

제한된 공간에서의 얼굴인식

이영식* · 배철수**

Face Recognition in a Meeting Room

Young-Sik Lee* · Cheol-Soo Bae**

요 약

한정된 공간에서는 낮은 해상도, 적은 조명, 예상할 수 없는 머리의 움직임, 연속적으로 변화하는 얼굴의 표정과 색 등으로 인하여 사람의 얼굴인식에 많은 문제점을 가지고 있다.

이러한 문제점들을 해결하기 위하여 동적 공간 변화(DSW)라는 새로운 알고리즘을 제안한다

알고리즘은 공간의 제약이 따르는 환경에서의 국부적 특징들을 조합하는 것이며 모아진 자료를 기초로 고유한 얼굴 이미지와 DSW를 비교하고, 정면과 프로파일 얼굴 이미지, 두 단계의 색 변화를 가지는 이미지를 가지고 테스트하였다.

실험 결과로 PCA 알고리즘이 82.7% DSW가 89.4%로 DSW를 적용한 경우가 PCA 방법보다 6.9%더 좋은 결과를 나타내었으며 고유한 얼굴이미지에 접근함을 알 수 있었다.

ABSTRACT

In this paper, we investigate recognition of human faces in a meeting room. The major challenges of identifying human faces in this environment include low quality of input images, poor illumination, unrestricted head poses and continuously changing facial expressions and occlusion. In order to address these problems we propose a novel algorithm, Dynamic Space Warping (DSW). The basic idea of the algorithm is to combine local features under certain spatial constraints. We compare DSW with the eigenface approach on data collected from various meetings. We have tested both front and profile face images and images with

two stages of occlusion. As a result from the experiment, we obtained 82.7% for PCA algorithm, and 89.4% for DSW. We get to obtain 6.9% better result from conductive DSW approach rather than PCA. It turned out to be that it shows more original and unique facial image.

키워드
PCA, DSW

1. 서 론

얼굴 인식 시스템[1]은 많은 발전을 보이고 있는 반면, 응용분야에 있어서는 아직도 상당히 미흡하다.

영상에서 자동적으로 얼굴 부분을 결정하는데 있어서는 얼굴 특징과 얼굴 특징점은 가장 기본적인 요소이다.

얼굴 이미지는 타원형안에 눈, 입, 코 같은 특징들

의 집합으로 볼 수 있으며, 확실한 모델로 사용할 수 있다.

복잡한 배경, 자세의 변화, 그리고 안경이나 모자 등에 의해 왜곡된 영상에서도 얼굴과 얼굴 특징점을 검출할 수 있어야 한다.

얼굴인식에 대한 연구는 검출된 얼굴이 수집된 얼굴 데이터베이스의 얼굴과 맞는지에 대한 인식 문제에 집중되어 있다. 갤러리에 수천 개의 다른 사람들

*경동대학교

**관동대학교

의 얼굴을 모아 놓았을 때, 각 모델은 몇 장의 얼굴 영상으로 이루어지게 된다. 최근 연구들은 비디오 시퀀스에서부터 사람을 인식하는 시스템[2][3]을 개발하고 있다.

본 논문에서는 새로운 얼굴인식 알고리즘을 개발하기 위해서 회의실이라는 제한된 공간을 설정하고 회의 도중 나타나는 얼굴 표정을 인식하는 알고리즘을 제안하였다. 2장에서는 실험을 통해서 얻어지는 회의실 환경을 살펴볼 것이고, 3장은 얼굴 인식의 접근법인 PCA(Principal Component Analysis)를 기반으로 한 새로운 동적 공간 워핑(Dynamic Space Warping)을 설명할 것이다.

그리고 4장에서는 실험 결과와 회의실에서 촬영된 얼굴 영상 DB를 설명하고, 마지막으로 5장에서는 결론 및 향후과제를 제시한다.

II. 회의실 환경

얼굴을 마주보고 하는 일반적 회의는 보통 말하기, 몸동작, 쓰기, 각 개인 식별 등 몇몇 특징을 가지고 있다. 이런 각각의 특징들에 대한 인식과 통합은 나중에 참고하기 위한 정확한 회의 기록을 만드는 데 중요하다. 본 논문은 JANUS 음성인식 장치[4]와 다양한 사람식별 모듈[5]을 이용하여 회의의 중요한 측면을 지속적으로 추적하고 캡처, 통합하기 위해 다양한 회의실 환경에서 개발하였다. 회의 참석자의 식별은 음성 식별과 색채 식별의 혼합으로 결정된다. 연구 목적은 얼굴 인식을 추가함으로써 사람 식별 시스템의 인식을 증대를 이루는 것이다. 특히 자동 얼굴 인식은 까다로운 패턴인식작업을 요하게 되는데 조도, 방향, 크기, 얼굴 표정 등의 다양한 변화를 갖기 때문이다. 그러므로 여러 사람이 있는 회의실에서의 지속적인 얼굴 인식 작업은 다양한 변수를 만들게 된다.

II-1. 저품질 비디오 입력

전체 회의실 화면을 담기 위한 카메라가 하나 있어야 한다. 이 카메라로부터 얻어진 얼굴 영상은 상당히 저해상도의 이미지가 된다. 고해상도의 얼굴 영상을 촬영하기 위해서는 사람을 가까이에서 촬영하

는 카메라가 있어야 하지만, 각 사람 당 카메라를 할당하는 것은 현실적인 문제에 있어서 비능률적이다

II-II. 조도

얼굴의 조도는 조명이 위에서 비추고 사람들의 위치와 머리의 위치에 때문에 급격히 변화한다. 회의 전 과정을 통해서 회의실 조명은 일정하지만 음영의 변화는 대단히 크다.

II-III. 머리의 위치와 얼굴 표정

회의라는 특성상 다양한 머리의 위치와 얼굴 표정 변화가 생길 수 있다.

II-IV. 얼굴 가림 현상

회의를 하는 동안 회의 참석자들은 머리와 손 등을 끊임없이 움직이게 되고, 얼굴 전체나 일부분이 손이나 종이 등 다른 물건으로 가려지기도 한다. 그리고 카메라의 개수와 위치가 일정하기 때문에 인식 장치는 다른 사람에 의해서 차단되고 가려지는 것들을 처리해야 한다.

그림 1은 회의하는 동안 위와 같이 발생하는 장면을 촬영된 얼굴 영상이다.

III. 얼굴인식 시스템

III-1. 국부적인 인식과 전체적인 얼굴 인식

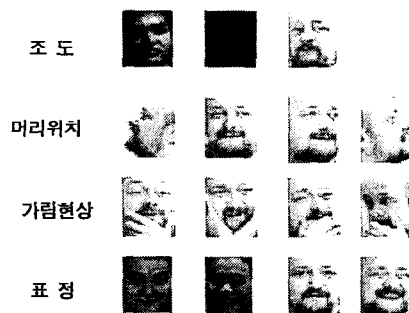


그림 1. 회의실에서의 얼굴 확대 영상

초기의 얼굴 인식을 위한 시스템은 얼굴 영상에서

기하학적인 특징을 측정하고 이전에 저장된 패턴과 산출된 벡터와 비교하는 것이었다[6]. 최근에는 이런 특징 기반의 국부적 인식은 템플릿 기반 알고리즘인 전체적인 인식으로 넘어가고 있다.

III-II. PCA 기반 얼굴 인식

최근 발표되는 다수의 전체적 얼굴 인식 알고리즘 중에서 Turk & Pentland[7]에 의해서 제안된 eigenface 알고리즘이 가장 효율적인 알고리즘 중의 하나이다. 이 알고리즘은 저차원 특징 공간에 고차원의 영상을 선형적으로 투영하기 위해서 PCA(Principal Component Analysis) 알고리즘을 사용한다. 특징 공간을 측정하는 분산 매트릭스의 고유벡터가 결정되면, 특징 공간에서 테스트 영상과 참조 영상간의 유클리드 거리를 계산함으로써 인식을 하게 된다. 그러나 eigenface 알고리즘은 얼굴만 존재하는 영상에서는 좋은 결과를 보이지만, 얼굴가림현상이 있는 영상에서는 좋은 결과를 보이지 못하는 단점이 있다. 이 알고리즘은 특징 공간에서 단일 포인트로 입력 영상을 인코딩하기 때문에 다른 물건이나 손등으로 가려지게 되면 디코딩 할 수 없게 된다. 그 결과는 그림 8에 나타내었다.

III-III. DSW: Dynamic Space Warping

알고리즘은 우선, 특징 공간에서 단일 포인트에 얼굴 영상을 투영하는 대신 포인트 시퀀스를 만들기 위해서 그림 2에 나타냈듯이 이동 윈도우(사각 점선)를 이용하게 되는데 이동하는 경로는 일반 주사방식과 유사하게 영상의 왼쪽 상단에서 우측 하단 점으로 이동하게 된다.

훈련 영상 수열 $\Gamma_1, \Gamma_2, \dots, \Gamma_m$ 으로부터 각각의 얼굴 영상 Γ_i 에 대해 서브영상 $\Gamma_{v(i)} = \Gamma_i^1 \Gamma_i^2 \dots \Gamma_i^{n_i}$ 를 생성하고, 서브영상 $\Gamma_i^j, j=1, \dots, m, j=1, \dots, n_i$ 을 기반으로 PCA를 수행하여 고유공간의 각각 벡터 $\Gamma_{v(i)}$ 를 투영한다. 본 논문의 실험에서 서브영상의 개수는 전체 훈련 영상 ($n_i=c$)에서 일정하다. 서브영상 Γ_i^j 는 PCA에서 두 가지 방법으로 이용된다. 단일 고유공간에서 모든 서브 영상 Γ_i^j 의 결합뿐만 아니라 영상

$(\Gamma_1^1, \Gamma_1^2, \dots, \Gamma_1^m), (\Gamma_2^1, \Gamma_2^2, \dots, \Gamma_2^m), \dots, (\Gamma_c^1, \Gamma_c^2, \dots, \Gamma_c^m)$ 을 이용하여 다른 고유공간 c 를 만들 수 있게 한다.



그림 2. 영상에서의 윈도우 이동경로

그러나 본 연구의 실험은 이들 변수간의 현저한 차이를 보이지는 않는다. 특징 공간에서 포인트의 경과 시퀀스는 주어진 훈련 영상에 대한 참조 시퀀스로 저장되어진다. 모르는 얼굴 영상은 음성인식에서 사용되는 동적 시간 워핑(dynamic time warping)[2] 같은 기술로 만들어진 동적 프로그래밍을 이용하여 저장된 참조시퀀스와 비교된다. 이런 유사성 때문에 이동적 프로그래밍을 DSW라 한다. 테스트 시퀀스의 서브 윈도우 i 와 각 템플릿 k 의 서브윈도우 j 는 격자 포인트 (i, j, k) 의 수열로 정의되고, 각각의 격자 포인트는 각각의 윈도우간의 거리 $d(i, j, k)$ 와 결합하게 된다. 이 알고리즘은 테스트 패턴과 참조 패턴간에 최상의 매칭을 취하는 격자 포인트를 통과하는 경로를 검색하게 된다. 격자 포인트 (i, j, k) 에 리드되는 경로를 따라 최소 축적 거리를 $D(i, j, k)$ 라 정의하면 $D(i, j, k)$ 는 다음과 같다.

$$D(i, j, k) = \sum_{n=1}^j d(1, n, k)$$

위 식을 다시 쓰면 다음과 같아진다.

$$D(i, j, k) = d(i, j, k) + \min \{ D(i-1, j, k), D(i-1, j-1, k), D(i, j-1, k) \}$$

최상의 참조 시퀀스는 $\min_k D(n, j, k)$ (n 은 테스트 시퀀스의 길이에 존재한다.)으로 주어진다. 그림 3은 알고리즘의 단계를 나타내었다. 이동 윈도우와 수평수직 오프셋의 크기는 입력 영상의 크기에 따라 자동적으로 결정된다. 실험상 나타난 것은 아홉개의 오버랩핑 영역에 얼굴을 분할하는 것으로 다른 국부적 방법과는 대조적으로 DSW 알고리즘은 눈, 코, 입과 같은 얼굴 특징들의 위치측정을 할 필요가

없다는 것이다.

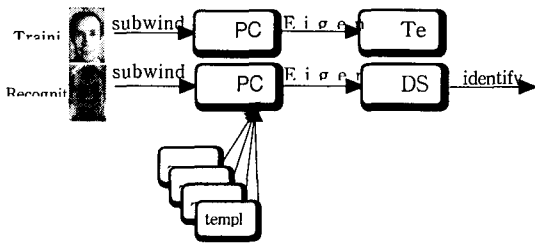


그림 3. DSW 알고리즘의 처리과정

IV. 실험 및 고찰

알고리즘을 평가하기 위해 이미지를 6개의 그룹으로 나누고 회의 때마다 기록하였고 얼굴위치, 오리엔테이션, 독자성과 얼굴 가림 정도에 따라 수동으로 분류하였다.

이 결과 데이터 세트의 이미지는 15x20과 40x54 사이에서 다양한 크기로 나타났다.

수집된 이미지의 크기를 표준화하기 위한 전처리 과정(영평균과 단위변화, 조도조절, 히스토그램평화)을 수행하였다

모델을 생성하기 위해 6개의 그룹에 대하여 세 가지의 각기 다른 모습(정면과 양측면)을 사용하였고, 모든 학습 이미지는 하나의 고유공간에 결합[8]되었다.

회의 중에 참가자의 위치가 움직임에 따라 변화하는 모습과 조명조건을 위해 우리는 데이터베이스를 대략 1200장의 그림으로 구성하였고, 이를 평균화하여 약 60여장의 사진을 모델화 하였다.

임의의 이미지를 학습시키고 사진들의 모델을 만들어 이를 테스트하였다. 신뢰성 있는 결과를 위해 이 과정을 되풀이하고 결과의 평균값을 얻을 수 있었다 또한, 그림 4는 기본적인 방법의 고유얼굴의 접근과 DSW의 승인 비율을 학습 이미지를 바꾸기 위해 비교한다.

양쪽 알고리즘의 비교를 위해 각각 두 가지의 변화 PCA1, PCA2와 DSW1, DSW2를 실험하였다.

첫번째 버전에서 1 모델을 위해 훈련 이미지 위에서 PCA 또는 DSW의 적용에서 생기고 있는 패턴 벡터는 평균치가 되고 두 번째 변화를 위해 모든 벡터는 저장되었다.

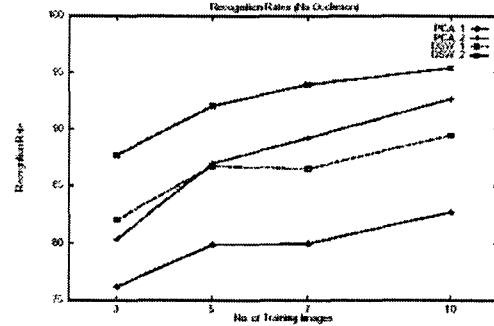


그림 4. PCA와 DSW알고리즘의 인식율

변화된 DSW 접근은 더 높은 인식율을 보여준다. 보통의 얼굴 이미지와 더불어 2명의 얼굴 가림 현상을 대략 150 가지로 나타내었다.

그림 5는 양쪽 카테고리에 대한 예를 보여준다.

DSW 접근은 eigenface 알고리즘을 능가하는 성능을 보여주었다.

표 1은 DSW의 결과이다.

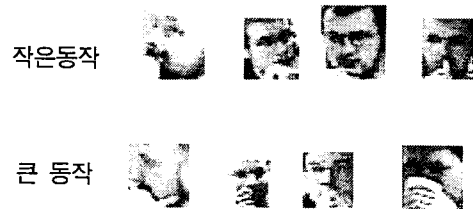


그림 5. 얼굴가림현상의 예

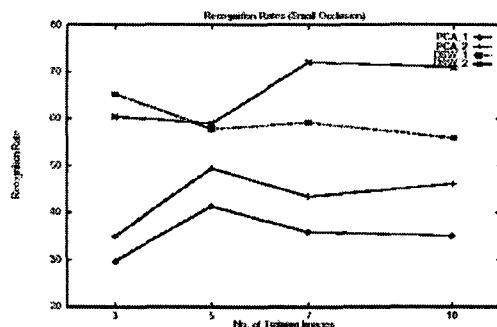


그림 6. 작은 동작의 얼굴 가림 현상에서의 인식을

표 1. PCA와 DSW에서의 인식율 비교

Train image	3	5	7	10
PCA1 w/o occl	76.2%	79.9%	80.0%	82.7%
DSW1w/o occl	82.0%	86.8%	86.5%	89.4%
PCA1 sml occl	29.7%	41.4%	35.9%	35.0%
DSW1sml occl	65.4%	57.9%	59.2%	55.9%
PCA1 lrg occl	25.3%	31.6%	29.0%	30.8%
DSW1 lrg occl	45.5%	49.9%	47.5%	48.6%

표준 PCA는 부분적 얼굴 가림 현상을 더 잘 나타내 준다 그림 8은 재건된 얼굴이미지를 통해 이를 보여 주고 있다.

이 과정에서 최초의 이미지는 첫번째로 고유영역에 적용되고, 그 다음 특징 공간의 고유공간 표현과 얼굴의 기초를 재건하였다.

색이 없는 얼굴을 재건하는 것은 독립적인 얼굴과 유사점의 상을 만든다.

만일 얼굴의 부분이 손 또는 펜으로 막히면 DSW로부터 얻게 되는 이미지가 유지되는 동안 PCA 고유 공간에서 재건되는 얼굴 이미지는 강한 왜곡을 보여주게 된다. 따라서 이러한 단계는 회의 동안 발생하는 상황을 처리할 능력이 있는 효과적 얼굴 식별 시스템에 대한 첫번째 단계이다. 얼굴 추적 알고리즘은 실시간에 보여지는 영 역에서 다양한 얼굴을 추적할 수 있다. 따라서 이 같은 학습 방법과 얼굴 추적 알고리즘과 통합하게 될 때 제한 된 공간에서 나온 사용 가능한 많은 영상 데이터를 효과 적인 모델로 만들 수 있다.

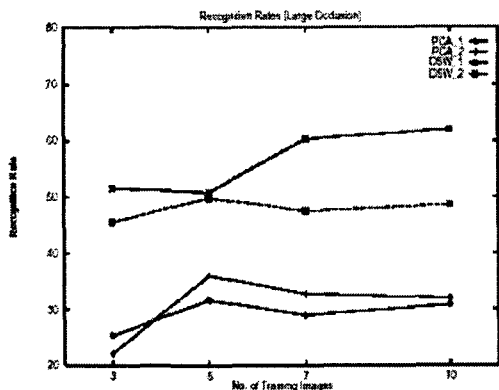


그림 7. 큰동작의 얼굴가림현상에서의 인식율

V. 결 론

얼굴 인식을 위한 새로운 알고리즘을 제안하였고, 고유한 얼굴의 인식에 관하여 실험에 의해 개선된 점을 보여 주었다.

변화하는 조명과 낮은 해상도의 입력 이미지라는 조건에서 시스템은 얼굴 가림 현상에 접근하였다.

낮은 해상도의 이미지를 입력으로 사용한 결과 이러한 이미지는 삼차원 모델의 표준화에는 적합하지 않음을 실험으로써 알 수 있었으며, PCA 알고리즘이 82.7% DSW가 89.4%로 DSW를 적용한 경우가 PCA방법보다 6.9% 더 좋은 결과를 나타내었으며 고유한 얼굴 이미지에 접근함을 알 수 있었다.

그러므로 향후 삼차원 모델을 표준화 할 수 있는 알고리즘에 대해 더 연구해야 할 것이다.

참고문헌

- [1] P. Phillips, H. Moon, P. Rauss, and S. Rizvi. The FERET evaluation methodology for face-recognition algorithms. In *CVPR'97*, 1997.
- [2] A. Howell and H. Buxton. Towards unconstrained face recognition from image sequences. In *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, 1996.
- [3] S. McKenna and S. Gong. Face recognition from sequences using models of identity. In *Proc. Asian Conference on Computer Vision*, 1998.
- [4] H. Yu, M. Finke, and A. Waibel. Progress in automatic meeting transcription. In *Proceedings of the Eurospeech '99*, 1999.
- [5] J. Yang, X. Zhu, R. Gross, J. Kominek, and A. Waibel. Multimodal people ID for a multimedia meeting browser. In *Proceedings of ACM Multimedia*, 1999.
- [6] T. Kanade. Picture processing by computer complex and recognition of human faces. Technical report, Dept. of Information Science, Kyoto University, 1973.
- [7] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*,

3(1), 1991.
 [8] A. Pentland, B. Moghaddam, and T. Starner.
 View-based and modular eigenspaces for face
 recognition. In *IEEE Conference on Computer
 Vision & Pattern Recognition*, 1994..

2003년 1월~현재 한국통신학회 국내저널 편집부위
 원장
 2003년 1월~현재 대한전자공학회 이사
 1981년~현재 관동대학교 정보통신공학부 교수
 ※ 관심분야: 영상처리, 신호처리시스템, 영상압축

저자소개



이영식(Young-Sik Lee)

1986년 2월 한국항공대학교 통신
 정보공학과 졸업(공학사)
 1996년 2월 경희대학교 정보통신
 공학과 졸업(공학석사)
 2001년 3월~현재 관동대학교 대
 학원 전자통신공학과 박사과정 재학중
 1985년 12월~1992년 2월 삼성전자 정보통신부문 통신
 종합연구소
 1992년 3월 경북대학 전자계산과 전임강사
 2001년 3월 경동대학교 정보통신공학부(부교수)/교무처
 장
 ※ 관심분야: 영상처리, 신호처리시스템, 영상압축



배철수(Cheol-Soo Bae)

1979년 2월 명지대학교 전자공
 학과 졸업(공학사)
 1981년 2월 명지대학교 대학원
 전자공학과졸업(공학석사)
 1988년 8월 명지대학교 대학원
 전자공학과졸업(공학박사)
 1991년 3월~현재 산업자원부 공업기술기반 평가위
 원 및 심의위원
 1998년 3월~현재 강원도 정보화위원회 위원
 1999년 3월~2001년 5월 관동대학교공과대학 학장
 2000년 3월~2002년 2월 관동대학교 양양캠퍼스 창
 업보육센터 소장
 2001년 3월~현재 해양정보통신학회 강원지부장
 2001년 11월~현재 기초과학연구소 운영위원
 2002년 5월~현재 중소기업 정보화 혁신 컨소시엄
 추진위원
 2002년 6월~현재 관동대학교 평생교육원 원장