

얼굴 특징 추출 및 클러스터링 기반의 사진 자동 분류 시스템

추승오, 이승엽, 석진훈, 이강민, 윤태상, 유흥석(교신저자)

경운대학교 소프트웨어학부

e-mail: {sam014789, syeop9905, wlsngns00724, vbvbm4, se_ie}@naver.com, hsyoo@ikw.ac.kr

Automatic Photo Classification System Based on Face Feature Extraction and Clustering

Seung-oh Choo, Seung-yeop Lee, Jin-hoon Seok, Gang-min Lee,

Tae-sang Lee, Hongseok Yoo(Corresponding Author)

Dept. of Aeronautical Software Engineering, Kyungwoon University

● 요약 ●

맞벌이 가정이 증가함에 따라 영유아, 장애인, 노인 등의 사회적 약자를 낮시간 동안 보육/보호하는 데이케어 센터의 수요가 증가하고 있다. 데이케어 센터는 센터 경쟁력 확보 및 보호자 만족도 제고를 위해서 피보호자의 일상 사진을 제공하는 곳이 대부분이다. 하지만 데이케어 센터의 직원이 다수의 사람에 대한 사진을 촬영 및 선별해서 메시지를 전송하는 일은 데이케어 센터 본연의 업무를 방해할 수 있다. 따라서 본 논문에서는 사진 선별을 업무 부담을 완화시키는데 도움을 줄 수 있는 얼굴 특징 기반 사진 자동분류하는 시스템을 개발한다. 제안한 방법에서는 얼굴 특징 추출 기법과 클러스터링 알고리즘인 DBSCAN을 이용하여 얼굴 기준 사진 분류시스템을 설계하였다. 특히, OpenCV와 face recognition 라이브러리를 이용하여 카메라로 촬영된 사진 속의 얼굴 객체를 인식하고 얼굴사진을 저장한 후 얼굴의 특징을 추출한다.

키워드: 안면인식(Face Recognition), 군집화(Clustering), DBSCAN

I. Introduction

맞벌이 가정이 증가함에 따라 영유아, 장애인, 노인 등의 사회적 약자를 낮시간 동안 보육/보호하는 데이케어 센터의 수요가 증가하고 있다. 데이케어 센터는 센터 경쟁력 확보 및 보호자 만족도 제고를 위해서 피보호자의 일상 사진을 제공하는 곳이 대부분이다. 하지만 데이케어 센터의 직원이 다수의 사람에 대한 사진을 촬영 및 선별해서 메시지를 전송하는 일은 데이케어 센터 본연의 업무를 방해할 수 있다. 따라서 본 논문에서는 사진 선별 업무의 부담을 완화시키는데 도움을 줄 수 있는 얼굴 특징 기반 사진 자동분류하는 시스템을 개발한다.

사람 객체 인식 및 얼굴 검출과 얼굴의 특징 추출을 위해 OpenCV와 Face recognition 라이브러리를 이용하였다. 인식된 얼굴에 대한 특징값은 임베딩 벡터형식으로 저장하였다. 얼굴 기준으로 사진을 분류하기 위해서 클러스터링 알고리즘을 적용하였다. 사진 속에 등장하는 분류 대상인 사람의 수가 정해진 것이 아니기 때문에 군집의 개수가 미리 정해져서 하는 K-means 알고리즘 대신에 DBSCAN 알고리즘을 이용하여 군집화를 실시하였다. 성공적으로 군집화가 처리되면 해당 얼굴이 포함된 원본사진을 클러스터 인덱스로 명명되는

폴더로 자동 복사처리가 이루어지도록 프로그램을 구성하였다.

II. Preliminaries

제안한 시스템은 Window 10 환경의 Python용 통합 개발 환경인 PyCharm, 머신러닝 프레임워크와 라이브러리를 제공하는 Python 기반의 Anaconda, 사람 객체 인식 및 얼굴 검출과 얼굴의 특징 추출을 위해 OpenCV와 Face recognition 라이브러리를 통해 개발하였다.[1]

클러스터링 알고리즘은 밀도 방식의 클러스터링을 사용하는 DBSCAN(Density-Based Spatial Clustering of Application with Noise)을 사용한다. DBSCAN은 최종적인 군집의 갯 수를 입력 파라미터로 요구 받지 않고 노이즈(noise)에 영향을 받지 않으며, K-means 군집화에 비해 연산량은 많지만 K-means 군집화가 잘 처리하지 못하는 비선형적 데이터를 처리하는 데 유용하다고 알려져 있다.

