

줌 카메라를 통해 획득된 거리별 얼굴 영상을 이용한 원거리 얼굴 인식 기술*

문 해 민,[†] 반 성 범[‡]
조선대학교

The Long Distance Face Recognition using Multiple Distance Face Images Acquired from a Zoom Camera*

Hae-Min Moon,[†] Sung Bum Pan[‡]
Chosun University

요 약

지능형 서비스를 제공하는 로봇에서 특정 사람을 인지하거나 구별하는 인식 기술은 매우 중요하다. 기존 단일 거리 얼굴 영상을 학습으로 사용한 얼굴 인식 알고리즘은 원거리로 갈수록 얼굴 인식이 떨어지는 문제점이 있다. 실제 거리별 얼굴 영상을 이용한 방법은 얼굴 인식이 향상되지만, 사용자 협조가 요구되는 단점이 있다. 본 논문에서는 줌 카메라를 통해 거리별 얼굴 영상을 획득하여 학습으로 사용하는 LDA 기반 원거리 얼굴 인식을 제안한다. 제안하는 방법은 기존 단일거리 얼굴 영상을 학습으로 이용한 방법에 비해 7.8% 향상된 성능을 보였고, 거리별 얼굴 영상을 학습으로 이용한 방법과 비교했을 때 8.0% 저하된 성능을 보였다. 그러나 거리별 얼굴 영상을 취득하기 위해 추가적인 시간과 사용자 협조가 요구되지 않는 장점이 있다.

ABSTRACT

User recognition technology, which identifies or verifies a certain individual is absolutely essential under robotic environments for intelligent services. The conventional face recognition algorithm using single distance face image as training images has a problem that face recognition rate decreases as distance increases. The face recognition algorithm using face images by actual distance as training images shows good performance but this has a problem that it requires user cooperation. This paper proposes the LDA-based long distance face recognition method which uses multiple distance face images from a zoom camera for training face images. The proposed face recognition technique generated better performance by average 7.8% than the technique using the existing single distance face image as training. Compared with the technique that used face images by distance as training, the performance fell average 8.0%. However, the proposed method has a strength that it spends less time and requires less cooperation to users when taking face images.

Keywords: Long Distance Face Recognition, Multiple Distance Face Image, Zoom Camera, Training Face Image

접수일(2014년 8월 6일), 수정일(2014년 11월 10일),
게재확정일(2014년 12월 9일)

* 본 연구는 산업통상자원부와 한국 산업기술진흥원의 지역
특화산업 육성사업과 2014학년도 조선대학교 학술연구비

의 지원을 받아 수행된 연구결과입니다.

[†] 주저자, bombilove@gmail.com

[‡] 교신저자, sbpan@chosun.ac.kr(Corresponding author)

I. 서 론

과거 로봇이 단순히 대량생산을 위한 산업용으로 이용됐다면, 최근에는 지능형 서비스를 기반으로 한 자동 로봇 산업으로 발전되고 있다. 지능형 로봇에 있어서 중요한 점은 인간과 로봇 간 상호작용을 통해 로봇이 스스로 판단하고 행동하는 것이다. 따라서 인간과 상호작용 기능을 가지며 외부 환경의 변화를 인식하고 스스로 판단하여 자율적으로 동작할 수 있는 지능형 로봇에 대한 관심이 증가하고 있다. 로봇환경에서 특정 사람을 인지하거나 구별하는 사용자 인식 기술은 보안 및 감시 서비스를 위해서 반드시 필요하다 [1]. 사용자 인식을 통한 로봇 서비스는 특정한 사용자에 대해 적절한 개인 서비스를 빠르고 정확하게 제공할 수 있게 된다[2]. 특히 얼굴 인식은 비접촉 및 비협조에도 인식할 수 있고 원거리에서도 인식이 가능하므로 얼굴을 이용한 원거리 인식에 대한 연구가 진행 중이다[3][4].

로봇 환경에서의 얼굴 인식은 로봇의 카메라로부터 입력된 영상에 존재하는 사람의 얼굴을 검출하여 신원을 인증하는 기술이다. 기존 얼굴 인식 방법은 크게 기하학적 특징을 기반으로 하는 EBGM(Elastic Bunch Graph Matching)[5]이 있고, 얼굴 전체의 통계적인 값을 특징으로 인식하는 PCA(Principal Component Analysis)[6]나 LDA(Linear Discriminant Analysis)[7] 등이 있다.

일정한 거리에서 협조적인 자세로 인식을 수행하는 기존 얼굴 인식 방법과 달리, 로봇환경에서는 로봇과 사람의 거리가 유동적이기 때문에 사용자에게 협조적인 자세를 요청하기 어렵다. 기존에 사용된 얼굴 인식 기술을 로봇환경에 이용할 경우 근거리에서는 얼굴 인식 성능이 우수할 수 있으나, 거리가 멀어질수록 얼굴 인식 성능이 떨어지는 문제점이 발생할 수 있다. 그러므로 로봇환경에 얼굴 인식을 적용하기 위해서는 원거리로 갈수록 저하되는 영상의 화질을 고려해야 한다.

최근에는 원거리에서도 고화질 영상을 취득할 수 있는 고성능의 줌 카메라를 이용해 원거리 얼굴 인식을 하는 기술이 연구되고 있다[8-9]. 고성능 카메라를 이용한 얼굴 인식의 경우 초기 설치 및 관리에서 많은 비용이 요구되기 때문에 보편적으로 이용되기 어려운 문제점이 있다. 그러므로 기존 저해상도 기반 영상에 적용 가능한 원거리 얼굴 인식 알고리즘의 개

발이 필요하다. 최근 Moon 등이 원거리 얼굴 인식 성능을 향상하기 위해 학습용 얼굴 영상을 거리별 얼굴 영상으로 이용한 연구를 진행했다[10]. 이 방법은 단일거리 얼굴 영상을 등록으로 사용한 방법에 비해 원거리 얼굴 인식률이 증가하지만, 학습용 얼굴 영상을 취득하기 위해 사용자가 1m~5m를 직접 이동해야 하는 불편함이 존재한다.

본 논문에서는 사용자 등록을 위한 얼굴 영상 취득방법을 개선한 LDA 기반 원거리 얼굴 인식 방법을 제안한다. 제안하는 방법은 사용자 등록에 사용되는 얼굴 영상 취득을 위해 줌 카메라를 이용한다. 줌 카메라를 이용할 경우 사용자는 직접 1m~5m를 이동하지 않고, 고정된 위치에서 줌 기능을 이용해 거리별 얼굴 영상을 취득한다. 원거리 얼굴 인식 실험을 위해 검증 영상은 고정 카메라로부터 획득된 1m~5m의 거리별 얼굴 영상을 사용하고, 원거리 저해상도 영상의 정규화를 위해 양선형 보간법을 사용한다[11]. 유사도 측정 방법으로는 Euclidean Distance 거리척도 방법을 사용한다[12]. 실험결과, 제안하는 방법은 기존 단일거리 얼굴 영상을 학습으로 이용한 방법에 비해 7.8% 향상된 성능을 보였다. 실제 거리별 얼굴 영상을 학습으로 이용한 방법과 비교했을 때 8% 저하된 성능을 보였지만 사용자 협조가 적게 요구되는 장점이 있다.

본 논문의 구성은 다음과 같다. II장에서는 제안하는 원거리 얼굴 인식 알고리즘과 줌 카메라를 이용한 학습 영상 취득 방법을 설명한다. III장에서 실험결과를 분석하고, 마지막으로 IV장에서 결론을 맺는다.

II. 제안하는 원거리 얼굴 인식 알고리즘

2.1 얼굴 인식 알고리즘

Fig. 1(a)는 제안하는 LDA 기반 원거리 얼굴 인식의 순서도를 나타낸다. 얼굴 인식 알고리즘의 전체적인 흐름은 기존의 LDA 기반 얼굴 인식과 동일하다. 먼저 클래스 간 분산(between-class scatter) 행렬 S_B 는 다음 식 (1)과 같다. 여기서 N_k 는 클래스 k 내의 데이터 수이고, $mean_k$ 는 클래스 k 내에서 평균 영상이다. c 는 클래스 수이다.

$$S_B = \sum_{k=1}^c N_k (mean_k - mean)(mean_k - mean)^T \quad (1)$$

클래스 간 분산(between-class scatter) 행렬 S_W 는 다음 식 (2)와 같다. 여기서 x 는 학습 영상이다. 만약 S_W 가 정칙이라면, S_B 와 S_W 의 비율이 최대가 되는 행렬 W_{opt} 를 다음 식 (3)을 통해 정의할 수 있다.

$$S_W = \sum_{k=1}^c \sum_{x_i \in c_k} (x_i - mean_k)(x_i - mean_k)^T \quad (2)$$

$$W_{opt} = \operatorname{argmax} \left| \frac{W^T S_B W}{W^T S_W W} \right| = [w_1, w_2, \dots, w_m] \quad (3)$$

최적의 투영 W_{opt} 를 계산하기 전에 S_W 가 정칙이 되도록 해야 한다. S_W 가 정칙이 되기 위해 먼저 PCA 방법을 사용하여 특징 벡터의 차원을 N-c로 줄이고, LDA에 의해 c-1까지 차원을 줄인다. 이를 수식으로 표현하면 다음과 같다.

$$\begin{aligned} W_{opt}^T &= W_{lda}^T W_{pca}^T \\ W_{pca} &= \operatorname{argmax} |W^T S_T W| \\ W_{lda} &= \operatorname{argmax} \left| \frac{W^T W_{pca}^T S_B W_{pca} W}{W^T W_{pca}^T S_W W_{pca} W} \right| \end{aligned} \quad (4)$$

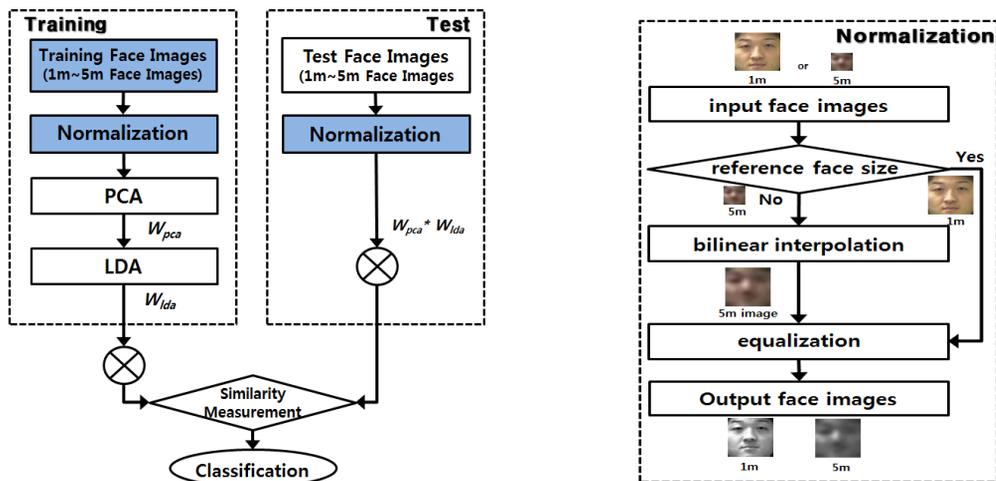
LDA 기반 얼굴 인식은 클래스내 분산 행렬의 특이점에 의해 소표본 문제가 발생하기 때문에 학습 영

상 샘플의 크기가 검증 영상 샘플의 크기보다 크거나 작다면 LDA를 적용할 수 없다. 즉, 다양한 거리에서 추출된 얼굴 영상은 거리에 따라 얼굴 영상 샘플의 차원이 다르게 추출되는 문제가 발생한다. 이러한 문제를 해결하기 위해 본 논문에서는 Fig. 1(b)와 같이 정규화 방법을 사용한다. 거리별 얼굴 영상의 정규화 과정은 다음과 같다. 다양한 크기의 거리별 얼굴 영상이 입력되면 입력 영상은 학습에 사용된 얼굴 영상의 참조 얼굴 크기에 맞게 스케일링 된다. 만약 학습에 사용된 참조 얼굴 크기가 1m를 기준으로 한다면, 참조 얼굴 영상의 크기는 70×70이 된다. 입력된 얼굴 영상의 크기가 70×70이면 다음 단계인 평활화를 하고, 영상의 크기가 70×70보다 작거나 크다면 양선형 보간법을 통해 70×70으로 스케일링을 한다. 그리고 다음 단계인 평활화를 수행한다. 이 과정을 통해 입력되는 모든 얼굴 영상은 현재 참조 얼굴 크기인 70×70으로 정규화 된다.

2.2 줌 카메라를 이용한 거리별 얼굴 영상 획득

얼굴 인식에서 학습에 사용되는 영상의 수는 제한적이고, 많은 양의 영상을 사용하기 위해서는 사용자의 추가적인 협조가 요구된다. 본 논문에서는 사용자 협조가 적은 환경에서 거리별 얼굴 영상을 획득하기 위해 줌 카메라를 사용하는 방법을 제안한다.

Fig.2 는 사용자 등록에 사용되는 거리별 얼굴 영상을 취득하는 방법이다. Fig. 2(a)는 사용자가



(a) face recognition

(b) face size normalization

Fig. 1. The overall flow of proposed LDA- based face recognition

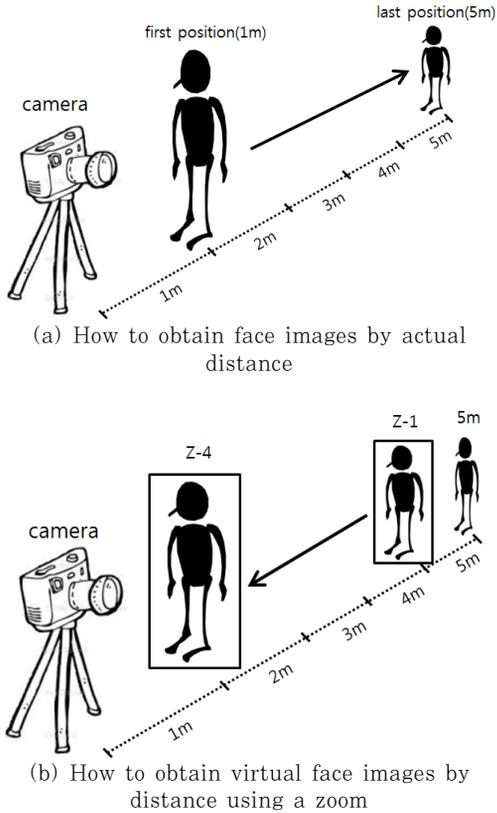
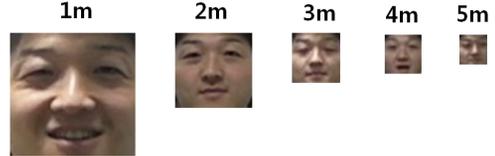


Fig. 2. Acquisition methods of training face images

직접 이동하면서 얼굴 영상을 취득하는 방법이다. 이 방법은 사용자의 협조적인 자세와 단일거리 얼굴 영상 기반 얼굴 인식과 비교하면 시간이 많이 소요되는 단점이 있다. Fig. 2(b)는 줌 카메라를 이용해 거리별 얼굴 영상을 취득하는 방법이다. 제안하는 방법은 사용자가 직접 이동하지 않고, 정해진 위치에 정지해 있기 때문에 단일거리 기반 얼굴 인식과 같은 수준의 사용자 협조만 요구된다.

Fig. 3은 학습 영상 취득 방법에 따라 취득된 얼굴 영상으로 정규화 되지 않은 실제 얼굴 영상이다. Fig. 3(a)는 사람이 직접 1m~5m를 이동하며 취득한 영상이고, Fig. 3(b)는 5m에 정지한 사람으로부터 줌을 이용해 거리별 얼굴 영상을 취득한 영상이다. 두 영상을 비교했을 때, 추출된 얼굴 영상의 해상도가 거리별로 유사함을 확인할 수 있다.



(a) How to obtain face images by actual distance



(b) How to obtain virtual face images by distance using a zoom

Fig. 3. Face images depending on acquisition methods of training face image

III. 실험 및 결과

제안하는 얼굴 인식 방법은 기존에 존재하는 얼굴 DB를 이용해 실험이 어려우므로 본 연구진이 제작한 자체 DB를 사용한다. DB의 구성은 Table 1과 같다. 실험용 DB는 가정용 로봇환경으로 가정하여 한 사람당 150장(1m~5m: 각 30장)의 얼굴 영상을 총 10명으로부터 취득하여 구성되었다. 원본 영상의 크기는 640*480 영상으로 줌이 사용되지 않은 동일한 배율에서 촬영된 영상이다. 실제 거리별 얼굴 영상의 크기는 각각 거리별로 추출된 얼굴 영상의 평균을 분석하여 기준 얼굴 영상 크기로 정했다. 줌 카메라를 통해 획득된 얼굴 영상의 크기는 실제 거리별 얼굴 영상에 기준영상의 크기를 바탕으로 획득했다. 이때, 줌 카메라를 통해 획득된 얼굴 영상은 오직 학습영상으로 사용된다.

본 논문에서는 실험 영상을 기준으로 1m~2m에서

Table 1. Face database

Total persons	10	
Face images per person	150	
Face size by actual distance	- 1m : 50×50	- 2m : 25×25
	- 3m : 16×16	- 4m : 12×12
	- 5m : 10×10	
Face size by distance using a zoom	- Z-4 : 50×50	- Z-3 : 25×25
	- Z-2 : 16×16	- Z-1 : 12×12
	- 5m : 10×10	
Environment of obtained face images	- various lighting change	
	- 1m~5m distance change	
	- face position change	

Table 2. Face recognition depending on training image

Case	Test		
1	Training image per person	: 1m	: 5 images
	Test images per person	: 1m~5m	: 30 images each
2	Training image per person	: 1m~5m	: 1 image each
	Test images per person	: 1m~5m	: 30 images each
3*	Training image per person	: 줌4~5m	: 1 image each
	Test images per person	: 1m~5m	: 30 images each

* Training image is the virtual face image by distance obtained with a zoom

추출된 얼굴 영상은 근거리로, 3m~5m에서 추출된 얼굴 영상은 원거리로 정의한다. 얼굴 인식은 1:1 인종이 아닌 1:N 검색 방법으로써 데이터베이스에 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 추출하기 때문에, 자동 얼굴 검출 방법보다 정교하게 얼굴 영역이 추출된다. 추출된 얼굴 영상은 틀어짐이나 회전은 고려하지 않고 그대로 사용한다.

LDA를 통한 얼굴 인식 방법은 학습에 사용되는 얼굴 영상의 정보나 데이터의 수에 따라 인식률의 변화가 발생하기 때문에 학습 영상의 선택이 매우 중요하다. LDA 기반 얼굴 인식 방법에서 얼굴 인식률을 향상하는 방법은 학습에 사용되는 얼굴 영상 데이터의 수를 증가시키는 방법이 있다. Moon 등은 원거리 얼

굴 인식 성능을 향상하기 위해 실제 거리별 얼굴 영상을 취득해 학습 영상으로 사용한 방법을 사용한다. 본 논문에서는 거리별 얼굴 영상을 취득하기 위해 줌 카메라를 이용한다.

Table 2는 학습 영상의 구성 방법에 따른 얼굴 인식 성능을 분석하기 위한 조건이다. Case1은 기존에 단일거리 영상만을 학습으로 이용한 방법으로 1m에서 추출한 얼굴 영상만을 학습으로 사용한다. 1인당 학습 영상의 수는 5장이다. Case2는 사람이 직접 이동하며 거리별 얼굴 영상을 취득한 경우이며, 1m~5m의 거리별 얼굴 영상을 학습으로 사용한다. 1인당 학습 영상의 수는 각 거리별로 1장씩 총 5장이다. 마지막으로 Case3은 줌 카메라를 통해 거리별 얼굴 영상을 취득한 경우이며, 1인당 학습 영상의 수는 총 5장이다. 이때, 실험인원 10명에 대한 총 학습 영상의 수는 Case1부터 Case3까지 모두 동일하게 50

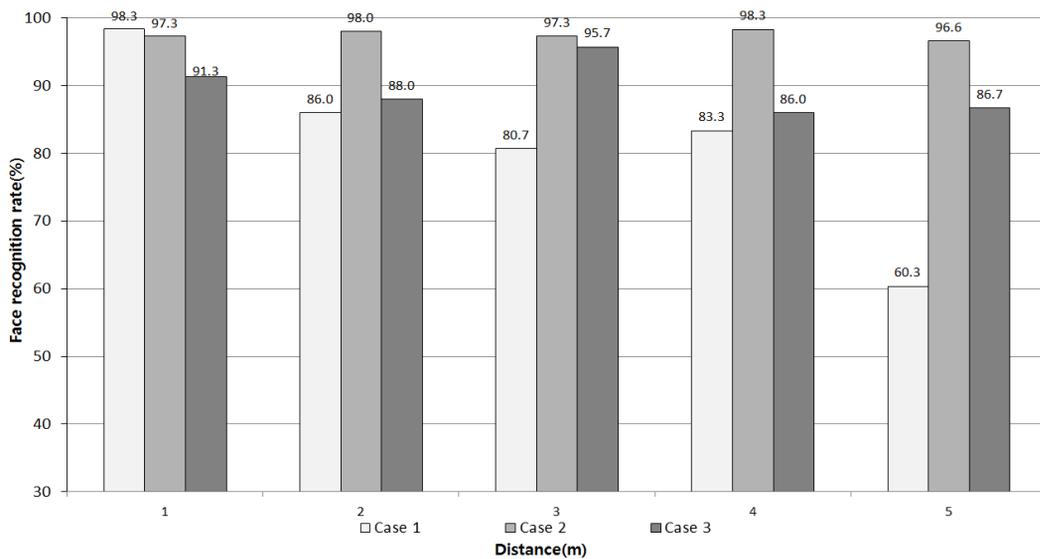


Fig. 4. Face recognition rate depending on training face image

장이다. 검증 영상의 수는 Case1부터 Case3까지 모두 동일하게 1인당 거리별로 각각 30장씩 사용한다. 1인당 총 검증 영상의 수는 150장이고, 10명에 대한 총 검증 영상의 수는 1500장이다.

Fig. 4는 학습 영상 구성에 따른 얼굴 인식을 변화를 나타낸다. 실험결과, 기존 단일거리 얼굴 영상을 학습으로 이용한 방법(Case 1)은 근거리에서 평균 92.2%, 원거리에서 평균 74.8%의 인식을 보였다. 실제 거리별 얼굴 영상을 학습으로 이용한 방법(Case 2)은 근거리에서 평균 97.6%, 원거리에서 평균 97.4%의 인식을 보였다. 마지막으로 제안하는 방법(Case 3)은 근거리에서 평균 89.7%, 원거리에서 평균 89.4%의 인식을 보였다. 전체 평균은 각각 81.7%, 97.5%, 89.5%로 실제 거리별 얼굴 영상을 학습으로 이용한 경우가 가장 우수결과를 나타냈다.

실제 거리별 얼굴 영상을 학습으로 사용한 경우가 가장 인식이 높은 이유는 각 거리별 얼굴 영상과 가장 유사한 특징을 지니고 있는 얼굴 영상을 학습으로 사용하기 때문이다. 이는, 각 거리별 조명 변화 및 거리 변화에 따른 데이터 부족 등과 같은 문제를 포함한 얼굴 영상을 학습으로 사용하기 때문에 동일한 환경에서 입력되는 검증 영상을 인식할 수 있게 된다. 줌 카메라를 이용한 거리별 얼굴 영상은 거리에 의한 영상 데이터 문제는 해결했지만, 거리별 조명 변화 문제를 해결하지 못해 실제 거리별 얼굴 영상을 학습으로 이용한 방법에 비해 인식이 떨어졌다. 결과적으로 제안한 방법은 실제 거리별 얼굴 영상을 이용한 방법에 비해 8% 떨어지는 성능을 보였지만, 줌 카메라를 이용한 방법은 1m~4m의 얼굴 영상을 취득하기 위해 별도의 사용자 협조가 요구되지 않는 장점이 있다.

IV. 결 론

본 논문에서는 가정용 로봇 및 감시카메라 시스템 환경에 적용 가능한 LDA 기반 원거리 얼굴 인식 방법을 제안한다. 제안하는 방법은 기존의 사용자 등록 과정에 발생하는 불편함을 개선하기 위해 줌 카메라를 통해 거리별 얼굴 영상을 취득하고, 취득된 거리별 얼굴 영상을 학습 영상으로 이용한다. 실험결과, 제안하는 방법은 기존 단일거리 얼굴 영상을 학습으로 이용한 방법에 비해 7.8% 향상된 성능을 보였다. 실제 거리별 얼굴 영상을 학습으로 이용한 방법과 비교했을 때 8% 저하된 성능을 보였지만 사용자 협조가 적게 요구되는 장점이 있다. 향후에는 줌 카메라를 통해 거

리별 얼굴 영상을 취득하는 과정의 고속 및 자동화를 위해 하드웨어 설계를 할 것이다.

References

- [1] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, vol. 83, no.5, pp. 705-741, May 1995.
- [2] D.H. Kim, "A vision-based user recognition in robot environments by using semi-biometrics," Ph.D. Thesis, Pusan National University, Feb. 2009.
- [3] H.M. Moon, S.H. Chae, D.S. Moon, Y.W. Chung, and S.B. Pan, "Intelligent video surveillance system using two-factor human information," *Telecommunication Systems*, vol. 52, no. 4, pp. 2249-2257, Apr. 2013.
- [4] K.D. Ban, J.Y. Lee, D.H. Kim, J.H. 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.
- [5] L. Wiskott, J.M. Fellous, N. Krueger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 776-779, July 1997.
- [6] M.A. Turk and A.P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.
- [7] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces : Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.
- [8] B.D. Thang, V. Nam, and G. Medioni, "High resolution face sequences from a

- PTZ network camera,” Proceedings of 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, pp. 531-538, Mar. 2011.
- [9] U.S. Park, H.C. Choi, A.K. Jain, and S.W. Lee, “Face tracking and recognition at a distance: A coaxial and concentric PTZ camera system,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 10, pp. 1665-1677, Oct. 2013.
- [10] H.M. Moon and S.B. Pan, “Performance Analysis of Face Recognition by Distance according to Image Normalization and Face Recognition Algorithm,” Journal of The Korea Institute of Information Security & Cryptology, 23(4), pp. 737-742, Oct. 2013.
- [11] R.C. Gonzalez and R.E. Woods, Digital Image Processing, Prentice Hall, 3rd ed., 2008.
- [12] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, John Wiley & Sons, USA, 2004.

〈저자 소개〉



문 해 민 (Hae-Min Moon) 학생회원
 2009년: 조선대학교 공학사
 2010년: 조선대학교 공학석사
 2010년~현재: 조선대학교 정보통신공학과 박사과정
 관심분야: 영상압축, 영상처리, 워터마킹



반 성 범 (Sung Bum Pan) 종신회원
 1999년: 서강대학교 공학박사
 2005년: 한국전자통신연구원 정보보호연구단 생체인식기술연구팀 팀장
 2005년~현재: 조선대학교 전자공학과 교수
 관심분야: 바이오인식, 영상처리, VLSI 신호처리