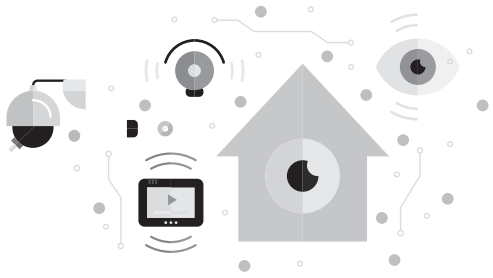
지능형 영상 보안 시스템의 얼굴 마스킹 기술



김동칠 전자부품연구원 스마트미디어연구센터 선임연구원
박성주 전자부품연구원 스마트미디어연구센터 센터장

1. 머리말

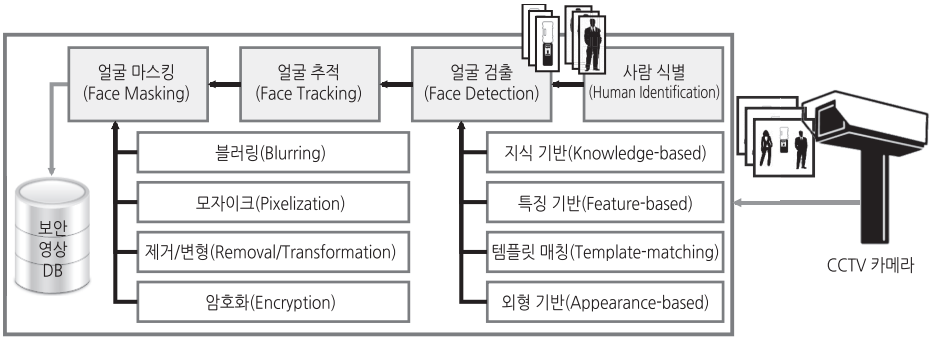
영상 보안 시스템은 범죄 예방, 교통 감시, 시설 관리 및 재난 관리 등을 주요 목적으로 공공 기관 및 민간 기업에서 설치와 활용이 증가하고 있다. 최근, 영상 보안 시스템에 지능형 영상 분석 기술이 결합하면서 방범 및 감시 영역을 넘어 소방, 국방, 교육 등 다양한 영역에서 국민 보호 및 생활 안전 향상에 기여하고 있다. 지능형 영상 보안 시스템은 영상을 취득하는 CCTV 카메라를 포함하여 전송 장치, 저장 장치, 영상 분석 장치 및 관리자에게 다수의 영상을 표출 및 제어하기 위한 관제 솔루션으로 구성된다[1]. 정부에서도 2017년도까지 229개의 시군구에 CCTV 관제센터 구축 계획을 세우고 있으며, 2018년부터 전국 지방자치단체 CCTV 관제 서비스를 지능형으로 전환하기 위해, CCTV 관제 서비스 체계 구축을 위한 정보화전략계획(ISP, Information Strategy Planning) 수립 작업에 착수하고 있다.

지능형 영상 보안 시스템을 구성하는 핵심 요소인 CCTV 카메라의 급격한 성능 향상으로 인해, CCTV 영상을 이용한 얼굴 인식 정확도가 증가하고 있으며, 해당 기술을 활용한 범죄 수사는 점차 확대되고 있다[2]. 그러나 CCTV 촬영 영상이 무분별하게 유출 및 배포되는 사고가 발생함에 따라, 촬영된 영상에 개인의 프라이버시 보호를 위해 얼굴 등 특정 부위를 마스킹하여 배포 및 저장할 필요성이 증가하고 있다[3]. [그림 1]과 같이 얼굴 마스킹은 CCTV 카메라로부터 획득한 영상에서 사람을 식별하고 관심 영역(ROI, Region of Interest)인 얼굴을 검출한다. 얼굴 추적은 얼굴 검출과 함께 필수적인 기술로 동일한 사람에 대한 반복적인 연산을 피하고자 이용된다. 이후, 얼굴 마스킹 기술에 의해 프라이버시 보호가 되고 해당 영상은 보안 영상 DB에 저장되어 시스템에 접근하는 사용자에 따라 다양한 형태로 제공된다.

본고에서는 지능형 영상 보안 시스템에서 개인 영상 정보 보호를 위해 CCTV 카메라로부터 획득한 개인의 얼굴 정보를 식별할 수 없도록 하기 위한 얼굴 마스킹 기술을 소개하고자 한다.

2. 얼굴 검출 기술

얼굴 검출 기술은 얼굴 마스킹 이전에 실행되어야 하는 핵심 과정으로 CCTV 영상에서 얼굴이 있



[그림 1] 영상 보안 시스템에서 얼굴 마스킹 기술

는 영역을 찾아내는 기술이다. 얼굴 검출 기술은 검출하는 방법에 따라 지식 기반 방법(Knowledge-based Methods), 특징 기반 방법(Feature-based Methods), 템플릿 매칭 방법(Template-matching Methods), 외형 기반 방법(Appearance-based Methods) 등이 있다[4][5].

2.1 지식 기반 방법

지식 기반 방법은 사람의 얼굴을 구성하는 눈, 코, 입, 눈썹 등을 특징 요소로 인식하고, 각 특징 요소 간의 거리와 위치 관계를 규칙 기반으로 분석하여 얼굴 영역을 검출하는 방법이다. 주로 입력 이미지의 히스토그램을 이용하여 얼굴의 좌우 특성 및 눈, 코, 입, 눈썹 등의 위치 정보를 찾아낸다. 지식 기반 방법은 얼굴 특징 요소간의 상관관계 규칙을 간단하게 정의할 수 있지만, 얼굴의 각도, 표정 등과 같은 다양한 얼굴 특징 요소들의 변화가 있는 영상에서는 얼굴의 검출이 어려운 단점이 있다.

2.2 특징 기반 방법

특징 기반 방법은 얼굴의 특징 요소, 피부색(Skin Color), 질감(Texture), 외곽선(Outline) 정보와 이들 성분의 조합된 형태의 정보를 이용해서 얼굴 영역을 검출하는 방법이다. 특징 기반 방법은 얼굴 영역 검출 처리 시간이 빠르고 자세나 얼굴 방향에 민감하지 않아 다른 방법에 비해 비교적 쉽게 얼굴을 찾을 수 있는 장점이 있다. 그러나 피부색과 유사한 배경에서의 객체에 대한 인식 오류, 조명에 따른 피부색 변화, 카메라 위치, 얼굴의 기울어짐에 따른 모양 변화 등으로부터 얼굴 검출에 많은 영향을 받는다.

2.3 템플릿 매칭 방법

템플릿 매칭 방법은 얼굴 영상의 부분 영역이나 외곽선을 기반으로 미리 정의된 규칙에 따라 표준 템플릿을 생성한 후, 입력 영상과 유사상관도를 비교하여 얼굴 영역을 검출하는 방법이다. 템플릿 매칭 방법은 미리 정의된 템플릿을 이용하는 검출 방법과 변형 템플릿을 이용하여 검출하는 방법이 있다. 템플릿 매칭 방법은 얼굴의 특징 성분을 찾는 과정이 필요 없으므로 조명 변화나 배경의 영향을 덜 받게 되고, 복잡한 배경에서도 얼굴의 검출이 가능한 장점이 있다. 그러나 개개인의 얼굴 크기 차이, 얼굴 회전각도, 기울어짐 등에 민감하여 표준 템플릿의 패턴에 따라 영향을 받게 되며, 지식기반 방법과 같이 각각 다른 자세에 대한 템플릿을 정의하기 어려운 단점이 있다.

2.4 외형 기반 방법

외형 기반 방법이란 영상 학습을 통해 학습된 모델을 이용해서 입력 영상으로부터 얼굴 영역을 검출하는 방법이다. 복잡한 영상에서 얼굴 영역을 검출하기 위해, 외형 기반 방법은 얼굴 영역과 얼굴이 아닌 배경 영상을 입력받아 두 영상의 차이를 잘 나타낼 수 있는 특징들을 찾고, 이를 학습하여 입력 영상에 대해 학습된 정보를 이용하여 얼굴 영역을 검출한다. 외형 기반 방법은 현재 얼굴검출 분야에서 가장 많이 사용하는 방법의 하나이며, 다른 검출 방법들에서 언급된 여러 제약 조건들이 학습을 통해 극복되기 때문에 인식률이 높은 방법이다. 그러나 학습 및 얼굴 특징 추출 과정에 따라 공간적 비용과 많은 시간이 필요하게 되고, 데이터베이스가 변경되면 재학습을 해야 하는 단점이 있다.

3. 지능형 영상 보안 시스템의 얼굴 마스킹 기술

지능형 영상 보안 시스템에서 CCTV 카메라로부터 취득되는 영상이나, 이미 취득되어 저장된 영상에서 개인의 프라이버시를 침해할 수 있는 소지가 가장 높은 것이 얼굴 영상이다. 또한, 영상 처리를 이용하여 특정인을 감시 추적하는 경우에, 그 외 불특정 다수의 얼굴 영상은 육안으로나 컴퓨터를 이용하여 식별되지 않도록 얼굴 마스킹 기법을 이용하여 처리해야 한다. 얼굴 마스킹 기술은 CCTV 카메라에서 획득된 사용자 얼굴 정보를 얼굴 영역 검출 알고리즘을 통해 효율적으로 얼굴 영역 정보를 추출한 후, 검출된 얼굴 정보에 대해 프라이버시를 제공하기 위한 얼굴 마스킹 기술을 적용해야 한다. 대표적인 얼굴 마스킹 기술은 마스킹하는 방법에 따라 블러링(Blurring), 모자이크(Pixelization), 제거 및 변형(Removal/Transformation), 암호화(Encryption) 등이 있다[6].

3.1 블러링 기법

블러링 기법은 입력 영상에서 얼굴 영역을 흐리게 만들어 얼굴을 식별하기 어렵게 함으로써 개인의 프라이버시를 보호하는 기술이다. 블러링 기법은 다른 기법에 비해 비교적 구현이 쉬울 뿐만 아니라, 한번 마스킹을 하면 추가적인 유출에 따른 프라이버시 피해가 없다. 영상을 블러링하기 위한 대표적인 방법으로는 가우시안 저역 필터를 이용하여 다음의 식과 같이 주어진 분산값(σ)에 따라 가우시안 함수로 컨벌루션(Convolution)하여 영상을 흐리게 처리한다. 현재 Google Street View 등의 상업용 시스템에 적용되어 사용되고 있다[7].

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

[그림 2]는 가우시안 필터를 적용하여 블러링 된 얼굴 영상을 나타낸다. 블러링 기법의 단점은 한번 마스킹 된 영상으로부터 원본 영상을 복원해야할 필요성이 발생하는 경우 이를 완벽하게 복구할 수 없다. 또한, 얼굴 인식 기술의 발전으로 인해 블러링을 얼굴 영역에 적용했어도 얼굴 인식이 가능한 사례가 증가하고 있어, CCTV 영상의 품질에 따라 분산값을 어떻게 설정해야 얼굴 인식할 수 없는 형태로 블러링이 되는지에 대한 연구가 필요하다.

3.2 모자이크 기법

모자이크 기법은 주어진 블록사이즈(B×B)에 대해 블록의 평균값을 이용하여 같은 밝기 값으로 대체함으로써 개인의 얼굴에 대한 프라이버시를 제공한다. 모자이크 기법은 텔레비전 뉴스 및 다큐멘터리에서 흔히 사용되는 기법이며 용의자, 증인, 또는 일반 사람들의 얼굴을 가려서 익명성을 유지하는데 사용된다. [그림 3]은 얼굴 영상에 블록사이즈를 변화시켰을 때의 모자이크 정도를 나타낸다. 그러나 모자이크 기법의 단점은 시간 경과에 따른 픽셀

(a) 원본 영상 (b) 가우시안 저역 필터링 적용

[그림 2] 얼굴 영역 블러링 예시(Akiyo, 352×288 pixels/frame)

(a) 원본 영상 (b) 블록 사이즈(B=5) (c) 블록 사이즈(B=10)

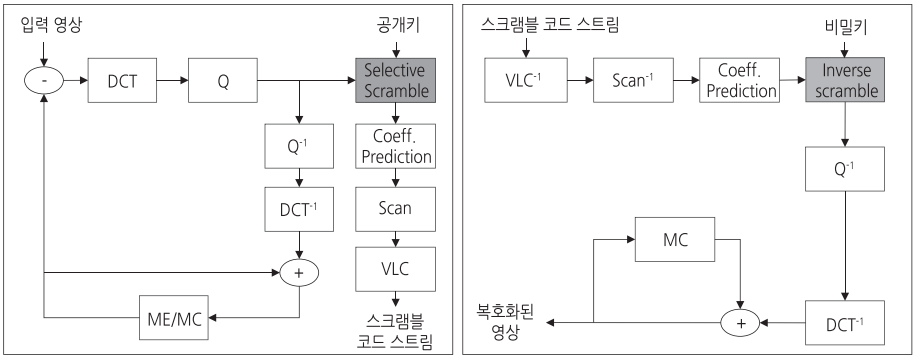
[그림 3] 얼굴 영역 모자이크 예시(Akiyo, 352×288 pixels/frame)

(a) 원본 영상 (b) 얼굴 영역 제거 영상

[그림 4] 얼굴 영역 제거 예시(Grandma, 176×144 pixels/frame)

(a) 원본 영상 (b) 얼굴 영역 변형 영상

[그림 5] 얼굴 영역 변형 예시(Akiyo, 352×288 pixels/frame)



(a) 인코딩 과정 (b) 디코딩 과정

[그림 6] MPEG-4에서의 영상 스크램블링 구조

(a) 원본 영상 (b) 암호화된 영상 (c) 복원 영상

[그림 7] H.264에서 FMO 기반 암호화 기법 예시(Miss America, 176×144 pixels/frame)

을 통합하면 은폐된 정보를 부분적으로 복구할 수 있다.

3.3 제거 및 변형 기법

제거 및 변형 기법은 감시 대상의 얼굴 영역을 검출한 후 얼굴 영역을 영상에서 완전히 제거하거나 변형하여 개인의 프라이버시를 보호하는 방법이다. [그림 4]는 얼굴 영역을 완전히 제거함으로써 프라이버시를 보호한 영상이며, [그림 5]는 변형 기법을 이용하여 얼굴의 특징 정보들을 변형하여 얼굴을 확인할 수 없도록 하였다[8].

3.4 암호화 기법

지능형 영상 감시 시스템에서 CCTV 카메라로부터 취득된 얼굴 영상에 대한 블러링이나 모자이크 방법을 통해 개인의 얼굴 식별을 어렵게 하여 개인 정보 보호를 하였더라도, 사후에 원영상으로 복원하여 얼굴을 식별해야 하는 상황이 있다. [그림 6]은 MPEG-4에서 스크램블링을 하기 위한 인코딩 및 디코딩 과정을 나타내며, 스크램블링 기법은 암호화 키를 통해 암호화하여 개인 얼굴 이미지 등을 마스킹하고 이를 허가된 관리자만 암호화 키를 이용한 언마스킹을 통해 원영상을 복원할 수 있다[9]. 스크램블링 기법은 프라이버시 영역인 얼굴 영역을 검출하여, 마스킹하더라도 합법적인 목적에 의해서 언마스킹이 가능하다는 장점이 있다. 또한, 스크램블링의 강도를 조절할 수 있어, 단순 모니터링 용도에서, 정밀 인물 추정 용도까지 사용될 수 있다. 그러나 스크램블을 위한 암호화 키가 유출될 수 있으므로 이를 위한 보안이 추가로 필요하다.

H.264는 영상의 품질은 그대로 유지하면서 비트율이 낮아 영상 저장이나 전송 측면에서 높은 효율성을 보여 CCTV 분야에서 주목을 받고 있다. H.264 비디오에서 관심 영역의 프라이버시 보호를 위해 플렉서블 매크로블럭 순서기법(FMO, Flexible Macroblock Ordering)을 이용한다[10]. FMO 기법은 총 7가지의 매칭 타입을 가진다. FMO 타입 2는 관심 영역 코딩에 사용되지만, 직사각형 영역에만 적합하여 불규칙한 얼굴 영역을 잘 표현하지 못하는 단점이 있어, 일반적으로 FMO 타입 6을 주로 이용한다. [그림 7]과 같이 FMO 기법은 현재 널리 사용되고 있는 CCTV 표준 포맷에 대해 실시간으로 특정 얼굴 영상 영역을 암호화하고 필요에 따라 복호화가 가능하다.

4. 맺음말

지능형 영상 보안 시스템은 사회 질서 유지, 범죄 예방 등을 위해 공공 기관 및 민간 기업에서 설치 및 활용이 급증하고 있다. 그러나 CCTV 카메라로부터 획득한 개인 영상이 무분별하게 유출 및 배포되어 프라이버시 침해라는 새로운 문제점을 야기하고 있다. 본고에서는 지능형 영상 보안 시스템에서 CCTV 카메라로부터 취득되는 영상이나, 이미 취득되어 저장된 영상에서 개인의 프라이버시를 침해할 수 있는 소지가 가장 높은 얼굴 영상을 식별할 수 없게 하기 위한 얼굴 마스킹 기술을 소개하였다. 지능형 영상 보안 시스템에서 얼굴 마스킹 기술을 통해 CCTV 영상으로부터 개인 프라이버시를 보호할 수 있다고 생각된다.

※본 연구는 2017년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임[No. 2014-0-00077(B0101-17-0525), 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황예측 기술 개발].

[참고문헌]

[1] 박성주, '통합관제시스템과 외부 영상 보안 시스템 간의 연동 인터페이스', TTA Journal, Vol. 169, 2017.01.

[2] 신용녀, 전명근, '영상감시 시스템에서의 얼굴 영상 정보보호를 위한 기술적·관리적 요구사항', 정보보호학회논문지, Vol. 24, No. 1, 2014.02.

[3] 문해민, 반성범, '지능형 영상 감시 시스템에서 프라이버시 보호를 위한 De-identification 기술 분석', 한국정보기술학회논문지, Vol. 9, No. 7, 2011.07.

[4] 한국인터넷진흥원, 'CCTV기반 얼굴 검출 및 인식 시스템 보안 프레임워크에 관한 연구', 2009.07.

[5] M. Yang, D. Kriegman and N. Ahuja, 'Detecting Faces in Images: A Survey', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, pp. 34-58, 2002.01.

[6] F. Dufaux and T. Ebrahimi, 'A Framework for the Validation of Privacy Protection Solutions in Video Surveillance', IEEE International Conference on Multimedia and Expo, pp. 66-71, 2010.07.

[7] A. Frome, et. al., 'Large-scale Privacy Protection in Google Street View', IEEE International Conference on Computer Vision, pp. 2373-2380, 2009.05.

[8] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. Nayar, 'Face Swapping: Automatically Replacing Faces in Photographs', ACM Transactions on Graphics, pp. 1672-1675, 2008.08.

[9] F. Dufaux and T. Ebrahimi, 'Scrambling for Privacy Protection in Video Surveillance Systems', IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 8, pp. 1168-1174, 2008.07.

[10] F. Peng, X. Zhu, and M. Long, 'An ROI Privacy Protection Scheme for H.264 Video Based on FMO and Chaos', IEEE Transactions on Information Forensics and Security, Vol. 8, No.10, pp. 1688-1699, 2013.04.