

레터논문 (Letter Paper)

방송공학회논문지 제25권 제1호, 2020년 1월 (JBE Vol. 25, No. 1, January 2020)

<https://doi.org/10.5909/JBE.2020.25.1.105>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

보안시스템을 위한 실시간 저해상도 얼굴 인식 알고리즘

권 오 설^{a)*}

Real-time Low-Resolution Face Recognition Algorithm for Surveillance Systems

Oh-Seol Kwon^{a)*}

요 약

본 논문은 초고해상도 기법을 이용한 실시간 저해상도 얼굴 인식 시스템을 제안한다. 기존의 비대면 얼굴인식은 거리에 따라 해상도가 저하되면서 얼굴인식의 성능이 저하되는 한계가 있다. 이러한 문제를 해결하기 위해서 초고해상도 기법에 대한 연구도 진행되었으나 비대면 얼굴인식 전 과정에 대한 통합적인 설계에 관한 연구는 미흡하다. 제안한 비대면 얼굴인식은 저해상도 영상으로 키프레임 검출, 얼굴검출, 초고해상도 기법, 특징추출 및 얼굴인식 결과까지 약 2초 이내에 수행함으로써 먼 거리에서도 비대면 얼굴인식의 성능을 향상하였다. 다양한 형태의 영상에 대한 실험을 통해 제안한 방법은 기존 방법에 비해 실시간 및 성능측면에서 저해상도 얼굴 인식이 우수함을 확인하였다.

Abstract

This paper presents a real-time low-resolution face recognition method that uses a super-resolution technique. Conventional face recognition methods are limited by low accuracy resulting from the distance between the camera and objects. Although super-resolution methods have been developed to resolve this issue, they are not suitable for integrated face recognition systems. The proposed method recognizes faces with low resolution using key frame selection, super resolution, face detection, and recognition on real-time processing. Experiments involving several databases indicated that the proposed algorithm is superior to conventional methods in terms of face recognition accuracy.

Keyword : Face recognition, CCTV, and Surveillance

a) 창원대학교 전기전자제어공학부(School of Electronical Electronics and Control Engineering, Changwon National University)

* Corresponding Author : 권오설(Oh-Seol Kwon)

E-mail: osk1@changwon.ac.kr

Tel: +82-55-213-3669

ORCID: <http://orcid.org/0000-0002-1077-9615>

* This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning No.2019R1F1A1058489) and LINC+ Program funded by the Ministry of Education.

· Manuscript received October 7, 2019; Revised November 15, 2019; Accepted November 15, 2019.

I . Introduction

최근 딥러닝 기반의 얼굴 인식 기법들이 연구되면서 인식의 정확도 향상에 초점을 맞춘 연구들이 많이 진행되고 있다. 일부 데이터를 기반으로 한 조사에서는 얼굴 인식의 성능이 99%에 이르렀다는 내용도 발표되었다. 그러나 이러한 연구에도 불구하고 이 결과들을 통해 실제 시스템으로 구현하는 것은 별도의 과정이 요구된다. 일반적으로 얼굴 인식의 성능을 저하시키는 요인으로는 전처리과정인 정면화(Frontalization) 및 정렬(Alignment)과 핵심요소인 자세(Pose), 표정(Expression), 조명(Illuminant) 등의 변수들에 의해서 인식의 성능은 매우 밀접하게 연관되어 결정된다. 특히 CCTV를 포함한 대부분의 보안 시스템에서와 같이 획득된 영상의 거리가 멀어짐에 따른 저해상도 영상이 입력되는 경우에는 정보의 손실이 많기 때문에 때로는 인식 자체가 불가능해질 수 있다. 이를 해결하기 위해서 저해상도 영상의 초고해상도화 기법을 통한 얼굴인식 향상 기법이 연구되었다. 기존의 저해상도 영상의 초고해상도 기법(Low-Resolution Face Recognition, LRFR)으로는 양선형 보간법(Bicubic)을 포함하는 주변화소를 이용한 해상도 확대 기법이 있다. 최근에는 딥러닝 기반의 초고해상도 기법들이 연구되고 있다. Yang^[1]은 희소표현에 기반한 사전지식을 통해 초고해상도화를 제안하였으며, Dong et al.^[2]은 딥(Deep) 구조 형태의 신경망을 통해 초고해상도를 복원하였다. 또한 Li et al.^[3]은 다양한 신경망 알고리즘을 분석함으로써 저해상도의 초고해상도화를 통한 실험적 및 통계적 방법에 대한 결과를 제시하였다. 그러나 이러한 실험들은 기존의 데이터베이스 즉, 이미 준비된 얼굴 영상에서 인식의 성능을 검증하는 형태로 분석하였기 때문에 실제 영상의 입력 및 출력에 해당하는 전 과정을 실시간으로 수

행하는 것에는 한계가 있다. 한편 Moon et al.^[4]은 기존의 다양한 저해상도의 입력에 대한 문제를 해결하기 위해 자동 영상 생성 기법을 이용하여 전체 얼굴 인식 시스템에 적용하는 방법을 제안하였다. 그러나 이 방법 또한 실시간 처리에는 한계가 있다. 본 논문에서는 실시간 저해상도 얼굴 인식의 문제를 해결하기 위하여 CCTV 입력 신호로부터 키프레임 결정 및 초고해상도 기법을 적용함으로써 얼굴 인식 과정에서 실시간이 가능한 얼굴 인식 기법을 제안하고자 한다.

II . 제안한 실시간 저해상도 얼굴 인식 기법

일반적으로 실시간 얼굴인식 과정은 입력영상으로부터 얼굴검출 단계와 특징추출 단계 그리고 데이터베이스와의 매칭을 통한 인식단계 등으로 구성된다. 한편 카메라와의 거리에 따라 해상도의 저하로 인해 얼굴 검출이 어려운 경우에는 전처리과정을 통해서 특징 정보의 복원이 필요하다. 이를 위해서는 저해상도 영상을 초고해상도로 변환하는 과정이 필요하다. 그러나 기존의 저해상도 얼굴인식의 분야는 초고해상도로 변환하는 과정에서 많은 시간이 소요되기 때문에 실시간 처리에 대한 연구가 부족하였다. 이에 본 논문에서는 전처리의 최소화 및 초고해상도 방법의 실시간 기법 적용을 통한 비디오 입력으로부터 실시간이 가능한 저해상도 영상의 얼굴인식 방법을 제안한다. 그림 1은 제안한 실시간 저해상도 얼굴 인식 기법의 흐름도이다. 먼저 비디오 신호로부터 모든 프레임에 대해 얼굴 검출을 수행한다면 실시간 처리가 어렵게 된다. 따라서 키 프레임(Key-frame) 선택을 통해서 프레임간의 일정 이상의 변화가 있을 경우에만 얼굴 검출을 수행한다. 이 과정을 통해서 이전 프

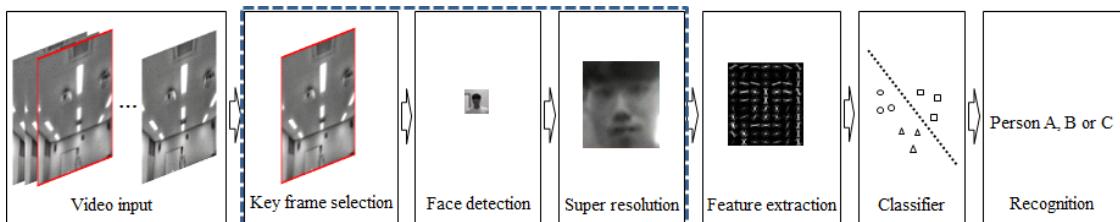


그림 1. 제안한 실시간 저해상도 영상의 얼굴 인식 기법의 흐름도
Fig. 1. Flowchart of the proposed LRFR algorithm

프레임이 얼굴 검출이 필요 없는 경우라면 다음 프레임이 키 프레임이 아닐 경우 다음 프레임 또한 얼굴 검출을 수행하지 않기 때문에 프로그램의 복잡도(Complexity)를 줄일 수 있는 장점이 있다. 다음으로 현재 영상이 키 프레임이라면, 양선형 보간법(Bicubic interpolation)을 통한 전처리 과정을 수행 후 얼굴 검출을 시도한다. 이 과정을 통해 인식이 가능한 저해상도 영상의 유효 거리를 확장할 수 있는 효과가 있다. 얼굴이 검출된 프레임에서는 얼굴 영역만 추출하여 초고해상도 알고리즘을 수행한다. 초고해상도 기법을 적용한 영상은 특징 추출 기법을 통해 얼굴에 대한 특징을 생성하고 이를 분류기에 적용하여 최종적으로 얼굴 인식의 결과를 도출한다.

제안한 방법은 먼저 실시간 처리를 위하여 비디오 신호로부터 키 프레임(Key frame)을 선택한다. 기존의 키 프레임 선택은 영상에서 주요 컨텐츠를 자동적으로 주석을 넣는 형태로 사용되어 왔다^[5]. 본 논문에서는 모든 프레임에 대하여 얼굴 인식과정을 수행하면 효율성이 저하되기 때문에 프레임간의 변화가 큰 프레임으로 정의한다. 좁게는 얼굴 영역에서의 변화가 많을 경우에는 얼굴 인식을 적용함으로써 유사한 프레임의 경우에는 얼굴인식을 수행하지 않음으로써 효율적인 프로그램을 구현할 수 있다. 키 프레임을 선택하는 수식은 다음과 같다. 추가적으로, 기존의 키 프레임 선택은 카메라의 이동, 줌, 흔들림 등이 고려되었으나, 본 연구에서는 다음 수식과 같이 수평과 수직의 속도(Speed)만을 고려함으로써 보다 간편하게 키 프레임을 결정하였다.

$$\bar{V}_x(t) = \frac{1}{W} \sum_{s=t-W/2}^{t+W/2-1} H(s) \quad (1)$$

이때 속도는 정지와 동적 패턴 사이의 차별화를 나타내는 기본적인 특징이다.

$$\bar{V}_y(t) = \frac{1}{W} \sum_{s=t-W/2}^{t+W/2-1} V(s) \quad (2)$$

여기에서 W는 원도우의 길이로써 정의된다. 다음으로 선택된 키 프레임에 대하여 얼굴 검출을 위해서는 얼굴 검출 기법이 필요하다. 사용된 얼굴 검출로는 Viola John 방

법으로써 Haar 특징과 Adaboost 기반의 기계학습을 통해 영상 내에서 특징을 검출하는 방법이다. 다만, 본 논문에서는 매우 작은 해상도에서도 검출이 가능하고자 전처리를 통해 입력 영상에 Bicubic 보간법을 미리 적용하였다. 이를 통해 보다 멀리에서도 얼굴 검출이 가능하도록 하였다. 얼굴 검출이 수행되면 수행된 얼굴 영역만을 잘라내고 이 영역에 대해서 초고해상도 알고리즘을 적용한다.

본 논문에서는 기존의 최신 알고리즘을 성능 및 실시간 처리 측면에서 비교 분석함으로써 최적의 초고해상도 알고리즘을 결정하였다. 최종적으로 본 논문에서 적용된 초고해상도 알고리즘은 희소코딩 기반의 신경망 기법^[6]이다. 이 방법은 비선형 신경 계층을 기반으로 새로운 감쇠 함수를 적용하여 기존의 희소표현 기반의 신경망을 설계하는 방법이다. 더 나아가 객체의 크기 변화에 강인하도록 캐스캐이드 구조를 적용함으로써 보다 빠르게 학습과 초고해상도의 결과가 생성된다. 이 방법은 기존의 활성 함수를 간단히 구현하기 위해서 적당한 문턱치를 갖는 기존의 뉴런은 두 개의 선형 스케일 계층과 하나의 뉴런으로 나누어진다.

$$\begin{aligned} [h_\theta(a)]_i &= \text{sign}(a_i)\theta_i(|a_i|/\theta_i - 1)_+ \\ &= \theta_i h_1(a_i/\theta_i) \end{aligned} \quad (3)$$

두 개의 스케일 계층의 계수는 θ 에 의해 정의된 대각 매트릭스이고 각 요소별로 역수이다. 다음 단계에서 희소 계수가 고화질 사전에 곱해지고 고화질 패치가 생성된다. 제안한 신경망 기법의 학습을 위한 비용함수로 평균 승수 오차 방법을 사용하고 최적화 목표는 다음과 같이 표현된다.

$$\min_{\Theta} \sum_i \| SCN(I_y^{(i)}; \Theta - I_x^{(i)}) \|_2^2 \quad (4)$$

여기에서, $I_y^{(i)}$ 와 $I_x^{(i)}$ 는 저해상도와 고해상도 훈련 데이터의 i 번째 쌍이고 $SCN(I_y^{(i)}; \Theta)$ 는 변수집합 Θ 을 가지는 희소코딩 기반 신경망을 사용하여 예측된 $I_y^{(i)}$ 에 대한 초고해상도 영상이다. 다음으로 초고해상도 기법이 적용된 얼굴 영상에 지역 서술자 기반의 특징 검출과 다양한 분류기의 성능을 평가함으로써 최종적으로 얼굴 인식을 수행하였다.

표 1. 제안한 방법과 기존 방법의 PSNR 결과 비교

Table 1. The performance of PSNR for proposed and conventional algorithms

Zooming	Bicubic	Sparse coding [1]	Deep convolutional network [2]	Proposed
x4	33.15	33.65	32.06	34.32
x9	30.01	30.94	29.06	31.92
x16	30.37	31.48	26.67	31.97

표 2. Extended Yale B에서 SVM 분류기를 이용한 얼굴 인식률 비교

Table 2. Face recognition rate (%) for the Extended Yale B images with SVM classifier

Zooming	LR image	Bicubic	Sparse coding [1]	Deep convolutional nework [2]	Proposed
x4	56.32	74.90	79.14	76.44	89.47
x9	30.03	56.85	66.52	65.11	85.96
x16	21.32	41.03	51.76	54.02	78.71

III. 실험 및 결과

제안한 알고리즘의 성능을 평가하기 위하여 Extended Yale B 데이터베이스를 이용하였다. Extended Yale B 데이터는 다양한 주변 광원에 따른 조건하에서 10개의 객체로 구성된다. 이때 각 객체는 특별한 자세(Pose)와 주변 광원에 의해서 획득된다. 본 논문에서는 해상도에 따라 3 개의 저해상도 영상(1/4, 1/9, 1/16)을 정의하였으며, 총 2,400 장의 테스트 영상을 획득하였다. 표 1은 영상복원 결과에 대한 PSNR의 성능을 비교한 결과이다. 해상도에 관계없이 제안한 방법이 우수한 PSNR의 결과를 나타내었다.

표 1은 기존 방법과 제안한 방법의 초고해상도 영상 복원에 대한 PSNR 결과이다. 이때, 특징 추출 방법은 지역 차분 패턴 방법을 이용하였다. 이 방법은 이웃화소 및 방향성을 이용하는 기법으로 반복적인 방법을 통해 고차의 특징을 쉽게 추출하는 장점이 있다. 제안한 방법은 대부분의 해상도에서 좋은 결과를 나타냄을 확인할 수 있다. 표 2는 Extended Yale B 데이터에서 SVM 분류기를 사용하였을 때 얼굴인식의 성능을 나타내고 있다. 기존 저해상도 영상을 포함한 기존의 방법에 비해 제안한 방법은 가장 우수한 얼굴 인식 정확도를 나타내고 있다.

IV. 결 론

본 논문은 저해상도 비디오 입력으로부터 초고해상도 기

법을 적용한 얼굴인식 성능을 개선하는 방법을 제안하였다. 제안한 방법은 기존의 거리에 따른 저해상도 영상의 얼굴 인식의 성능저하 및 불가능한 상태를 해결하였으며, 실제 시스템에 적용할 수 있는 실시간 처리에 가까운 성능을 나타내었다. 이를 위해 다양한 형태의 데이터베이스의 영상에 대해 실험한 결과 제안한 방법의 성능이 기존의 문제인 거리에 대한 한계를 극복할 수 있음을 확인하였다.

참 고 문 헌 (References)

- [1] J. Yang, J. Wright, T. Huang, and Y. Ma, "Imaging super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp.3467-3478, 2012.
- [2] C. Dong, C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," In *Proceedings of European Conference on Computer Vision*, pp. 184-199, 2014.
- [3] P. Li, L. Prieto, D. Mery, and P. Flynn, "On low-resolution face recognition in the wild: Comparisons and new techniques," *IEEE Trans. on Information Forensics and Security*, vol. 14, no. 8, pp. 2000 - 2012, 2019.
- [4] M. Haemin and P. Sungbum, "Long distance face recognition system using the automatic face image creation by distance," *Journal of The Institute of Electronics and Information Engineers* vol. 51, no. 11, Nov. 2014.
- [5] G. Ivan, M. Tomas, G. Ascension, and D. Fernando, "Temporal segmentation and keyframe selection methods for user-generated video search-based annotation," *Expert systems with Applicaitons*, vol. 42, no. 17, pp. 488 - 502, 2015.
- [6] K. Seol, "Face recognition based on super-resolution method using sparse representation and deep learning," *Journal of Korea Multimedia Society*, vol. 21, no. 2, Feb. 2018.