

Room Occupancy Measurement Using Low-Resolution Infrared Cameras

Martin Berger[†], Alistair Armitage^{*}

*School of Computing
Edinburgh Napier University,
UK*

E-mail: [†]martinberger85@gmx.de

^{*}a.armitage@napier.ac.uk

Abstract — Measuring room occupancy has always been a desirable endeavour. This could be for security reasons, to save energy or simply for statistical reasons. Recently the use of thermal imaging systems in this area has increased as these systems became cheaper and therefore more attractive. This paper investigates room occupancy measurement using a low resolution (16x16 pixel) infrared camera. After giving a short overview of application areas, the chosen people detection algorithm is outlined. This includes the introduction of a novel method for blob classification, using morphological area adjustment.

Keywords — People Detection, Occupancy Measurement, Thermal Imagery, Blob Detection, Low-Resolution, Infrared

I INTRODUCTION

Knowing how many people occupy a certain room or building is a key factor for its energy and security management. The detection of people is increasingly done by thermal imaging systems as they are independent of light and therefore changing illumination. Also, people are normally hotter than their background; this simplifies detection methods in thermal imagery.

A lot of research has been done for detection of people in high-resolution thermal imagery [1, 2, 3]. Most of this research aims towards pedestrian detection in larger, outdoor areas, e.g. streets or parks. Moreover, high resolution infrared cameras are still quite expensive. On the other hand, studies have been done for simple motion detection using very-low-resolution infrared sensors [4, 5]. These sensors are mainly passive infrared (PIR) sensors; they are not able to determine still objects. Also, individual PIR detectors are not sufficiently effective: several have to be combined into an array to cover an adequate area.

A small amount of work has been done using low-resolution (16x16 pixel) PIR sensors to detect people's trajectories [6]. However, for this work

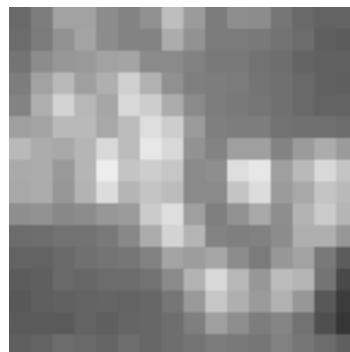


Fig. 1: Example image for low-resolution infrared

we used a 16x16 pixel infrared camera which obtains absolute temperatures using a chopper wheel. This camera provides significantly more information than individual PIR sensors, yet costs a fraction of high resolution infrared cameras. Figure 1 shows an example image of a group of people in a hallway viewed from above.

This paper first presents an overview of the proposed application areas in section II. Section III describes the camera used. In section IV we introduce the proposed algorithm and give a short

overview of it. Section V describes the applied background subtraction approach, whereas the actual person detection is outlined in section VI. Finally we present our results and a conclusion in sections VII and VIII.

II APPLICATION AREAS

Most intelligent building systems use motion detectors to obtain occupancy information, generally for energy control. By using low-resolution cameras we are no longer dependent on the movement of a person as they measure the absolute temperatures of the observed scene. Areas where movement is infrequent or little, e.g. libraries or computer centres, could be monitored more accurately. In addition, simple motion detectors, also known as burglar detectors, can not count people as they do not provide position information.

For security applications, visible CCTV cameras are still very common. However, by using low resolution IR imagery, we do not have to worry about privacy issues because people are not distinguishable by their gender, age or other personal attributes. Also, people detection in low-resolution infrared imagery is less complex than in visible imagery [7].

Another area of application is in the health care sector, where elderly people in assisted living accommodation can be monitored, searching for unusual behaviour [5]. For example, an alarm might be given when the observed person is detected in the middle of a room and does not move for a designated time.

In our investigations, we focused on energy saving applications for indoor areas. For example, an intelligent lighting system can use the occupancy data to automate lighting control. Heating control is also conceivable as we can measure absolute temperatures rather than temperature differences. The occupancy information can then contribute to a demand-oriented heating control. However, the system is just as suitable for security and surveillance applications as the sensitivity can be adjusted accordingly.

III CAMERA PROPERTIES AND LIMITATIONS

Our recordings have been done using an IRI 1011 thermal imager from Irisys¹. The key part of this is the 16 by 16 focal plane array (FPA) detector, which delivers images with a frame rate of 8 fps. An accuracy was given with $\pm 5^{\circ}C$, and a spectral response range between 8 to 14 micrometers, with a temperature detection range of $-10^{\circ}C$ to $300^{\circ}C$.

The major limitation for our purpose was the narrow field of view, with a lens of only 20° . For observing broader scenes, which is necessary for occupancy measurement, we had to do recordings

¹Irisys: <http://www.irisys.co.uk>

from large distances. For example, a distance of 10m observes a plane area of 3.5m by 3.5m.

For this prototype study, we decided to keep the camera in a fixed position, which drastically decreases the background subtraction's complexity. Moreover, the angle between scene and camera was kept near to a top view position, i.e. close to 90° .

IV PEOPLE DETECTION APPROACH

For people detection in high resolution images, pattern recognition can be used to detect parts of the human body. However, in low-resolution thermal imagery we can not see detailed characteristics of people. People appear to have a round "blob" shape. Thus, we decided to use blob detection methods for people detection.

A lot of work has been done to detect blobs in thermal or visual imagery [8, 9]. Area, centre point or direction are only some attributes of blobs which can help to classify them as persons. In our work we use the centre point to obtain peoples' location, whereas the area is used for classification. Figure 2 shows an overview of our approach.

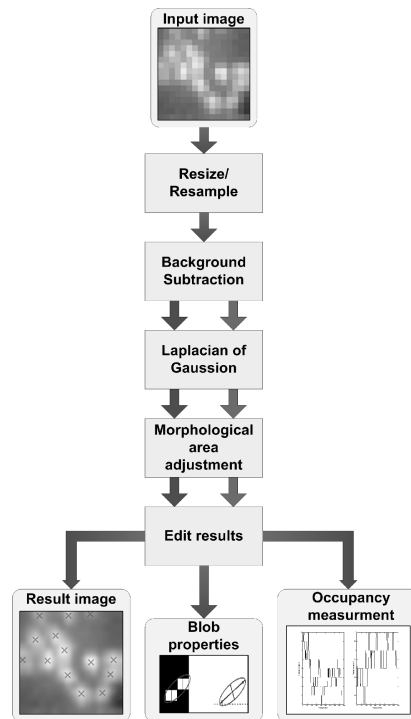


Fig. 2: People detection algorithm

The low-resolution input image needs to be pre-processed before we actually apply blob detection. Thus, the first step in our algorithm is a rescaling operation which increases the size and quality. For rescaling we used bicubic interpolation, which increased the image quality significantly, by reducing the effects of aliasing. Rescaling is followed by the background subtraction explained in section

V. Afterwards, the Laplacian of Gaussian blob detection method is applied, which is then followed by our area adjustment algorithm (see section VI and VII). Finally we extract the detection results and superimpose them on the rescaled image; the blob attributes are also available. Moreover, we measure the occupancy in different image regions by dividing the image into areas of interest and counting people in each area. The algorithm and its testing environment has been implemented using Matlab² with its image processing toolbox.

V BACKGROUND SUBTRACTION

To define people or objects in images one has to determine whether a pixel belongs to the background or to the foreground. In this algorithm we can choose between two different methods to identify the background. The first is simply by choosing background frames manually. The second method is a background estimation using morphological operations. Having estimated the background we then subtract it from the input image. The output is then given by two images, one containing objects warmer than the background and another one consisting of objects colder than the background.

a) Method for choosing frames

Determining the background manually by choosing frames is a fairly easy and fast way to estimate the background. Because of image noise and also changes of the background itself, an average over a range of frames $B_k(x, y)$ should be used to estimate the background $B(x, y)$. Equation 1 shows the calculation of the estimated background using N manually chosen background frames.

$$B(x, y) = \frac{\sum_{k=1}^N B_k(x, y)}{N} \quad (1)$$

However, this method requires a scene where object free frames are obtainable. This might cause problems in constantly occupied scenes. Also, the background changes from time to time, e.g. when the room temperature drops or computers get switched on or off. Therefore, a more dynamic method is introduced in the next section.

b) Morphological background subtraction

This approach uses morphological opening to build a background model. Thus, it uses the geometrical shape informations of persons. Seen from an angle close to top view, this shape appears to have a round, blob shaped boundary, especially in low-resolution images.

In general, morphological opening is often used to eliminate objects which are smaller than their

structuring element S . The same applies for its use for background estimation. By choosing the structuring element to be a similarly shaped object, circular in this case, we can eliminate objects from the image and obtain the background. Therefore, the size of the structuring element has to be greater or equal than the actual object. To delete both warm and cold objects we have to use opening twice. At first we apply it to the original image f_r which gives us the interim result B_a . As morphological operators we used \circ for opening, \ominus for erosion, \oplus for dilation and C for building the complement.

$$B_a = f_r \circ S = (f_r \ominus S) \oplus S \quad (2)$$

The interim result still contains the cold objects as only positive blobs were cropped. To erase cold objects as well, we need to build the complement of the interim result and apply opening again. See equations 3 and 4, with n denoting the image's bit depth.

$$B_a^C(x, y) = -B_a(x, y) + (2^n - 1) \quad (3)$$

$$B^C = B_a^C \circ S = (B_a^C \ominus S) \oplus S \quad (4)$$

As this background image is still in its complement version we need to build the complement once more to obtain the estimated background B .

$$B(x, y) = -B^C(x, y) + (2^n - 1) \quad (5)$$

Because this method calculates the background for each frame, previously calculated backgrounds would be lost. Therefore, a running average, adopted from [10], was implemented; this builds the background estimation using a fraction of the newly calculated background and the previously estimated background. The remembrance rate α defines to what extent the new or the previous background is used for building the current background frame. The range of α is defined as $0 \leq \alpha \leq 1$. Thus, the background of frame i can be determined as seen in equation 6.

$$B_{(i)} = \begin{cases} \tilde{B}_{(i)} & \text{if } i = 1 \\ \alpha \tilde{B}_{(i)} + (1 - \alpha) B_{(i-1)} & \text{otherwise} \end{cases} \quad (6)$$

Alternatively, we can express this in a mathematical manner with all frames taken into account.

$$B_{(i)} = \sum_{k=0}^{i-1} \alpha (1 - \alpha)^k * \tilde{B}_{(i-k)} \quad (7)$$

VI BLOB ENHANCEMENT AND DETECTION

a) Fixed-scale Laplacian of Gaussian (LoG)

People standing close together are hard to distinguish in low-resolution imagery as they merge together and build a new single blob that is obviously

²Mathworks: <http://www.mathworks.com>

bigger. By using the Laplacian of Gaussian (LoG) we can separate merged blobs to some extent.

For this purpose all blobs are assumed to have a circular shape. Thus, applying the LoG with the right scale creates a maximum in the centre of each smaller blob, which is part of the merged blob [9].

During our work various person blobs have been analysed with a range of scale factors to find a reasonable diameter value for an average person. However, the maximum response and therefore the scale factor have been determined experimentally and have to be adjusted if the camera properties, especially the camera height, change. We found out that it is more accurate to determine the scale factor first and to derive an estimated blob diameter from that scale, instead of measuring person diameters and calculating the scale factor from them. The connection between estimated blob diameter d_b and scale factor σ can be seen in equation 8.

$$d_b = 2r_b = 2\sqrt{2}\sigma \quad (8)$$

The normalized LoG function, as seen in equation 9, was chosen as convolution kernel.

$$h_n(x, y, \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^2} \exp^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (9)$$

b) Morphological area adjustment

Whereas the Laplacian of Gaussian was used for initial separation of merged blobs, we now actually detect and measure blobs. Therefore, we are especially interested in a blob's centre point position but also in its area, which is normally used to validate blobs.

We can estimate a value for an average area of a person A_b , by using the estimated blob diameter d_b introduced earlier.

$$A_b = \left(\frac{d_b}{2}\right)^2 \pi \quad (10)$$

The requirements for a blob evaluation are based on how blobs can appear. In the following the three major cases are summarized.

- The easiest case is described by a single blob belonging to a single person with no other blob in contact. We can immediately say that the blob belongs to a single person.
- The second case occurs if people are close together: they can be merged into a single blob. We would have to attempt separating them as far as possible. The main characteristic of this case is that each of the blobs, contributing to the merged blob, belongs to a person and should have the area of A_b .
- The third case describes the phenomenon that a single person results in multiple blobs. For

example, if we imagine a person wearing a thick sweater with outstretched arms, the hands can be detected as extra blobs. These blobs are then separated from the body's blob. In this case we need to try to merge them together.

The main problem we have is that case two and case three are contradictory. In case two we try to separate blobs and in case three we aim to bring them together. If we try to consider both cases we can not simply apply an algorithm to the whole image. Therefore, we had to find a way to evaluate blobs individually.

The found solution is called "area adjustment". The basic idea is, when we adjust the area of each blob to the estimated blob area A_b of a person, some of the small blobs belonging to case three will be enlarged to a typical person size and therefore merge together. On the other hand blobs with a greater area will be narrowed, with the aim of separating merged blobs belonging to multiple persons. Due to the actual implementation of the algorithm it is advantageous to use a slightly higher area value than A_b , which we introduce with A_{max} . Also a minimum value A_{min} was determined to erase unusual small blobs.

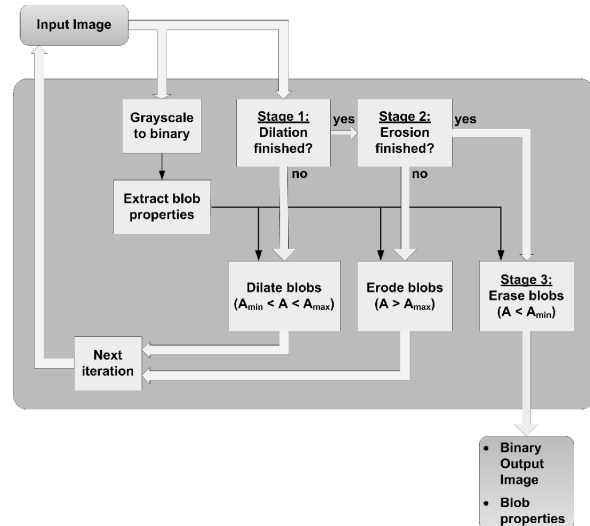


Fig. 3: Overview of the area adjustment algorithm

The algorithm contains three stages:

- In the first stage all blobs with an area A of $A_{min} < A < A_{max}$ are dilated, until their area is greater than A_{max} .
- In the second stage all blobs with an area A of $A_{max} < A$ are eroded, until their area is below A_{max} . This includes also the dilated blobs from the first stage.
- The final erosion for a blob can cause disruption of blobs, which can then result in very

small, invalid blobs. Therefore, all blobs with an area A of $A_{min} > A$ are erased during the last stage.

By applying this three-stage algorithm we obtain adjusted blobs with an area value which is definitely within A_{min} and A_{max} . The majority have an area close to A_{max} .

For the morphological operations, we used a circular shaped structuring element. For every iteration all blob properties are determined from the binary description of the input image. These properties are then provided for the three different stages. After the final deletion of small blobs, the output image is in its binary form. Also, the final blob properties are given, i.e. blob area and centre point. Note that the algorithm has to be applied for the separated warm and cold objects. Figure 3 shows an algorithm overview.

VII RESULTS

We use this section to outline the results achieved in relation to people detection and occupancy measurement.

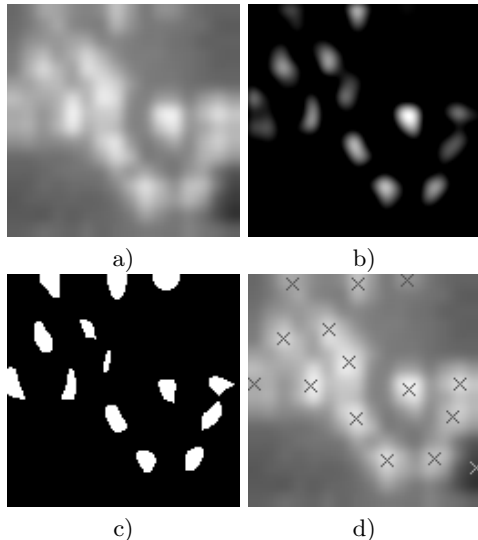


Fig. 4: a) Rescaled input image; b) Warm objects after background subtraction and LoG; c) Binary image after the area adjustment; d) Result image

In figure 4a) we can see the rescaled input image from figure 1 using bicubic interpolation. After subtracting the background using morphological background estimation and applying the Laplacian of Gaussian, we can see a reasonable separation of warm objects in figure 4b). Figure 4c) shows the output of the area optimisation algorithm. In this example all persons have been separated successfully, where the similar size of each blob highlights the principle of the area optimisation. After determining the blobs' centre points, we can output the analysed image, figure 4d), by superimposing crosses at each centre point

onto the input image.

We also extract the room occupancy by determining people counts for the observed area. Figure 5a) shows an occupancy measurement of a crowded hallway, whereas figure 5b) shows the occupancy results obtained after the main University entrance. Smoothing the people count can decrease the failure rate, but also decreases detection rate for fast moving people. In this examples we smoothed the results using a moving average over 8 elements, which equals one second.

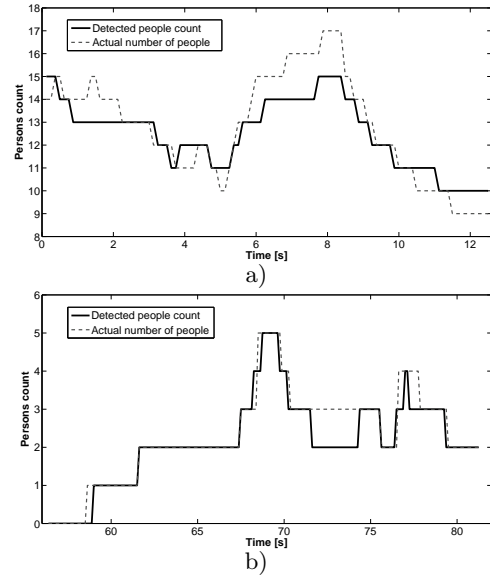


Fig. 5: a) Occupancy measurement of a crowded hallway; b) Occupancy measurement after the University's main entrance

Although the work was focused on post examination of the recordings, we achieved a simulation time that was near to real time for images rescaled with the factor 4 (64x64 pixels). Even though the scene was crowded, and therefore the iterative morphological operations took up significantly more time.

Due to noise and background subtraction deviations, thresholding is applied for warm and cold objects after subtracting them from the background. The threshold values describe the omitted temperature range around the estimated background. Accurate results have been obtained using a threshold of $\pm 2^{\circ}C$, which means that persons need to appear at least $2^{\circ}C$ warmer or colder than the background for a successful detection. Lowering this threshold would increase the detection and failure rate. Alternatively, increasing the threshold would decrease detection and failure rate.

In figure 6, the two circles highlight undetected persons. We can see they are hard to recognise for a human observer as their temperature is close to the background temperature. Thus, they are

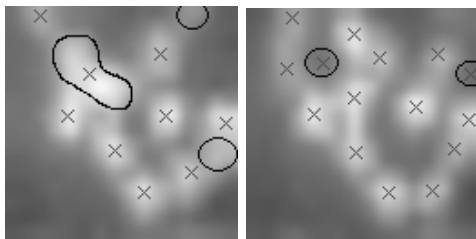


Fig. 6: Not detected and merged persons

Fig. 7: Background estimation fault

neglected due to the threshold explained previously. The marked area in the left of the image shows two people being detected as one, because they are close together and neither the Laplacian of Gaussian nor the area adjustment separated them. If we lower the estimated area for the area adjustment algorithm we can enhance separation of blobs. However, we would also increase the erasure rate and therefore increase the number of undetected persons.

Figure 7 shows detections, based on image noise and inaccuracies of the background subtraction. In this case the threshold value was too low to eliminate the influence of noise and the gap between estimated and real background.

VIII CONCLUSION AND FUTURE WORK

In this paper we present a state of the art solution for room occupancy measurement in low-resolution thermal imagery. We outline possible application areas; we focus on intelligent building systems. A people detection algorithm is presented in detail, where we discuss the implemented background subtraction methods as well as blob detection and evaluation. We introduce a novel approach for blob classification, using iterative morphology to adjust area properties. An extensive set of Matlab code was programmed to implement and evaluate the algorithm. We kept the algorithm's sensitivity adjustable to allow adaptation, based on the requirements of the desired application area. Even though we were using calculation intensive methods, the simulation time could be held close to real time as a result of the limited resolution. We have also outlined the sources and effects of failed detections.

Experiments with various recorded scenes showed promising results, including the separation of close persons. For further improvements of the results, we will enhance the background subtraction methods, to reduce the necessary thresholds and therefore increase detection rate. Furthermore, we will investigate to what extent people tracking can be realised. Also, we plan to evaluate results of recordings, taken with a reasonable lens angle to widen the observed scene.

REFERENCES

- [1] James W. Davis and Vinay Sharma. Robust background-subtraction for person detection in thermal imagery. *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04.*, page 128, July 2004.
- [2] James W. Davis and Mark A. Keck. A two-stage template approach to person detection in thermal imagery. *Seventh IEEE Workshops on Application of Computer Vision (WACV/MOTION'05)*, 1(7):364–369, January 2005.
- [3] Paul Viola, Michael J. Jones, and Daniel Snow. Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision*, 63(2):153–161, July 2005.
- [4] Robert H. Dodier, Gregor P. Henze, Dale K. Tiller, and Xin Guo. Building occupancy detection through sensor belief networks. *Energy and buildings*, 38(6):1033–1043, 2006.
- [5] A. Kaushik, N. Lovell, and B. Celler. Evaluation of pir detector characteristics for monitoring occupancy patterns of elderly people living alone at home. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 3802–3805, August 2007.
- [6] Alistair Armitage, David Binnie, Jon Kerridge, and L Lei. Measuring pedestrian trajectories with low cost infrared detectors: Preliminary results. *Pedestrian and Evacuation Dynamics: 2nd International Conference held at the Old Naval College, University of Greenwich, London, 20-22*, pages 101–110, August 2003.
- [7] J. Kerridge, R. Kukla, A. Willis, A. Armitage, D. Binnie, and L. Lei. In *Traffic and Granular Flow 03*, pages 383–391. Springer, Berlin Heidelberg, 2005.
- [8] Stefan Hinz. Fast and subpixel precise blob detection and attribution. *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, pages 11–14, 2005.
- [9] Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.
- [10] Massimo Piccardi. Background subtraction techniques: a review. In *SMC (4)*, pages 3099–3104, 2004.