**스테레오 타입으로 구성된 음성데이터의 발원 방향을 추정하는 모델 개발**

팀원 : 20151041990 이수인

**개 요**

스테레오 타입으로 구성되어 있는 음성 데이터에서 노이즈를 제거하여 실제 사람의 음성 데이터만을 추출한 뒤, 해당 음성이 어느 방향으로부터 발원되어오는지를 추정하는 딥러닝 모델 개발

1. **서론** 
   1. **연구 배경**

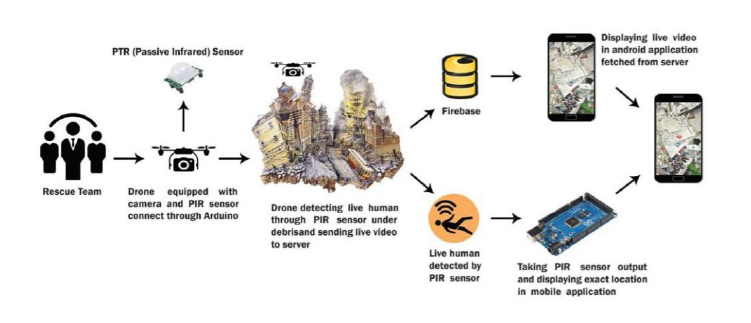
현대 사회에서는 대형복합재난이 지속적으로 발생하고 있다. 대형복합재난은 동시 또는 순차적으로 두 가지 이상의 자연, 또는 사회적인 재난이 발생하는 것이며 그 영향이 복합화되어 피해가 극심하기 때문에 국가적 위협으로 분류하여 범부처의 통합적 대응이 필요한 재난이다. 이러한 대형복합재난의 모든 현장이 사람이 접근 가능한 현장이라고 확신하기는 어렵다. 본 연구는 그러한 상황에서 더욱 더 신속한 인명구조가 이루어질 수 있도록 노이즈가 극심한 음성데이터에서 사람의 음성 만을 추출 하여 해당 발원지를 추정하고자 한다.

* 1. **연구목표**

연구에 사용되는 데이터는 드론에서 취득된 스테레오 음성으로 사람의 구조요청 소리와 극심한 노이즈 소리로 이루어져있다. 본 연구의 목표는 해당 음성 데이터로부터 필요한 사람의 구조요청 소리만을 분리해내고, 해당 소리가 어디로부터 발생되었는지 그 방향을 20도 간격으로 추정하는 것을 목표로 한다.

1. **기존 연구**
   1. **기존 연구 1 - DronAID**

사람을 탐지할 수 있는 실시간 자율 무인 기술 시스템인 ‘DronAID’에 관한 연구로서, 가장 빠른 시점에 생존자의 위치를 정확히 파악하고 구조할 수 있게 하는 것을 목표로 한다. 무인 항공기 기반 시스템이기 때문에 쉽게 동원과 제어가 가능할 뿐더러, 카메라 모듈 및 센서장치가 포함되어 있어 잔해 아래에 묻혀있는 인간의 존재 역시 확인이 가능하다.



PIR 센서를 탑재해 해당 센서가 탐지가 가능한 반경 안에 사람이 있는 경우, 사람이 방출하는 방사선을 감지하고, 그 위치를 파악할 수가 있다. 드론의 시야를 통해 사람을 탐색하는 것이 아닌 방사선을 감지하는 방식이기 때문에 건물 아래에 사람이 깔렸거나 사람이 외부에서 보이지 않는 경우에도 탐지가 가능하다.

* 1. **기존 연구 2 - 드론으로 취득된 음성 정보에서 구조요청 소리 및 방향 감지 딥러닝 모델 개발**

본인이 저번 학기에 진행을 했었던 동일한 연구로써, 드론으로 추출된 스테레오타입의 음성 데이터를 사용해 해당 사람이 어디서부터 구조요청을 발원했는지를 추정하는 연구이다.

* 1. **기존 연구의 문제점**

첫 번째 기존 연구 같은 경우, 희생자를 탐지할 때 사용하는 PIR 센서는 일정한 적외선을 가진 물체가 움직이는 것을 감지하는 것이므로, 사람의 움직임이 없는 경우는 감지가 불가능하다는 단점이 있다. 또한 미세한 정도의 움직임 역시 감지를 하지 못하기 때문에 대형복합재난의 현장에서 희생자가 움직이기가 여의치 않은 경우에는 성능을 기대하기가 어렵다. 더 나아가 센서 자체가 급격한 온도 변화와 같은 주변 환경의 영향을 민감하게 받아들이기 때문에 야외에서 사용할 때는 정확한 탐지가 어렵다.

두 번째 기존 연구 같은 경우, 동일한 주제로 이미 진행했었던 연구이지만, 만족할 만한 결과가 나오지 못했다. 결과적으로 사람의 목소리를 추출하여 해당 목소리의 발원방향을 추정하는 것까지는 성공하였지만, 해당 추정의 범위는 60도 정도로 실제 현장에 해당 각도를 적용시켜본다면 꽤나 넓은 범위가 선택되어 희생자의 정확한 위치 파악이 어렵다는 단점이 있다.

1. **프로젝트** 
   1. **기존 연구와 차이점 및 해결방안’**

데이터는 기존 연구와 동일한 데이터를 이용해 연구를 진행하지만, 해당 데이터로부터 추출할 특징의 개수가 달라질 것이다. 기존에는 MFCC와 GCC-PHAT 두 가지 특징을 사용해 연구를 진행하였지만, 이번에는 그것보다 더욱 더 많은 양의 특징을 추출해 연구를 진행할 예정이다. 또한, 기존 연구에서는 아주 기본적이고 간단한 CNN 모델을 사용하여 음성의 발원방향을 추정하였지만, 이번에는 더욱 더 성능을 향상시키기 위하여 더욱 깊고 복잡한 신경망 모델을 사용할 예정이다.

* 1. **프로젝트 내용**

드론에서 취득된 스테레오 음성을 데이터로 사용하여 프로젝트를 진행한다. 해당 데이터는 사람의 구조요청 소리와 극심한 노이즈 소리로 이루어져있기 때문에 우선은 해당 음성 데이터로부터 필요한 사람의 구조요청 소리만을 분리해낸다. 그 다음 분리된 음성 데이터로부터 모델을 학습시킬 다양한 특징을 추출하고, 궁극적으로는 해당 소리가 어디로부터 발생되었는지 그 방향을 20도 간격으로 추정하는 모델을 개발한다.

1. **진행 일정**

|  |  |  |
| --- | --- | --- |
| 진행 주차 | 담당 | 내용 |
| ~4 | 이수인 | 연구 주제 선정 |
| 5~7 | 이수인 | 데이터 셋 확인 및 노이즈 제거 |
| 8~9 | 이수인 | 특징 추출 |
| 10~11 | 이수인 | 딥러닝 모델 구현 |
| 12~14 | 이수인 | 모델 성능 최적화 및 최종 모델 선정 |
| 15 | 이수인 | 모델 추정 결과 시각화 |
| 16 | 이수인 | 발표 |

1. **결론**

드론에서 입력 받은 구조요청 소리 데이터에서 노이즈를 제외한 사람의 음성 데이터만을 추출한 후, 해당 음성의 발원방향을 딥러닝 모델을 통해 20도 간격으로 추정한다. 이러한 딥러닝 모델은 대형복합재난에서 사람이 직접 들어가지 못하는 재난현장에서의 드론의 구조 역할과 효율을 매우 높이게 되며 결과적으로 많은 인명을 효과적으로 구조할 수 있게 된다.

**참고 문헌**

[1] Rameesha Tariq , Maham Rahim , Nimra Aslam , Narmeen Bawany , Ummay Faseeha ,

DronAID : A Smart Human Detection Drone for Rescue. In IEEE, 2018

[2] 이수인, 이민혜, 박준혁, 이영구. 드론으로 취득된 음성 정보에서 구조요청 소리 및 방향 감지 딥러닝 모델 개발. KCC, 2019