Occupancy monitoring with image recognition methods for low-resolution thermal sensors

Graduate

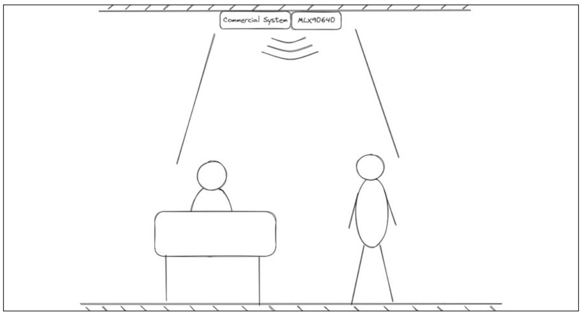
Remo Pfister

Introduction: Demand-driven control of heating, ventilation, air conditioning, lighting and other large energy-consuming systems in facility management is becoming increasingly important in the era of climate change and energy crisis. A key component of demand-driven control are smart sensors. These sensors measure and process relevant information in realtime and trigger appropriate reactions. One of the most important required information for demand-driven control of applications is room occupancy. However, commercial smart sensors that can measure room occupancy are typically expensive because they are based on high-resolution camera systems and other powerful hardware, making them unsuitable for equipping an entire building. In addition, these solutions do not protect people’s privacy, at least not at the hardware level.

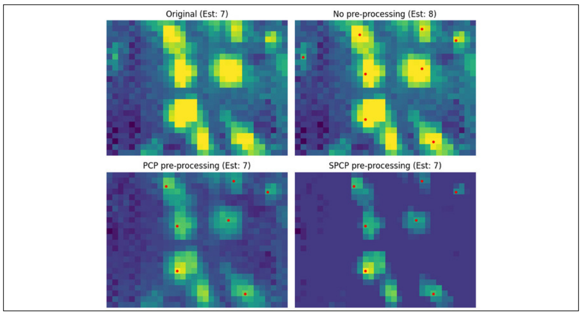
Approach / Technology: In this work, we evaluate various methods from the field of computer vision for their suitability to estimate room occupancy based on low-resolution (32x24 pixel) infrared images. The sensors collecting these images are cheap and protect privacy at the hardware level. We consider both state-of-the-art and tried-and-tested methods. By state-of-the-art methods we mean different neural networks architectures and by tried-and-tested methods we mean maximum filters, blob detectors and other ”classical” image processing solutions. Since the solution is supposed to run on an embedded device, we consider only neural network architectures that are proven to be compatible with the resource constraints of embedded devices. Additionally, we evaluate three data-driven methods for foreground extraction that we use for data pre-processing. These methods filter out the background from the infrared images and thus simplify the occupancy estimation problem. Finally, we compare the performance as well as the number of parameters and the computational complexity of the evaluated solutions. The images were recorded in a breakout room during operation with occupancies ranging from 0 to 15 people.

Result: The results show that the best neural network and the best classical image processing solution performed comparably good. The best neural network architecture, a simple CNN, reached an accuracy of 81.22%, while the best classical solution, a maximum filter, reached 77.43%. However, the maximum filter requires only 23 parameters while the CNN uses 444’880 parameters – an enormous difference. In addition to the significantly reduced number of parameters, the maximum filter does not require extensive training and can be immediately generalized to other settings, making it our preferable solution. Using foreground extraction as a pre-processing step can help reducing the complexity of the occupancy estimation problem. All investigated algorithms were able to successfully extract the background, while stable principal component pursuit (SPCP) was also able to remove most of the noise. This capability makes SPCP our clear favorite as it allows the use of simpler estimators and flexible temperature thresholds.

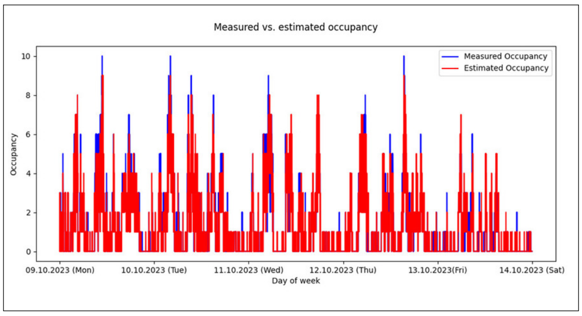
Illustration showing the data collection setup. The two sensors are mounted to the ceiling of the room.
Own presentment



Example of maximum filter solution identifying people in low-resolution thermal images.
Own presentment



Measured vs. estimated occupancy (maximum filter).
Own presentment



Advisor
Prof. Dr. Lin Himmelmann

Co-Examiner
Jürg Dietrich, AXA, Winterthur, ZH

Subject Area
Data Science

Project Partner
Volkher Scholz, Belimo Automation AG, Hinwil, ZH