Robust Human Detection under Visual Degradation with Low-cost mmWave Radars and Thermal Cameras

Abstract

The detection of humans in general is a crucial task in many applications that rely on perception such as surveillance, autonomous driving, or search and rescue. Despite the fact that significant progress was made in developing computer vision algorithms for automatic human detection in recent years, the problem is still far from being solved especially under challenging visual degraded environments such as nighttime without sufficient illumination, adverse weather conditions like rain and fog, or fire scenes full of smoke. Such conditions, however, are highly relevant for the already mentioned applications and thus it becomes apparent that there is still both potential and need for improvement.

Human detection neural network models using vision-based methods have been shown to achieve outstanding performance. However, RGB cameras cannot work in a challenging visually degraded environment because of their high dependence on visual conditions, such as fire fighting situations, where optical sensors are impaired by the presence of airborne obscurants (i.e., smoke and dust). Compared to the RGB camera, the thermal camera is robust against visual conditions and has the advantage to be deployed in such environments. But the thermal camera has its own limitations, e.g. lack of robust features due to only temperature profile is captured, low contrast due to its low resolution, and periodical suspension to perform Non-Uniformity Correction. While mmWave radar has a low spatial resolution and strong indoor multipath reflections, it is another sensor that can work in this situation because of its robustness to adverse environmental conditions (e.g., smoke, fog, and dust). Recent studies have shown the effectiveness of single-source human detection using thermal cameras or mmWave radars but still have some challenges, such as noisy mmWave radar data in the narrow indoor environment and the low contrast between humans and the background in the fire scene for thermal cameras. Fusing such perception sources with the goal of leveraging their strengths while compensating for their weaknesses to achieve overall higher accuracy and robustness is very promising.

In our work, we propose a thermal camera and mmWave radar fusion model to solve the challenges above. We consider the general problem of multi-modal human detection, taking input raw data from the thermal camera and mmWave radar. Because there is no public dataset for our case, we self-collect the datasets from indoor environments to evaluate our model performance for real-world challenging human detection tasks. Our current evaluation results show +14% accuracy compared with single-source thermal human detection.

Keyword

human detection, sensor fusion, deep learning