Privacy-compliant People Tracking with a Thermopile Array

Graduate

Gian-Reto Wiher

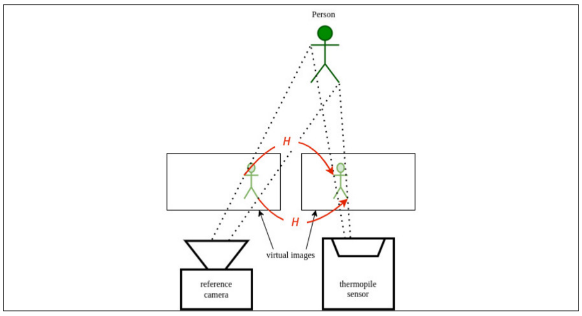
Introduction: Tracking technology is commonly used to gather information from an environment where people reside. However, it raises serious concerns about privacy and potential misuse of personal information as individuals may be identified and tracked over a large network of cameras. The collected data can be used for purposes other than those for which it was intended. It is crucial to use this technology responsibly and there is a growing demand for privacy-compliant surveillance solutions that can provide benefits without compromising the privacy of individuals. A safe method to implement this is by limiting the hardware capabilities of such a system. Thermal imaging cameras containing a low-resolution sensor are one option for privacy-compliant people tracking. The thermal images do not contain color information therefore only shapes can be recognized. The low sensor resolution makes sure the shape can not be recognized precisely. Another advantage of thermal sensors is that they are less dependent on the weather and light conditions compared to visible-light cameras.

Approach: This thesis comes up with a system that records reference images and low-resolution thermal images. The images were taken in different environments and at different distances from people. The dataset consists of 1929 images, of which 276 are used to test the network performance. An existing deep learning network was used to detect people from high-resolution images. With the intrinsic calibration, the location of the virtual image of both cameras can be determined. The extrinsic calibration defines the relative location of both virtual images in the real world. The homography H of the camera pair is determined and therefore detections in the high-resolution reference images can be transferred to the thermal images. With this approach, a dataset containing thermal images with labels, containing information about the location of people, has been generated. This dataset was used to train a deep-learning network and to test its performance.

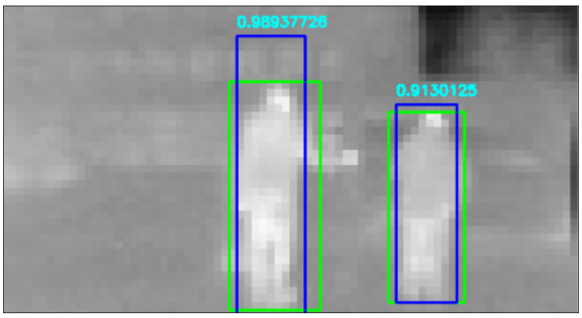
Result: Two different networks were trained. A small YOLO network with reduced size was trained with the recorded dataset. The performance is similar to a larger-sized model, but the inference speed is increased by a factor of seven. YOLO detects people with a size of 6 by 25 pixels with a probability that is typical for the network. In the case that the background contains heated parts, it is possible that the network may not detect a person that stands in front of or very close to it. The thermopile sensor outperformed the reference camera under poor lighting conditions. The YOLO model was successfully retrained to process thermal images and reached a detection rate of over 90% on the test dataset. In addition to increasing the detection success rate, the accuracy of predictions also improved. Thermopile arrays are relatively slow sensors and are therefore not suited for dynamic sceneries. Both versions of the network have difficulties detecting persons in a special pose like kneeling down or stretching out one or both arms. The reason for this weakness is that the training data did not contain many images of such poses. Adding more images with people in different poses will increase the model's performance.

The recording unit used to gather the dataset
Own presentment

Transferring information from the reference images to the thermal images.
Own presentment



The network detects people in thermal images.
Own presentment



Advisor
Hannes Badertscher

Co-Examiner
Gabriel Sidler, Teamup Solutions AG, Zürich, ZH

Subject Area
Sensor, Actuator and Communication Systems