Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 19 (2013) 678 – 685

Procedia Computer Science

The 3rd International Conference on Sustainable Energy Information Technology (SEIT 2013)

# Using a thermopile matrix sensor to recognize energy-related activities in offices

Luis Ignacio Lopera Gonzalez, Marc Troost, Oliver Amft

*ACTLab, Signal Processing Systems, TU Eindhoven, Netherlands*

## Abstract

Various installations and appliances used by building occupants are manually operated, including office devices, kitchen appliances, washing basins, etc. By monitoring appliances usage and thus energy consumption, office occupants could received feedback on their energy needs, which is considered vital to spur energy conservation. In this work, we investigate a novel generation of 2D-matrix thermopile sensors for recognising objects and object-occupant interactions from their heat patterns for a total of 21 activities using a single sensor installation. The activities were chosen according to their relevance for appliance energy consumption. We present a processing concept adapted for thermopile matrix sensors to detect and track objects. Furthermore, detected objects were classified according to object state and occupant interaction categories. In scripted and real-life datasets using a ceiling mounted matrix sensor, we demonstrate that a single sensor installation can provide information on various activities, rather than instrumenting many devices and appliances with individual sensors. We show that activities with a clear thermal signature can be recognized with more than 96% accuracy. We also show experimental results for activities that have a thermal signature closer to the ambient temperature.


Selection and peer-review under responsibility of Elhadi M. Shakshuki

*Keywords:* Activity Recognition, Thermopile sensor, Energy profile

## 1. Introduction

Energy conservation while maintaining occupant comfort is a critical optimisation tradeoff in commercial and residential buildings. Although modern building energy management systems (BEMS) can control lighting and heating/ventilation systems, various installations and appliances used by occupants during the day are manually operated, including office devices, kitchen appliance, washing basins, etc. By providing feedback on appliances usage and thus energy consumption, occupant awareness on energy needs can be improved. To provide accurate feedback, usage patterns and occupant activities could be recognised from ambient sensors.

*Email addresses:* `l.i.lopera.gonzalez@tue.nl` (Luis Ignacio Lopera Gonzalez), `m.a.troost@student.tue.nl` (Marc Troost), `o.amft@tue.nl` (Oliver Amft)


Selection and peer-review under responsibility of Elhadi M. Shakshuki
doi:10.1016/j.procs.2013.06.090

Ambient sensor modalities that have been successfully used for activity recognition include video cameras and microphones, e.g. [1, 2]. However, these modalities are often perceived by occupants as privacy intrusive. Moreover, cameras may require regular maintenance to ensure their robust operation. Previous investigations on activity recognition in buildings considered passive infrared (PIR) detectors too, e.g. [3]. While PIR sensors are optimal for motion detection, are low cost, and require minimal maintenance, the obtained motion information is often too limited for activity recognition. In particular, PIR sensors can detect an initial motion in their field of view, but cannot detect constant presence of a heat source, such as a person. This effect is a result of the PIR's operation principle that works by detecting heat-differences. In contrast, thermopile sensors that exploit the Seeback effect to detect temperature differences could continuously detect heat differences. Thus, occupants in the sensor's field of view could be continuously recognised, not only while moving. Moreover, objects that show a temperature difference compared to its surroundings could be identified, such as a coffee pot or sink used with warm or cold water. Activities recognised with a single thermopile sensor could be attributed to energy consumption, e.g. to provide feedback on actual consumption due to using warm or cold water, or using the kettle.

In this work, we investigate a novel generation of thermopile sensors constructed in a 2D-matrix for recognising objects and occupant-object interactions from a single sensor installed in an office pantry area. We used a ceiling mounted sensor matrix to detect heat distributions captured by the matrix' elements and subsequently recognise objects and occupants. With this approach we can show that a single sensor installation can provide information on various activities, rather than instrumenting many devices and appliances with individual sensors.

In particular, this paper provides the following contributions:

1. We introduce the thermopile sensing concept and our processing framework to process the sensor matrix data. The framework detects and tracks objects in the sensors field of view, and classifies the detected objects according to state and interaction categories. The framework provides concurrent responses for all configured and detected objects, thus can process multi-user scenarios.
2. We present our evaluation study comprising (1) a scripted set of 21 activities used as training dataset, and (2) a uncontrolled, real-life dataset used for testing. During the training analysis, optimal features and parameter sets for the classifiers were determined.
3. We evaluate our approach and processing framework using the real-life study dataset and determine classification performances for all concurrent state and interaction classifiers.

## 2. Related Work

PIR sensors have been used to recognize activities, in [4] PIR sensors where used to keep track of how many people where in a room, and in [5] PIR's where used to detect activity as a series of activations of certain areas in the home. Although both efforts presented promising results, their approaches depended on a gateway and assume that activities are performed when entering or exiting a coverage area. If applied to a constrained area i.e. a bathroom, sensor need to be placed on all areas or objects of interest, as presented in [6]. These works required to place sensors in locations where they could interfere with the activity being performed. Multiple devices also means that maintenance requires more effort, even if the sensors are *"tape and forget"*, as argued in [6].

Infra-red cameras provide thermal images that can be conveniently used to tracking people, as they usually shine against cooler backgrounds. In [7] it was shown that thermal images provide advantages for problems like identifying pose and thus, inferring the activity a person. Although well suited for activity recognition, the cost of thermal cameras is higher than that of a thermopile sensor grid. Furthermore, since infra-red cameras look like standard visual light cameras may yield similar privacy concerns, from the perspective of end users.

## 3. Approach

This section details the thermopile sensor concept and particular device choice. Moreover, the processing architecture used to recognise object state and interaction classes are described.

### 3.1. Thermopile sensor

Thermopile sensors are capable of measuring the thermal radiation absorbed on their active area. They belong to the category of thermal detectors, which generate a small thermoelectric voltage proportional to the detected radiation. Their operation principle is based on the Seebeck effect. The Seebeck effect describes the electric current in a closed circuit composed of two dissimilar materials when their junctions are maintained at different temperatures [9, 10]. When several thermopile sensors are arranged in a matrix, a scene image can be constructed from the heat radiation. These temperature differences between the sensing elements (pixels), can be interpreted as objects by measuring how the pixel's values differ from the ambient temperature. In this work, we considered the Panasonic GridEye sensor [8], which is an array of 64 thermopile sensors arranged in a 8 x 8 matrix. Example images and processing steps performed based on the thermal images obtained from the sensor are detailed in the following section.

### 3.2. Processing architecture

The proposed architecture for detecting fine-grained interactions between people and objects of interest in a scene consists of three main modules: sensor layer, object detection layer, and classification layer. See Figure 1 for a diagram of the architecture.

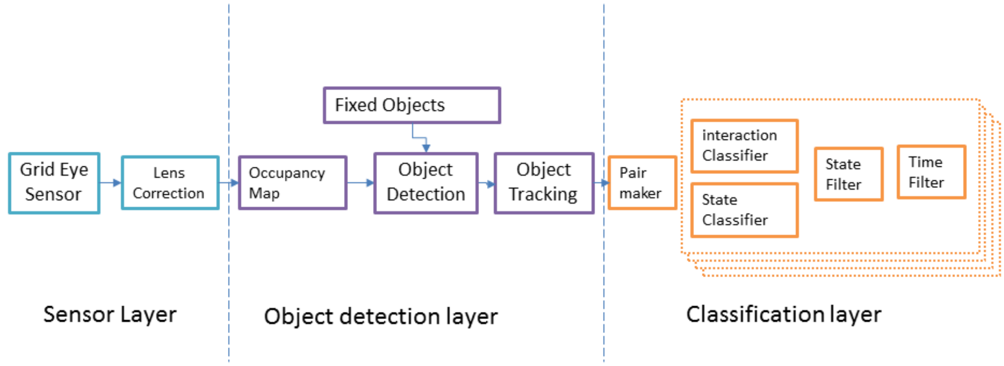



Fig. 1: Detailed diagram of the processing architecture to process thermopile sensor images in this work. Please refer to the main text for more details regarding the functionalities inside each block.

***Sensor layer***. The sensor layer uses raw data from the thermopile sensor matrix and applies a Brown's lens correction (see Eq. 1) to fix the barrel distortion due to the sensor's construction. Here, $r_c$ and $r_u$ are corrected and uncorrected distances of pixels with respect to the optical axis. $K_n$ are radial distortion coefficients, here $K_1 = 7.4 \times 10^{-3}$ and $K_2 = 0.17 \times 10^{-3}$ are used.

$$r_c = r_u + K_1 r_u^3 + K_2 r_u^5 \tag{1}$$

$$T_c = CA_u(T_u - T_{amb}) + T_{amb} \quad \text{and} \quad C = \frac{1}{4d^2 \tan\theta^2} \tag{2}$$

To compensate for the sensor mounting angle, an area correction was applied as described by Eq. 2. Here, $T_c$ and $T_u$ are corrected and uncorrected temperature of a pixel respectively, $T_{amb}$ is the ambient temperature, $A_u$ is the uncorrected projected area of a pixel and $C$ is the area normalization constant. These corrections provide a grid where features appearances are independent of their location in the matrix. The parameters $K_n$ and $A_u$ where fitted according to [8].

***Object detection layer.*** By taking the corrected matrix from the sensor layer, an occupancy map was derived by assigning the most probable state given the pixel's temperature difference with the ambient temperature. The states can be one of empty, cold or hot occupant. The resulting occupied pixels are then grouped by searching the surrounding pixels for the ones in the same state. This approach does not allow for an object to be partially hot and partially cold with respect to ambient temperature. If adjacent pixels present hot-cold behaviour, separate objects will be detected.

Subsequently, information about the scene context was added. We used here prior knowledge on the stationary objects located in the sensors field-of-view. Such objects might not be visible by the sensor, e.g. objects that are overshadowed or at room temperature. In Figure 2 (c) it can be seen how the inclusion of the prior object location knowledge allows to split object blobs into two separate objects. Also, this process allowed us to make a first classification: objects created from intersecting blobs and considering prior object location knowledge are considered static objects, while the remaining ones are considered dynamic objects. The final step in the object detection layer was to keep track of each object across frames. This is needed to keep a consistent history of each object as required by the following steps.

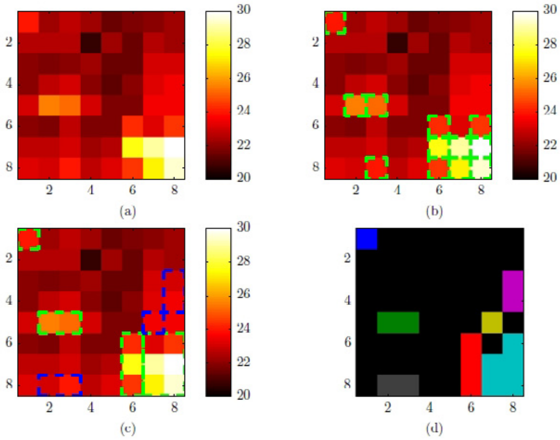



Fig. 2: Output object detection layer for a scene with one dynamic and four static objects; (a) lens corrected sensor output, (b) resulting occupancy map with occupied pixels outlined in green, (c) detected objects (green) and fixed objects (blue), (d) resulting labelled objects.

***Classification layer.*** We arranged all objects identified in an image scene into possible pairs as follows: (1) static object paired with dynamic objects, and (2) pairs of two dynamic objects. For each pair, a reference object was selected. For static-dynamic pairs, the static object was always used as reference. Object processing queues were then created per object pair and results grouped according to the reference object. As a result, the classification layer will provide the current activity for each reference object.

An activity is defined as the state an object or the interaction this object with another one in the scene. We classified the reference object state and its interaction with the other object in the couple. State and interaction results were then fed into a state filter to remove unlikely or impossible states given the sequence in which activities have occurred.

At this point an activity interest list was employed. This list contained all possible interactions with the reference object ordered by their relevance. The current activity for an reference object was determined by selecting from all the detected interactions the one ranking higher in the list.

As a final step, a temporal filtering was applied, equivalent to a low pass filter, to remove short transient states that last very short time periods only.

| Object | Activity | | Scripted dataset | Real-life dataset |
|---|---|---|---|---|
| Coffee Pot | State | Off | 1 | 0 |
| | | On | 3 | 1 |
| | Interaction | Away | 15 | — |
| | | Present | 17 | — |
| | | Serving | 7 | 27 |
| Faucet | State | Off | 8 | 39 |
| | | Hot | 3 | 38 |
| | | Cold | 3 | — |
| | Interaction | Away | 8 | 913 |
| | | Present | 6 | 305 |
| Microwave | State | Off | 6 | 10 |
| | | On | 3 | 9 |
| | Interaction | Away | 9 | — |
| | | Present | 6 | — |
| Refrigerator | State | Closed | 8 | 7 |
| | | Open | 6 | 6 |
| | Interaction | Away | 8 | — |
| | | Present | 6 | — |
| | | Interacting | 12 | 12 |
| People | Interaction | Single | 34 | 24 |
| | | Meeting | 8 | 23 |

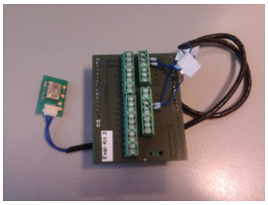
Table 1: Overview on objects, interactions, and occurrence instances in scripted and real-life datasets.

## 4. Study Methodology and Implementation Details

This section describes the study methodology and implementation details of the recordings with the thermopile sensor.

***Test installation and data recording***. The sensor was placed in an office building pantry area, overseeing several static objects and dynamic objects. In particular, the pantry area contained a faucet with hot and cold water, a coffee pot, and a microwave. In this space several activity scenarios are regularly performed, where static objects interact with dynamic objects. As an example, a person (dynamic object) uses the faucet (static object), or two persons are talking (dynamic objects). Table 1 shows the states and interactions considered for each static object, and the interaction considered between dynamic objects. It also lists the activity occurrences in both scripted and real-life dataset.

After classifier training using the scripted dataset, a validation was performed using recordings of ($4.9 Hours$) on a regular working day. All recordings were performed in the pantry area. Ground truth for the activity dataset was obtained using manual annotations from a video recorded at 1 fps. The annotations were up-sampled after the recordings to match the sensor data rate.



(a) Thermopile sensor used: Panasonic GridEye.

(b) Office pantry area used for evaluation.

Fig. 3: Illustrations of the thermopile sensor and placement in the studies. The sensor was placed at the celling, capturing the table corners, microwave, and counter with the faucet, refrigerator, and the coffee pot.

***Object and interaction classifiers***. Classifiers were used to determine state and interaction classes. Table 3 presents the features selected and used for classifications per pair of objects. For state classifiers, only features for the reference object were considered. For interaction classifier, the complete set was used. Since multiple dynamic objects could exist in a scene at any given time, the interaction classifiers ran for all object pairs containing the same static object as reference. In the pantry area and corresponding to the

| Reference object | Coffee pot | | Faucet | | Microwave | | Refrigerator | | People | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Index | Activity | Index | Activity | Index | Activity | Index | Activity | Index | Activity |
| | 3 | Away | 2 | Away | 2 | Away | 3 | Away | 2 | Single |
| | 2 | Present | 1 | Present | 1 | Present | 2 | Present | 1 | Meeting |
| | 1 | Serving | | | | | 1 | Interacting | | |

Table 2: Activity interest list per reference object as used in our evaluation. The lower the index, the higher the interest (relevance) for the recognition. The interest can be adjusted according to application needs.

| Reference objects | | Paired objects | |
|---|---|---|---|
| Index | Feature | Index | Feature |
| 1. | Area temperature product | 2. | Area temperature product |
| 3. | Temperature variance | 4. | Temperature variance |
| 5. | Area | 6. | Area |
| — | | 7. | Distance to object 1 |
| 8. | Gradient X direction | 12. | Gradient X direction |
| 9. | Gradient Y direction | 13. | Gradient Y direction |
| 10. | Gradient magnitude | 14. | Gradient magnitude |
| 11. | Gradient phase | 15. | Gradient phase |
| 16. | Temperature | 17. | Temperature |
| — | | 18. | Position variance |

Table 3: Feature set considered for the object and interaction classification. The list is structured into features for state classifiers (Reference objects) and for interaction classifiers (reference & paired objects).

number of objects, a total of eight classifiers were used. We used support vector machines (SVM), where parameters had been tuned with a grid search method as suggested by [11] on training data.

The state filtering was implemented using HMMs and fitted on the training dataset using hmm-estimate from Matlab. The function calculates the maximum likelihood estimate for transition and emission probabilities given the sequence and known states extracted from the training set. Subsequently the activity interest list refinement was applied. Table 2 shows the activity interest lists defined per reference object. For example, if there are four dynamic objects, and the *Coffee Pot* as static object, then the classified interactions are: $[Away, Away, Serving, Present]$. After applying the activity interest list, the recognised result is *Serving*.

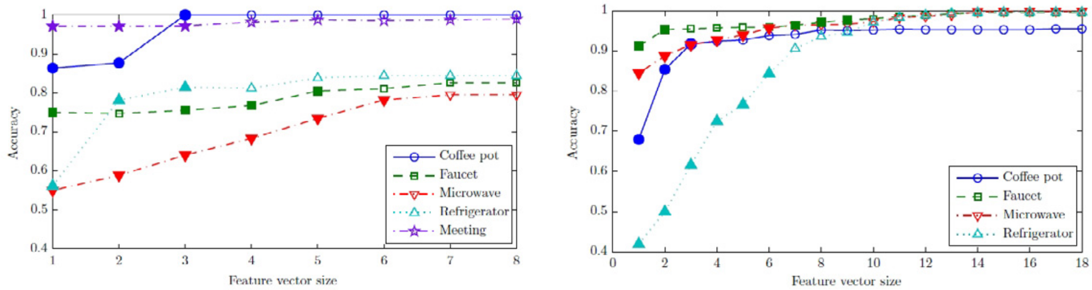
***Evaluation procedure***. To evaluate our approach, we initially determined the relevance of features presented in Tab. 3 for both classifiers. We used a variation of the approach presented in [12] to determine relevance. Instead of using the complete feature set jointly, each feature was evaluated individually since a high degree of correlation could be expected. In result, this yielded a much smaller feature vector size. Subsequently, accuracy performance measurements were obtained using the real-life dataset.

## 5. Results

The feature relevance analysis results are shown in Figure 4 for state and interaction classifiers. As the diagrams indicate, the six best features were sufficient to achieve high accuracy for state classifiers. For interaction classifiers, eight features were needed to obtain high accuracy. Choosing additional features did not improve performance for any of the two classifier groups.

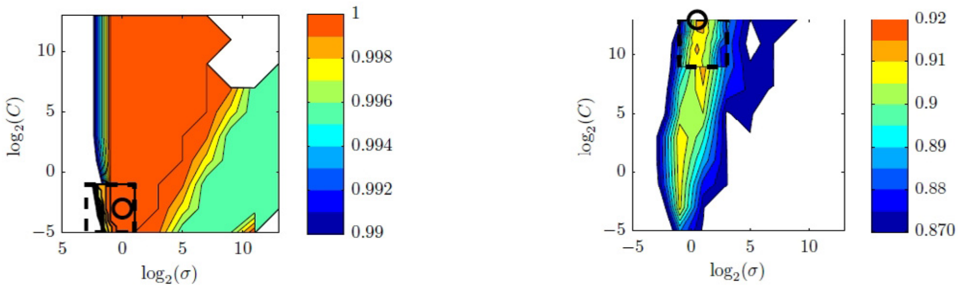
Figure 5 shows an example of the grid search results obtained for SVN parameters $C$ and $\sigma$. Here, a smooth surface near the RBF kernel center with small $\sigma$ could be observed. This indicates that the SVM should perform well with the test data, which is confirmed by the results shown in Table 4. Similar plots where obtained for the other seven classifiers.

Table 4 summarizes classifier modelling performances on the training data. As the result shows, the classifiers can sufficiently model most cases, except for the microwave state classifier. We observed that the microwave states did not result in sufficient temperature difference between *On* and *Off*. Rather, a sequence of interaction events involving the microwave could provide sufficient discriminatory power. This issue




(a) (b)

Fig. 4: Accuracy vs. feature vector size for (a) state classifiers, (b) interaction classifiers. The analysis was performed in a 5-fold cross validation using the training dataset.




(a) (b)

Fig. 5: SVM parameter grid search for *Coffee Pot* classifiers: (a) state classifier, (b) interaction classifier.

becomes more pronounced for the real-life dataset, shown in Table 5. As people will simply stand in front of the microwave, e.g. to read the billboard, the object state cannot be reliably determined.

## 6. Discussion and Conclusion

The 2D-matrix thermopile sensor provides information for detecting complex activities, like serving coffee in a pantry area. Our test scenario showed that the matrix configuration simplified monitoring relatively complex areas. While our approach required to obtain a map of static objects, there was no need to carefully measure overlap of each thermal device with the sensor's field of view. This issue was mentioned by Wren and Tapia [3] as limiting factor for classifying activities using ambient sensors. Although not all interactions described in [3] where tested in this work, e.g. for meetings (corresponding to split and join activities) similar performances were achieved in our approach. Nevertheless, our method required a simpler sensor installation.

The hight at which a sensor is placed determines the tradeoff between coverage and resolution of the scene image. The tradeoff can be observed in Tabs. 5 and 4, where some good performance ratings obtained during training did not hold for validation. We consider that the performance reduction was due to proximity

| Classifier | Coffee pot | Faucet | Microwave | Refrigerator | Meeting |
|---|---|---|---|---|---|
| State | 1.00 | 0.70 | 0.56 | 1.00 | |
| Interaction | 0.82 | 0.97 | 0.86 | 0.73 | 0.85 |

Table 4: Normalized average training set accuracies per classifier. The result shows sufficient modelling capability of the approach for most state and interaction classifiers.

| Object | Activity | | Accuracy |
|---|---|---|---|
| Coffee Pot | State | Off | NaN |
| | | On | 100,00% |
| | Interaction | Away / Present | 96,55% |
| | | Serving | 96,43% |
| Faucet | State | Off | 80,00% |
| | | Cold / Hot | 45,45% |
| Microwave | State | Off | 34,48% |
| | | On | 32,14% |
| Refrigerator | State | Closed | 66,67% |
| | | Open | 50,00% |
| | Interaction | Away / Present | 24,44% |
| | | Interacting | 11,36% |
| People | Interaction | Single | 65,52% |
| | | Meeting | 60,71% |

Table 5: Overview on recognition performance using the real-life dataset for testing classifiers. In the real-life dataset, not all activities were performed, shown here as combined states for some classes.

and size of the areas of interest. For example, the normal use of the faucet, makes people invade the refrigerator area, resulting in erroneous class responses.

Thermopile sensors allow us to recognize multiple activities with one sensing device in places where multiple sensor modalities would have been required. It can be expected that every recognized activity can be mapped to an energy cost, which in turn, can be fed back to the user to guide energy consumption awareness. This energy consumption feedback can be given instantaneously when the activity is being performed. Although our test showed that the sensor can be used to identify complex activities in a scene, the processing was done offline. In further work, the processing architecture could be implemented in the thermopile sensor device.

## 7. Acknowledgments

This work was kindly supported by the EU FP7 project GreenerBuildings, contract no. 258888, and the Netherlands Organisation for Scientific Research (NWO) project EnSO, contract no. 647.000.004.

## 8. References

[1] N. Oliver, A. Garg, E. Horvitz, Layered representations for learning and inferring office activity from multiple sensory channels, Comput Vis Image Und 96 (2) (2004) 163–180, special issue on event detection in video. doi:10.1016/j.cviu.2004.02.004.

[2] D. Kawanaka, T. Okatani, K. Deguchi, Hhmm based recognition of human activity, IEICE Transactions on Information and Systems E89-D (7) (2006) 2180–2185. doi:10.1093/ietisy/e89-d.7.2180.

[3] C. R. Wren, E. M. Tapia, Toward scalable activity recognition for sensor networks, in: LoCA 2006: Proceedings of the 4th International Symposium on Location, Vol. 3987 of Lecture Notes in Computer Science, Springer, 2006, pp. 168–185. doi:10.1.1.65.6272.

[4] F. Wahl, M. Milenkovic, O. Amft, A distributed pir-based approach for estimating people count in office environments, in: EUC 2012:Proceedings of the 10th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, 2012.

[5] K. Murao, T. Terada, A. Yano, R. Matsukura, Detecting room-to-room movement by passive infrared sensors in home environments, in: AwareCast 2012: Workshop on Recent Advances in Behavior Prediction and Pro-active Pervasive Computing, 2012.

[6] E. M. Tapia, S. S. Intille, K. Larson, Activity recognition in the home using simple and ubiquitous sensors, in: Pervasive, 2004, pp. 158–175.

[7] D. Ramanan, D. A. Forsyth, A. Zisserman, Strike a pose: Tracking people by finding stylized poses, in: In CVPR, 2005, pp. 271–278.

[8] Panasonic Electric Works Corporation, Infrared array sensor: Grid-eye. (June 2012).
URL `http://pewa.panasonic.com/assets/pcsd/catalog/grid-eye-catalog.pdf`

[9] C. Kittel, Thermal physics, W. H. Freeman, San Francisco, 1980.

[10] J. Fraden, Handbook of modern sensors physics, designs, and applications, AIP Press/Springer, New York, 2004.

[11] C.-w. Hsu, C.-c. Chang, C.-j. Lin, A practical guide to support vector classification, Bioinformatics 1 (1) (2010) 1–16.

[12] L. Hermes, J. Buhmann, Feature selection for support vector machines, in: Pattern Recognition, 2000. Proceedings. 15th International Conference on, Vol. 2, 2000, pp. 712 –715 vol.2. doi:10.1109/ICPR.2000.906174.