1)RGBD 이미지 기반 핸드제스처 분류 딥러닝 기법의 연구

박종찬 이연 신병석
인하대학교 컴퓨터공학과
e-mail: marin0625@naver.com, {leeyeon,bsshin}@inha.ac.kr

A Study on Hand Gesture Classification Deep learning method device based on RGBD Image

Jong-Chan Park, Yan Li, Byeong-Seok Shin
Dept of Computer Science, Inha University

요 약

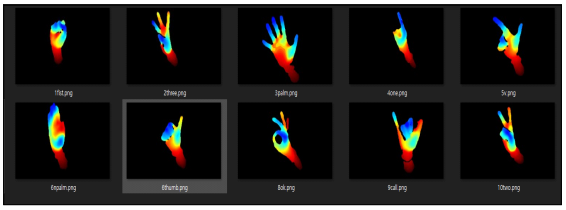
소음이 심하거나 긴급한 상황 등에서 서로 다른 핸드제스처에 대한 인식을 컴퓨터의 입력으로 받고 이를 특정 명령으로 인식하는 등의 연구가 로봇 분야에서 연구되고 있다. 그러나 핸드제스처에 대한 전처리 과정에서 RGB데이터를 활용하거나 또는 스켈레톤을 활용하는 연구들이 다양하게 연구되었지만, 실생활에서의 노이즈가 많아 분류 정확도가 높지 않거나 컴퓨팅 파워의 사용이 과다한 문제가 발생했다. 본 논문에서는 RGBD 이미지를 사용하여 Hand Gesture를 트레이닝 받은 Keras 모델을 통해 입력받은 Hand Gesture을 분류하는 연구를 진행하였다. Depth Camera를 통하여 입력받은 Hand Gesture Raw-Data를 Image로 재구성하여 딥러닝을 진행하였다.

1. 서론

실생활 중 소음이 심하거나 긴급한 상황 속에서 기기의 버튼을 누를 수 없는 경우 장비에 최대한 손을 대지 않고 명령을 내리는 방법[1][2]이 가장 위생적이며 편리한 방법이라 할 수 있다. 이러한 방법들 중 하나로 Hand Gesture를 입력받아 이를 판별하여 명령을 내리는 등의 연구가 진행되고 있다[3][4]. 그러나 일반 RGB포맷의 Hand Gesture 데이터는 손과 배경 색을 혼동하여 정확히 판별하지 못할 수 있고, 다른 기구 및 장비가 있는 경우 RGB 데이터들은 정확도가 크게 떨어지는 모습을 보인다. Depth 자료를 이용하는 방법으로는 손을 3D화하여 사용을 하거나 Skeletal Data를 사용하는 방법으로 문제를 다루고 있었다. 이는 높은 정확도를 가지지만 입력받은 데이터의 처리속도가 매우 느리다는 단점이 있었다. 본 논문에서는 Depth 값을 이용해 상대적으로 가까운 손 부분을 추출하는 방법[5]을 딥러닝 과정에 도입하여 2D 흑백 이미지로 전처리하여 딥러닝 결과의 정확도를 높였다.

1) 본 논문은 정부(정보통신 기술 진흥원)의 재원으로 AR 기반 수술 툴킷 및 애플리케이션 개발(과제번호: 2017-0-018715)의 지원을 받아 수행된 연구입니다.

2. DataSet 촬영과 전처리



(그림 1) Gesture 이미지 모양

Depth 자료 촬영 카메라는 Intel real Sense D415를 이용했다. 촬영 프로그램은 Intel에서 제공하는 Depth Quality Tool을 이용하여 촬영했다.

(그림 1)은 본 연구에서 사용된 10가지 Hand Gesture들이다. 각 Gesture들은 600장~1000장들로 구성되었다. 촬영 결과는 16진수 binary file로 저장되었다.

촬영 결과는 2D 흑백 이미지로 변환되었고, 흑백 이미지를 색에 따라 구분하지 않도록 손 부분 전체를 하얗게 만들어 오직 손의 모양만으로 이미지를 구분하도록 설계하였다. 이 방법은 Depth 카메라의 문제인 그림자 발생 또한 차단해주었다.

3. Training과 Model

conv2D(180*320*32[relu])->Maxpool(180*320*32)->
conv2D(90*160*256[relu])->Maxpool(90*160*256)->
conv2D(22*80*256[relu])->Maxpool(22*80*256)->
conv2D(7*26*256[relu])-> Maxpool(7*26*256)

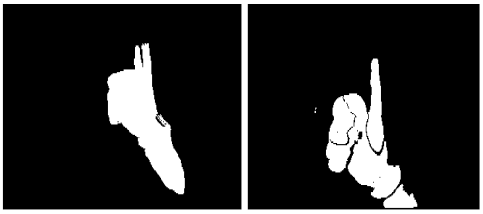
학습은 위의 Keras CNN 모델로 진행되었다. Optimizer는 rmsProp, loss는 categorical cross entropy을 사용하였다. training에는 8482장, validation set 1060장 test에는 1061장이 사용되었다.

	loss	accuracy
last training (validation)	0.0080	0.9972
Test	0.0043	0.9983

(표1) 학습 및 TEST 결과

(그림 2) 실제 상황 실험

Training 완료 후 Test 후 (표1)의 결과들이 나왔다. 실제 사용 시의 정확도를 확인하기 위하여 (그림2)와 같이 세팅을 하였고, 의사들의 수술 영상을 확인한 결과 수술에 따라 다르나 외과 수술들은 약 30cm ~ 50cm 사이의 거리를 두고 수술을 한다는 것을 확인하여 비슷한 상황을 설정하고 촬영하였다. (그림2)의 환경에서 촬영된 이미지들의 실제 test에서는 three 동작과 two 동작, 그리고 fist 동작과 thumb 동작을 서로 혼동하는 모습을 보여주었다.



3-1 3-2

(그림 3) 잘못된 데이터

이유로는 그림3-1과 같이 데이터 촬영 도중 발생한 손 흔들림, 그리고 그림3-2)와 같이 알 수 없는 검은 선들이 Training, Test 데이터 이미지에 들어간 것이 주된 이유로 확인되었다.

4. 실험평가

실험 결과 처음 목표로 삼았던 속도의 개선은 확실히 발전된 것을 확인하였다.

	cpu	memory
python	1%	318.6MB
total	46%	85%

(표 2-1) 컴퓨터 자원 사용량(대기 상태)

	cpu	memory
python:during transform	49%	653.4MB
total	100%	88%

(표 2-2) 컴퓨터 자원 사용량(동작 중)

이미지 처리 중 메모리 사용량의 변화는 크지 않았으나 CPU 점유율이 크게 상승함을 확인하였다. (표2-2)의 상황에서 992장의 640*320 해상도 Raw Data 변환에 33초가 걸렸다. 약 3FPS의 속도를 보여주었다. 기존 Skeletal 시스템을 사용할 때보다 속도가 크게 빨라졌다. 당시 Skeletal 시스템에서의 테스트는 이번 실험 내내 사용한 랩탑 PC(GPU: GTX 1050/CPU: i5-8250u)보다 고성능의 제품인 데스크탑 PC(GPU: GTX 1080ti/CPU: i7-4470)에서 진행하였음에도 불구하고 전처리 과정이 매우 느려 초당 1프레임의 속도를 보여주었다. 하지만 RGBD 데이터를 받아 Depth로 실험을 진행한 결과 (표2-2)에서 처럼 다른 프로그램을 사용 중 진행하여도 초당 3프레임의 속도를 보여주었다. RGB 데이터를 활용한 실험에서의 문제의 예로는, 수술 시 시야 내 환경에 피가 뒤덮여 환자의 몸과 손을 색에 따라 구분할 수 없다는 것이 있었다. 그로 인해 배경과 손의 구분이 어려워져 정확도가 매우 떨어지는 현상이 발생한다. 이러한 문제는 Depth 값을 통하여 상대적으로 가장 가까이 있는 손을 포착하도록 하여 해결했다.

5. 결론

본 연구에서는 색상의 영향을 많이 받는 환경에서 핸드제스처 분류 네트워크가 다양한 색상 노이즈의 영향을 받는 문제점을 해결하기 위해 RGBD 이미지를 사용하는 핸드제스처 분류용 딥러닝 네트워크를 연구하였다. 실제 실험 결과 정확도는 70~85% 수준의 fist, two, three 동작을 제외하면 96~99% 사이의 정확도들을 보여주었다. 이는 잘못된 dataset 제거와 다양한 각도 및 위치를 가진 dataset 추가학습을 통해 고칠 수 있는 문제이다.

참고문헌

[1] Rong Wena, Wei-Liang Tayb, Binh P. Nguyena, Chin-Boon Chnga, Chee-Kong Chui "Hand gesture guided robot-assisted surgery based on a direct augmented reality interface"(38 pages)
[2]Zhou Ren, Jingjing Meng, Junsong Yuan "Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction" (5pages)
[3] Guillaume Devineau, Wang Xi, Fabien Moutarde, Jie Yang(2018)"Deep Learning for Hand Gesture Recognition on
Skeletal Data" (9 pages)
[4] Ebrahim Nasr-Esfahani1, Nader Karimi1, S.M. Reza Soroushmehr, M. Hossein Jafari1, M. Amin Khorsandi, Shadrokh Samavi, Kayvan Najarian "Hand Gesture Recognition for Contactless Device Control in Operating Rooms" (13 pages)
[5]SANJAY MEENA "A Study on Hand Gesture Recognition Technique"(65 pages)