

결정 트리를 사용한 강인한 얼굴 영역의 획득

장석우
안양대학교 소프트웨어학과

Robust Facial Area Acquisition Based on Decision Tree

Seok-Woo Jang
Department of Software, Anyang University

요약 다양한 실내외의 환경 조건에서 고속의 카메라로 촬영된 컬러 영상으로부터 사람의 얼굴 영역을 방향 및 주요한 특징점을 포함해 정확하게 검출하는 것은 어려운 작업이다. 본 연구에서는 받아들인 컬러 영상으로부터 개인 정보를 나타내는 사람의 얼굴 영역을 방향을 고려하여 강인하게 획득하는 기법을 기술한다. 본 연구에서는 먼저 시스템에 받아들인 컬러 영상으로부터 레티나 얼굴 모델을 사용하여 얼굴 영역을 1차적으로 검출한다. 그런 다음, 랜덤 포레스트를 적용하여 방향 및 주요한 특징점을 고려하여 얼굴 영역을 최종적으로 획득한다. 실험 결과에서는 본 연구에서 제안한 접근 방법이 여러 가지 종류의 영상 콘텐츠로부터 랜덤 포레스트 모델을 기반으로 노출된 사람의 얼굴 영역을 방향을 고려하여 강인하게 획득한다는 것을 제시한다. 본 연구에서 제안된 랜덤 포레스트 기반의 얼굴 영역 획득 알고리즘은 생체 인식을 통한 사용자의 신원 인증, 동영상 콘텐츠에서 얼굴 인식 및 추적, 그리고 실내외 건물의 모니터링과 같은 컴퓨터 비전과 연관 있는 다양한 실생활 분야에서 중요한 요소 기술로 활용될 것으로 기대된다.

Abstract It is difficult to accurately detect a person's face, including directions and feature points, from a color image captured by a high-speed camera in various environmental conditions. In this study, we describe a technique for robustly obtaining the face area of a person that represents personal information in the received color image by considering the direction. A retina face model is used to roughly detect a face region from a color image received by the system. Then, by applying the random forest method, a face region is obtained in consideration of the direction and feature points. Experimental results show that the proposed approach robustly acquires face areas from various types of images. The proposed algorithm could potentially be used in various fields related to computer vision, such as authentication of a user's identity through biometric recognition, face recognition and tracking in videos, and monitoring of indoor and outdoor buildings.

Keywords : Random Forest, Labeling, Estimated Direction, Image Processing, Facial Area

1. 서론

초당 수십 프레임을 촬영할 수 있는 일반적인 평속 카메라와는 달리, 초당 수백에서 심지어는 수천 프레임씩 영상 촬영이 가능한 초고속의 카메라가 이제는 저렴하게 보급되기 시작했다[1]. 따라서 일반 사용자들도 초고속의

영상 데이터를 인터넷을 통해 손쉽게 획득할 수 있게 되었다. 이러한 초고속의 컬러 영상 데이터는 정교한 물체 모델링이나 미세한 동작 인식과 같은 실제적인 산업 분야에서 값어치 있게 사용될 수 있다[2].

이와 같이 여러 가지 종류의 다양한 초고속의 영상 데이터를 일반인들도 손쉽게 획득할 수 있는 환경이 조성

*Corresponding Author : Seok-Woo Jang(Anyang Univ.)

email: swjang@anyang.ac.kr

Received June 15, 2021

Accepted July 2, 2021

Revised July 1, 2021

Published July 31, 2021

되었다는 측면에서는 긍정적이지만, 반대로 부정적인 효과도 나타나고 있다. 다시 말해, 얼굴이나 신체의 일부와 같이 사람의 개인적인 정보를 포함하고 있는 영역들이 그대로 노출되어 있는 초고속의 영상 데이터가 자유롭게 유통되고 있어서 많은 사회적인 문제를 일으키고 있다.

따라서 빠른 속도로 들어오는 초고속의 컬러 영상으로부터 개인 정보[3]를 대표하는 영역들을 정확하게 검출한 다음, 검출된 대상 영역들을 모자이크나 블러링 기법 등을 통해 효과적으로 커버링하는 연구가 필요하다. 그리고 이런 연구를 통해서 사람의 중요한 개인 정보가 본인의 승인 없이 외부로 노출되는 것을 방지할 수 있다.

받아들인 여러 가지 종류의 컬러 영상 콘텐츠로부터 얼굴이나 신체의 노출된 일부와 같이 사적인 정보를 대표하는 부분을 획득하기 위하여 전형적으로 실시되었던 연구는 관련된 참고문헌들에서 조사해 볼 수 있다. [4]에서는 비디오 시퀀스에서 사람의 얼굴을 정확하게 추적하기 위해 얼굴 감지 모델(SENResNet)과 얼굴 추적 모델(RNFT)을 포함하는 답러닝을 기반으로 하는 효율적인 얼굴 감지 및 추적 프레임워크를 제안하였다. 여기에서 얼굴 감지 모델은 각 프레임의 얼굴 정보를 감지하고 대상 얼굴의 위치를 추출하여 얼굴 추적을 위한 초기화 창을 제공한다. 그런 다음 얼굴 추적 모델은 인접 프레임에서 얼굴 특징을 추출하고 다음 프레임에서 대상 얼굴의 위치를 예측한다. 실험결과로는 제안된 모델이 성능 측면에서 기존의 방법보다 우수하다는 것을 보여주었다.

[5]에서는 주의(attention) 메커니즘을 가지는 새로운 감지 모델을 제안하여 폐쇄된 얼굴 영역을 감지하는 방법을 제시하였다. 제안된 방법은 얼굴 인식 작업에서 영상 생성 방법을 사용해 특징을 강화하였다. 그리고 가려진 객체의 간섭을 해결하기 위해 특징 추출을 위한 지역 생성 분기 SG-net을 제안하였다. 다양한 데이터 세트에 대한 실험은 제안된 알고리즘이 보다 넓은 데이터 세트에서 개선된 성능을 보여주며, 자체 라벨링된 모니터링 데이터 세트에서 높은 평균 정밀도를 달성한다는 것을 보여주었다.

[6]에서는 모바일 환경에서 얼굴의 세그먼트를 기반으로 부분적인 얼굴을 검출하는 방법을 제시하였다. 일반적으로 얼굴 검출 방법은 가려져 있고 부분적으로만 보이는 얼굴 영역들이 상당히 있기 때문에 모바일 환경에서 잘 동작하기가 어렵다. 이런 문제를 해결하기 위해서 제안된 방법은 제안(proposal) 기반 검출과 엔드-투-엔드 회귀에 의한 검출이라는 두 가지의 접근 방식을 소개하였다. 실험에서는 제안된 기법이 다른 방법에 비해 우

수함을 보였다.

[7]에서는 정교하게 설계된 컨볼루션 신경망에서 추출한 구분적인 특징을 기반으로 한 빠른 얼굴 검출 방법을 제안하는데, 여기에서 얼굴 검출은 완전한 특성 맵에서 직접 수행된다. 제안된 방법에서의 특징은 빠른 속도와 유망한 성능으로 얼굴을 감지하는데 도움이 되는 크기 불변적인 능력을 보여주었다. 따라서 제안된 방법에서는 기존의 방법에서 사용된 영상 피라미드에서 다중 스케일 특징을 추출할 필요가 없으므로 얼굴 검출의 효율을 크게 향상시킬 수 있다. 앞에서 언급한 다양한 접근 방법들 외에도 컴퓨터 비전 기법을 적용해 사람의 얼굴이나 피부와 같이 노출된 개인 정보를 나타내는 영역들을 강인하게 획득하는 방법들이 지속적으로 소개되고 있다[8].

앞에서 기술한 기존에 나와 있는 방법들은 보통의 환경에서는 비교적 잘 동작하지만 초고속 영상이 아니라 일반 영상을 주요한 대상으로 한다. 그리고 얼굴이 바라보는 방향이 정면이 아니라 좌측 또는 우측 측면 등일 경우에는 검출된 얼굴 영역 내에 위치한 특징점들이 올바른 위치에 있지 않는 경향이 빈번하게 발생한다. 다시 말해, 얼굴의 방향을 고려하지 않아 부정확하게 특징 점들이 검출되는 편이다.

따라서 본 논문에서는 레티나 얼굴(RetinaFace) 모델을 적용하여 받아들인 컬러 영상 안에 들어있는 사람의 얼굴 영역을 1차적으로 추출한 다음, 랜덤 포레스트 모

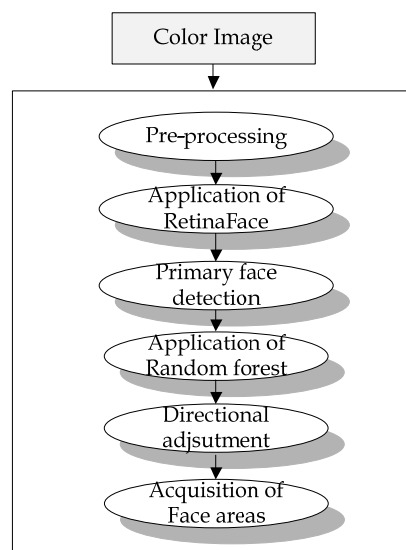


Fig. 1. Overall flow of the suggested algorithm

델을 사용하여 사람의 얼굴이 위치한 방향에 따라 주요한 특징점(landmark)을 조정하여 얼굴 영역을 정확하게 획득하는 알고리즘을 제시한다. Fig. 1은 본 논문에서 기술하고 있는 랜덤 포레스트 모델을 사용한 사람의 얼굴 영역 획득 접근 방법의 기본적인 흐름도를 그림으로 제시하고 있다.

Fig. 1에서 이해할 수 있듯이 본 논문에서 언급한 접근 방법에서는 우선 래티나 얼굴 모델을 사용하여 입력 받은 컬러 영상으로부터 얼굴 영역을 1차적으로 검출한다. 그런 다음, 1차적으로 검출된 영역에 랜덤 포레스트를 적용해 방향 및 주요한 특징점까지 고려하여 얼굴 영역을 정확하게 획득한다.

1장에서는 본 논문을 진행하게 된 기본적인 동기와 개요에 대해 서술하였다. 2장에서는 래티나 얼굴 모델을 적용해 영상으로부터 얼굴 영역을 1차적으로 추출하는 기법에 대해 서술한다. 3장에서는 랜덤 포레스트 기법을 사용해 컬러 영상으로부터 얼굴 영역을 방향을 고려하여 추출하는 기법에 대해 설명한다. 그리고 4장에서는 소개된 랜덤 포레스트 기반의 얼굴 추출 기법의 성능을 정량적으로 확인하기 위해 수행한 실험 결과에 대해 설명한다. 5장에서는 본 논문의 결론과 후속의 연구 계획에 대해 설명한다.

2. 초기의 얼굴 검출

본 연구에서는 빠른 속도로 입력되는 초고속의 영상으로부터 개인 정보를 포함하고 있는 목표 영역을 정확하게 추출하고자 한다. 본 연구에서는 사람의 얼굴 영역을 개인 정보를 대표하는 영역으로 설정한다. 그리고 사람의 얼굴 영역을 검출하기 위해서 다섯 개의 랜드마크(landmark)를 사용하고 있는 래티나 얼굴 모델을 적용한다[9]. 본 연구에서 선택한 사람의 얼굴 영역에 대한 다섯 개의 랜드마크는 두 개의 눈동자 위치, 코끝 위치, 양쪽의 입가 위치에 해당한다.

본 연구에서 적용하는 래티나 얼굴 모델은 단일 단계(stage)의 밀집 얼굴 지역화 기법을 이용하고 있다. 이 기법은 결합 추가(extra) 감독(supervised)과 자율(self) 감독 다중 작업 학습을 기반으로 여러 가지 스케일의 얼굴 영역에 관한 픽셀 기반의 얼굴 검출을 수행한다. 본 연구에서 사용하는 래티나 모델 학습 기법의 중요 실험 과정에 대한 기술은 다음과 같다.

첫째, 다중 작업 손실에 관한 기술이다. 다시 말해, 래

티나 모델은 어떤 훈련 앵커(anchor) i 에 대해 손실이 식 (1)과 같이 최소화된다.

$$L = L_{ds}(p_i, p_i^*) + \lambda_1 p_i^* L_{box}(t_i, t_i^*) + \lambda_2 p_i^* L_{pts}(l_i, l_i^*) + \lambda_3 p_i^* L_{pixel} \quad (1)$$

위의 식 (1)에서 $L_{ds}(p_i, p_i^*)$ 는 얼굴 범주 손실을 의미한다. p_i 는 앵커 i 가 얼굴이 되는 예측된 확률을 의미한다. 그리고 p_i^* 는 긍정적인 앵커의 경우에는 1로 설정되고, 부정적인 앵커에 대해서는 0으로 설정된다. 범주 손실은 이진 클래스에 관한 소프트맥스(softmax) 손실이다. 사람의 얼굴에 대한 회귀 손실인 $L_{box}(t_i, t_i^*)$ 에서 긍정적인 앵커와 관련되어 예측된 박스의 좌표와 그라운드 트루스 박스의 좌표를 나타내는 $t_i = \{t_x, t_y, t_w, t_h\}_i$ 과 $t_i^* = \{t_x^*, t_y^*, t_w^*, t_h^*\}_i$ 가 각각 이용된다. 그리고 얼굴 박스 영역에 관해서는 얼굴의 센터 위치, 가로의 길이와 세로의 길이를 활용해 대상 영역을 정규화한다. 얼굴 영역 회귀 손실인 $L_{pts}(l_i, l_i^*)$ 은 긍정적인 앵커와 연관되어 예상된 다섯 개의 얼굴 랜드마크, 그리고 그라운드 트루스인 $l_i = \{l_{x_1}, l_{x_2}, \dots, l_{x_5}, l_{y_5}\}_i$ 와 $l_i^* = \{l_{x_1}^*, l_{x_2}^*, \dots, l_{x_5}^*, l_{y_5}^*\}_i$ 를 이용한다. 얼굴 박스 영역과 유사하게 다섯 개의 얼굴 랜드마크 회귀 역시 앵커의 센터를 기반으로 한 목표 정규화가 수행된다. 마지막으로 밀집 회귀 손실인 L_{pixel} 은 손실 균형 파라미터인 $\lambda_1 - \lambda_3$ 를 0.25, 0.1, 0.01로 각각 설정한다. 이것은 감독 신호에서 보다 좋은 상사 및 랜드마크 위치의 중요성을 높인다는 의미를 나타낸다.

둘째, 밀집 회귀 분기에 관한 기술이다. 본 연구에서 사용하는 래티나 얼굴 모델은 국소화되고 빠른 스펙트럴 필터링을 기반으로 한 그래프 콘볼루션 방법을 사용해 매쉬(mesh) 디코더를 완성한다. 그리고 추가적인 가속을 위해서 결합 모양 및 텍스처 디코더를 채택한다. 또한, 전형적인 콘볼루션과는 틀리게 그래프 콘볼루션은 유클리드 그리드 수신 필드를 이웃과의 거리를 추출하여 이용한다. 즉, 두 개의 정점을 연결하는 최소 개수의 에지를 그래프 상에서 세어서 거리를 계산한다. 그리고 래티나 모델에서 그래프 정점(vertex)은 결합 모양 및 텍스처 정보를 담고 있는 색상 얼굴 매쉬를 이용해 생성되며, 희소 인접(sparse adjacency) 행렬을 사용하여 정점 사이의 연결 상태가 인코딩된다.

셋째, 영상 화질의 구분에 관한 기술이다. 본 연구에서 제시하는 래티나 모델에서는 다양한 종류의 얼굴 영

상을 영상의 품질을 바탕으로 다섯 가지의 종류로 나눈다. 본 연구에서 영상의 품질을 나누는 기법은 사람의 얼굴 영역 위에 다섯 개의 랜드마크를 표시하는 것이 얼마나 어려운지를 바탕으로 수행된다.

본 논문에서는 앞에서 기술한 처리를 사용해 훈련된 모델을 초고속의 카메라로 촬영된 영상에 통합하여 이용한다. 다시 말해, 일반적인 환경에서 초고속의 카메라로 촬영된 영상은 비 균일한 조명, 예기치 않게 포함되는 잡음으로 인해 품질이 좋지 못할 수 있다. 하지만 래티나 모델은 얼굴의 범주, 밀집도 조사, 랜드마크 간의 그래프 개념이 사용된 기법을 채택하여 품질이 안 좋은 상황에서도 상대적으로 좋은 성능을 발휘할 수 있다.

3. 랜덤 포레스트 모델

본 연구에서 사용한 래티나 모델을 사용하여 얼굴일 확률이 아닌 랜드마크의 관점에서 접근해 보면 부분적으로 정확도가 높지 않은 경우도 확인할 수 있다. 다시 말해, 이 모델에서 학습된 랜드마크의 위치를 확인했을 때 올바른 위치가 아닌 다른 위치에 존재하는 것을 일부 확인할 수 있다. 이는 초고속 카메라의 잡음으로 인해 랜드마크가 올바르게 표시되지 않은 지점에 위치하기도 한다. 또한, 얼굴의 랜드마크가 보이지 않는 각도에서 촬영된 영상임에도 래티나 모델을 적용할 경우 얼굴의 랜드마크가 표시되어 있는 것을 확인할 수도 있었다.

본 연구에서는 우선 정확하지 않은 얼굴의 랜드마크를 보정하기 위해서 얼굴의 형태를 구분할 수 있는 모델을 생성한다. 여기에서 생성할 구분 모델은 추출된 얼굴 영역을 입력으로 받은 다음 얼굴의 방향을 출력해 주도록 한다. 본 연구에서 사용하는 구분 모델은 영상 분류에 많이 이용되는 랜덤 포레스트(random forest)[10]인데, 그 정의는 다음과 같이 기술될 수 있다.

보통 랜덤 포레스트의 기본 모델인 결정 트리(decision tree)는 학습이 수행될 때 입력되는 데이터의 특징을 기반으로 분기가 발생한다. 그런데 만일 입력 데이터의 특징이 매우 많으면 분기가 빈번하게 발생하므로 오버피팅(overfitting)이 일어난다.

따라서 오버피팅을 효과적으로 극복하기 위하여 가지치기(pruning)라고 알려진 방법과 엔트로피(entropy)라는 개념을 이용한다. 일반적으로 가지치기는 분기가 발생하는 최대 깊이, 터미널 노드의 최대 개수, 또는 하나의 노드가 분기하는데 필요한 최소 데이터의 개수를 한

정하는 것이다. 그리고 엔트로피는 불순도라는 개념과 연관이 있는데, 불순도는 해당하는 범주 내에 동일하지 않은 데이터가 얼마나 혼합되어 있는가를 의미한다. 다시 말해, 엔트로피는 불순도를 수량적으로 표현한 척도로서 식 (2)와 같이 정의될 수 있다.

$$Entropy = -\sum_i (p_i) \log(p_i) \quad (2)$$

식 (2)에서 p_i 는 하나의 영역 내에 있는 데이터 중에서 부류 i 에 포함되는 데이터의 비율을 나타낸다. 보통 엔트로피가 크다는 의미는 불순도가 높다는 것을 나타내는데, 하나의 부류 내에 동일하지 않은 데이터가 많이 존재한다는 의미이다.

결정 트리 기법은 보통 정보 획득(gain)이 최대화되는 방향으로 학습이 수행된다. 다시 말해, 이 알고리즘은 받아들인 영상이 어떤 특징의 어떤 분기점에서 정보 획득 값이 최대치가 되는지를 판단하여 분기가 수행된다. 본 연구에서 사용하는 정보 획득은 식 (3)과 같이 정의될 수 있다.

$$Information\ gain = entropy(parent) - weighted\ average \times entropy(children) \quad (3)$$

식 (3)에서 $entropy(parent)$ 는 분기 이전의 엔트로피 값이고, $entropy(children)$ 는 분기 이후의 엔트로피 값이다. 그리고 결정 트리에서 분기를 수행하면 범주가 두 개 이상으로 분할되므로 분기 이후의 엔트로피 값들을 가중 평균하여 사용한다.

랜덤 포레스트는 앞에서 만들어진 결정 트리에 임의성을 제공함으로써 결정 트리가 서로 다른 특징을 가지게 만든다. 랜덤 포레스트의 이런 특성은 결정 트리의 예측이 비상관성(decorrelation)을 보유하게 만들며, 결국 받아들인 데이터에 대한 일반화 성능을 고도화한다.

결정 트리에서 임의성을 주는 작업은 각 결정 트리의 훈련 과정 중에 수행되며, 무작위 학습 데이터의 획득 기법을 사용하는 앙상블 학습 기법인 배깅(bagging)을 이용한다. 배깅은 받아들인 훈련 데이터로부터 중복을 허락하면서 원래 데이터 셋과 동일한 크기의 데이터 셋을 생성하는 처리인 부트스트랩(bootstrap)을 이용하여, 약간씩 상이한 데이터에 대해 훈련된 기본 분류기들을 결합하는 기법이다. 앞에서 기술한 배깅을 통해 만들어진 트리는 바이어스는 동일하게 유지하면서, 분산은 축소시키므로 잡음에 강인해지며 성능은 증가시킨다.

본 연구에서는 검출된 얼굴의 방향에 적합한 랜드마크를 추출하기 위해서 기존에 적용된 모델인 래티나 얼굴 모델과 랜덤 포레스트를 결합하여 이용하였다. 다시 말해, 래티나 모델을 적용해 사람의 얼굴 영역을 획득한 다음, 학습된 랜덤 포레스트 모델을 이용하여 얼굴의 방향을 추출하고, 추출된 방향에 따라 랜드마크를 설정한다.

4. 실험 결과

본 연구에서 소개된 접근 방법의 구현을 위해 이용된 데스크탑 컴퓨터는 Intel Core i7-6700 3.4 GHz의 CPU, 16GB의 램(RAM), 갤럭시 지포스 GTX 1080 Ti 그래픽스 카드, 256GB의 SSD로 구성되어 있다. 그리고 개발을 위하여 이용된 개인용 컴퓨터 내에는 윈도우즈 10 OS가 인스톨되어 있다. 또한 본 논문에서 제시된 프로그램의 구현을 위한 통합 개발 환경으로는 마이크로소프트사의 Visual Studio 2017이 인스톨되었다. 본 논문에서는 소개된 접근 방법을 보다 효과적으로 구현하기 위하여 공개된 컴퓨터 비전 라이브러리인 OpenCV도 이용되었다.

Fig. 3은 소개된 랜덤 포레스트 모델 기반의 알고리즘을 사용하여 입력된 컬러 영상으로부터 개인 정보에 해당하는 사람의 얼굴 영역을 추출한 결과를 제시한다. Fig. 3에서 얼굴 영역을 감싸고 있는 사각형은 제안된 방법으로 추출된 얼굴 영역을 나타내는 최소 인접 사각형(MER: minimum enclosing rectangle)을 표현한다. 그리고 최소 인접 사각형 내에 존재하는 5개의 점은 얼굴 모델을 형성하는 주요 특징점인 랜드마크를 의미한다. 또한 최소 인접 사각형 내부에 표현된 아라비아 숫자는 추출된 사각형 영역이 실제로 사람의 얼굴일 확률을 의미한다. Fig. 2에서 확인할 수 있듯이 정면 방향뿐만

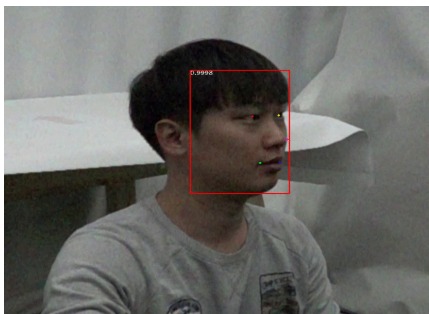


Fig. 2. Facial area extraction

아니라 측면 방향에서 촬영된 컬러 영상에 대해서도 사람의 얼굴 영역 및 5개의 랜드마크를 비교적 정확하게 추출한다.

본 논문에서는 소개된 랜덤 포레스트 모델 기반의 얼굴 획득 알고리즘의 성능을 정확성 관점에서 정량적으로 측정 및 비교하였다. 본 논문에서는 받아들인 영상 콘텐츠로부터 제시된 알고리즘을 적용하여 올바르게 획득된 사람의 얼굴 영역의 수, 그리고 테스트에 이용되는 전체적인 영상 콘텐츠에 원래부터 들어있는 얼굴 영역의 수와의 비율로 제시되는 식 (4)와 같은 정확도 지표를 적용하였다. 식 (4)에서 $FR_{extracted}$ 는 제안된 알고리즘을 이용하여 올바르게 추출된 얼굴 영역의 개수를 의미한다. 그리고 FR_{whole} 은 본 연구의 테스트에 이용된 전체 영상 콘텐츠에 들어있는 얼굴 영역의 총 개수를 의미한다. 식 (4)에서 알 수 있듯이 본 논문에서 생성하여 적용하는 얼굴 추출 정확성 지표는 백분율 단위로 제시된다.

$$\Omega_{correctness} = \frac{FA_{extracted}}{FA_{whole}} \cdot 100 (\%) \quad (4)$$

Fig. 3은 기존의 피부 색상 기반의 접근 방법과 본 논문에서 소개된 접근 방법을 사용해 입력된 컬러 영상으로부터 사람의 얼굴 영역을 검출한 결과를 막대 그래프로 비교하여 제시한다. Fig. 3에서 확인할 수 있듯이, 본 논문에서 소개한 알고리즘이 랜덤 포레스트 모델을 적용하여 개인 정보로 대표되는 얼굴 영역을 방향을 고려하여 보다 정확하게 획득한다.

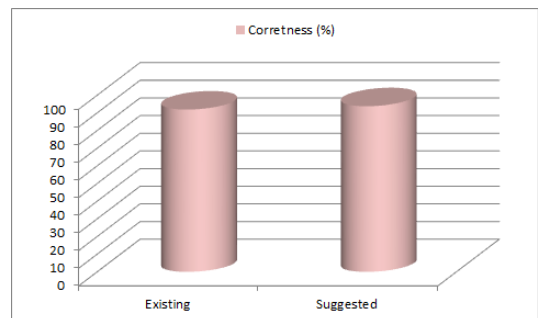


Fig. 3. Performance comparison

실험 결과를 분석해 보면, 기존의 얼굴 추출 알고리즘은 입력받은 컬러 영상으로부터 얼굴이 위치한 방향을 고려하지 않고 단순히 사람의 얼굴 검출만을 시도한다. 따라서 검출된 특징 점들의 위치 및 개수가 부정확한 경

우가 상당히 많아 얼굴 영역의 오 검출이 빈번히 발생하였다.

반면에, 본 논문에서 소개한 알고리즘은 1차적으로 얼굴 영역을 개략적으로 검출한 다음, 얼굴 영역이 위치하는 방향을 고려하여 랜드마크 포인트들의 위치와 개수를 적절하게 조율하므로 보다 정확하게 얼굴 영역을 추출할 수 있었다.

5. 결론

일반적으로 연속적으로 입력되는 컬러 영상을 처리할 때 사람의 얼굴 영역은 매우 중요한 요소이다. 왜냐하면 얼굴 영역은 해당하는 컬러 영상 안에 포함된 사람이 누구인지를 확인할 수 있는 개인적인 정보를 제공하고 있기 때문이다. 최근 들어, 이러한 개인 정보의 외부로의 노출이 사생활 침해라는 논란을 불러일으키고 있어서 사회적으로 커다란 문제가 되고 있는 실정이다. 따라서 영상 콘텐츠로부터 노출된 개인 정보 영역 중의 하나인 얼굴을 검출하고 블로킹하는 처리가 중요하다.

본 논문에서는 일반적인 실내외의 자연스러운 환경에서 고속의 카메라로 캡처되어 입력되는 컬러 영상으로부터 노출된 개인 정보를 포함하고 있는 얼굴 영역을 강인하게 획득하는 알고리즘을 기술하였다. 기술된 알고리즘에서는 먼저 받아들인 영상 데이터로부터 래티나 얼굴 모델을 사용해 사람의 얼굴 영역을 1차적으로 검출하였다. 그런 다음, 얼굴의 형태를 효과적으로 구분할 수 있는 랜덤 포레스트를 적용하여 입력된 컬러 영상으로부터 노출된 개인 정보를 포함하고 있는 얼굴 영역을 방향을 고려하여 정확하게 추출하였다. 실험 결과에서는 본 연구에서 기술한 랜덤 포레스트 모델 기반의 알고리즘이 입력받은 다양한 종류의 컬러 영상으로부터 얼굴 영역을 정면의 방향과 그 이외의 방향에서 비교적 강인하게 추출한다는 것을 제시하였다.

추후에는 본 논문에서 소개한 랜덤 포레스트 모델 기반의 얼굴 영역 추출 방법의 고도화를 진행하여 다양한 실내외의 환경에서 촬영된 컬러 영상에서 얼굴 영역 추출이 보다 강인하게 수행될 수 있도록 개선할 계획이다. 그리고 추출된 사람의 얼굴 영역을 영상 블러링이나 모자이크 삽입 기법 등을 통해 효과적으로 블로킹하여 개인 정보 영역들이 본인의 동의 없이 외부에 노출되지 않도록 보호할 예정이다. 또한 성능평가의 지표를 보다 다양화할 계획이다.

References

- [1] J. Javh, J. Slavic, and M. Boltezar, "High Frequency Modal Identification on Noisy High-Speed Camera Data," *Mechanical Systems and Signal Processing*, Vol.98, pp. 344-351, January 2018.
DOI: <https://doi.org/10.1016/j.ymssp.2017.05.008>
- [2] R. Menaka, R. Janarthanan, and K. Deeba, "FPGA Implementation of Low Power and High Speed Image Edge Detection Algorithm," *Microprocessors and Microsystems*, Vol.75, pp. 1-7, February 2020.
DOI: <https://doi.org/10.1016/j.micpro.2020.103053>
- [3] S. Zafeiriou, C. Zhang, and Z. Zhang, "A Survey on Face Detection in the Wild: Past, Present and Future," *Computer Vision and Image Understanding*, Vol.138, pp. 1-24, 2015.
DOI: <https://doi.org/10.1016/j.cviu.2015.03.015>
- [4] G. Zheng and Y. Xu, "Efficient Face Detection and Tracking in Video Sequences Based on Deep Learning," *Information Sciences*, Vol.568, pp. 265-285, March 2021.
DOI: <https://doi.org/10.1016/j.ins.2021.03.027>
- [5] Q. Jin, C. Mu, L. Tian, and F. Ran, "A Region Generation Based Model for Occluded Face Detection," *Procedia Computer Science*, Vol.174, pp. 454-462, July 2020.
DOI: <https://doi.org/10.1016/j.procs.2020.06.114>
- [6] U. Mahbub, S. Sarkar, and R. Chellappa, "Partial Face Detection in the Mobile Domain," *Image and Vision Computing*, Vol.82, pp. 1-17, January 2019.
DOI: <https://doi.org/10.1016/j.imavis.2018.12.003>
- [7] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A Fast Face Detection Method via Convolutional Neural Network," *Neurocomputing*, Vol.395, pp. 128-137, July 2019.
DOI: <https://doi.org/10.1016/j.neucom.2018.02.110>
- [8] Z. Zhou, Z. He, Y. Jia, J. Du, L. Wang, and Z. Chen, "Context Prior-Based with Residual Learning for Face Detection: A Deep Convolutional Encoder-Decoder Network," *Signal Processing: Image Communication*, Vol.88, pp. 1-13, July 2020.
DOI: <https://doi.org/10.1016/j.image.2020.115948>
- [9] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, USA, pp. 5203-5212, June 2020.
DOI: <https://doi.org/10.1109/CVPR42600.2020.00525>
- [10] Y. Chen, W. Zheng, W. Li, and Y. Huang, "Large Group Activity Security Risk Assessment and Risk Early Warning Based on Random Forest Algorithm," *Pattern Recognition Letters*, Vol.13, pp. 1-5, January 2021.
DOI: <https://doi.org/10.1016/j.patrec.2021.01.008>

장 석 우(Seok-Woo Jang)

[종신회원]



- 1995년 2월 : 송실대학교 전자계산학과 (공학사)
- 1997년 2월 : 송실대학교 일반대학원 컴퓨터학과 (공학석사)
- 2000년 8월 : 송실대학교 일반대학원 컴퓨터학과 (공학박사)
- 2009년 3월 ~ 현재 : 안양대학교 소프트웨어학과 교수

<관심분야>

로봇비전, 증강현실, HCI, 비디오 색인 및 검색, 등