

텍스처 기술자들을 이용한 이질적 얼굴 인식 시스템

배한별, 이상윤*

Heterogeneous Face Recognition Using Texture feature descriptors

Han Byeol Bae, Sangyoun Lee*

요약 최근 많은 지능형 보안 시나리오 및 범죄수사에서는 사진이 아닌 얼굴 영상과 다수의 정면 사진과의 매칭을 요구한다. 기존의 얼굴 인식 시스템은 이러한 요구를 충분히 충족시킬 수 없다. 본 논문에서는 동일 인물의 스케치와 사진 간의 양식 차이를 줄임으로써, 이질적 얼굴 인식 시스템의 성능을 향상시키는 알고리즘을 제안한다. 제안하는 알고리즘은 텍스처 기술자들(그레이 레벨 동시 발생 행렬, 멀티스케일 지역 이진 패턴)을 통하여 영상의 텍스처 특징들을 각각 추출하고, 이를 바탕으로 고유특징 정규화 및 추출기법을 통해 변환 행렬을 생성하게 된다. 이렇게 생성된 벡터들 간 계산된 스코어 값은 스코어 정규화 방식들을 통하여 최종적으로 스케치 영상의 신원을 인식하게 된다.

Abstract Recently, much of the intelligent security scenario and criminal investigation demands for matching photo and non-photo. Existing face recognition system can not sufficiently guarantee these needs. In this paper, we propose an algorithm to improve the performance of heterogeneous face recognition systems by reducing the different modality between sketches and photos of the same person. The proposed algorithm extracts each image's texture features through texture descriptors (gray level co-occurrence matrix, multiscale local binary pattern), and based on this, generates a transformation matrix through eigenfeature regularization and extraction techniques. The score value calculated between the vectors generated in this way finally recognizes the identity of the sketch image through the score normalization methods.

Keywords : ERE(Eigenfeature Regularization and Extraction), GLCM(Gray Level Co-occurrence Matrix), HFR(Heterogeneous Face Recognition), MLBP(Multi-scale Local Binary Pattern), score level fusion

1. 서론

이질적 얼굴인식(Heterogeneous Face Recognition, HFR)에 관한 연구는 지난 10년 동안 꾸준히 연구되어 오고 있다. 이러한 연구는 대부분 viewed 스케치(전문가들이 얼굴 사진을 옆에 놓고 따라 그리는 방식)를 사용하여 얼굴 사진과의 양식(modality)의 차이를 줄이는 해결책을 제시하면서 시작 되었다. 이후로 근적외선(NIR), 열적외선(thermal IR), forensic 스케치(목격자의 진술을 토대로 그려진

스케치)영상과 같은 추가적인 다른 양식들을 통해 연구는 지속되어 오고 있다. ([그림 1] 참조)

이질적 얼굴인식의 핵심은 동일 인물 내에 서로 다른 양식들로 인해 발생하는 차이를 최대한 줄이고자 하는 것이며, 이러한 차이를 줄이고자 연구되어왔던 주요한 방법은 크게 합성방법[1],[2],[3],[4],[5],[6],[7]과 영상의 특징에 기반한 방법[8],[9],[10],[11],[12],[13]이 제시되었다.

*Corresponding Author : Department of Electrical and Electronic Engineering Yonsei University(syleee@yonsei.ac.kr)
 Received June 14, 2021 Revised June 14, 2021 Accepted June 15, 2021

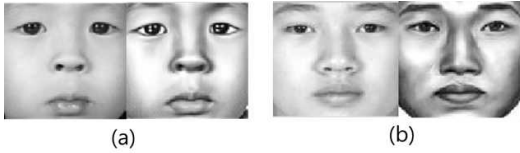


그림 1. 연세대학교 생체인식연구센터(BERC) 데이터베이스 (a) 동일 인물의 사진과 viewed 스케치, (b) 동일 인물의 사진과 forensic 스케치

Fig. 1. Yonsei University Biometrics Engineering Research Center (BERC) database (a) Photo and viewed sketch of same person, (b) photo and forensic sketch of same person

본 논문에서는 동일 인물의 스케치(viewed, forensic)와 사진 간의 양식 차이를 줄임으로써, 얼굴 인식 성능을 향상시키는 알고리즘을 제안한다. 제안하는 알고리즘은 텍스처 기술자로 활용되는 동시 발생 행렬(Gray-Level Co-occurrence Matrix, GLCM)[14]과 멀티스케일 지역 이진 패턴(Multiscale Local Binary Pattern, LBP)[13]을 통하여 영상의 텍스처 특징들을 각각 추출하고, 이를 바탕으로 고유특징 정규화 및 추출기법[15]을 통해 변환 행렬을 생성하게 된다. 이렇게 생성된 벡터들 간 계산된 스코어 값은 스코어 정규화 방식들[16]을 통하여 최종적으로 스케치 영상의 신원을 인식하게 되는 시스템이다.

본 논문은 다음과 같이 기술되어있다. 2절 : 제안하는 알고리즘에 대한 설명, 3절 : 실험 내용 및 구현 결과, 4절: 결론.

2. 본 론

2.1 멀티스케일 지역 이진 패턴

영상의 텍스처 특징을 잘 표현해 내는 기술자 중의 하나인 지역 이진 패턴(Local Binary Pattern, LBP) 기술자에 대해, 우리는 Klare et al.[13]가 제안한 알고리즘에서 언급하였던 멀티스케일 지역 이진 패턴(Multiscale Local Binary Pattern, MLBP)을 사용하였다. 사용하였던 MLBP는 [그림 2], 식(1)에서처럼 영상의 현재 픽셀(g_c)과 비교되는 8개($P=8$)의 주변 이웃 픽셀(g_p)들 사이의 거리 변화($R=1,3,5,7$)를 주어 각각에 대해 균일 패턴(uniform pattern, u_2)을 고려

하여 생성된 히스토그램들을 결합하여 특징벡터로 사용하였다.

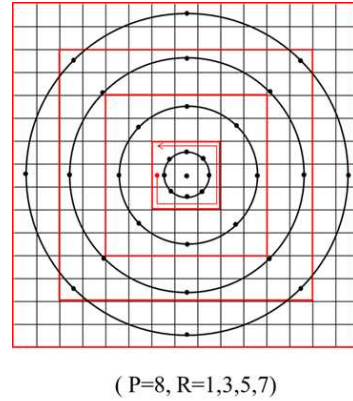


그림 2. 멀티스케일 지역 이진 패턴
Fig. 2. Multiscale local binary pattern

$$LBP_{P,R}^{u_2} = \begin{cases} T(LBP_{P,R}) & \text{if } U(LBP_{P,R}) \leq 2, \\ (P-1)P + 2 & \text{otherwise} \end{cases} \quad (1)$$

$$\text{Where, } U(LBP_{P,R}) = \sum_{k=1}^{P-1} |s(g_k - g_c) - s(g_{k-1} - g_c)|$$

2.2 그레이 레벨 동시 발생 행렬

LBP 기술자와 더불어 텍스처 특징을 잘 표현해 내는 또 다른 기술자로 그레이 레벨 동시 발생 행렬(Gray Level Co-occurrence Matrix, GLCM)이 많은 분야에 응용되어 사용된다. GLCM 기술자는 양자화(quantization) 과정을 거친 영상의 현재 픽셀과 그 이웃하는 픽셀의 밝기 값의 관계를 균일성(energy), 엔트로피(entropy), 대비(contrast), 동질성(homogeneity), 공관계(correlation)등과 같은 특징들을 추출하여 사용하게 된다. Verma et al.[14]은 이러한 픽셀들 간의 관계를 규정짓는 다양한 거리와 방향들을 적용하여 GLCM들을 생성해 냈다. 이를 통해 추출된 특징 벡터들을 적절한 조합을 찾아 실험에 사용하였다. 우리는 그레이 영상의 픽셀 값들을 32 레벨로 양자화 과정을 수행한 후 비교할 픽셀 간의 관계를 다양한 방향($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$)과 거리($d=2$ 고정)를 사용하여 GLCM들을 생성해 내었다. GLCM들로부터

추출 한 특징들로는 식 (2)~(6)을 사용하였으며, 이를 하나의 열벡터로 연결하여 최종 특징벡터로 실험에 사용하였다. (그림 3) 참조)

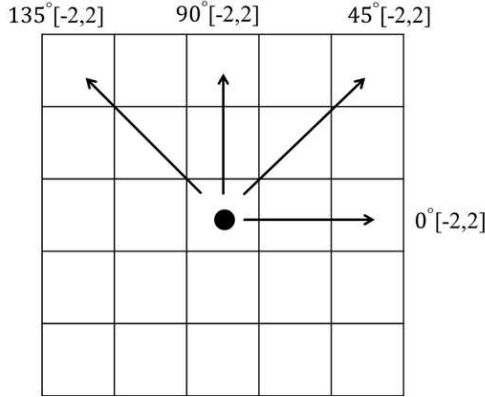


그림 3. 동시 발생 행렬에 적용한 거리 및 방향들
Fig. 3. Distances and directions applied to the co-occurrence matrix

$$\text{Energy (ene)} = \sum_i \sum_j g_{ij}^2 \quad (2)$$

$$\text{Entropy (ent)} = - \sum_i \sum_j g_{ij} \log_2 g_{ij} \quad (3)$$

$$\text{Contrast (con)} = \sum_i \sum_j (i-j)^2 g_{ij} \quad (4)$$

$$\text{Homogeneity (hom)} = \sum_i \sum_j \frac{1}{1+(i-j)^2} g_{ij} \quad (5)$$

$$\text{Correlation (cor)} = \frac{\sum_i \sum_j (ij)g_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (6)$$

2.3 고유특징 정규화 및 추출기법

텍스처 기술자들을 통해 생성된 학습 데이터베이스의 특징 벡터들은 고유특징 정규화 및 추출 기법 (Eigenfeature Regularization and Extraction, ERE)을 통해 최종적인 변환 행렬(projection matrix)을 생성하게 된다. ERE는 특징 차원 축소화 동시에 불안정성, 과적합, 일반화 실패 문제를 경감하는 알고리즘 [15]으로써 클래스 내 흠뜨림 행렬(within-class scatter matrix)의 고유공간(eigenspace)을 세 개의 부공간으로 분할하게 된다. 식 (7)과 같이 얼굴특징영역, 잡음 영역, 널 영역으로 나누어, 각 영역에 해당하

는 정규화된 고유값($\tilde{\lambda}_k^w$: eigenvalue)을 아래와 같이 계산한다.

$$\tilde{\lambda}_k^w = \begin{cases} \lambda_k^w, & k < m \text{ (얼굴특징영역)} \\ \frac{\alpha}{k + \beta}, & m \leq k \leq r \text{ (잡음 영역)} \\ \frac{\alpha}{r + 1 + \beta}, & r < k \leq n \text{ (널 영역)} \end{cases} \quad (7)$$

Where, $m = \arg \max \{ \forall \lambda_k^w < (2\lambda_k^w - \lambda_r^w) \} - 1$

$$\alpha = \frac{\lambda_1^w \lambda_m^w (m-1)}{\lambda_1^w - \lambda_m^w} \quad \beta = \frac{m \lambda_m^w (m-1)}{\lambda_1^w - \lambda_m^w}$$

다음으로 학습 데이터베이스를 가중화된 고유공간으로 변환시키고 변환시킨 데이터의 전체 흠뜨림 행렬 (total mixture scatter matrix)을 구한다. 전체 흠뜨림 행렬에서 차원 축소를 통해 선택된 고유벡터들은 앞서 구한 가중화된 고유벡터와의 곱을 통해 최종 변환 행렬을 생성하게 된다.

2.4 스코어 레벨 결합

학습 및 검증 데이터베이스는 서로 다른 특징 기술자 MLBP, GLCM에 의해 각각 독립적으로 영상에서 특징 벡터들을 추출 하여 사용 하였다. 따라서, 검증 데이터베이스의 스케치와 사진 사이에 유클리디안 거리 (euclidean distance) 연산을 통해 나온 각각의 스코어 값은 결합하기 이전에 공통된 영역으로 값들을 정규화 시켜야 한다. 이에 Jain et al.[16]은 스코어 정규화 방식에 대한 다양한 방법들을 정리하였다. 본 논문에서는 식 (8), (9)에 나와있는 “min-max”, “median and MAD(Median Absolute Deviation)” 스코어 정규화 방법들을 사용하였다.

[Min-max normalization]

$$S_k^c = \frac{S_k - \min}{\max - \min} \quad (8)$$

[median and MAD]

$$S_k^c = \frac{S_k - \text{median}}{MAD} \quad (9)$$

Where, $MAD = \text{median}(|S_k - \text{median}|)$

3. 실험 결과 및 분석

3.1 실험 구현

[그림 4]는 본 연구에서 제안하는 얼굴 스케치 인식 시스템의 전체 개요이다. 실험에 쓰인 데이터베이스는 눈의 위치(양안 거리)를 기반으로 영상 크기를 90×90 으로 기하학적으로 정규화 하였으며, 데이터베이스는 앞서 [그림 1]에 제시한 연세대학교 생체인식연구센터(BERC)에서 자체 제작된 데이터베이스가 사용되었다. BERC 데이터베이스의 경우, 크게 두 개의 데이터베이스로 나뉘어져 있다. 첫 번째 데이터베이스는 390쌍의 viewed 스케치와 사진으로 구성되어 있으며, 두 번째 데이터베이스는 157쌍의 forensic 스케치와 사진으로 구성 되어있다. 데이터베이스의 성능 신뢰도를 높이기 위해 K-묶음 교차 검증법을 사용하였으며, 본 논문에서는 5-묶음 교차 검증법을 사용하였고, 5차례 반복 실험을 통해 진행 하였다.

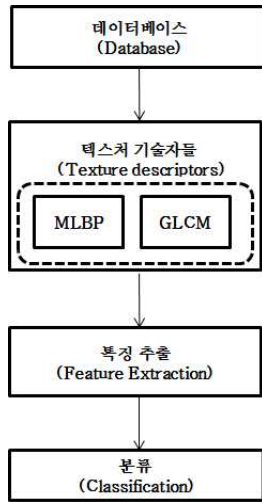


그림 4. 제안 알고리즘 개요
Fig. 4. Overview of the proposed algorithm

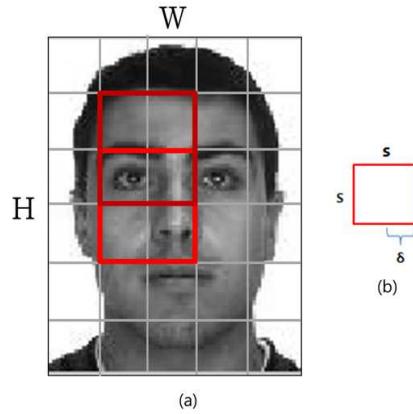


그림 5. (a) 패치가 적용된 영상 (b) 적용한 패치
Fig. 5. (a) image applied to the patch (b) applied patch

생성되는 MLBP의 특징 벡터의 경우 [그림 5]와 같이 패치 단위의 윈도우가 영상 전체를 슬라이딩(sliding)해가며 국소 영역의 텍스처 특징들을 추출해 내게 된다. 영상 하나에서 식 (10), (11)에 따라 $M \times N$ 개의 패치를 구해낼 수 있다.

$$M = ((H - S) / \delta) + 1 \quad (10)$$

$$N = ((W - S) / \delta) + 1 \quad (11)$$

따라서, 하나의 패치로부터 MLBP 경우 균일 패턴 59개에 다양한 픽셀간 거리($R = 1, 3, 5, 7$)를 결합한 236차원의 특징 벡터를 구할 수 있으며, 최종적으로 영상 하나에서 $236 \times M \times N$ 차원의 열벡터가 생성 되게 된다. 또한 GLCM의 특징 벡터의 경우는 영상 전체에서 픽셀간 다양한 방향($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$)에 따른 GLCM들을 생성하여 각각 5개의 특징들(energy, entropy, contrast, homogeneity, correlation)을 생성하게 된다. 이를 하나의 열벡터로 만들어 최종적으로 20 차원의 열벡터를 생성하게 된다. 이를 바탕으로 ERE를 통해 변환 행렬을 생성하게 되며, 이렇게 생성된 벡터들 간 계산된 스코어 값은 스코어 레벨 결합 방식들을 통하여 최종적으로 스케치 영상의 신원을 인식하게 된다. 제안한 알고리즘 성능 평가를 위해서 Klare et al.[13]의 국소 특징에 기반한 분별력 있는 분석법(Local Feature-Based Discriminant Analysis,

LFDA)과 비교 하였다.

3.2 실험 결과

표 1. BERC viewed 스케치 데이터베이스의 상위 순위 변화에 따른 인식률(%), F-: 텍스처 기술자들

Table 1. Recognition rate (%) according to the rank of the BERC viewed sketch database, F-: Texture descriptors

"min-max normalization"				
방법	Rank-1	Rank-10	Rank-20	Rank-30
LFDA	86.82	96.61	97.64	98.30
F-ERE	97.64	99.58	99.94	100
"median and MAD normalization"				
방법	Rank-1	Rank-10	Rank-20	Rank-30
LFDA	94.76	99.02	99.69	99.89
F-ERE	98.87	99.79	99.84	100

표 2. BERC forensic 스케치 데이터베이스의 상위 순위 변화에 따른 인식률(%), F-: 텍스처 기술자들

Table 2. Recognition rate (%) according to the rank of the BERC forensic sketches database, F-: Texture descriptors

"min-max normalization"				
방법	Rank-1	Rank-10	Rank-20	Rank-30
LFDA	8.76	52.60	75.87	85.72
F-ERE	13.65	68.19	90.47	99.36
"median and MAD normalization"				
방법	Rank-1	Rank-10	Rank-20	Rank-30
LFDA	8.50	52.59	76.95	87.64
F-ERE	13.24	66.26	89.00	99.23

[표 1], [표 2]에 BERC viewed 스케치 인식 결과와 BERC forensic 스케치 인식 결과를 상위 n순위 변화에 따른 인식률 결과를 제시하였다. 여기서 상위 n순위 인식률이란 검증 데이터베이스의 probe(스케치 영상)와 gallery(사진 영상)간 유사도가 높은 n개의 후보군 안에 속 할 확률이다. 상위 순위 변화에 따른 인식률 성능 비교를 위해서 다른 조건들은 동일하게 적용시켰다. 패치 크기는 30×30으로 선택하였고, 고유 에너지(고유 값의 누적 합)는 90% 될 때까지 정하였다. 또한 실험에 쓰인 스코어 정규화 방법들 각각에 대한 결과도 포함하였

다. [표 1], [표 2]에 제시되어 있는 실험 결과 모두 텍스처 기술자들을 바탕으로 ERE를 적용한 F-ERE의 결과가 LFDA 알고리즘을 적용한 결과 보다 성능이 모두 우수 하였다. 상위 순위가 증가할수록 인식 성능이 향상된 이유는 상위 순위가 증가할수록 probe(스케치 영상)에 대응하는 gallery(사진 영상)가 있을 확률이 높아지기 때문이다. 뿐만 아니라 forensic 스케치의 인식 성능이 전반적으로 viewed 스케치의 인식 성능 보다 현저히 떨어지는 이유는 forensic 스케치는 viewed 스케치처럼 사진을 옆에 놓고 그리는 방식이 아니라 목격자의 진술만의 토대로 그리는 방식이기 때문에 사진에 대한 객관적이고 심세한 표현이 떨어지기 때문이다. 그러므로 양식의 차이를 줄이는 것도 어려울 뿐 만 아니라 실제 인식 또한 어려운 문제점이 있다.

4. 결론

본 논문에서는 스케치 영상과 사진 간 발생하는 양식의 차이를 줄임으로써 스케치의 인식 성능을 개선시키는 방법에 대해 제안하였다. 서로 다른 양식을 가진 사진과 스케치 영상의 특징들을 각 텍스처 기술자들인 MLBP와 GLCM 각각에 대하여 공통의 특징 공간으로 특징 벡터들을 재 표현 하게 된다. 이는 사진과 스케치 간의 직접적인 접근을 가능케 만들었으며, 이렇게 생성된 특징 벡터들은 ERE 기법에 따라 고유 특징을 고유 스펙트럼에 기반하여 서로 다르게 정규화 시킴과 동시에 차원 수를 줄임으로써 최종적인 분별력 있는 특징 정보를 담게 하였다. 더불어 검증 데이터베이스 벡터들 간 계산되어 나오는 스코어 값들을 스코어 정규화 방법을 통하여 실험에 적용 하였다. 본 연구에 이어 다양한 사진이 아닌 영상들에 대해 효과적으로 인식할 수 있는 이질적 얼굴 인식 시스템을 모색할 것이다.

REFERENCES

[1] Tang, Xiaoou, and Xiaogang Wang. "Face sketch recognition." Circuits and Systems for Video Technology, IEEE Transactions on 14.1 (2004): 50-57.
 [2] Liu, Qingshan, et al. "A nonlinear approach for face sketch synthesis and recognition."

Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.

[3] Wang, Xiaogang, and Xiaoou Tang. "Face photo-sketch synthesis and recognition." Pattern Analysis and Machine Intelligence, IEEE Transactions on 31.11 (2009): 1955-1967.

[4] Tu, Ching-Ting, and J-JJ Lien. "Facial sketch synthesis using direct combined model." Multimedia and Expo (ICME), 2010 IEEE International Conference on. IEEE, 2010.

[5] Gao, Xinbo, et al. "Face sketch synthesis algorithm based on E-HMM and selective ensemble." Circuits and Systems for Video Technology, IEEE Transactions on 18.4 (2008): 487-496.

[6] Zhang, Wei, Xiaogang Wang, and Xiaoou Tang. "Lighting and pose robust face sketch synthesis." Computer Vision-ECCV 2010. Springer Berlin Heidelberg, 2010. 420-433.

[7] Li, Yung-hui, Marios Savvides, and Vijayakumar Bhagavatula. "Illumination tolerant face recognition using a novel face from sketch synthesis approach and advanced correlation filters." Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on. Vol. 2. IEEE, 2006.

[8] Liao, Shengcai, et al. "Heterogeneous face recognition from local structures of normalized appearance." Advances in Biometrics. Springer Berlin Heidelberg, 2009. 209-218.

[9] Klare, Brendan, and Anil K. Jain. "Heterogeneous face recognition: Matching NIR to visible light images." Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010.

[10] Yi, Dong, et al. "Partial face matching between near infrared and visual images in mbgc portal challenge." Advances in Biometrics. Springer Berlin Heidelberg, 2009. 733-742.

[11] Lei, Zhen, and Stan Z. Li. "Coupled spectral regression for matching heterogeneous faces." Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009.

[12] Bhatt, Himanshu S., et al. "On matching sketches with digital face images." Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on. IEEE, 2010.

[13] Klare, Brendan F., Zhifeng Li, and Anil K. Jain. "Matching forensic sketches to mug shot photos." Pattern Analysis and Machine Intelligence, IEEE Transactions on 33.3 (2011): 639-646.

[14] Verma, Manisha, and Balasubramanian Raman. "Center symmetric local binary co-occurrence pattern for texture, face and bio-medical image retrieval." Journal of Visual Communication and Image Representation 32 (2015): 224-236.

[15] Jiang, Xudong, Bappaditya Mandal, and Alex Kot. "Eigenfeature regularization and extraction in face recognition." Pattern Analysis and Machine Intelligence, IEEE Transactions on 30.3 (2008): 383-394.

[16] Jain, Anil, Karthik Nandakumar, and Arun Ross. "Score normalization in multimodal biometric systems." Pattern recognition 38.12 (2005): 2270-2285.

저자약력

배 한 별(HanByeol Bae)

[정회원]



〈관심분야〉

- 2010년 연세대학교 전기전자공학과 학사 졸업
 - 2015년 연세대학교 전기전자공학과 석사 졸업
 - 2020년 연세대학교 전기전자공학과 박사 졸업
 - 현재 : 연세대학교 박사 후 연구원
- 얼굴 인식, 패턴인식, 이미지 변환, 이미지 분류 등

이 상 윤(Sangyoun Lee)

[정회원]



- 1987년 연세대학교 전기전자공학과 학사 졸업
- 1989년 연세대학교 전기전자공학과 석사 졸업
- 1999년 조지아공과대학 전기 및 컴퓨터 공학과 박사
- 현재 : 연세대학교 전기전자공학 교수

〈관심분야〉

얼굴인식, 패턴인식, 비디오 코덱 등