

실시간 동영상에서의 인물 선별 송출 시스템

우채운, 박나형, 백지윤, 정유진, 김명주

서울여자대학교 정보보호학과

eogkr2017@daum.net, nhpark326@naver.com, pottery030@naver.com,

yujinj96@naver.com, mjkim@swu.ac.kr

Person Selectable Transmission System in Real-Time Video Conference

Chae-Yoon Woo, Na-Hyung Park, Ji-Yoon Beak, Yu-Jin Jung, Myuhng-Joo Kim
Dept. of Information Security, Seoul Women's University

요약

본 논문에서는 실시간 얼굴인식 기술을 활용한 인물 선별 송출 시스템을 제안한다. 실시간 동영상 안에 등장하는 다수의 인물 객체 얼굴을 검출하고 인식하기 위해 Haar 특징 정보 기반의 다단계 (Cascade) 학습 알고리즘을 사용한다. 이 시스템은 다수의 인물을 인식하고 학습할 수 있으며 인식된 각 인물의 송출 여부를 사용자가 직접 선택할 수 있는데 이 모든 선택 송출 과정을 실시간으로 처리할 수 있다. 여기에서 제시한 기술은 다자간 화상 채팅이나 다자간 화상 회의에서 특정인의 프라이버시 보호를 위한 기술로 활용될 수 있다.

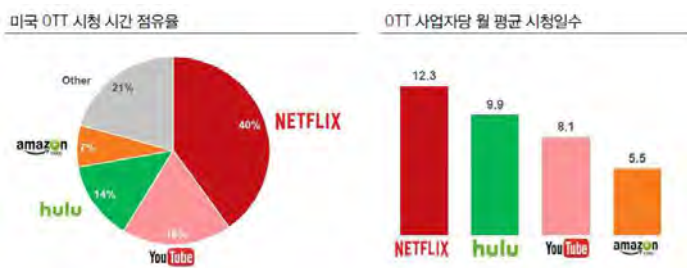
1. 개발배경

정보통신기술의 발달과 인터넷의 대중화가 이뤄지면서, 개인이 직접 다양한 콘텐츠를 생산 및 공유하는 1인 미디어 사업이 확산되고 있다. 그중에서도 스마트폰이 현대인들의 필수품이 되면서, 스트리밍 방식으로 언제 어디서든 영상 콘텐츠를 볼 수 있는 서비스인 OTT(Over The Top)의 시장 규모가 계속해서 성장하고 있다.



(그림2) 디지털 보안 및 프라이버시 안전 수준(DMC MEDIA M.U.D 연구팀, “디지털 미디어 정보 침해에 대한 소비자 보고서”, p.6, 2013)

특히 1인 미디어 플랫폼 중 하나인 실시간 스트리밍 서비스에서의 개인 정보 노출 사고가 문제 되고 있다. 사용자 본인이 아닌 타인의 얼굴이 노출되는 등 초상권 침해에 대한 우려가 커지고 있는 상황이다[1]. 본 논문에서는 Python 3.6, opencv, PyQt5 을 사용하여 실시간 동영상에서의 인물 선별 송출 시스템 프로그램을 구현하였다. 이를 통하여 개인의 프라이버시와 타인의 초상권을 보호해 줄 수 있다. 본 논문에서 구현한 프로그램은 windows 10 (64비트), windows 7 (64비트)에서의 운영환경을 지원하고 있다.



(그림1) 미국 OTT 시장점유율 및 사업자당 월 평균 시청일수(정보통신산업진흥원 디지털콘텐츠산업기획팀, “2017년 CG/VFX 산업 주간 이슈페이퍼 8월호”, p.13, 2017)

하지만 이러한 1인 미디어와 관련된 개인정보 보호에 대한 인식과 규제가 아직 명확하지 않아, (그림 2)와 같이 이에 대한 소비자의 보안위협 및 사생활 침해에 대한 불안함을 가지고 있는 것으로 분석됐다.

2. 유사 프로그램과의 비교



(그림3) 기존 프로그램과의 비교

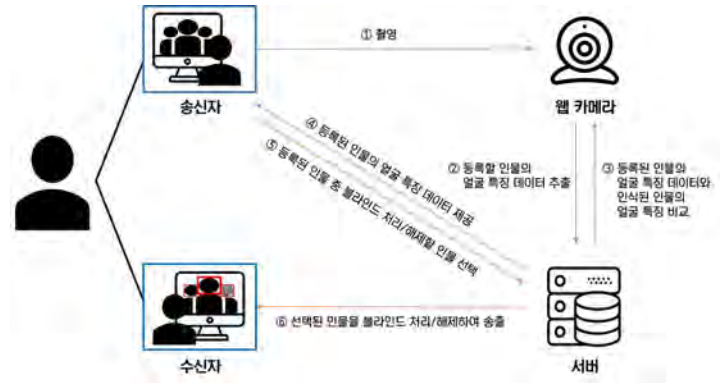
(그림3)은 실시간 스트리밍에서 사용되는 프로그램들을 비교한 것이다. 기존 프로그램들은 실시간 통신 서비스만 제공하는 반면 본 논문에서는 프라이버시 보호에 초점을 맞춰 보안이 강화된 실시간 통신 서비스를 제공한다. 등록된 사용자 이외의 모든 인물을 보이지 않도록 하여 타인의 프라이버시를 보호하고, 사용자에게 객체의 송출 여부를 선택할 수 있도록 하는 자율성을 제공한다. 따라서 본 연구는 사용자에게 자율성을 부여하고 보안 측면을 강조함으로써 기존 프로그램들과 차별성을 제시한다.

3. Haar based Cascade의 활용

실시간 컴퓨터 비전을 목적으로 한 프로그래밍 라이브러리인 OpenCV에서는 Haar Feature를 기반으로 한 Cascade Classifiers를 제공하고 있다. Haar 기능 기반 캐스케이드 분류기를 사용한 객체 탐지는 2001년 Paul Viola와 Michael Jones가 "단순한 기능의 강화된 캐스케이드를 사용한 빠른 객체 탐지"에서 제안한 객체 탐지 방법이다[2].

영역 간의 밝기 차이인 haar 모양의 특징으로 특징값을 계산하고, 적분 영상을 적용한다. 그리고 AdaBoost를 이용하여 약 분류기 중 물체 검출에 크게 기여하는 강 분류기를 만들고, 여러 개의 강 분류기를 사용하여 Cascade 구조로 구성한다. 그리고 조건에 만족하는 이미지를 분류하여 조건에 맞지 않을 경우 버리고, 조건이 맞다면 다음 조건으로 넘어간 뒤 최종 검출로 출력한다[3]. 검출기 분류를 통하여 검색 속도를 향상시킬 수 있는 장점을 제공한다. 본 논문에서는 이상에서 제시한 내용을 포함하여 구현을 진행하였다.

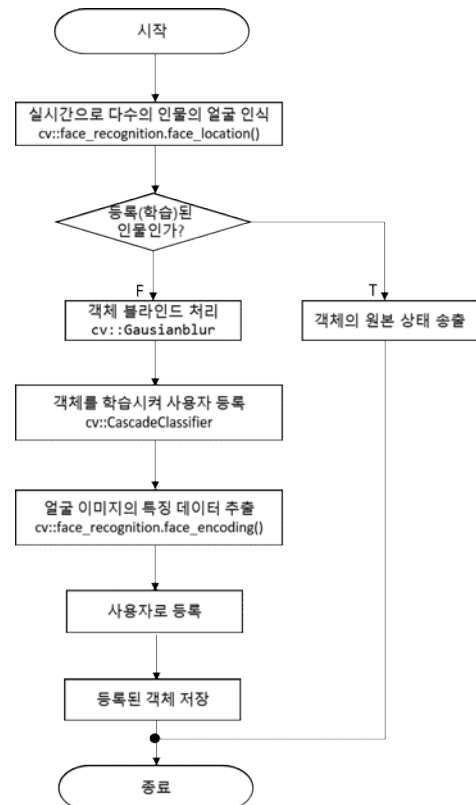
4. 시스템 구성도



(그림4) 시스템 구성도

본 논문에서 개발한 프로그램의 사용자는 (그림4)가 보여주는 것과 같이 송신자 역할과 수신자 역할을 모두 할 수 있다. 송신자는 웹캠으로 촬영 중인 영상에서 송출할 인물 객체를 선택적으로 등록할 수 있으며 이 기능은 서버에서 실행된다. 서버는 등록된 인물 객체의 데이터를 송신자에게 반환하며, 등록된 객체 중 송신자가 선택한 객체를 블라인드 또는 해제 처리한다. 수신자는 서버에서 처리된 영상을 수신한다.

5. 얼굴인식 과정



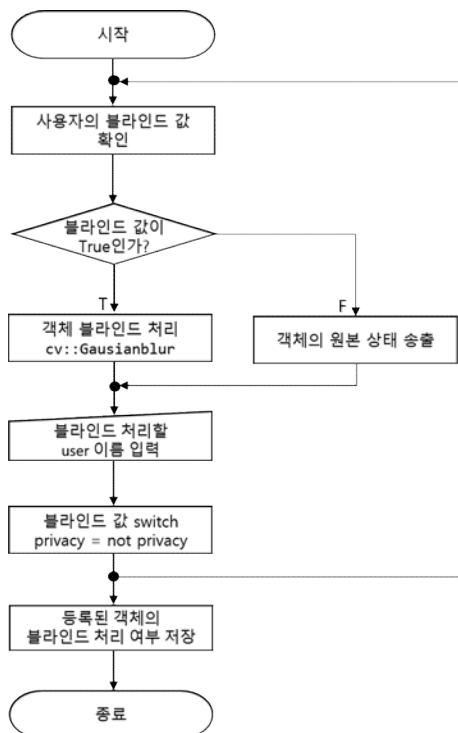
(그림5) 얼굴인식 및 학습 알고리즘

얼굴인식 및 학습 과정은 (그림5)와 같이 진행된다. 먼저 프로그램이 실행되면 OpenCV에서 제공하는 cv::face_recognition 클래스의 face_location 메서드를 사용하여 카메라 촬영 영상 frame에서 얼굴 영역과 특징을 추출한다.

다음으로 cv::face_recognition 클래스의 distance 메서드를 사용하여 frame에서 추출한 얼굴 특징을 기존에 등록(학습)된 얼굴 특징과 비교하여 거리 척도를 환산한다. 거리가 0.6 이하일 시 동일 인물로 인식하여 원본 상태의 frame을 송출하며, 등록된 인물이 아닐 시 해당 인물을 블라인드 처리하여 송출한다.

haarcascade 방식으로 cv::CascadeClassifier 메서드를 이용하여 객체의 얼굴 영역을 식별하며 cv::face_recognition.encoding 메서드로 식별된 얼굴 영역에서 68개의 얼굴 특징 위치를 분석하여 데이터로 추출한다. 추출된 데이터는 등록된 객체의 정보로 저장된다.

6. 블라인드 처리 과정



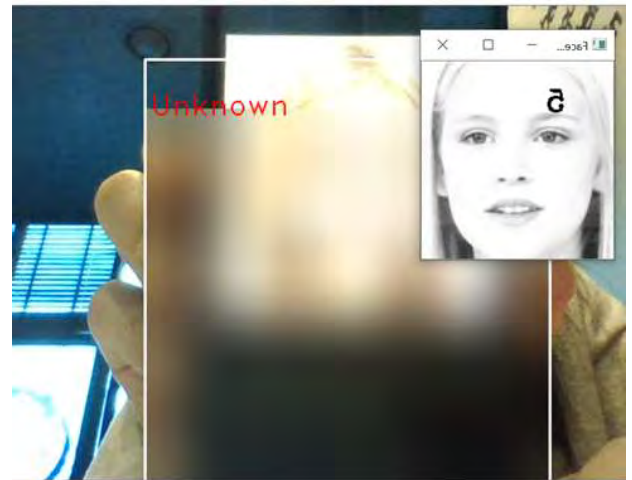
(그림6) 블라인드 처리 알고리즘

인식된 객체에 대한 블라인드 처리 과정은 (그림6)과 같이 진행된다. cv::rectangle 메서드를 이용하여 인식된 객체의 얼굴에 박스를 그리고 사용자가 등록된(학습된) 객체의 블라인드 처리 여부를 결정한다. cv::Gaussianblur 메서드로 해당 객체의 박스를 블라

인드 처리 또는 해제한다. 등록된 사용자에 대한 블라인드 값은 기본적으로 False로 되어있어, 사용자의 원본 상태를 출력한다. 이때, 사용자가 블라인드 처리하고 싶은 사용자의 이름을 입력하면 해당 사용자의 블라인드 값을 True로 전환시킨다. 사용자가 블라인드 처리 및 해제를 적용하고 싶을 때마다 이와 같은 과정을 반복한다.

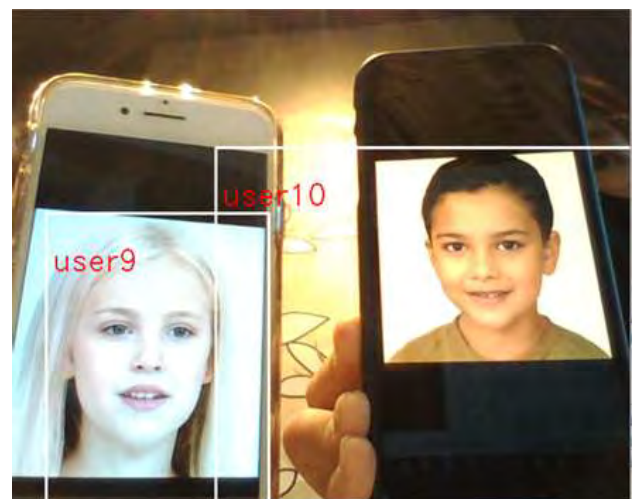
7. 실행 결과

각 얼굴 사진은 Generated Photos 사이트에서 제공하는 무료 AI 생성 인물사진을 활용했다[4].



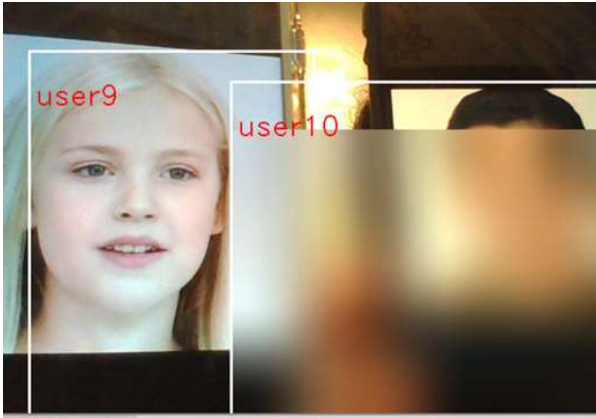
(그림7) 인물 학습 모듈

(그림7)은 학습되지 않은 인물을 학습하는 과정이다. 새로운 인물의 사진을 학습하여 사용자로 인식시킨다.



(그림8) 다수 객체 인식 모듈

(그림8)은 동시에 다수의 객체를 각각 다른 인물로 인식하는 것을 확인할 수 있다.



(그림9) 개별 블라인드 모듈

(그림9)에서와 같이 인식된 사용자 각각을 구별하고, 사용자의 송출 여부 선택에 따라 송출되지 않는 인물은 블라인드 처리가 적용된다.

8. 보안 요소

1) 사용자 프라이버시 보호

프로그램이 처음 실행될 때 사용자의 얼굴을 학습하여 등록한다. 등록된 사용자만 보여주고, 등록되지 않으면 모두 블라인드 처리해 송출한다. 이로써 별도 처리 과정 없이 타인의 얼굴이 노출되는 것을 방지한다. 만약 등록되지 않은 사용자의 얼굴을 송출하고 싶다면 이를 등록하고 송출한다. 등록된 인물의 얼굴 송출 여부는 사용자가 변환할 수 있다. 이처럼 등록되지 않은 객체는 모두 블라인드 처리하고, 등록된 객체의 블라인드 여부를 결정하게 하여 사람들의 프라이버시를 보호한다. 또한, 학습된 얼굴 데이터는 프로그램이 종료될 때 삭제하여 개인 정보를 파기한다.

2) 암호화 통신

Frame을 jpg 이미지로 인코딩을 한 후 raw images를 통해 연속적으로 반환시켜서 다른 사용자에게 프레임을 송출한다. raw images를 보낼 때, 암호화 과정을 거쳐서 영상을 송출한다. 따라서 수신자는 영상을 수신할 때, 복호화 과정을 거친 후 프레임을 볼 수 있다. 실시간으로 통신을 해야 하므로 상대적으로 속도가 빠른 AES256 알고리즘을 사용했다. 시스템 상에 임의로 key 값을 설정한다. 송신자는 raw images를 시스템 상 정의되어있는 비밀키로 암호화하여 수신자에게 실시간으로 보낸다. 수신자 역시 시스템 상에 정의되어있는 비밀키로 복호화하여 실시간으로 동영상을 수신한다.

9. 기대 효과

본 논문에서는 실시간으로 송출해야 할 얼굴을 사용자가 직접 선별함에 따라 사용자들에게는 개인의 프라이버시를 보장해주며, 타인의 초상권 또한 보호해 주는 시스템을 제안하였다. 이를 통해 시스템 카메라 화면에 들어오는 모든 사람의 개인정보를 보호하는 핵심 기술로 자리 잡을 수 있다. 또한, 이 기술을 사용하면 스트리밍 서비스 등에서 타인의 초상권 보호를 위한 별도의 영상 편집 과정을 거치지 않아도 되므로 경제적으로 인적 자원 비용이 절감되는 기대효과를 제공한다.

영상 콘텐츠 시장이 점점 확대되고 있는 1인 미디어 시장뿐 아니라 화상 회의, 방송 영상 등 다양한 분야로의 발전 가능성이 무한하다고 볼 수 있다. 원하는 객체를 선별적으로 송출해야 하는 사람들을 대상으로 선제적인 시장 창출을 할 수 있을 것이라 예상된다.

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음(2016-0-00022)

참고문헌

- [1] “노숙인도 방송소재 삼는 1인 미디어들…서울시 “초상권 침해는 범죄”, 스포츠경향. 2019년 3월 5일 수정, 2020년 4월 7일 접속, http://m.sports.khan.co.kr/view.html?art_id=201903050831003&sec_id=561101#c2b.
- [2] opencv, 2019, https://docs.opencv.org/4.1.0/d7/d8b/tutorial_py_face_detection.html
- [3] 유제훈, 심귀보, Cascade 안면 검출기와 컨볼루션 신경망을 이용한 얼굴 분류, 한국지능시스템학회 논문지, 26(1), 70-75. (2016).
- [4] “Generated Photos”, 2020년 5월 9일 접속, <https://generated.photos/faces>