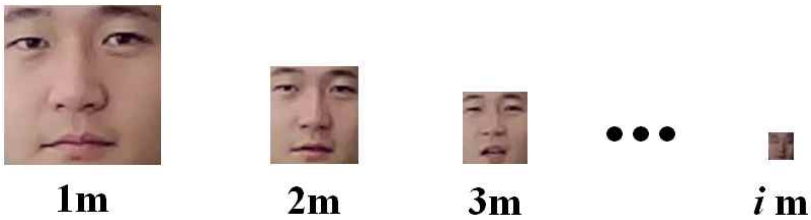


그림 14. 근거리와 원거리 얼굴 영상의 비교

그림 15는 1m 거리에서 획득한 영상에서 추출한 단일 영역의 얼굴 영상을 이용

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

해 학습 데이터를 구성하는 예이다. 학습용 데이터는 지상으로부터 약 1.8m 높이에 설치된 카메라를 기준으로 사용자가 1m 떨어진 위치에서 획득된다. 그림 16은 2m에서 추출된 얼굴 영상을 학습 데이터로 사용하는 예이고, 그림 17은 im에 해당하는 거리에서 추출된 얼굴 영상을 학습 데이터로 사용하는 예로 학습 데이터의 정의는 다음 식 (22)와 같다.

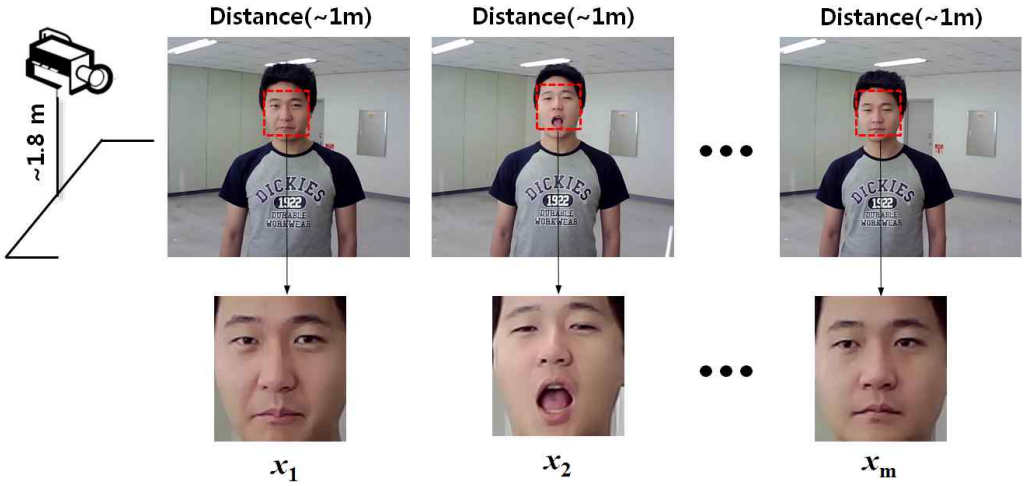


그림 15. 1m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법

$$\Gamma_n = [x_1, x_2, \cdots, x_k], \quad k = 1, 2, \cdots, m$$
$$S_j = \{\Gamma_1, \Gamma_2, \cdots, \Gamma_n\}, \quad j = 1, 2, \cdots, i \tag{22}$$

한 후보당 학습 데이터용 얼굴 영상이 m개라고 할 때, Γ_n는 각 후보별 학습 데이터를 구성하는 얼굴이고 n은 얼굴 인식 시스템에 등록되는 총 후보의 수이다. 전체 후보의 학습 데이터용 얼굴 영상의 집합을 S_j라고 할 때, j는 학습 데이터로 사

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

용되는 얼굴이 취득되는 거리를 나타내고 i는 해당 카메라로부터 얼굴이 획득될 수 있는 최대 거리다. 여기서, 학습 데이터는 S_1부터 S_i까지 각각 구성되어 별개의 얼굴 인식 시스템에 적용된다.

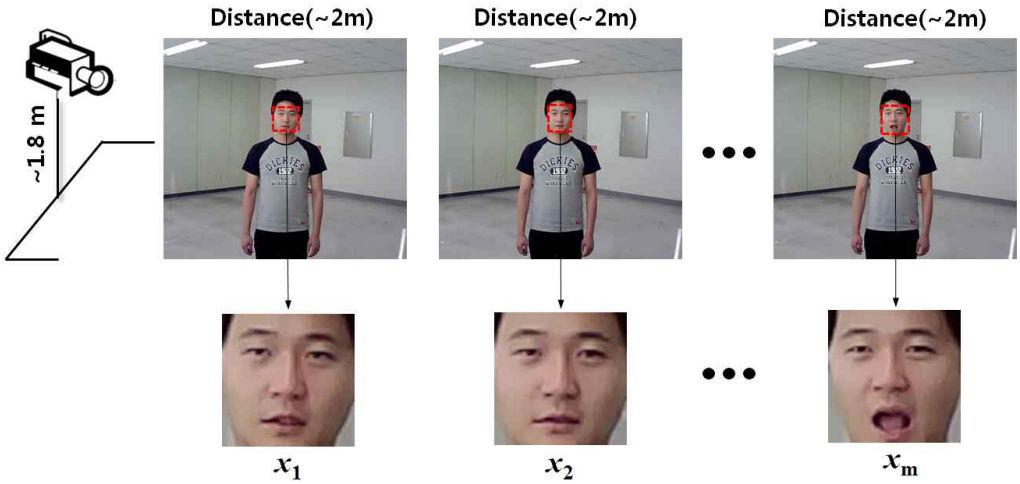


그림 16. 2m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법

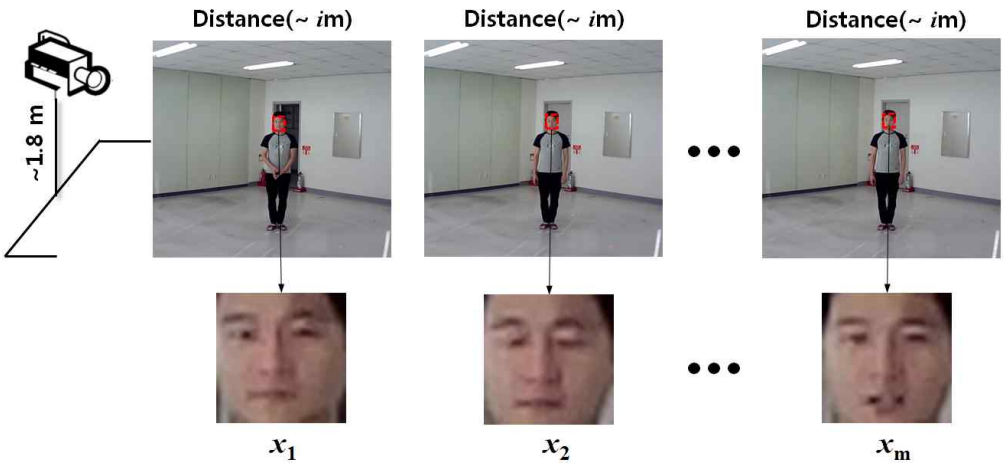


그림 17. i m 얼굴의 단일 영역을 이용한 학습 데이터 구성 방법

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

2. 다중 영역기반 얼굴 영상

전영역 방법을 이용한 원거리 얼굴 인식은 영상의 전체적인 특징을 이용한 인식 방법으로 학습에 사용되는 영상과 검증에 사용되는 얼굴 영상의 차이가 적을수록 인식률이 향상된다. 그림 18은 추출된 얼굴 영역에 따른 후보 간 유사도의 특징의 예를 나타낸다. 전영역 방법을 이용할 경우 그림에서와 같이 추출된 얼굴 영역이 유사한 후보 A와 B의 유사도는 추출된 얼굴 영역의 차이가 심한 동일한 후보 A의 유사도보다 높은 특징이 나타날 수 있다. 실제 근거리의 동일한 거리에서 추출되는 얼굴 영상은 원본 영상의 크기가 크기 때문에 추출되는 얼굴 영역의 차이가 크게 나지 않는다. 그림 19와 같이 원거리 얼굴 영상의 경우 얼굴을 표현하는 해상도가 근거리 얼굴 영상보다 훨씬 낮기 때문에 근거리에서와 같이 정교한 얼굴 영역의 추출이 어렵게 된다. 즉, 1m에서 추출한 얼굴 영상을 학습 데이터로 한다면 원거리에서 추출된 얼굴 영상이 검증으로 들어왔을 때 얼굴 인식에 실패할 확률이 높아지게 된다.

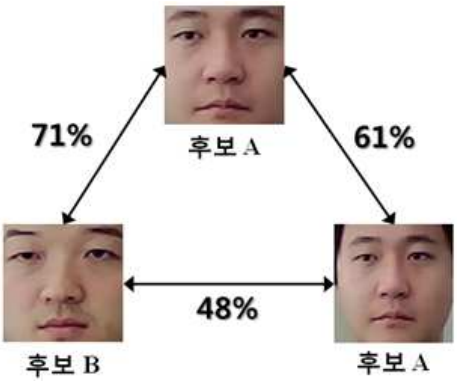


그림 18. 얼굴 영역에 따른 후보 간 유사도 특징

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

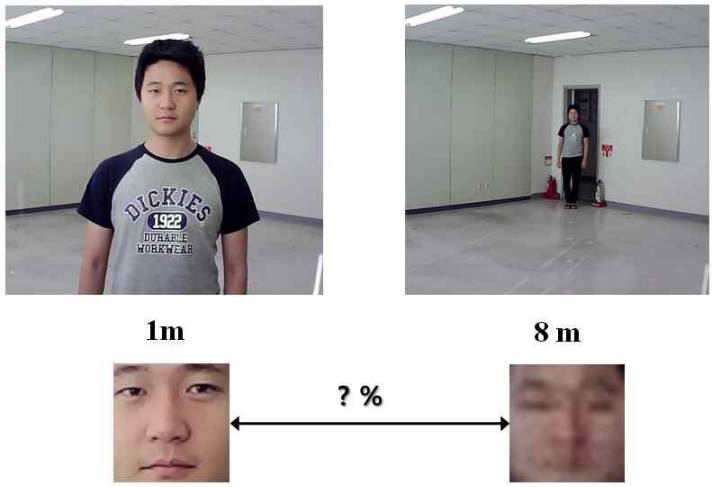


그림 19. 근거리와 원거리에서 검출된 얼굴 영역 비교

이와 같은 문제를 해결하기 위해 본 논문에서는 단일 거리 내에서 각 프레임별로 얼굴 영역의 범위를 다르게 하는 다중 영역기반 얼굴 추출 방법을 사용한다. 그림 20은 1m에서 프레임별로 얼굴 영역의 범위를 변화시켜 얼굴을 추출하는 예를 나

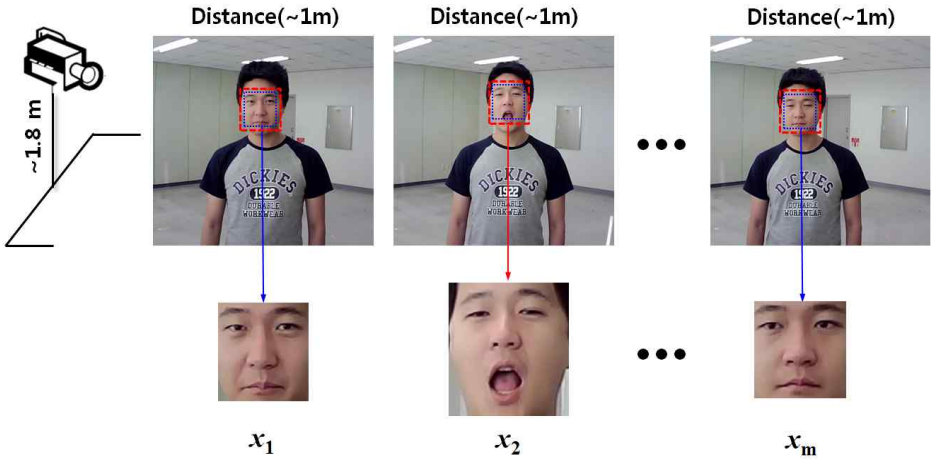


그림 20. 1m 얼굴의 다중 영역을 이용한 학습 데이터 구성 방법

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

타낸다. 이와 같이 불규칙한 얼굴 영역으로 추출된 얼굴 영상을 학습으로 사용할 경우 원거리에서 불규칙하게 추출되는 얼굴 인식에 강인할 수 있게 된다. 다중 영역기반 얼굴 추출 방법도 식 (22)와 같이 단일 영역을 이용한 방법과 동일하게 1m ~ im까지 학습 데이터를 구성한다.

제2절 제안하는 거리별 얼굴 영상을 이용한 얼굴 인식

제안하는 거리별 얼굴 영상을 이용한 방법은 거리별 얼굴 영상의 취득 방법에 따라 총 3가지로 나뉜다. 첫째는 CCTV로부터 획득된 실제 거리별 얼굴 영상을 이용해 학습 영상으로 사용하는 방법이고, 둘째는 광학식 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 이용하는 방법이다. 마지막으로 셋째는 단일 거리 얼굴 영상을 이용해 거리별 얼굴 영상을 자동으로 생성해 학습 영상으로 이용하는 방법이다.

1. 실제 거리별 얼굴 영상

실제 거리별 얼굴 영상을 학습 데이터로 사용하는 방법은 1m부터 카메라를 통해 얼굴 검출이 가능한 최대 거리까지 사용자가 직접 이동하며 얼굴을 획득하는 방식이다. 그림 21은 실제 거리별 얼굴 영상을 획득하는 방법으로 카메라의 위치는 고정되어 있고 사람이 직접 1m ~ im까지 1m 간격으로 이동하며 얼굴 영상을 취득하는 과정을 나타낸다. 실제 거리별 얼굴 영상을 이용한 방법은 정확한 근거리 및 원거리 얼굴 영상을 취득할 수 있는 장점이 있다. 그러나 단일 거리 얼굴 영상을 이용하는 방법에 비해 사용자 협조가 추가적으로 요구되는 단점이 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

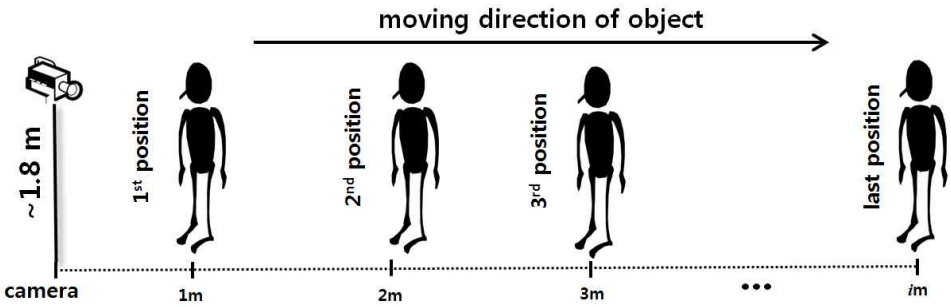


그림 21. 실제 거리별 얼굴 영상 획득 방법

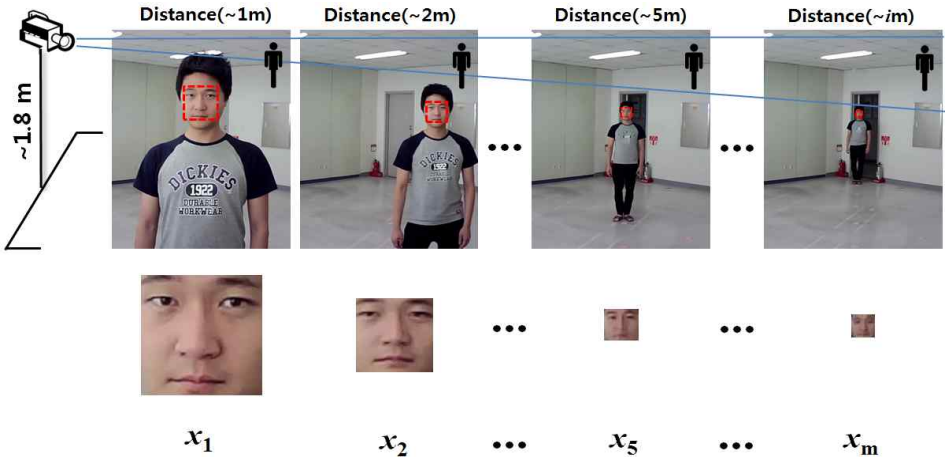


그림 22. 실제 거리별 얼굴 영상을 이용한 학습 데이터 구성 방법

그림 22는 1m ~ im에서 거리별 얼굴 영상을 획득하는 예이다. 이때, 카메라의 위치는 지상으로부터 약 1.8m 높이에 설치됐다. 각 거리별로 얼굴 영상 x_1부터 x_m을 비교하면 추출된 얼굴 영상의 크기가 다름을 확인할 수 있다. PCA나 LDA와 같이 전영역 방법을 이용한 얼굴 인식은 클래스내 분산 행렬의 특이점 때문에 학습 영상 샘플의 크기가 검증 영상 샘플의 크기보다 크거나 작다면 얼굴 인식을 적용할 수 없다. 즉, 다양한 거리에서 추출된 얼굴 영상은 거리에 따라 얼굴 영상 샘플의

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

차원이 다르게 추출되는 문제가 발생한다. 이러한 문제를 해결하기 위해 영상 크기 정규화 과정이 필요하다. 이때, 얼굴 영상은 정규화 방법에 따라 복원되는 얼굴의 정보가 달라지기 때문에 원거리 얼굴 인식 상황에 적합한 보간법의 선택이 요구된다. 거리별 얼굴 영상의 크기 정규화 과정은 다음 식 (23)과 같이 정의된다. 여기서, $F_{intp}(\,)$는 보간 함수로 본 논문에서는 최근접 이웃 보간법, 양선형 보간법, 양3차회선 보간법, 이중 선형-3차회선 보간법을 사용한다. x_k는 영상 크기가 정규화되지 않는 원본 얼굴 영상이고, X_k는 제안하는 보간법에 의해 스케일링된 얼굴 영상이다. R_s는 스케일링 배율 값으로 식 (24)을 통해 구해진다. 여기서, x_s 기준이 되는 거리의 얼굴 영상이다.

$$X_k = F_{intp}(R_s, x_k) \qquad k = 1, 2, 3, \cdots, m \tag{23}$$

$$R_s = \frac{x_s}{x_k} \tag{24}$$

실제 거리별 얼굴 영상을 이용한 학습 데이터 정의는 다음과 같다. 한 후보당 학습 데이터용 얼굴 영상이 m개라고 할 때, Γ_n는 각 후보별 학습 데이터를 구성하는 얼굴이고 n은 얼굴 인식 시스템에 등록되는 총 후보의 수이다. 전체 후보의 학습 데이터용 얼굴 영상의 집합 S는 다음 식 (25)와 같이 정의할 수 있다.

$$\begin{aligned} \Gamma_n &= [X_1, X_2, X_3, \cdots, X_k] \\ S &= \{\Gamma_1, \Gamma_2, \cdots, \Gamma_n\} \end{aligned} \tag{25}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

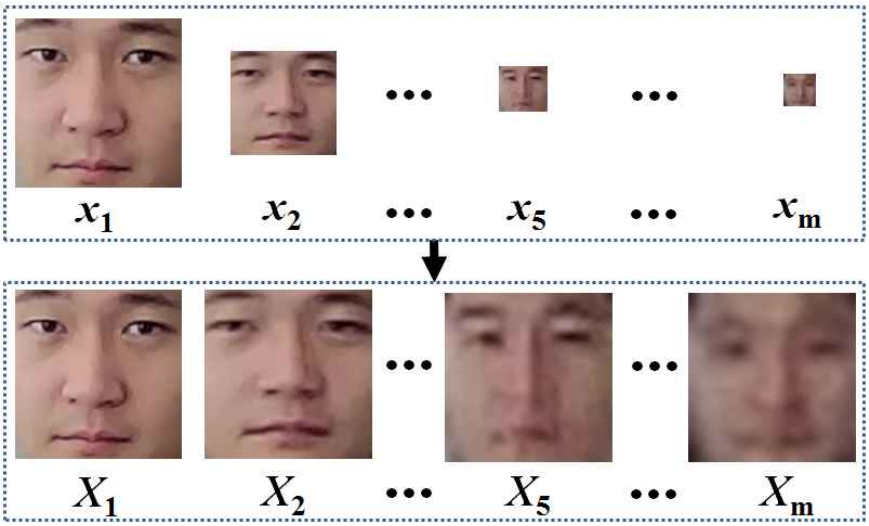


그림 23. 보간법을 통해 정규화된 얼굴 영상

그림 23은 보간법을 이용해 실제 거리별 얼굴 영상을 1m 얼굴 영상의 크기에 맞춰 정규화한 예를 나타낸다. 이때, 기준 영상의 크기는 사용자의 요구에 따라 1m ~ im가 될 수 있다.

영상에서 공간주파수는 밝기 값의 변화량이 공간에 따라서 얼마나 자주 급격하게 변하느냐를 나타내는 기준이 된다. 물체의 에지 부분은 고주파 성분을 많이 포함하고, 배경이나 비슷한 밝기 값을 가지는 물체 영역에서는 저주파 성분을 많이 포함한다. 그림 28에서 영상 크기가 정규화된 근거리 얼굴 영상 X_1의 경우 고주파 성분을 많이 포함하고, 원거리 얼굴 영상 X_m의 경우 저주파 성분을 많이 포함하는 것을 확인할 수 있다.

본 논문에서는 근거리에서 추출된 얼굴 영상에 저역 통과 필터링을 통해 고주파 성분을 제거함으로써 원거리에서 추출된 얼굴과 유사하게 만드는 화질 열화 방법을 사용한다. 고해상도 영상의 화질 열화를 위해 가우시안 저역 통과 필터링을 사용한다. 가우시안 저역 통과 필터링은 가우시안 함수를 이용해 영상에 나타나는 패턴

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

영역의 대표적인 일치성을 구해 이미지를 평활화하는 방법이다. 2차원 이미지 상에서 가우시안 저역 통과 필터링이 적용될 때, 이 필터는 가우시안 분포의 윤곽선이 중심점으로부터 단일 돌출 부분을 가지는 동심원의 형태로 나타난다. 이 분포들의 값들은 원본 이미지에 적용될 행렬을 만드는데 사용되어지며, 중앙에 위치한 화소와 먼 거리에 있는 이웃 화소들을 가중치로 감소시켜, 가중한 이웃의 평균값으로 대치하는 특징을 가진다. 즉, 각 화소의 새로운 값은 화소 주변 값들의 가중 평균값으로 설정되어 진다. 이때, 가중치를 결정하는 함수는 다음 식 (26)과 같다.

$$G(w,z) = \frac{1}{2\pi\sigma^2} e^{\left(-\frac{w^2+z^2}{2\sigma^2}\right)} \tag{26}$$

식 (26)에서 가우시안 필터의 성능을 결정하는 주요 파라미터는 σ이다. 일반적인 컨볼루션은 무한대 영역으로 수행하지만 디지털 필터링에서는 무한이 넓은 윈도우를 설정하기가 불가능하므로, 영역을 제안하게 된다. 이때의 윈도우 사이즈를 결정하는 값이 σ이다. 그림 24는 가우시안 저역 통과 필터링이 적용된 거리별 얼굴 영상의 결과를 나타낸다. 이때, 가우시안 저역 통과 필터링은 모든 거리별 얼굴 영상에 적용하지 않고, 고주파 성분을 많이 포함한 근거리 얼굴 영상에만 적용된다. 거리별 얼굴 영상에서 고주파 성분을 많이 포함한 영상을 결정하는 방법은 다음 식 (27)과 같다.

$$X_k^G = \begin{cases} X_k \bullet G(w,z) & \text{if} \quad x_k(width) \geq I_{th} \\ X_k & \text{if} \quad x_k(width) < I_{th} \end{cases} \qquad k = 1, 2, 3, \cdots, m \tag{27}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

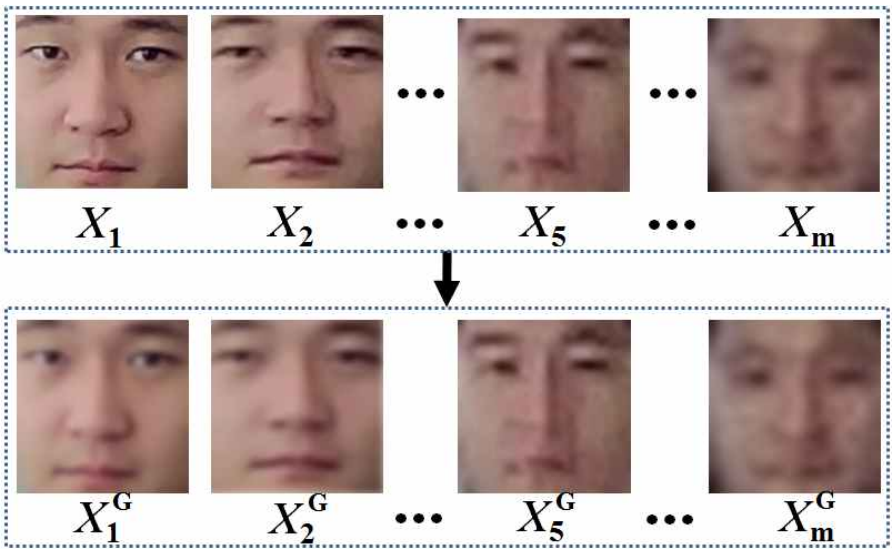


그림 24. 가우시안 저역 통과 필터링이 적용된 거리별 얼굴 영상

얼굴 영상의 크기가 $width \times height$인 영상에서 X_k^G는 얼굴 크기가 정규화된 거리별 얼굴 영상에 가우시안 저역 통과 필터링을 적용한 영상으로, 거리별 원본 얼굴 영상 x_k의 $width$ 값이 I_{th}보다 크거나 같으면 근거리에서 획득된 고해상도 영상으로 판단하고 저역 통과 필터링을 적용한다. 만약 x_k의 $width$ 값이 I_{th}보다 작으면 원거리에서 획득된 저해상도 영상이라고 판단하고 저역 통과 필터링을 적용하지 않는다. 여기서, I_{th}는 고해상도 영상과 저해상도 영상을 판단하는 기준으로 다음 식 (28)로 정의한다.

$$I_{th} = \frac{\frac{1}{m}\sum_{k=1}^{m} x_k^{dose}(width) + \frac{1}{m}\sum_{k=1}^{m} x_k^{far}(width)}{2} \tag{28}$$

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

여기서, x_i^{close}는 거리별 얼굴 영상에서 가장 근거리에서 추출된 얼굴 영상이고, x_k^{far}는 가장 원거리에서 추출된 영상이다. m은 동일한 거리에서 추출된 얼굴 영상의 수이다. 결과적으로, 실제 거리별 얼굴 영상을 방법에서 가우시안 저역 통과 필터링을 사용하기 때문에 학습 데이터를 정의하는 식 (25)는 다음 식 (29)로 나타낼 수 있다.

$$X_k^G = F_{intp}(R, x_k) \bullet G(w,z) \qquad k = 1, 2, 3, \cdots, m$$

$$\Gamma_n^g = \left[X_1^G, X_2^G, X_3^G, \cdots, X_k^G\right] \tag{29}$$

$$S^g = \left\{\Gamma_1^g, \Gamma_2^g, \cdots, \Gamma_n^g\right\}$$

2. 줌 카메라를 이용한 가상의 거리별 얼굴 영상

얼굴 인식에서 학습에 사용되는 영상의 수는 제한적이고, 많은 양의 영상을 사용하기 위해서는 사용자의 추가적인 협조가 요구된다. 거리별 얼굴 영상을 학습 데이터로 사용한 방법은 사용자 등록에 사용되는 거리별 얼굴 영상을 취득하기 위해 사용자가 직접 거리별로 이동해야 하기 때문에 추가적인 사용자 협조가 필요한 단점이 있다. 본 논문에서는 사용자 협조가 적은 환경에서 거리별 얼굴 영상을 획득하기 위해 줌 카메라를 사용하는 방법을 제안한다. 그림 25는 광학식 줌을 이용해 가상의 거리별 얼굴 영상을 취득하는 방법을 나타낸다. 사용자는 직접 해당 거리로 이동하지 않고, 가장 먼 거리에 위치하면 줌을 이용해 가상의 거리별 얼굴 영상을 취득하게 된다. 이 방법은 사용자 협조가 단일 거리 얼굴 영상을 이용하는 것과 동

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

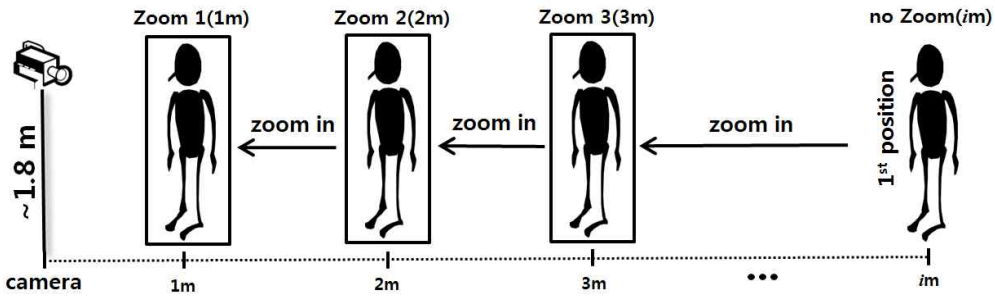


그림 25. 광학식 줌을 이용한 거리별 얼굴 영상 획득 방법

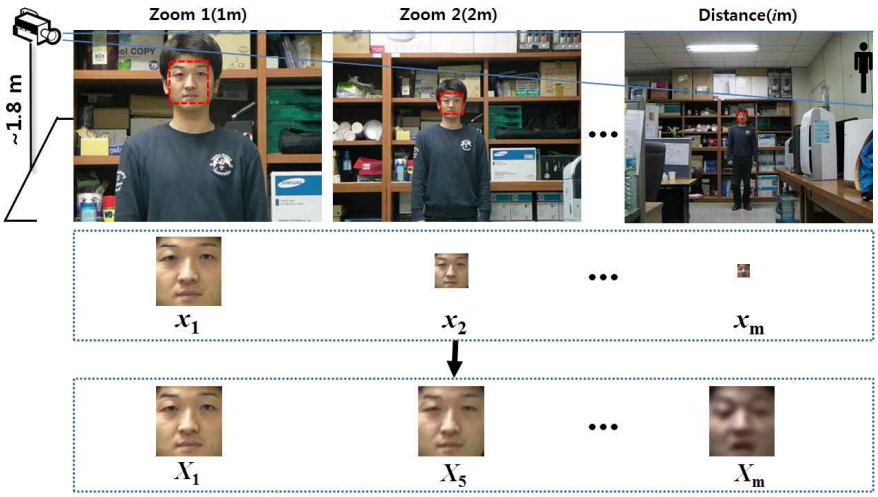


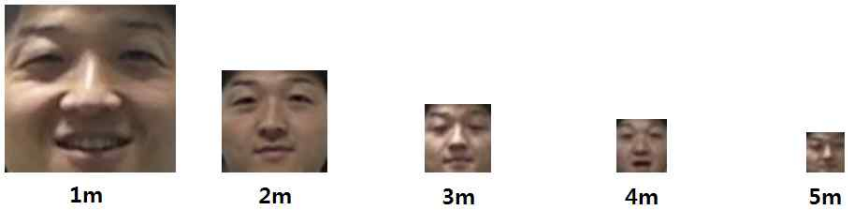
그림 26. 줌 카메라를 이용한 가상의 거리별 얼굴 영상 구성 방법

일한 장점이 있고 정확한 정보를 포함한 원거리 얼굴 영상과 비교적 정확한 정보를 포함한 근거리 얼굴 영상을 획득할 수 있다. 그러나 거리가 멀어질수록 고배율의 광학식 줌이 요구됨에 따라 비용이 늘어나는 단점이 있다. 이때, 줌 카메라는 오직 학습 데이터를 구성하는 때에 이용되고 실제 검증 과정에서는 이용되지 않기 때문에 학습 데이터 취득용 한 대만 요구된다.

그림 26은 광학식 줌을 이용해 가상의 거리별 얼굴 영상을 취득한 결과 영상을

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY



(a) 실제 거리별 얼굴 영상

(b) 줌 카메라를 이용한 가상의 거리별 얼굴 영상

그림 27. 실제 거리별 얼굴 영상과 가상의 거리별 얼굴 영상의 비교

나타낸다. 현재 영상의 배경을 통해 사용자는 고정된 위치에 정지해 있음을 확인할 수 있다. 그림 27은 실제 거리별 얼굴 영상과 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 비교한 결과로 매우 유사한 형태의 결과를 나타냄을 확인할 수 있다. 이때, 가상의 거리별 얼굴 영상도 획득되는 얼굴의 크기가 일정하지 않기 때문에 보간법을 통해 얼굴 크기 정규화를 실시한다. 또한, 가우시안 저역 통과 필터링을 통해 근거리 얼굴 영상은 화질 열화를 시킨다. 학습 데이터의 정의는 실제 거리별 얼굴 영상을 사용한 방법과 동일하다.

3. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상

학습에 의존적인 얼굴 인식 방법에서 인식률을 향상시키는 기본적인 방법은 검증용 데이터와 가장 유사한 얼굴을 학습 데이터로 구성하는 것과 학습 데이터의 수를

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

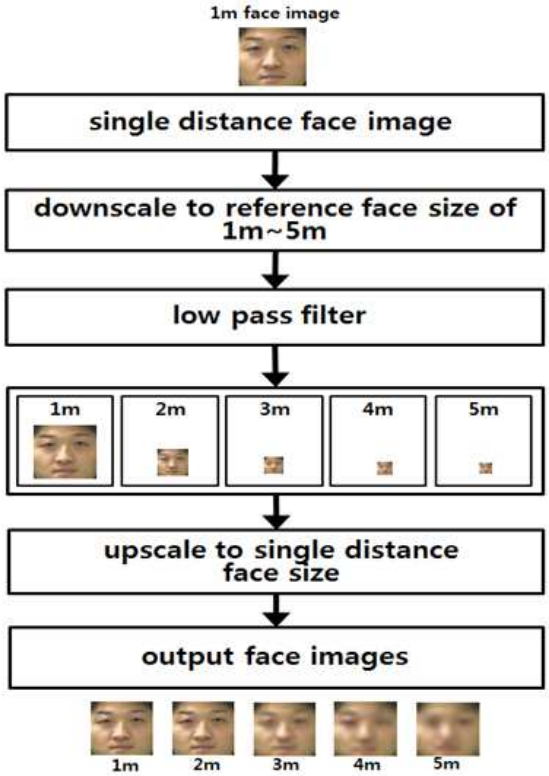


그림 28. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상 생성 과정

증가시키는 것이다. 실제 거리별 얼굴 영상을 이용한 방법은 우수한 학습 데이터를 획득할 수 있지만 추가적인 사용자 협조가 요구되고 줌 카메라를 이용한 가상의 거리별 얼굴 영상을 이용한 방법은 사용자 협조는 적지만 고가의 줌 카메라가 추가로 요구되는 단점이 있다. 본 논문에서는 사용자 협조가 적고 추가적인 카메라를 사용하지 않는 상황에서 단일 거리에서 획득된 얼굴 영상을 이용해 가상의 거리별 얼굴 영상을 자동으로 생성하는 방법을 제안한다.

그림 28은 1m 단일 거리 얼굴 영상을 이용하여 1m~5m에 해당하는 가상의 거리별 얼굴 영상 5장을 생성하는 과정을 보여준다. 학습용 얼굴 영상이 1m에서 촬영된 영상이라면, 1m 얼굴 영상을 이용해 5개의 영상을 복사 생성 후 각각 거리별

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

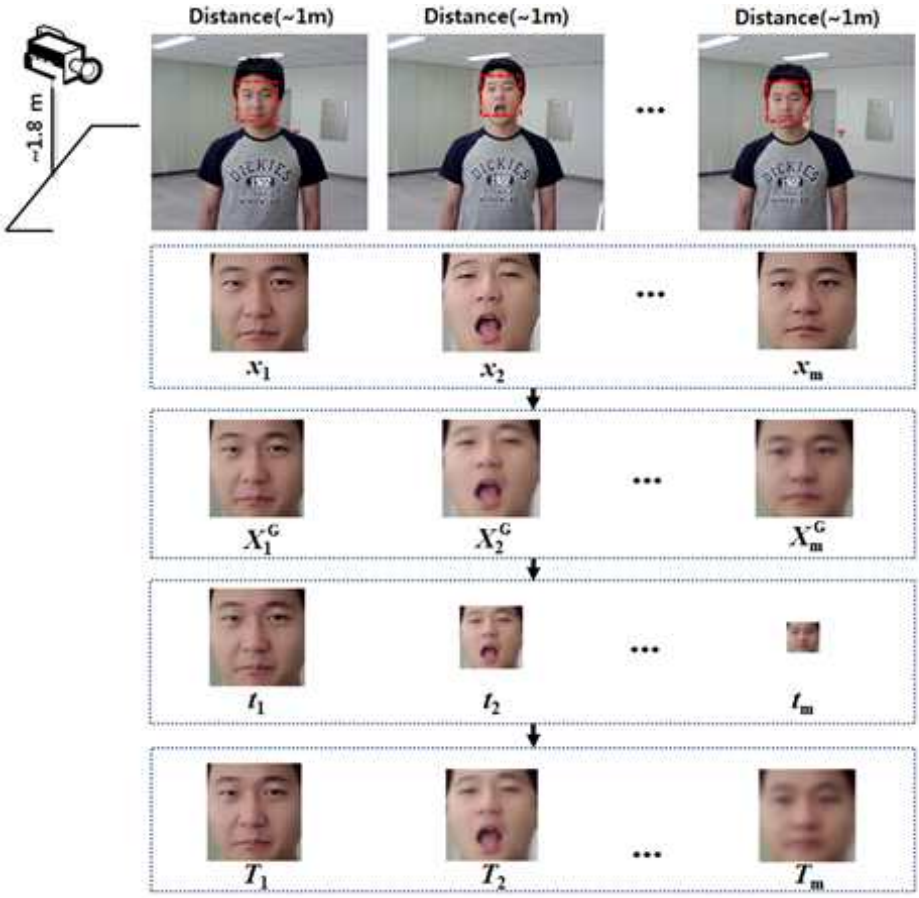


그림 29. 자동 생성 방법을 이용한 가상의 거리별 얼굴 영상 생성 예

에 맞게 축소를 시킨다. 5개의 거리별 영상은 무작위 크기로 축소되지 않고, 통계적으로 실제 1m~5m의 거리에서 추출되는 얼굴 영상의 크기를 고려하여 축소된다. 이때, 근거리 얼굴에 해당하는 영상은 저역 통과 필터링을 통해 화질 열화를 시킨다. 각각 축소된 5장의 얼굴 영상은 다시 학습 기준영상의 크기에 맞게 동일한 크기로 확대된다. 자동 생성 방법을 이용한 방법은 1m의 얼굴 영상을 통해 서로 다른 데이터를 포함하는 5장의 거리별 얼굴 영상을 생성하게 된다.

그림 29는 단일 거리에서 획득한 얼굴 영상을 이용해 가상의 거리별 얼굴 영상

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

을 생성하는 예이다. 이때, 기준이 되는 단일 거리는 1m이고 사용자의 요구에 따라 1m ~ im로 변경할 수 있다. 이때, 카메라의 위치는 지상으로부터 약 1.8m 높이에 설치됐다. 거리별 얼굴 영상 자동 생성 방법을 이용한 학습 데이터의 정의는 다음 식 (30)과 같다.

$$\Gamma_n = [T_1, T_2, T_3, \cdots, T_k] \qquad k = 1, 2, 3, \cdots, m$$
$$S = \{\Gamma_1, \Gamma_2, \cdots, \Gamma_n\} \tag{30}$$

한 후보당 학습 데이터용 얼굴 영상이 m개라고 할 때, Γ_n는 각 후보별 학습 데이터 T_k로 구성되어 있는 집합이고, n은 얼굴 인식 시스템에 등록되는 총 후보의 수이다. S는 전체 후보의 학습 데이터용 얼굴 영상의 집합이다. 여기서, T_k는 1m에서 추출된 k번째 얼굴 영상으로 다음 식 (31)에 의해 얼굴 크기가 정규화된 결과이다.

$$T_k = F_{intp}(R_s, t_k) \qquad k = 1, 2, 3, \cdots, m$$
$$t_k = \begin{cases} F_{intp}(R_{ds}, (x_k \bullet G(w,z))) & \text{if } x_k(width) \geq I_{th} \\ F_{intp}(R_{ds}, x_k) & \text{if } x_k(width) < I_{th} \end{cases} \tag{31}$$

여기서, R_s는 최종 기준 영상의 스케일링 배율이고, R_{ds}는 해당 거리별 영상의 스케일링 배율이다. 거리별 영상의 스케일링 배율 R_{ds}를 결정하는 방법은 다음 식 (32)와 같다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$R_{ds} = \frac{\frac{1}{m}\sum_{k=1}^{m} x_k^{ds}}{t_x}, \qquad ds = 1, 2, 3, \cdots i \tag{32}$$

그림 30은 실제 거리별 얼굴 영상과 자동 생성 방법을 통해 만들어진 가상의 거리별 얼굴 영상을 비교한 결과이다. 제안하는 방법의 주요 목표는 자동 생성 방법을 통해 만들어진 가상의 거리별 얼굴 영상을 실제 거리별 얼굴 영상과 유사하게 만드는 것이다. 1m를 기준 영상의 크기로 했을 때, 그림 30(a)는 1m의 단일 거리

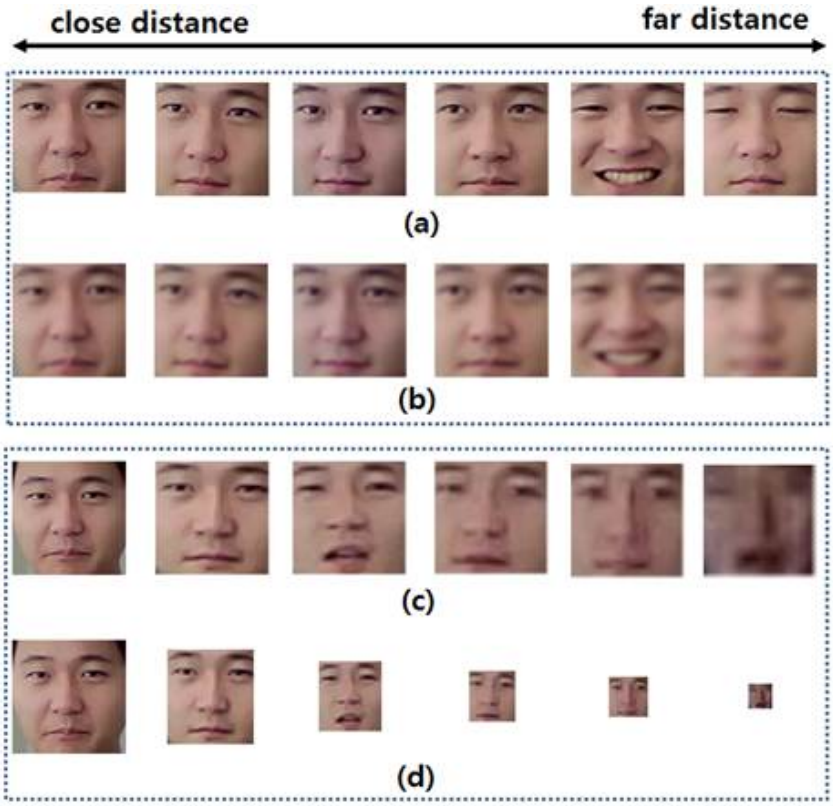


그림 30. 가상의 거리별 얼굴 영상과 실제 거리별 얼굴 영상의 비교

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

에서 추출된 6장의 영상에서 추출된 얼굴 영상이고 그림 30(b)는 자동 생성 방법을 통해 생성된 가상의 거리별 얼굴 영상이다. 그림 30(d)는 실제 다양한 거리에서 취득된 거리별 얼굴 영상이고, 그림 30(c)는 각 거리별 얼굴 영상을 1m의 기준 얼굴 영상 크기에 맞게 정규화된 영상이다. 실험 결과, 1m의 단일 거리의 얼굴 영상보다 자동 생성된 가상의 거리별 얼굴 영상이 실제 거리별 얼굴 영상과 유사함을 확인할 수 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제4장 실험 조건별 원거리 얼굴 인식 성능 분석

본 장에서는 제안하는 원거리 얼굴 인식 방법을 평가하기 위해 학습 영상의 구성 조건, 얼굴 크기 정규화 방법 및 얼굴 밝기 정규화 방법의 변화에 따른 실험을 진행하고 그 결과를 분석한다. 또한, 실제 카메라의 해상도와 얼굴 인식 조건에 따라 변화하는 인식률의 분석을 통해 상황별 최적의 원거리 얼굴 인식 방법을 제안한다.

제1절 실험 방법

본 논문에서 얼굴 인식은 1:1 인증이 아닌 1:N 검색 방법으로써 DB에 저장된 얼굴 영상 중 가장 유사한 첫 번째 얼굴 영상을 검증 영상에 대한 결과로 분류하는 방법을 사용한다. 또한, 본 실험에서는 거리와 관계없이 입력 영상에서 얼굴이 모두 검출된다는 가정으로 직접 얼굴 영역을 추출해 DB를 구성한다. 수동으로 얼굴을 추출하는 경우 자동 얼굴 검출 방법보다 정교하게 얼굴 영역이 추출되는 특징이 있다. 또한, 본 실험에서는 추출된 얼굴 영상의 틀어짐이나 회전은 고려하지 않고 원본 영상 그대로 사용한다.

1. 얼굴 DB

전통적으로 얼굴 인식 실험에는 AT&T DB, Yale DB, MIT Face DB, FERET DB 등과 같이 주로 사용되는 얼굴 DB가 있다. 기존 얼굴 DB의 특징을 간략하게 소개하면, AT&T DB는 후보들의 표정에 제한을 두지 않고 약간의 측면 움직임과

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

머리 회전을 허용한다[77]. 모든 후보들은 다른 시간대에서 다른 조면 조건들을 갖고 촬영되었고 총 후보의 수는 40명이다. Yale DB는 15명의 후보들에 대하여 각각 11개의 다양한 조건으로 만들어진 조건이다[45][78]. 조명과 표정 변화를 허용하고 조명은 좌·우, 중앙·좌, 우·좌 변화를 하고 표정은 무표정, 행복함, 슬픔, 졸림, 놀람, 윙크 등을 포함한다. MIT Face DB는 10명의 후보들로 구성된 얼굴 DB로 앉아 있는 상태에서 다양한 각도로 머리 방향이 변화하는 얼굴 영상을 획득했다. 이때, 조명과 카메라의 줌도 함께 변화하는 특징이 있다[79]. FERET DB는 1993년부터 1996년까지 다수에 걸쳐 획득된 얼굴 DB로써 1199명으로부터 14,126장의 영상을 획득했다. 동일한 인물에 한해서 적게는 여러 날부터 길게는 2년의 시간을 거쳐 획득한 특징이 있고 조명 및 얼굴 변화를 포함하고 있다[80].

얼굴 인식에 대표적인 DB들을 살펴보면 외부 변화로써 조명, 얼굴 틀어짐, 얼굴 표정 변화 등을 포함하고 있지만 거리 변화에 따른 얼굴 변화를 고려한 DB는 존재하지 않는다. 그러므로 본 논문에서는 얼굴 인식 실험을 위한 거리별 얼굴 DB를 직접 구성하여 사용한다.

가. ETRI DB

ETRI 얼굴 DB는 u-로봇 테스트 베드 환경에서 10명의 후보에 대한 얼굴 영상으로 구성되어 있다. 각 후보들은 조명 변화와 거리 변화를 포함하고 있다. 조명 변화는 실내조명을 이용해 연출했고, 거리는 1m~5m까지 변화한다. 후보 1인당 검증 영상은 거리 당 30장씩 총 150장이고, 전체 후보에 대한 검증 영상은 1500장이다. 이때, 학습용 영상은 전체 검증 영상에 포함되지 않는다. 표 3은 실험에 사용된 ETRI DB의 구성을 나타내고, 그림 31은 후보 1인에 대해 1m~5m까지의 실제 거리별 얼굴 영상으로 1m 얼굴 크기를 기준으로 정규화된 영상이다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 3. ETRI 얼굴 DB

사용 DB	ETRI DB
취득 환경	• 실내 환경 – 조명 변화 – 거리 변화 – 얼굴 변화
취득 카메라	–
취득 방법	• 동영상을 통한 프레임 분할
영상 해상도	• 320(W) × 240(H)
취득 거리	• 1m~5m (1m 간격)
총 후보 수	• 10명
1인당 얼굴 수 (평균 얼굴 크기)	• 1m : 30장 (50×50) • 2m : 30장 (25×25) • 3m : 30장 (16×16) • 4m : 30장 (12×12) • 5m : 30장 (10×10)

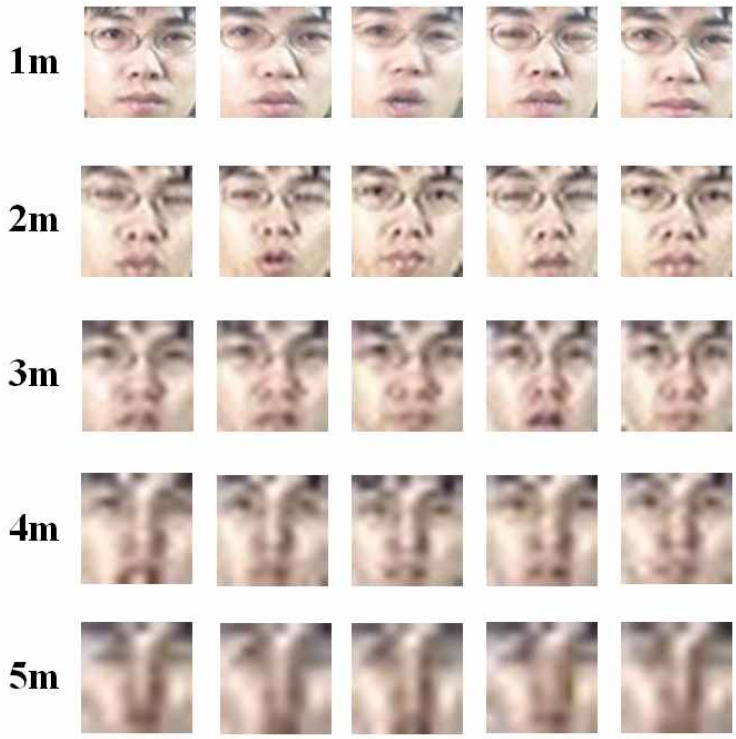


그림 31. ETRI DB의 예

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

나. IPES DB

IPES 얼굴 DB는 실내 환경에서 촬영된 얼굴 영상으로 IPES-640과 IPES-1280로 나뉜다. 표 4는 IPES DB의 구성을 나타낸다. IPES-640은 저해상도 기반 얼굴 인식 실험을 위해 640×480 해상도로 취득됐고 각 얼굴 영상은 동영상을 통한 프레임 분할 방식으로 취득됐다. 각 영상들은 거리 변화와 조명 변화를 포함하고 있고 거리는 1m~5m까지 변화한다. 조명 변화는 거리에 따라 달라지는 실내조명 변화를 포함한다. 후보 1인당 검증 영상은 거리 당 30장씩 총 150장이고, 전체 후보에 대한 총 검증 영상의 수는 1500장이다. 이때, 학습용 영상은 전

표 4. IPES 얼굴 DB

사용 DB	IPES-640 DB	IPES-1280 DB
취득 환경	• 실내 환경 - 얼굴 변화 - 거리 변화 - 거리에 의한 조명 변화	• 실내 환경 - 얼굴 변화 - 거리 변화 - 거리에 의한 조명 변화
취득 카메라	• HMX-R10B(캠코더)	• IPCAM-1000(IP 카메라)
취득 방법	• 동영상을 통한 프레임 분할 (MPEG-4 AVC1 -> JPEG)	• 동영상을 통한 프레임 분할 (H.264 -> JPEG)
영상 해상도	• 640(W) × 480(H)	• 1280(W) × 720(H)
취득 거리	• 1m~5m (1m 간격)	• 1m~9m (1m 간격)
총 후보 수	• 10명	• 12명
1인당 얼굴 수 (평균 얼굴 크기)	• 1m : 30장 (70×70) • 2m : 30장 (45×45) • 3m : 30장 (30×30) • 4m : 30장 (22×22) • 5m : 30장 (18×18)	• 1m : 30장 (120×120) • 2m : 30장 (70×70) • 3m : 30장 (50×50) • 4m : 30장 (40×40) • 5m : 30장 (30×30) • 6m : 30장 (25×25) • 7m : 30장 (22×22) • 8m : 30장 (19×19) • 9m : 30장 (16×16)
특이사항	• 광학식 줌을 통해 별도의 학습 데이터 추가 취득	-

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

체 검증 영상에 포함되지 않는다. 또한, IPES-640 DB는 광학식 줌 카메라를 통해 별도의 학습용 얼굴 영상을 추가적으로 취득한다. 이 영상은 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 학습으로 사용한 얼굴 인식 실험에 학습 데이터로 사용된다. 그림 32는 후보 1인에 대해 1m~5m까지 실제 거리별 얼굴 영상으로 1m 얼굴 크기를 기준으로 정규화된 영상이다.

IPES-1280은 고해상도 기반 얼굴 인식 실험을 위해 1280×720 해상도로 취득됐고 각 얼굴 영상은 동영상을 통한 프레임 분할 방식으로 취득됐다. 각 영상들은 거리 변화와 조명 변화를 포함하고 있다. 거리는 1m~9m까지 변화하고 조명은 거리에 따라 달라지는 실내조명 변화를 포함한다. 후보 1인당 검증 영상은 거리 당

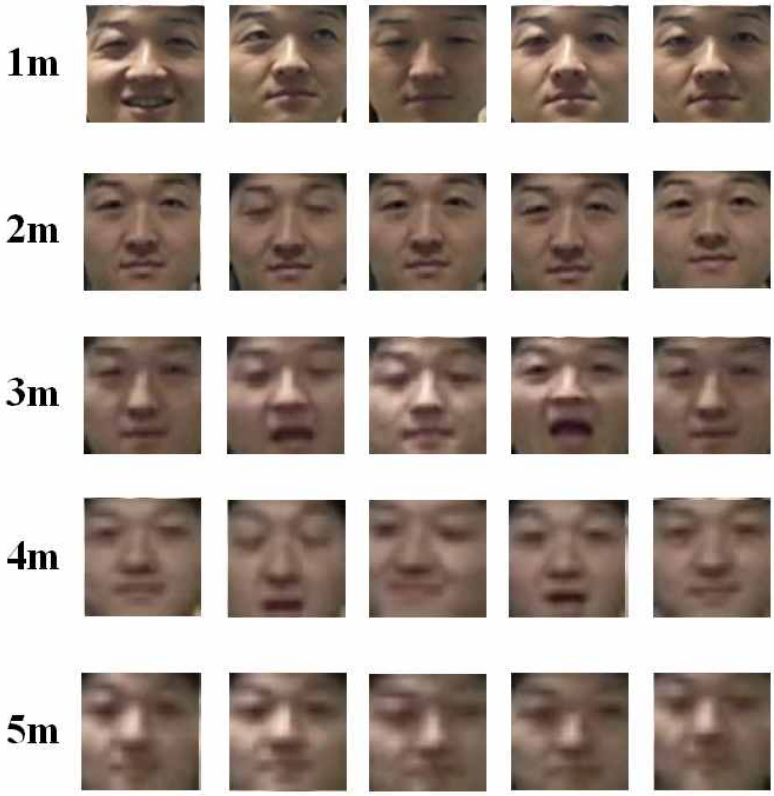


그림 32. IPES-640 DB의 예

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

30장씩 총 270장이고, 전체 후보에 대한 총 검증 영상의 수는 3,240장이다. 이때, 학습용 영상은 전체 검증 영상에 포함되지 않는다. 그림 33은 후보 1인에 대해 1m~9m까지 실제 거리별 얼굴 영상으로 1m 얼굴 크기를 기준으로 정규화된 영상이다.

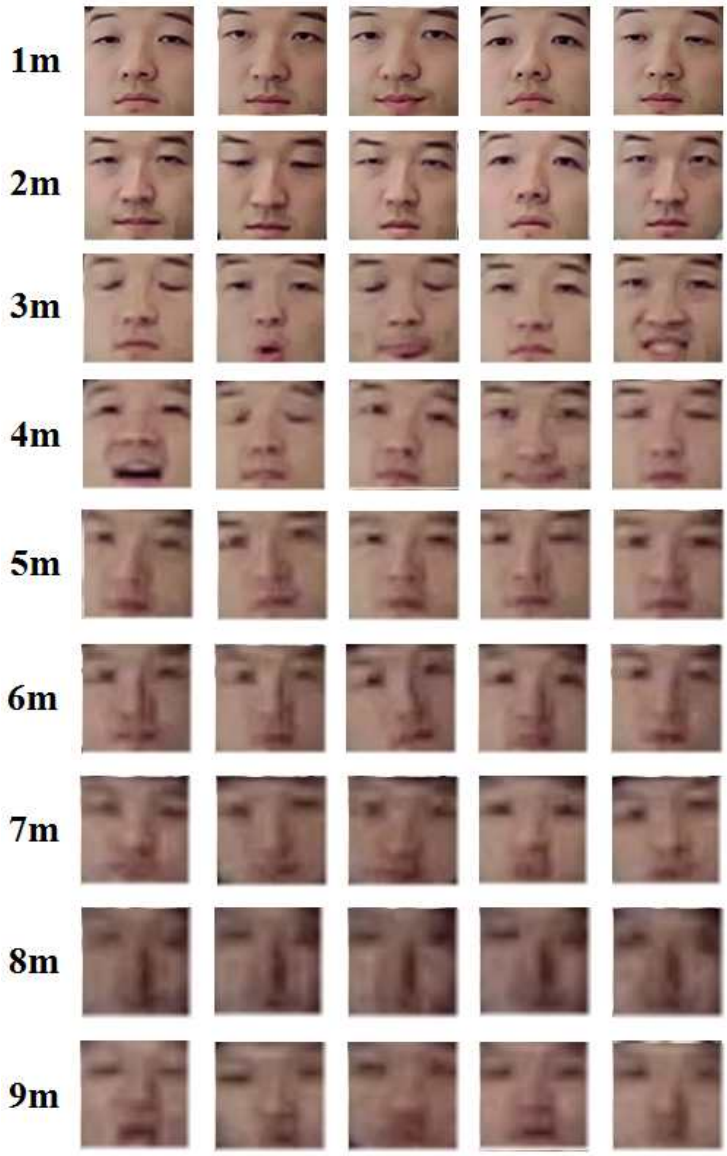


그림 33. IPES-1280 DB의 예

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

2. 성능 평가방법

본 논문에서 원거리 얼굴 인식 시스템의 인식률을 확인하기 위해서 특징행렬 사이의 유사도를 비교하여 인식률을 계산하는 방법을 사용한다. 두 개의 특징행렬 $X = (x_{ij})_{p \times q}$와 $Y = (y_{ij})_{p \times q}$가 있을 때 두 특징행렬 사이의 유사도는 다음 식 (33)을 통해 계산된다.

$$d(X, Y) = \sum_{j=1}^{q} \sqrt{\sum_{i=1}^{p} (x_{ij} - y_{ij})^2} \tag{33}$$

임의의 검증 영상 A_t를 특징행렬 X_t로 변환한 후 모든 훈련 데이터의 특징행렬과의 유사도를 측정한다. 그리고 식 (34)와 같이 최소의 유사도를 갖는 특징행렬 X_t를 찾고 이 특징행렬 X_t가 속한 클래스를 검증 영상의 최종 클래스로 결정한다. 최종 얼굴 인식률은 식 (35)와 같이 전체 검증 영상 수 T에 대하여 정확히 클래스가 검색된 검증 영상 수 T_s의 비율로 결정한다. 여기서, N은 전체 학습 영상의 수이다.

$$\min(d(X_t, X_k)), \quad k = 1, 2, \cdots N \tag{34}$$

$$recognition\ rate = \frac{T_s}{T} \times 100\,\% \tag{35}$$

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

제2절 거리별 얼굴 인식률 분석

본 절에서는 다양한 해상도의 거리별 얼굴 DB에서 제안하는 원거리 얼굴 인식 방법과 기존 원거리 얼굴 인식 방법의 성능을 비교한다. 그리고 제안하는 방법의 우수성을 증명하기 위해 근거리 얼굴 영상의 화질 열화, 얼굴 크기 정규화 및 얼굴 밝기 정규화 방법에 따른 얼굴 인식률을 분석한다.

1. 거리별 얼굴 영상을 이용한 원거리 얼굴 인식

본 실험에서는 다양한 해상도의 거리별 얼굴 DB에서 기존 원거리 얼굴 인식 방법과 거리별 얼굴 영상을 학습으로 이용한 방법의 비교를 통해 제안하는 원거리 얼굴 인식 방법의 우수성을 증명한다. 실험에 사용된 영상의 해상도는 QVGA, VGA, SXGA 급으로 각각 ETRI DB, IPES-640 DB, IPES-1280 DB를 사용한다.

실험에 사용된 기존 얼굴 인식 방법으로는 단일 거리 얼굴 영상을 이용한 방법, 구조적 특징을 이용한 저해상도 얼굴 인식 방법이다. 단일 거리 얼굴 영상을 이용한 방법은 얼굴의 다중 영역을 사용하고, 얼굴 인식 방법은 각각 PCA와 LDA를 사용한다. 구조적 특징을 이용한 방법은 CMs[32]와 CLPMs[33]을 사용한다. CLPMs 방법은 FERET DB[80]를 이용한 저해상도 얼굴 인식 방법에서 평균 90.1%로 가장 우수한 성능을 나타낸 방법이다[30]. 제안하는 원거리 얼굴 인식 방법은 실제 거리별 얼굴 영상을 이용한 방법, 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 이용한 방법, 자동 생성 방법을 통해 획득된 가상의 거리별 얼굴 영상을 이용한 방법을 사용한다. 얼굴 인식 방법은 각각 PCA와 LDA를 사용한다. 제안하는 원거리 얼굴 인식 방법은 1m~5m의 거리별 얼굴 크기 정규화를 위해 양

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

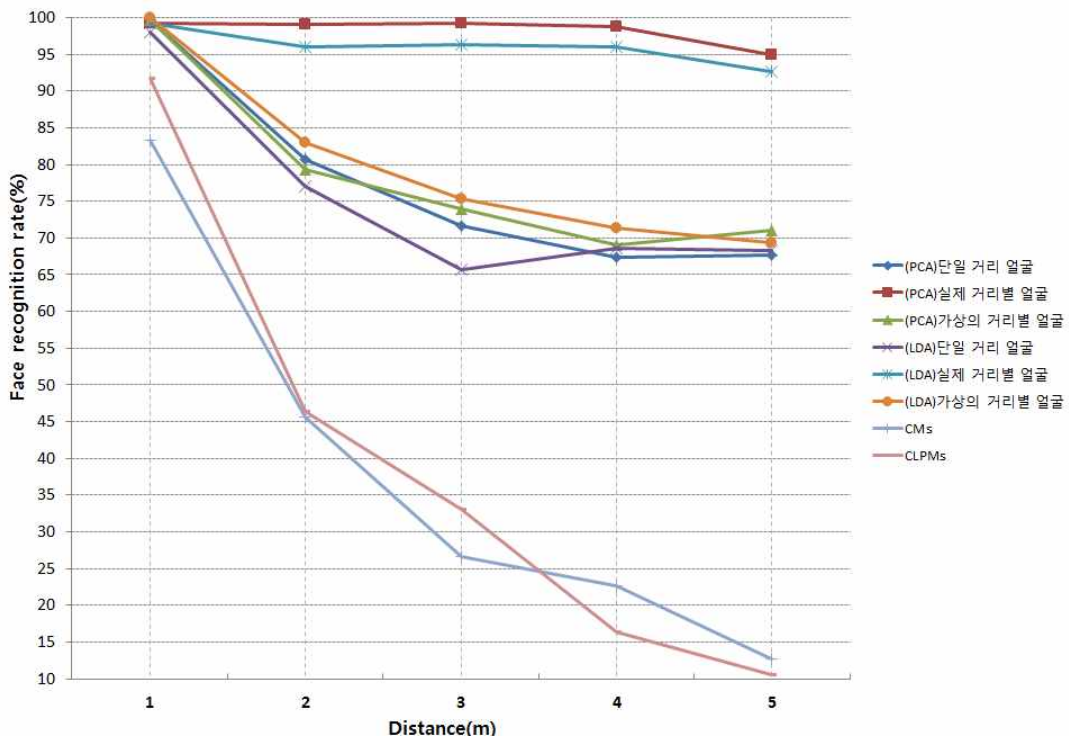


그림 34. QVGA급 영상에서의 원거리 얼굴 인식률 비교

선형 보간법을 사용하고 얼굴 밝기 정규화에는 히스토그램 평활화 방법을 사용한다. 이때, 얼굴 인식 방법에 관계없이 1인당 학습 영상의 수는 모두 동일하게 5장을 사용하고, 검증 영상의 수는 각 거리별로 30장씩 사용한다. 특징벡터의 차원 수는 모두 동일하게 50개로 한다.

그림 34는 QVGA급 영상에서 얼굴 인식 방법에 따른 거리별 얼굴 인식률을 나타낸다. 실험 결과, 기존 방법과 제안하는 방법이 1m에서는 모두 우수한 성능을 나타냈다. 단일 거리 얼굴 영상을 학습으로 사용한 방법과 기존 CMs, CLPMs 방법이 원거리로 갈수록 얼굴 인식률이 급격히 떨어지는 성능을 나타내는 반면, 제안하는 실제 거리별 얼굴 영상을 이용한 방법은 원거리에도 인식률이 높음을 확인할 수 있다. 전체 거리의 평균 얼굴 인식률을 비교했을 때는 제안하는 실제 거리별 얼굴

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

영상을 학습으로 사용한 PCA 기반 얼굴 인식 방법이 평균 98.3%로 가장 우수한 성능을 나타냈다. 제안하는 방법이 검증 단계에 입력되는 원거리 얼굴 영상의 특징과 가장 유사한 특징을 획득하기 위해 실제 거리별 얼굴 영상을 기반으로 추출한 반면 기존 방법은 가상의 특징 정보를 생성해 사용했기 때문이다. 즉, 제안하는 실제 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식 방법은 기존 QVGA급 저해상도 기반 CCTV 환경에서 우수한 성능을 나타냄을 증명했다.

그림 35는 VGA급 영상에서 얼굴 인식 방법에 따른 거리별 얼굴 인식률을 나타낸다. 실험에 사용된 기존 얼굴 인식 방법은 이전과 동일하고 제안하는 방법은 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 학습으로 사용한 방법을 추가한다. 실험 결과, 기존 단일 거리 얼굴 영상을 학습으로 사용한 방법이 1m에서 우수한 성능을 나타냈지만 원거리로 갈수록 인식률이 저하되는 특징을 보였다. 반면, 제

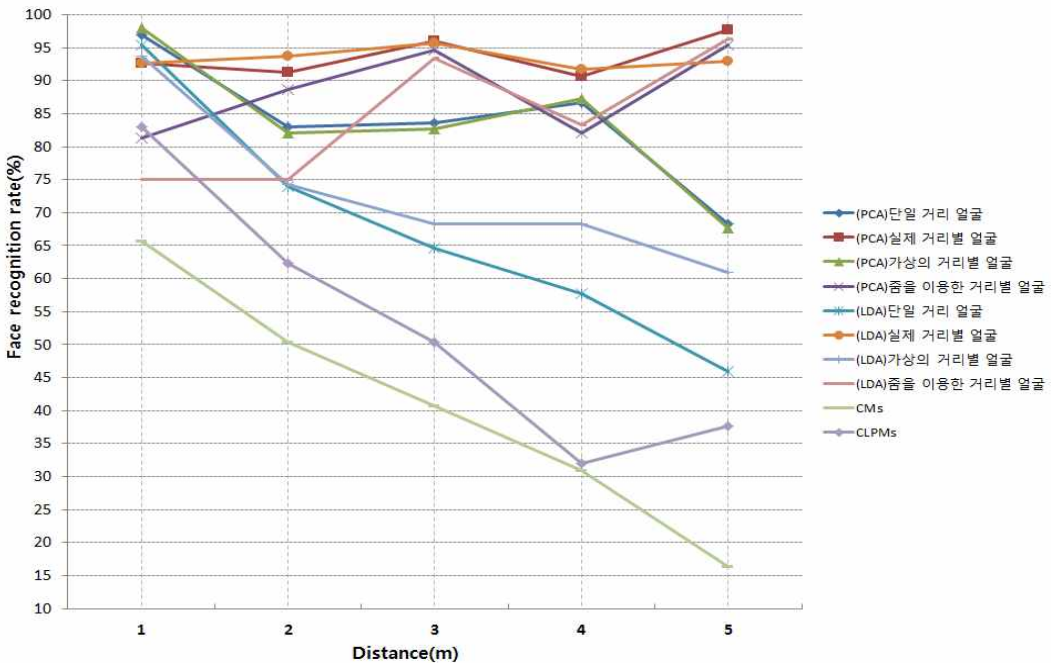


그림 35. VGA급 영상에서의 원거리 얼굴 인식률 비교

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

안하는 줌 카메라를 통해 획득된 가상의 거리별 얼굴 영상을 학습으로 이용한 방법은 원거리에서 우수한 성능을 나타냈지만, 근거리로 갈수록 인식률이 떨어지는 특징을 나타냈다. 이는, 줌 카메라를 통해 획득된 원거리 얼굴 영상의 특징이 실제 원거리 얼굴 영상의 특징과 매우 흡사하지만 줌을 통해 획득된 가상의 근거리 얼굴 영상은 실제 근거리 얼굴 영상과 차이가 발생하기 때문이다. 구조적 특징을 이용한 방법은 원거리로 갈수록 인식률 급격히 감소하는 성능을 나타냈다. 전체 평균 인식률을 비교했을 때는 실제 거리별 얼굴 영상을 학습으로 사용한 PCA 기반 얼굴 인식 방법이 평균 93.7%로 가장 우수한 성능을 나타냈다. 즉, 제안하는 실제 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식 방법은 기존 VGA급 저해상도 기반 CCTV 환경에서도 우수한 성능을 나타냄을 증명했다.

그림 36은 SXGA급 영상에서 얼굴 인식 방법에 따른 거리별 얼굴 인식률을 나타낸다. QVGA급과 VGA급 해상도에서 추출된 얼굴이 1m~5m의 거리로 구성되었다면 SXGA급 해상도는 거리별 얼굴 영상이 1m~9m까지 구성된 특징이 있다. 즉,

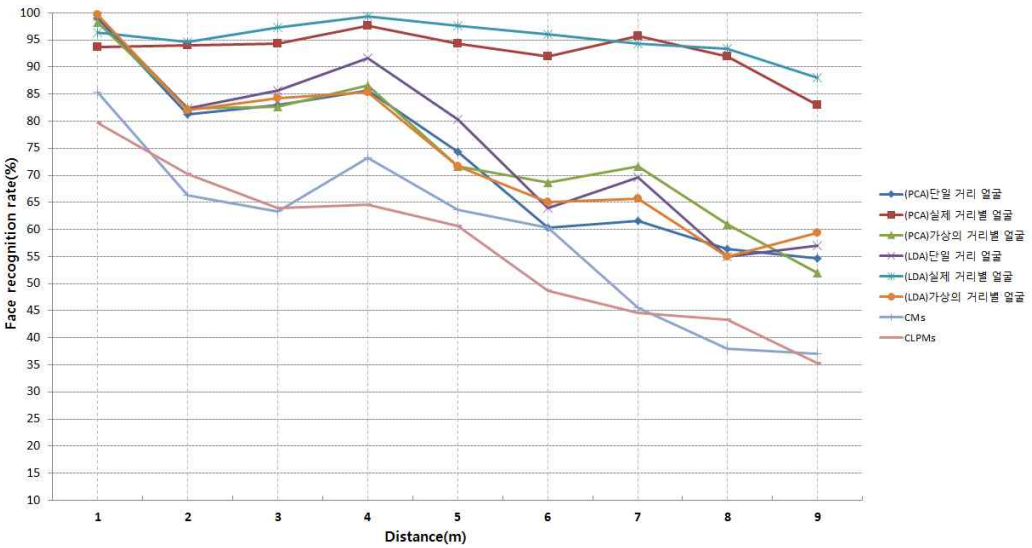


그림 36. SXGA급 영상에서의 원거리 얼굴 인식률 비교

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

원본 영상의 해상도가 높아질수록 원거리 얼굴 인식이 가능한 거리가 늘어나게 된다. 실험에 사용된 얼굴 인식 방법은 이전과 동일하다. 실험 결과, 모든 얼굴 인식 방법이 거리가 멀어질수록 인식률이 저하되는 특징이 있지만, 제안하는 실제 거리별 얼굴 영상을 이용한 방법은 원거리에서도 비교적 높은 인식률을 나타내고 있음을 확인할 수 있다. 전체 평균 인식률을 비교했을 때, 실제 거리별 얼굴 영상을 학습으로 사용한 PCA 기반 얼굴 인식 방법이 평균 92.9%, LDA 기반 얼굴 인식 방법이 평균 95.2%로 LDA를 사용했을 가장 우수한 성능을 나타냈다. PCA 기반 얼굴 인식을 사용한 경우 QVGA급과 VGA급 영상에서 우수한 성능을 나타냈다. 반면, SXGA급 영상에서 LDA 기반 얼굴 인식이 우수한 성능을 나타낸 이유는 SXGA급 영상은 거리별 얼굴이 1m~9m로 늘어난 만큼 학습에 사용된 얼굴 영상의 수가 증가했기 때문이다. 본 실험을 통해서 제안하는 실제 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식 방법은 기존 SXGA급 고해상도 기반 CCTV 환경에서도 우수한 성능을 나타냄을 증명했다.

2. 저역 통과 필터를 이용한 근거리 얼굴 영상의 화질 열화

원거리 얼굴 인식은 근거리에서 획득되는 고해상도 얼굴 영상과 원거리에서 획득되는 저해상도 얼굴 영상간의 차이 때문에 인식률이 떨어진다. 이와 같은 문제를 해결하기 위해 기존 원거리 얼굴 인식 방법은 저해상도 원거리 얼굴 영상을 고해상도로 복원해 최대한 근거리 얼굴 영상과 유사하게 만드는 방법을 사용했다.

본 논문에서는 기존 방법과 반대로 근거리에서 추출된 얼굴 영상에 저역 통과 필터링을 통해 고주파 성분을 제거함으로써 원거리에서 추출된 얼굴과 유사하게 만드는 화질 열화 방법을 제안한다. 제안하는 방법은 근거리 고해상도 영상의 화질 열화를 위해 가우시안 저역 통과 필터링을 사용한다. 표 5는 화질 열화 방법에 따라

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

표 5. 화질 열화에 따른 얼굴 인식 실험 조건

구분	실험 조건
Case 1	1 인당 학습 영상 수 – 1m : 5장 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	학습 영상 구성 : **실제 거리별 얼굴 영상**
Case 2	1 인당 학습 영상 수 – 1m : 5장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	학습 영상 구성 : **자동 생성된 가상의 거리별 얼굴 영상**
▪ 얼굴 DB : ETRI DB ▪ 얼굴 크기 정규화 : 양선형 보간법 ▪ 얼굴 밝기 정규화 : 히스토그램 평활화 ▪ 얼굴 인식 : PCA, LDA	

변호하는 얼굴 인식률을 분석하기 위한 실험 조건이다. 원거리 얼굴 인식 방법으로는 제안하는 실제 거리별 얼굴 영상을 이용한 방법과 자동 생성된 가상의 거리별 얼굴 영상을 이용한다. 얼굴 DB는 ETRI DB를 사용하고 얼굴 크기 정규화에는 양선형 보간법을 사용한다. 얼굴 밝기 정규화에는 히스토그램 평활화를 사용한다. 1인당 학습 영상의 수는 모두 동일하게 5장을 사용하고, 검증 영상의 수는 각 거리별로 30장씩 총 150장을 사용한다.

그림 37은 화질 열화에 따른 전체 거리 평균 인식률을 나타낸다. 실험 결과, LDA보다 PCA를 사용했을 때 전체 평균 10.6%로 우수한 성능을 나타냈다. 저역통과 필터링을 통한 화질 열화 방법에 따른 인식률을 비교했을 때는 PCA와 LDA 기반 얼굴 인식에서 각각 0.1%와 2.3% 향상된 성능을 나타냈다. 특히, 제안하는 방법은 가상의 거리별 얼굴 영상을 생성하는 방법에서 우수한 성능을 나타냈다. 즉, 가상의 거리별 얼굴 영상을 생성해 학습에 이용하는 상황에서는 제안하는 화질 열화 방법을 사용해 얼굴 영상을 생성하는 것이 효과적임을 확인했다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

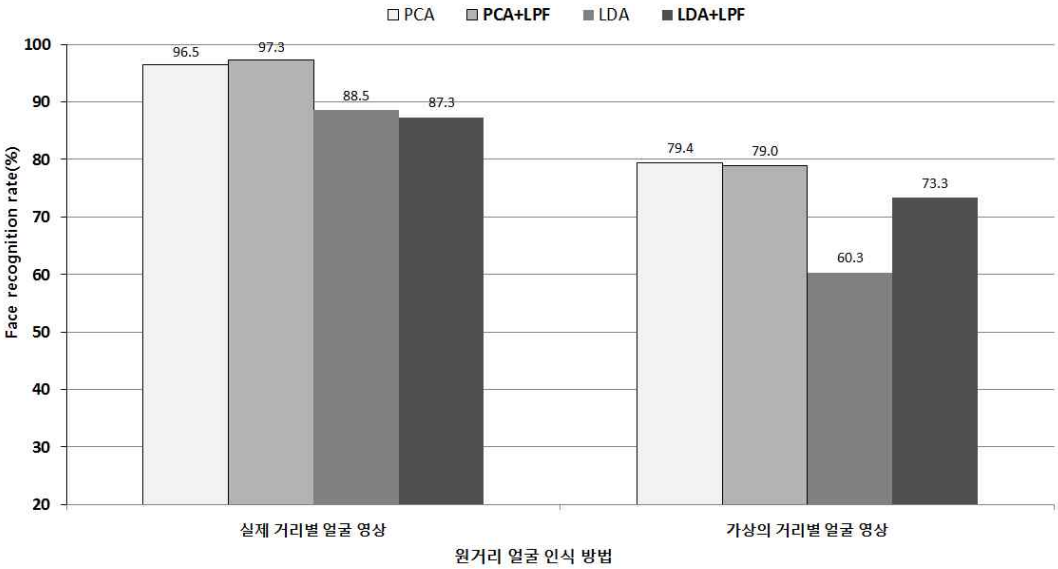


그림 37. 화질 열화에 따른 전체 거리 평균 인식률

3. 양선형 보간법을 이용한 원거리 얼굴 크기 정규화

본 실험에서는 얼굴 크기 정규화 방법에 따른 원거리 얼굴 인식 성능을 비교해 양선형 보간법을 이용한 원거리 얼굴 크기 정규화 방법의 우수성을 증명한다. 실험을 위한 얼굴 DB는 ETRI DB를 사용하고, 얼굴 인식은 실제 거리별 얼굴 영상을 학습으로 사용한 PCA와 LDA 기반 얼굴 인식 방법을 이용한다. 얼굴 밝기 정규화에는 히스토그램 평활화 방법을 사용한다. 특징벡터의 차원 수는 모두 동일하게 50개로 한다. 표 6은 얼굴 크기 정규화 방법에 따른 얼굴 인식 성능을 분석하기 위한 실험 조건이다. 이때, 학습 영상의 수는 모두 동일하게 5장을 사용했다. 검증 영상은 모두 동일하게 1m~5m의 거리별 얼굴 영상을 사용했고, 각 거리마다 30장씩 총 150장을 사용했다.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

표 6. 얼굴 크기 정규화 방법에 따른 얼굴 인식 실험 조건

구분	실험 조건
Case 1	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 크기 정규화 : **최근접 이웃 보간법**
Case 2	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 크기 정규화 : **양선형 보간법**
Case 3	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 크기 정규화 : **양3차회선 보간법**
Case 4	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 크기 정규화 : **이중 선형-3차회선 보간법**
▪ 얼굴 DB : ETRI DB ▪ 얼굴 밝기 정규화 : 히스토그램 평활화 ▪ 화질 열화 : 가우시안 저역 통과 필터링 ▪ 얼굴 인식 : PCA, LDA	

그림 38과 39는 얼굴 크기 정규화 방법에 따른 거리별 얼굴 인식률 변화를 나타낸다. 실험 결과, PCA 기반 얼굴 인식에서는 얼굴 크기 정규화에 최근접 이웃, 양선형, 양3차회선, 이중 선형-3차회선 보간법을 썼을 때, 각각 97.3%, 98.3%, 98.0%, 98.2%로 인식률의 변화가 크게 발생하지 않음을 확인할 수 있다. LDA 기반 얼굴 인식에서는 최근접 이웃 보간법을 사용 했을 때, 91.3%로 가장 낮은 성능을 나타냈고 양선형 보간법을 사용했을 때, 96.1%로 가장 우수한 성능을 나타냄을 확인할 수 있다. 결과적으로 실제 거리별 얼굴 영상을 학습으로 사용해 원거리 얼굴 인식을 수행할 시 양선형 보간법을 사용해 얼굴 크기를 정규화하는 것이 가장 적절함을 확인할 수 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

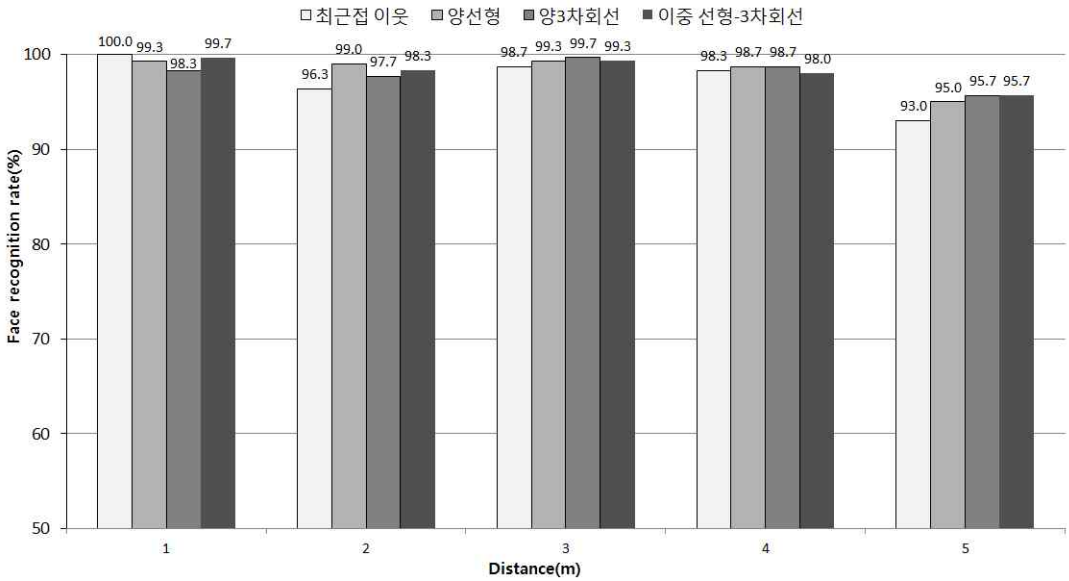


그림 38. 얼굴 크기 정규화 방법에 따른 PCA 기반 원거리 얼굴 인식률

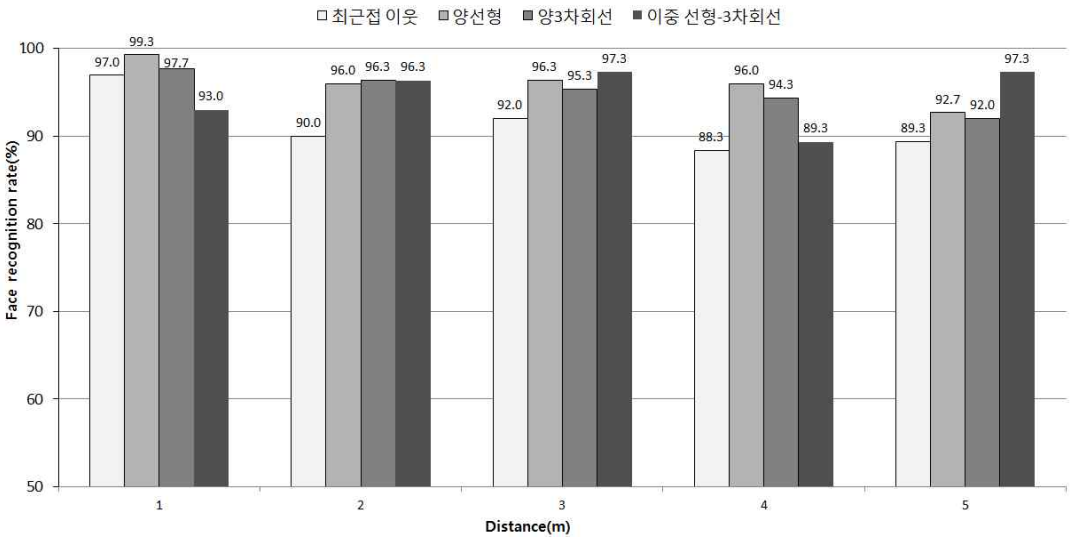


그림 39. 얼굴 크기 정규화 방법에 따른 LDA 기반 원거리 얼굴 인식률

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

4. 히스토그램 평활화를 이용한 원거리 얼굴 밝기 정규화

본 실험에서는 얼굴 발기 정규화 방법에 따른 원거리 얼굴 인식 성능을 비교해 히스토그램 평활화를 이용한 방법의 우수성을 증명한다. 실험을 위한 얼굴 DB는 ETRI DB를 사용하고, 얼굴 인식은 실제 거리별 얼굴 영상을 학습으로 사용한 PCA와 LDA 기반 얼굴 인식 방법을 이용한다. 얼굴 크기 정규화에는 양선형 보간법을 사용하고 특징벡터의 차원 수는 모두 동일하게 50개로 한다. 표 7은 얼굴 밝기 정규화 방법에 따른 얼굴 인식률의 변화를 분석하기 위한 실험 방법이다. 이때, 학습 영상의 수는 모두 동일하게 5장을 사용한다. 검증 영상은 모두 동일하게 1m~5m의 거리별 얼굴 영상을 사용하고, 각 거리마다 30장씩 총 150장을 사용한다. 그림 40과 41은 얼굴 밝기 정규화에 따른 원거리 얼굴 인식률의 변화를 나타낸다. 실험 결과, PCA와 LDA에 모두에서 히스토그램 평활화를 사용했을 때 각각 98.3%와 96.1%로 가장 우수한 성능을 나타냈다. 특히, 학습 영상의 수가 부족한

표 7. 얼굴 밝기 정규화 방법에 따른 얼굴 인식 실험 조건

구분	실험 조건
Case 1	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 밝기 정규화 : 히스토그램 평활화
Case 2	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 밝기 정규화 : MCT
Case 3	1 인당 학습 영상 수 – 1m~5m : 각 1장 (총 5) 1 인당 검증 영상 수 – 1m~5m : 각 30장 (총 150)
	얼굴 밝기 정규화 : LBP
▪ 얼굴 DB : ETRI DB ▪ 얼굴 크기 정규화 : 양선형 보간법 ▪ 화질 열화 : 가우시안 저역 통과 필터링 ▪ 얼굴 인식 : PCA, LDA	

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

상황에서 LDA 기반 원거리 얼굴 인식은 MCT와 LBP를 밝기 정규화로 사용할 경우 거리가 멀어질수록 얼굴 인식률이 급격히 떨어지는 특징을 보였다. 결과적으로

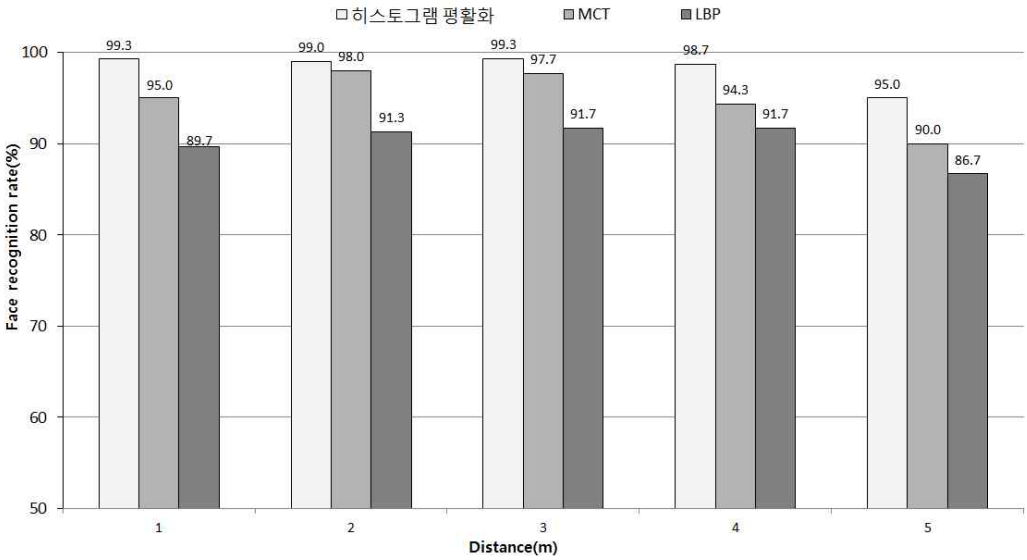


그림 40. 얼굴 밝기 정규화 방법에 따른 PCA 기반 원거리 얼굴 인식률

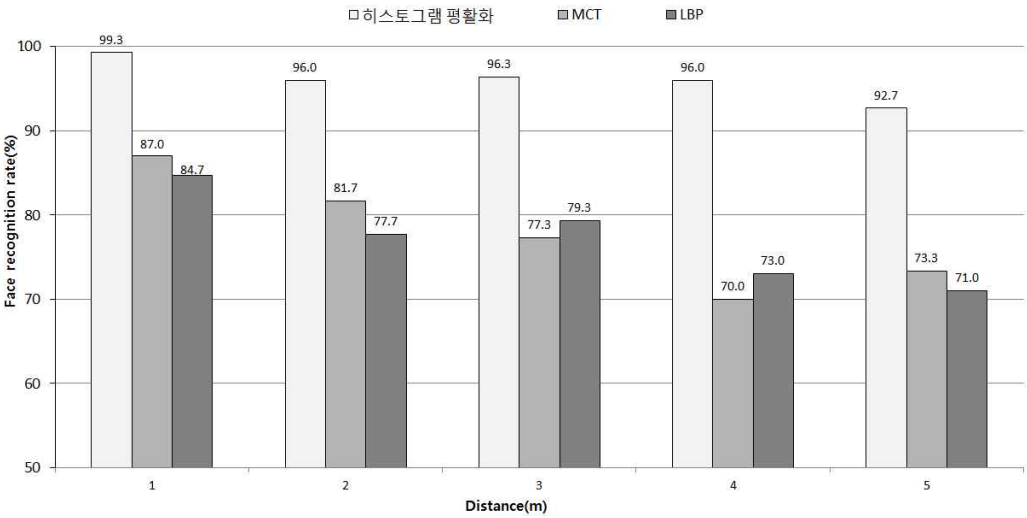


그림 41. 얼굴 밝기 정규화 방법에 따른 LDA 기반 원거리 얼굴 인식률

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

실제 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식에서는 얼굴 밝기 정규화로 히스토그램 평활화를 사용하는 것이 가장 적합함을 확인했다. 기존의 밝기 변화에 강인한 MCT나 LBP의 경우를 살펴보면 얼굴 밝기가 눈에 띄게 변하는 경우에 강인한 성능을 나타냈다. 본 실험에 사용된 DB는 조명 변화는 허용하지만 비교적 밝기의 변화가 심하지 않았다. 그렇기 때문에 MCT나 LBP를 이용한 방법에 비해 히스토그램 평활화를 이용한 방법이 우수한 성능을 나타냈다.

제3절 기준 영상 크기에 따른 얼굴 인식률과 수행 시간 관계

현재 설치된 CCTV는 VGA급 40만 화소 내외의 카메라가 보편적으로 사용되고 있는데, 최근 고해상도 CCTV에 대한 요구가 증가하면서 130만 화소 급의 CCTV의 사용이 증가하고 있는 추세이다. 일반적으로 영상 처리 기술들의 처리 속도는 알고리즘의 복잡도도 중요하지만 원본 영상의 해상도에 따라 처리 속도가 크게 변화한다. 즉, 40만 화소에서 원활히 수행되던 얼굴 인식 기술도 100만 화소 이상이 되면 실시간 수행을 장담할 수 없게 된다. 본 절에서는 SXGA급 100만 화소 해상도에서 기준 영상의 크기에 따른 인식률의 변화를 분석해 고성능과 고속 처리에 적합한 원거리 얼굴 인식 조건을 분석한다.

표 8은 동일한 조건에서 기준 영상의 크기만 변화 시켰을 때, 인식률과 처리 시간을 분석하기 위한 실험 조건이다. 실험을 위해 얼굴 DB는 IPES-1280 DB를 사용하고, 얼굴 인식은 단일 거리 얼굴 영상을 학습으로 사용한 방법으로 학습 영상은 1m에서 획득한 얼굴 영상 3장만을 사용한다. 이때, 기준 영상의 크기는 1m, 3m, 5m, 7m, 9m로 정의하고 각각 120×120, 50×50, 30×30, 22×22, 16×16으로 정규화한다. 검증 영상은 1m~9m별로 각 30장씩이다. 얼굴 크기 정규

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

표 8. 기준 영상 크기에 따른 얼굴 인식 실험 조건

구분	실험 조건
Case 1	1 인당 학습 영상 수 – 1m : 3장, 기준 영상크기 :120*120 1 인당 검증 영상 수 – 1m~9m : 각 30장
Case 2	1 인당 학습 영상 수 – 1m : 3장 기준 영상크기 :50*50 1 인당 검증 영상 수 – 1m~9m : 각 30장
Case 3	1 인당 학습 영상 수 – 1m : 3장 기준 영상크기 :30*30 1 인당 검증 영상 수 – 1m~9m : 각 30장
Case 4	1 인당 학습 영상 수 – 1m : 3장 기준 영상크기 :22*22 1 인당 검증 영상 수 – 1m~9m : 각 30장
Case 5	1 인당 학습 영상 수 – 1m : 3장 기준 영상크기 :16*16 1 인당 검증 영상 수 – 1m~9m : 각 30장
▪ 얼굴 DB : IPES-1280 DB ▪ 얼굴 크기 정규화 : 양선형 보간법 ▪ 얼굴 인식 : PCA	

화 방법으로는 양선형 보간법을 사용하고, 영상 밝기 정규화 방법은 히스토그램 평활화를 사용한다. 특징벡터의 차원 수는 모두 동일하게 50개로 한다.

그림 42는 기준 영상 크기에 따른 PCA 기반 원거리 얼굴 인식률을 나타낸다. 실험 결과, 기준 영상의 크기를 16*16으로 했을 때 전체 평균 인식률이 79.7%로 가장 우수한 성능을 나타냈고, 120*120으로 했을 때, 76.5%로 가장 낮은 성능을 보였다. 그림 43은 얼굴 인식 알고리즘의 주요 기능들에 대한 수행 시간을 나타낸다. PCA 처리 구간의 수행 시간을 비교했을 때, 기준 영상 120*120과 기준 영상 16*16의 경우 약 10배 이상의 시간을 줄일 수 있다. 그 밖에도, 전체적으로 기준 영상의 크기가 줄어들수록 수행 시간이 줄어듦을 확인할 수 있다. 즉, 원거리 얼굴 인식에서 동일한 특징벡터의 차원 수를 사용한다면 근거리 얼굴 영상의 크기를 기준으로 잡는 것보다는 원거리 얼굴 영상의 크기를 기준으로 잡는 것이 전체적인 인식률뿐만 아니라 수행 시간도 줄일 수 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

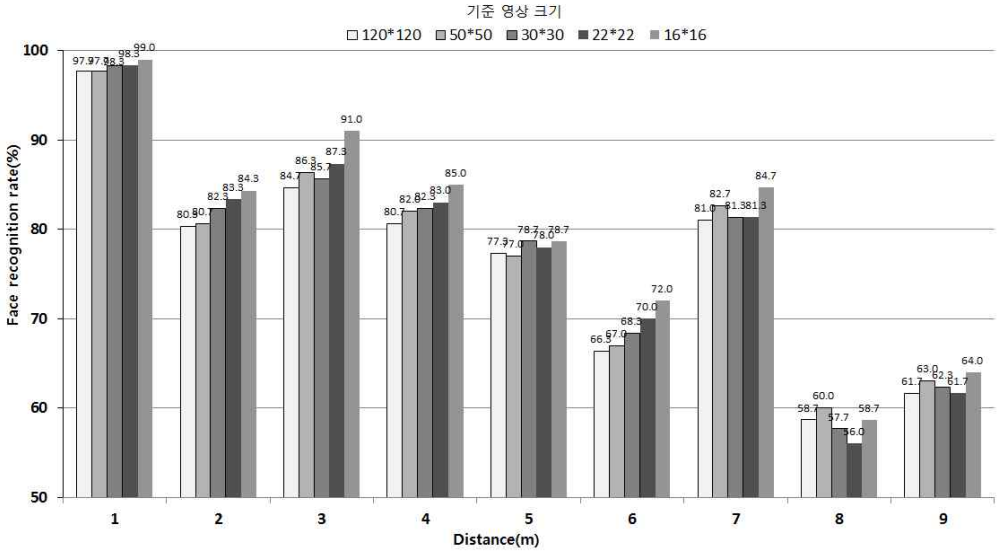


그림 42. 기준 영상 크기에 따른 PCA 기반 원거리 얼굴 인식률

그림 43. 기준 영상 크기에 따른 얼굴 인식 알고리즘 수행 시간

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제4절 제안하는 원거리 얼굴 인식 기술 분석

전체적인 실험을 통해 공통적으로 확인할 수 있는 부분은 PCA와 LDA 기반 얼굴 인식의 특성이다. 기존의 PCA와 LDA를 이용한 근거리 얼굴 인식에서 이미 알려진 바와 같이 원거리 얼굴 인식에서도 클래스당 학습 영상의 수가 적은 경우에는 PCA 기반 얼굴 인식이 우수했고, 클래스당 학습 영상의 수가 많은 경우에는 LDA 기반 얼굴 인식 방법이 우수한 성능을 나타냈다. 즉, 영상의 해상도가 커져 거리별 얼굴 영상을 많이 취득할 수 있다면 LDA 방법을 선택하고 그렇지 않다면 PCA 기반 원거리 얼굴 인식을 선택하는 것이 좋다.

제안하는 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식률의 변화를 분석하면 해상도와 얼굴 인식 방법의 구분 없이 실제 거리별 얼굴 영상을 학습으로 사용한 방법이 가장 우수한 성능을 나타냈다. 다음으로, 줌 카메라를 이용한 방법과 거리별 얼굴 영상 자동 생성 방법을 이용한 방법이 우수한 성능을 나타냈지만 실제 거리별 얼굴 영상을 이용하는 방법은 다른 방법에 비해 사용자 협조가 많이 요구되고, 줌 카메라의 경우 추가적인 비용이 발생되는 단점이 있다. 자동 생성 방법을 이용하는 방법은 가장 적은 사용자 협조와 비용이 들지만 다른 방법에 비해 인식률이 떨어지는 단점이 있다.

얼굴 크기 정규화 방법을 통한 원거리 얼굴 인식률 변화를 분석하면 전체적으로 양선형 보간법을 사용했을 때 가장 우수한 성능을 나타냈다. 즉, PCA나 LDA 기반 원거리 얼굴 인식에서는 객관적 화질 평가가 우수한 양3차회선 보간법이나 이중 선형-3차회선 보간법을 사용하는 것보다 얼굴 인식 성능 및 수행 시간 측면에서 봤을 때 계산 복잡도가 적은 양선형 보간법을 사용하는 것이 적합하다.

얼굴 밝기 정규화 방법에 따른 원거리 얼굴 인식률 또한 공통적으로 히스토그램

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

평활화를 사용했을 때 가장 우수한 성능을 나타냈다. 조명의 변화가 심하게 발생되지 않는 원거리 얼굴 인식 환경에서는 히스토그램 평활화를 사용하는 것이 가장 적합하다.

마지막으로 얼굴 영상의 기준 크기에 따른 원거리 얼굴 인식률과 수행 시간의 관계를 분석해보면, 특징벡터의 차원 수가 고정된 상황에서 기준 영상의 크기가 인식률에 미치는 영향은 크지 않음을 확인했다. 반면, 기준 영상의 크기가 줄어들수록 얼굴 인식 수행시간은 감소됨을 확인했다. 즉, 원거리 얼굴 인식에서 클래스 당 학습 영상의 수는 얼굴 인식률에 큰 영향을 미치지만 기준 영상의 크기는 비교적 영향을 적게 받기 때문에 기준 영상의 크기는 최소화 할수록 좋다.

본 논문에서는 기존 CCTV 환경에 적용 가능한 원거리 얼굴 인식을 위해 전영역 특징을 이용한 PCA와 LDA 기반 얼굴 인식 방법을 사용했고, 구조적 특징을 이용한 원거리 얼굴 인식 방법인 CMs와 CLPMs를 통해 성능 비교를 했다. 표 9는 해상도 변화와 얼굴 인식 방법에 따른 전체 거리 평균 인식률을 나타낸다. 실험 결과, 제안하는 실제 거리별 얼굴 영상을 이용한 방법이 모든 해상도에서 평균 94.9%로 가장 우수한 성능을 나타냄을 확인할 수 있다.

표 9. 원거리 얼굴 인식 방법에 따른 전체 거리 평균 인식률(%)

	기존 방법			제안하는 방법		
	단일 거리-PCA	CMs	CLPMs	실제 거리-PCA	줌 카메라	자동 생성
QVGA	77.4	38.19	39.56	98.3	-	78.6
VGA	83.7	40.8	53.1	93.7	88.4	83.5
SXGA	72.9	59.2	56.8	92.9	-	75.0

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 10. 원거리 얼굴 인식 방법 비교

비교 사항	초해상도 복원 기반	구조적 특징 기반	PTZ 카메라 기반	제안하는 방법
주목적	시각적 화질 개선	얼굴 인식률 향상	얼굴 인식률 향상	얼굴 인식률 향상
해상도 허용 범위	매우 작음	작음	매우 작음	작음
계산 복잡도	높음	낮음	낮음	낮음
필요 학습 영상 수	매우 많음	매우 적음	적음	적음
외부 변수	민감	매우 민감	민감	민감
비용	적음	적음	많음	적음
실험 환경	근거리 제한적인 환경	근거리 제한적인 환경	원거리 실제 환경	원거리 실제 환경
사용자 협조	적음	적음	적음	보통

표 10은 기존 원거리 얼굴 인식 방법들의 특징을 비교한 것으로 제안하는 방법은 원거리 얼굴 인식률의 향상에 초점을 둔 방법으로 영상에서 얼굴이 검출될 수 있는 최소 크기까지 얼굴 인식이 가능한 특징이 있다. 계산 복잡도의 경우 기존 PCA와 LDA 기반으로 이루어져 있기 때문에 기존 방법과 유사한 수준이다. 필요 학습 영상의 수는 많을수록 인식률이 증가하지만 적은 영상의 수로도 충분히 만족할 만한 인식 성능을 제공할 수 있다. 외부 변수에 대한 얼굴 인식률의 변화는 초해상도 복원 기반이나 PTZ 카메라 기반 방법에 비해 민감하게 작용한다. 얼굴 인식 실험이 수행된 환경 측면에서 비교했을 때, 제안하는 방법과 PTZ 카메라 기반 방법이 모두 실제 원거리 얼굴 인식 상황에서 수행되었다. 그러나 PTZ 카메라 기반 방법은 얼굴 인식 시스템에서 PTZ 카메라에 대한 추가 비용이 요구되는 단점이 있다. 사용자 협조 측면에서 비교했을 때는 제안하는 방법은 학습 영상 구성 단계에서 사용자의 협조를 통해 실제 거리별 얼굴 영상을 취득해야 하기 때문에 다른 방법에 비해 불편한 단점이 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제5장 결 론

최근 영상 감시 시스템은 효율적인 감시를 위해 영상 분석 기술을 기반으로 스스로 판단하고 능동적 대응이 가능한 지능형 시스템으로 진화하고 있다. 높은 수준의 지능형 영상 감시 시스템을 만족하기 위해서는 단순히 객체를 검출해서 분류하는 것뿐만 아니라 대상에 대한 정확한 신원 정보까지 확인할 수 있어야 한다. 얼굴 정보를 이용한 바이오 인식은 사용이 편리하고 시스템 적용이 쉽고 비교적 원거리에서 인식이 가능하다는 장점으로 많은 연구가 되고 있다.

얼굴 인식 응용분야의 다양화에 따라 얼굴 인식 기술 수요가 증가하고 있지만 원거리 상황이 고려되지 않은 기존 얼굴 인식 기술을 감시 시스템에 적용할 경우 사용자로부터 근거리 접근에 대한 협력이 요구되는 문제점 있다. 즉, 영상 감시 시스템에 얼굴 인식 기술을 적용하기 위해서는 실제 영상 감시 시스템의 환경을 고려한 원거리 얼굴 인식 기술이 필요하다. 또한, CCTV 기반 원거리 얼굴 인식의 경우 동일한 환경이라고 해도 카메라의 해상도에 따라 상이한 인식 성능이 나올 수 있다. 즉, 실제 카메라별 해상도에 따른 원거리 얼굴 인식 상황을 모두 만족할 수 있는 기술이 필요하다.

기존 얼굴 인식 방법을 이용해 원거리 얼굴 인식을 할 경우 다음과 같은 문제점이 발생된다. 첫째, 원거리 얼굴 인식의 경우 거리별로 추출되는 얼굴 영상의 크기가 모두 다르기 때문에 학습과 검증 영상간 얼굴 특징 차원의 부적합 문제가 발생된다. 둘째, 얼굴 영상 하나의 데이터는 크지만 클래스당 훈련 영상의 개수가 매우 한정되어 있을 때 소표본 문제가 발생된다. 셋째, 다양한 거리가 존재하는 원거리 얼굴 인식의 경우 자연스럽게 조명의 세기와 방향에 대한 얼굴의 밝기 차이가 발생한다. 넷째, 카메라의 해상도가 높아질수록 전처리 및 얼굴 인식 과정 등에서 처리

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

시간이 길어지는 문제가 발생한다.

본 논문에서는 위에서 언급한 네 가지의 고려사항을 해결하기 위해 다음과 같은 방법을 이용했다. 첫째, 학습과 검증 영상간의 얼굴 특징 차원의 부적합 문제를 해결하기 위해 영상 보간법을 사용했다. 둘째, 소표본 문제를 해결하기 위해 실제 거리별 얼굴 영상을 학습으로 사용하는 방법, 줌 카메라를 통해 가상의 거리별 얼굴 영상을 획득하는 방법, 거리별 얼굴 영상을 자동으로 생성하는 방법들을 사용했다. 셋째, 조명에 의한 얼굴의 밝기 차이 문제를 해결하기 위해 히스토그램 평활화, MCT, LBP 등을 통해 얼굴 밝기 정규화를 수행했다. 넷째, 고해상도 영상 기반 얼굴 인식에서 처리 속도 개선을 위해 학습 영상의 기준 영상의 크기를 변화시키는 방법을 사용했다.

본 논문에서는 실시간 원거리 얼굴 인식을 위해 사용자 얼굴 영상 등록 단계에서 실제 거리별 얼굴 영상을 이용하는 방법을 제안했다. 얼굴 영상 취득 과정에서 근거리와 원거리 얼굴 영상의 차이를 줄이기 위해 근거리 고해상도 얼굴 영상은 저역통과 필터링을 통해 화질 열화를 시켰고, 원거리 저해상도 영상은 양선형 보간법을 통해 복원했다. 거리 변화에 의한 얼굴 밝기 정규화에는 히스토그램 평활화를 사용했다. 본 논문에서 제안된 실제 거리별 얼굴 영상을 학습으로 사용한 원거리 얼굴 인식 방법의 성능을 평가하기 위해 QVGA, VGA, SXGA급 해상도로 구성된 ETRI, IPES-640, IPES-1280 얼굴 DB로 실험을 진행했다. 실험 결과, 제안하는 방법은 QVGA급 해상도에서 전체 거리 평균 98.3%로 기존 구조적 특징을 이용한 저해상도 얼굴 인식 방법에 비해 58.7% 높은 성능을 나타냈고, VGA급 해상도에서는 전체 거리 평균 93.7%로 기존 방법에 비해 40.6% 높은 성능을 나타냈다. 마지막으로 SXGA급 해상도에서도 제안하는 방법이 기존 방법에 비해 36.2% 우수한 성능을 나타냄을 확인했다.

본 논문의 결과물은 실제 CCTV를 통해 획득한 얼굴 DB를 통해 검증되었기 때

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

문에, CCTV가 설치된 9m 이내의 모든 환경에 적용할 수 있을 것으로 생각된다. 향후 과제로는 거리별 얼굴 영상 자동 생성 방법의 개선을 통해 실제 거리별 얼굴 영상과 유사한 얼굴을 생성하는 연구와 얼굴 틀어짐과 회전에 의한 인식률 저하 문제를 해결하는 연구가 진행되어야 한다. 또한, 정확한 조명 변화에 대한 외란을 극복하기 위해 조명의 위치, 방향, 세기 등의 조건을 제시하고 이를 통해 원거리 얼굴 인식 상황에서 조명 변화에 따른 인식률 저하 문제를 해결할 계획이다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

참고문헌

[1] 한종욱, 조현숙, “영상 보안시스템 기술 동향,” 정보보호학회논문지, 제19권, 제5호, pp. 29-37, 2009년 10월.

[2] 전범종, “국내외 지능형 CCTV 영상감시 산업동향,” 정보통신표준화소식, 제142권, pp. 50-55, 2012년 8월.

[3] 정치윤, 한종욱, “지능형 영상 분석 이벤트 탐지 기술동향,” 전자통신동향분석, 제27권, 제4호, pp. 66-73, 2012년 8월.

[4] 문현찬, “지능형 CCTV 기술 및 시장 동향,” 정보통신연구진흥원 주간기술동향, 통권1361호, pp. 13-27, 2008년 8월.

[5] 유장희, 문기영, 조현숙, “지능형 영상 보안 기술현황 및 동향,” 전자통신동향분석, 제23권, 제4호, pp. 80-88, 2008년 8월.

[6] L. Ma, X. Yang, and D. Tao, “Person re-identification over camera networks using multi-task distance metric learning,” *IEEE Trans. on Image Processing*, vol. 23, no. 8, pp. 3656-3670, Aug. 2014.

[7] X. Dai and S. Payandeh, “Geometry-based object association and consistent labeling in multi-camera surveillance,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 3, no. 2, pp. 175-184, June 2013.

[8] W. S. Zheng, S. Gong, and T. Xiang, “Reidentification by relative distance comparison,” *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 35, no. 3, pp. 653-668, Mar. 2013.

[9] J. Schiff, M. Meingast, D. K. Mulligan, S. Sastry, and K. Goldberg,

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

"Respectful cameras: detecting visual markers in real-time to address privacy concerns," *Protecting Privacy in Video Surveillance*, pp. 65-89, Oct. 2009.

[10] 오승근, 이종욱, 이한성, 정용화, 박대희, "감시시스템에서의 비정상 소리 탐지 및 식별," 정보과학회 논문지: 소프트웨어 및 응용, 제39권, 제2호, pp. 144-152, 2012년 2월.

[11] D. Bruckner, C. Picus, W. Herzner, and G. Zicker, "Hierarchical semantic processing architecture for smart sensors in surveillance networks," *IEEE Trans. on Industrial Informatics,* vol. 8, no. 2, pp. 291-301, May 2012.

[12] S. Ntalampiras, D. Arsic, M. Hofmann, M. Andersson, and T. Ganchev, "PROMETHEUS: heterogeneous sensor database in support of research on human behavioral patterns in unrestricted environments," *Signal, Image and Video Processing*, vol. 8, no. 7, pp. 1211-1231, Oct. 2014.

[13] A. Rabaoui, M. Davy, S. Rossignol, and N. Ellouze, "Using one-class SVMs and wavelets for audio surveillance," *IEEE Trans. on Information Forensics and Security,* vol. 3, no. 4, pp. 763-775, Dec. 2008.

[14] F. Peng, X. Zhu, and M. Long, "An ROI privacy protection scheme for H.264 video based on FMO and chaos," *IEEE Trans. on Information Forensics and Security*, vol. 8, no. 10, pp. 1688-1699, Oct. 2013.

[15] Y. C. Chen, C. T. Chu, J. N. Hwang, and J. H. Yoo, "A privacy-preserving human tracking scheme in centralized cloud based camera networks," in *Proc. 2014 IEEE Int. Conf. on Communications,* pp. 793-798, June 2014.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[16] H. M. Moon and S. B. Pan, "Implementation of the privacy protection in video surveillance system," in *Proc. IEEE Int. Conf. Secure Software and Reliability Improvement,* pp. 285-285, July 2010.

[17] H. M. Moon, S. H. Chae, and S. B. Pan, "The surveillance system based on RFID system for privacy protection," in *Proc. Int. Conf. Computer Science and its Applications,* pp. 726-731, Dec. 2011.

[18] 문해민, 반성범, "지능형 영상 감시 시스템에서 프라이버시 보호를 위한 De-identification 기술 분석," 한국정보기술학회논문지, 제9권, 제7호, pp. 289-200, 2011년 7월.

[19] 최봉기, "얼굴 인식 시스템의 새로운 도약-기술 개발 현황 및 시장 전망," KISTI Market Report, vol. 3, no. 5, pp.15-18, 2013년 9월.

[20] K. D. Ban, J. Lee, D. Kim, J. Kim, and Y. K. Chung, "Tiny and blurred face alignment for long distance face recognition," *ETRI Journal,* vol. 33, no. 2, pp. 251-258. Apr. 2011.

[21] H. C. Tsai, W. C. Wang, J. C. Wang, and J. F. Wang, "Long distance person identification using height measurement and face recognition," in *Proc, IEEE Region 10 Conf. TENCON 2009*, pp. 1-4, Jan. 2009.

[22] H. M. Moon and S. B. Pan, "The PCA-based long distance face recognition using multiple distance training images for intelligent surveillance system," *Lecture Notes in Computer Science,* vol. 7804, pp. 534-539, Mar. 2013.

[23] H. M. Moon and S. B. Pan, "Long distance face recognition for enhanced performance of internet of things service interface," *Computer Science and Information Systems,* vol. 11, no. 3, pp. 961-974, Dec. 2014.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[24] S. Baker and T. Kanade, "Hallucinating faces," in *Proc. IEEE 4th Int. Conf. on Automatic Face and Gesture Recognition,* pp. 83-88, Mar. 2000.

[25] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, Mar. 2002.

[26] W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Trans. on Image Processing,* vol. 21, no. 1, pp. 327-340, July 2012.

[27] C. Liu, H. Y. Shum, and C. S. Zhang, "A two-step approach to hallucinating faces: global parametric model and local nonparametric model," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition,* vol. 1, pp. 192-198, Dec. 2001.

[28] X. Wang and X. Tang, "Hallucinating face by eigentransform," *IEEE Trans. on Systems, Man and Cybernetics, Part C: applications and Reviews,* vol. 35, no. 3, pp. 425-434, Aug. 2005.

[29] J. Liu, J. Qiao, X. Wang, and Y. Li, "Face hallucination based on independent component analysis," in *Proc. IEEE Int. Symposium on Circuits and Systems*, pp. 3242-3245, May 2008.

[30] W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Trans. on Image Processing,* vol. 21, no. 1, pp. 327-340, July 2012.

[31] Z. Wang, Z. Miao, Q. M. J. Wu, Y. Wan, and Z. Tang, "Low-resolution face recognition: a review," *The Visual Computer,* vol. 30, no. 4, pp. 359-386, Aug. 2014.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[32] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Feature subspace determination in video-based mismatched face recognition," in *Proc. IEEE 8th Int. Conf. on Automatic Face and Gesture Recognition(FG)*, pp. 1-6, Sept. 2008.

[33] B. Li, H. Chang, S. G. Shan, and X. L. Chen, "Coupled metric learning for face recognition with degraded images," *Lecture Notes in Computer Science,* vol. 5828, pp. 220-233, Nov. 2009.

[34] B. Li, H. Chang, S. G. Shan, and X. L. Chen, "Low-resolution face recognition via coupled locality preserving mappings," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 20-23, Jan. 2010.

[35] F. Wheeler, R. Weiss, and P. Tu, "Face recognition at a distance system for surveillance applications," in *Proc. IEEE Int. Conf. on Biometrics: Theory Applications and Systems,* pp. 1-8, Sept. 2010.

[36] P. Salvagnini, L. Bazzani, M. Cristani, and V. Murino, "Person re-identification with a PTZ camera: an introductory study," in *Proc. 20th IEEE Int. Conf. on Image Processing,* pp. 3552-3556, Sept. 2013.

[37] R. Raghavendra, K. B. Raja, Y. Bian, and C. Busch, "Comparative evaluation of super-resolution techniques for multi-face recognition using light-field camera," in *Proc. 18th Int. Conf. on Digital Signal Processing,* pp. 1-6, July 2013.

[38] U. S. Park, C. H. Choi, A. K. Jain, and S. W. Lee, "Face tracking and recognition at a distance: a coaxial and concentric PTZ camera system," *IEEE Trans. on Information Forensics and Security,* vol. 8, no. 10, pp. 1665-1677, Oct. 2013.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[39] E. Hjelmas and B. K. Low, "Face detection: a survey," *Computer vision and image understanding*, vol. 83, no. 3, pp. 236-274, Sept. 2001.

[40] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: the state of the art," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 22, no. 12, pp. 1424-1445, Dec. 2000.

[41] R. Brunelli, T. Poggio, "Face recognition: features versus templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, Oct. 1993.

[42] 표창원, 박기남, "범죄취약지 CCTV 등 범죄감응장치 설치에 관한 연구," 치안논총, 제17권, pp. 579-689, 2001년 12월.

[43] Y. Zhang, J. Tian, X. He, and X. Yang, "Face recognition: a literature survey," *ACM Computing Surveys,* vol. 35, no. 4, pp. 399-458, Dec. 2003.

[44] 이홍주, "얼굴 인식 성능향상을 위한 조명 정규화와 영상 정합," 연세대학교 석사학위논문, 2009년 2월.

[45] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, Dec. 1991.

[46] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces : recognition using class specific linear projection," *IEEE Trans, on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.

[47] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks,* vol. 13, no. 4, pp. 411-430, June 2000.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[48] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based face recognition using mixture distance," in *Proc. IEEE Conf. on Computer Vision and Pattern recognition*, pp. 209-216, June 1996.

[49] L. Wiskoot, J. M. Fellous, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, July 1997.

[50] A. V. Nefian and M. H. Hayes Ⅲ, "Hidden Markov Models for face recognition," in *Proc. 1998 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 5, pp. 2721-2724, May 1998.

[51] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 84-91. June 1994.

[52] P. S. Penev and J. J. Atick, "Local feature analysis: a general statistical theory for object representation," *Network: Computation in Neural Systems*, vol. 7, no.3, pp. 477-500, Dec. 1996.

[53] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic face identification system using flexible appearance models," *Image and Vision Computing*, vol. 13, no. 5, pp. 393-401, June 1995.

[54] B. Heisele, T. Serre, M. Pontil, and T. Poggio, "Component-based face detection," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition,* vol. 1, pp. 657-662, Dec. 2001.

[55] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39,

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

no. 3, pp. 355-368, Sept. 1987.

[56] R. B. Paranjape, W. M. Morrow, and R. M. Rangayyan, "Adaptive-neighborhood histogram equalization for image enhancement," *Graphical Models and Image Processing*, vol. 54, no. 3, pp. 259-267, May 1992.

[57] J. A, Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. on Image Processing*, vol. 9, no. 5, pp. 889-896, May 2000.

[58] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Trans. on Consumer Electronics*, vol. 53, vol. 2, pp. 593-600, May 2007.

[59] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065-1076, Sept. 1962.

[60] V. A. Epanechnikov, "Nonparametric estimation of a multidimensional probability density function," *Theory of Probability & Its Applications,* vol. 14, no. 1, pp. 156-161, June 1969.

[61] B. Froba and A. Ernst, "Face detection with the modified census transform," in *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition,* pp. 91-96, May 2004.

[62] B. Kim and H. M. Park, "Efficient face recognition based on MCT and $I(2D)^2PCA$," in *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 2585-2590, Oct. 2012.

[63] K. Jeong, J. Choi, and G. Jang, "Semi-local structure patterns for

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

robust face detection," *IEEE Signal Processing Letters,* vol. PP, no. 99, pp. 1-11, Nov. 2014.

[64] Y. Fang, K. Wang, J. Cheng, and H. Lu, "A real-time hand gesture recognition method," in *Proc. IEEE Int. Conf. on Multimedia and Expo,* pp. 995-998, July 2007.

[65] 조광신, 박수경, 심동규, 이수연, "조명변화에 강인한 MCT와 프레임 연관성 기반 실시간 얼굴 인식 시스템," 전자공학회논문지-CI, 제45권, 제3호, pp. 123-134, 2008년 5월.

[66] 한동일, *VHDL 프로그래밍 바이블*, 생능출판사, 2011년.

[67] W. Chu, "Facial expression recognition based on local binary pattern and gradient directional pattern," in *Proc. IEEE Int. Conf. on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing,* pp. 1458-1462, Aug. 2013.

[68] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. on Image Processing,* vol. 19, no. 6, pp. 1635-1650, June 2010.

[69] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary pattern: a comprehensive study," *Image and Vision Computing,* vol. 27, no. 6, pp. 803-816, May 2009.

[70] C. K. Tran, T. F. Lee, L. Chang, and P. J. Chao, "Face description with local binary patterns and local ternary patterns: improving face recognition performance using similarity feature-based selection and classification algorithm," in *Proc. 2014 Int. Symposium on Computer, Consumer and Control,* pp. 520-524, June 2014.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[71] K. Meena and A. Suruliandi, "Local binary patterns and its variants for face recognition," in *Proc. Int. Conf. on Recent Trends in Information Technology*, pp. 782-786, June 2011.

[72] R. C. Gonzalez and R. E. Woods, *Digital image processing,* Prentice Hall, 2008.

[73] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. on Medical Imaging*, vol. 2, no. 1, pp. 31-39, Mar. 1983.

[74] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Trans. on Medical Imaging,* vol. 18, no. 11, pp. 1049-1075, Nov. 1999.

[75] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. on Acoustic, Speech, and Signal Processing,* vol. asp-29, no. 6, Dec. 1981.

[76] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Second Edition, Wiley-interscience, 2001.

[77] 문해민, 반성범, "선형 보간법과 3차회선 보간법을 결합한 디지털 영상 스케일러의 VLSI 구조," 전자공학회논문지, 제51권, 제3호, pp.112-118, 2014년 3월.

[78] A. H. FREng, "The database of faces," AT&T Lab. Cambridge, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html, 2002.

[79] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 23, no. 6, pp. 643-660, June 2001.

Collection @ chosun

조선대학교 CHOSUN UNIVERSITY

[80] B. Weyrauch, B. Heisele, J. Huang, and V. Blanz, "Component-Based face recognition with 3D morphable models," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition Workshop,* pp. 85-89, June 2004.

[81] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, Oct. 2000.

[82] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacements," in *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 742-746, June 1988.

[83] H. Demirel and G. Anbarjafari, "Discrete wavelet transform-based satellite image resolution enhancement," *IEEE Trans. on Geoscience and Remote Sensing,* vol. 49, no. 6, pp. 1997-2004, June 2011.

[84] S. H. Rhee and M. G. Kang, "Discrete cosine transform based regularized high-resolution image reconstruction algorithm," *Optical Engineering,* vol. 38, no. 8, pp. 1348-1356, Aug. 1999.

[85] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational bayesian super resolution," *IEEE Trans. Image Processing*, Vol. 20, no. 4, pp. 984-999, April 2011.

[86] K. Bernardin, F. V. D. Camp, and R. Stiefelhagen, "Automatic person detection and tracking using fuzzy controlled active cameras," in *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* pp. 1-8. June 2007.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

[87] Y. Xu and D. Song, "Systems and algorithms for autonomous and scalable crowd surveillance using robotic PTZ cameras assisted by a wide-angle camera," *Autonomous Robots,* vol. 29, no. 1, pp. 53-66, April 2010.

[88] N. Krahnstoever, T. Yu, S. Lim, K. Patwardhan, and P. Tu, "Collaborative real-time control of active cameras in large scale surveillance systems," in *Proc. Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications, Marseille,* France, pp. 1-12, Oct. 2008.

Collection @ chosun