

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.

III. Design and Development

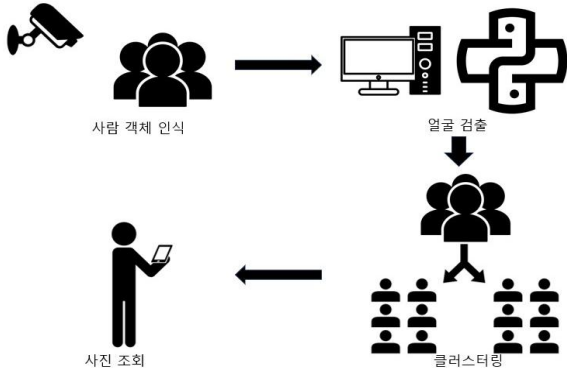


Fig. 1. system diagram

그림1와 같이 제안한 시스템에서 카메라는 화각 내에 사람 객체가 나타나면 정지 영상을 생성한 후 로컬서버로 전송한다. 인공지능 연산 처리가 가능한 컴퓨터에 구축된 로컬 서버에서는 Face recognition 프레임워크를 이용하여 얼굴 검출과 얼굴의 특징 추출을 수행한다. 얼굴 검출과 얼굴의 특징 추출이 완료되면 클러스터링 알고리즘인 DBSCAN을 이용하여 얼굴의 특징 데이터를 기준으로 얼굴 사진에 대한 클러스터링을 수행한다. 얼굴 클러스터링을 일정한 수량의 새로운 사진을 수집될 때마다 자동 실행되도록 스케줄을 구성하였다. 얼굴 클러스터링이 완료가 되면 이어서 각 얼굴 사진이 포함된 원본 사진을 클러스터 인덱스로 명명되는 폴더로 자동 복사가 이루어지도록 하였다. 사용자가 사진 조회를 요청하면 얼굴 썸네일 기준으로 클러스터링 된 원본 사진을 조회할 수 있도록 안드로이드 앱 및 웹기반 인터페이스 구축하였다.

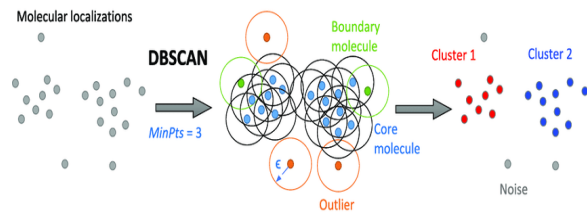


Fig. 2. DBSCAN

그림2는 DBSCAN 알고리즘의 군집화 과정으로 중심점을 기준으로 설정한 반경 epsilon 내에 설정한 MinPts의 값만큼 점이 있을 경우 하나의 군집으로 판단한다.

DBSCAN 알고리즘의 장점으로 클러스터의 밀도에 따라서 클러스터를 연결하기 때문에 기하학적인 모양을 갖는 군집도 잘 찾으며 K-Means 알고리즘과 같이 클러스터의 수를 정하지 않아도 된다는 장점이 있다.

제안한 방식에서는 다양한 epsilon 과 MinPts 값에 대한 클러스터링 성능을 분석하였다고 최대 ±10% 내외의 성능차가 발생하는 것을 확인하였고 최종적으로는 가장 좋은 성능을 제공하는 값을 확보하였다.

0.018	0.022	0.198	0.107	0.145	-0.00256	0.113	0.136	0.0387
-0.024	0.024	-0.09	-0.033	0.166	-0.0234	0.159	0.0218	0.0357
0.0117	-0.0261	-0.087	-0.033	-0.098	-0.118	0.094	-0.095	0.101
0.029	-0.0261	0.165	0.164	-0.0255	0.014			
-0.139	-0.2654	-0.0704	-0.01535	-0.09896	-0.5594-5	-0.0273	0.0405	-0.2401
-0.2523	0.08867	0.1076	0.2579	0.2146	-0.09403	0.04528	-0.1179	
0.0213	-0.1135	0.05627	0.1683	0.02371	0.07089	0.0183	0.0203	0.1176
0.1188	0.08178	-0.0893	-0.2115	0.08354	-0.1128	-0.0255	0.09825	
-0.1895	-0.1185	-0.3538	0.0561	0.0846	0.1085	-0.2129	0.0941	-0.01129
0.1586	0.08029	-0.0730	0.2125	-0.09846	0.01307	0.07082	0.0951	
-0.0114	-0.012-5	-0.1374	0.01697	0.0290	-0.0203	0.0203	0.1495	-0.1913
0.0594	-0.2018	0.1771	0.1453	-0.00291	0.0718	0.1325	0.0624	

Fig. 3. face feature embedding vector

그림3은 Face recognition 프레임워크를 이용하여 얼굴의 특징을 추출한 결과물이고 임베딩 벡터값과 원본 파일의 위치를 표시한다. 이 정보들을 이용하여 클러스터링 작업과 군집별로 사진을 저장한다. 시스템의 정확도를 테스트하기 위해 270장의 사진으로 테스트하였고 82%의 정확도를 보였다.

IV. Conclusions

본 논문에서는 사진 선별 업무의 부담을 완화시키는데 도움을 줄 수 있는 얼굴 특징 기반 사진 자동분류하는 시스템을 개발하였다. 제안한 시스템을 다양한 환경에서 활용하기 위해서는 DBSCAN의 하이퍼 파라미터를 영상 수집 환경에 따라 조절해야 하는 필요성을 확인하였고 향후 적응적 파라미터 조절기법을 연구할 계획이다.

REFERENCES

- [1] <https://pypi.org/project/face-recognition/>
- [2] <https://en.wikipedia.org/wiki/DBSCAN>