ICESF
Nottingham | UK | 9–11 September 2019

Building thermal output determination using visible spectrum and infrared inputs

G.Sfikas[1,2], A.Noula[1], J.Patacas[3], D.Ioannidis[1], D.Tzovaras[1]

[1]*Centre for Research & Technology, Information Technologies Institute, Thessaloniki, Greece*

[2]*Computer Science & Engineering Dpt., University of Ioannina, Ioannina, Greece*

[3] *School of Science, Engineering and Design, Teesside University, Middlesbrough, UK*

{sfikas,antigoni,djoannid,Dimitrios.Tzovaras}@iti.gr, J.Patacas@tees.ac.uk

Abstract

Accurate building thermal output determination is key for the development of energy use optimisation strategies, including demand response strategies. The analysis of thermal images of buildings presents the opportunity to estimate energy demand based on the actual as-built buildings, as opposed to the current assessment procedures used in the industry which are based on design values.

In this work, we present an image processing pipeline for calculating the thermal output of buildings, by identifying regions of interest given a dual modality (visible spectrum/RGB, infrared) input. The region of interest is assumed to be a building found approximately at the centre of the image field of view ('target building'). The visible spectrum/RGB input is first used to determine the position and outline of the target building in the field of view, and create a pixel-level binary mask with non-zero mask elements corresponding to the target. Subsequently, the produced mask is used to binarize the thermal imaging input and produce an intensity matrix containing only values that correspond only to the building / region of interest. With the proposed method, we are able to take into account only the thermal output of the region of interest, leaving out other image objects and other elements that act as 'noise' in this context. Once computed, the thermal signature of the target building can be subsequently used as input to an energy auditing process or as a component of urban energy planning. The proposed pipeline is evaluated on dual RGB/synthetic thermal image pairs captured on various buildings.

Keywords: Thermal imaging, thermal output determination, object detection, energy demand, Demand Response

1. Introduction

Accurate building thermal output determination presents a series of challenges, due to the difficulty in accounting for differences between the building design values that are used for estimation of energy consumption, and the actual performance of buildings, which is largely influenced by uncertainties associated with the quality of building materials, and the quality of the construction and installation process. Accurate methods for the estimation of energy demand in buildings are a key input for the optimisation of building energy use, including effective demand response (DR) programs that can be leveraged both by smaller energy consumers and energy providers in order to increase the amount of flexibility in DR. The analysis of thermal images for the estimation of energy demand in buildings has been proposed as an alternative to the current energy demand assessment methods, such as the Standard Assessment Procedure (SAP) in the UK [1], however there are still uncertainty issues regarding the proposed methods [2,3,4].

This research investigates how the use of combined thermal and visible-spectrum (RGB) digital images can be used to identify the demand response potential of building assets. In this paper, we propose a method for the identification of potential DR building assets through the application of a novel image processing pipeline for calculating the thermal output of a region of interest given a dual modality (visible spectrum/RGB, infrared) input. Our motivation of using a visible-spectrum input to detect the region of interest is that the infrared input in itself is not appropriate to be used to semantically differentiate objects; the infrared signature may be useful to determine areas of low or high thermal output, but in itself is insufficient to delineate a region of interest efficiently.

The proposed image processing pipeline will be used to identify and provide a range of temperature values for individual building features such as walls, windows, roofs and HVAC assets. Using the proposed image processing pipeline it is possible to isolate features from the building envelope such as the external walls and roof of the building and estimate their U values (i.e. overall heat transfer coefficient (W/m2K)), as well as the transmission heat losses [2,3]. The estimated transmission heat losses will be considered as an input to the estimation of energy demand, and will be used to estimate demand-side improvement potential, which will inform the DR potential estimation process.

Results from the proposed method are expected to provide improved baseline estimation and improved DR flexibility estimation to increase exploitation potential of building assets in DR programs. In particular, the resulting thermal imagery data can be analysed to provide more detailed information about users/customers behaviour.

2. Experimental

The proposed method involves correctly estimating the thermal signature of a target structure, given a thermal imaging camera input, along with an aligned visible spectrum / RGB input. In what follows, thermal signature is understood as the per-pixel statistics of the infrared image input, though the discussed pipeline is applicable to virtually any processing performed over an infrared input. The proposed processing pipeline involves the following steps: a) Use the RGB input to localise the region of interest / target building. This step provides a per-pixel mask that is used to b) determine the thermal output of the target building, using the IR input. The proposed processing pipeline can be examined at figure 1. We have performed our experiments at the ‘smart home’ site, situated at Pylaia-Thessaloniki, Greece and part of the facilities of the Centre for Research and Technology Hellas (CERTH). We have captured 5 visible-spectrum (i.e. colour) images of the target building taken from different poses. As hardware that would include an IR camera aligned with the visible-spectrum was unavailable, we have taken advantage of an available fixed-position IR camera to create synthetic IR images that are aligned pixel-to-pixel with the available RGB shots. Below, we discuss the aforementioned components in further detail.

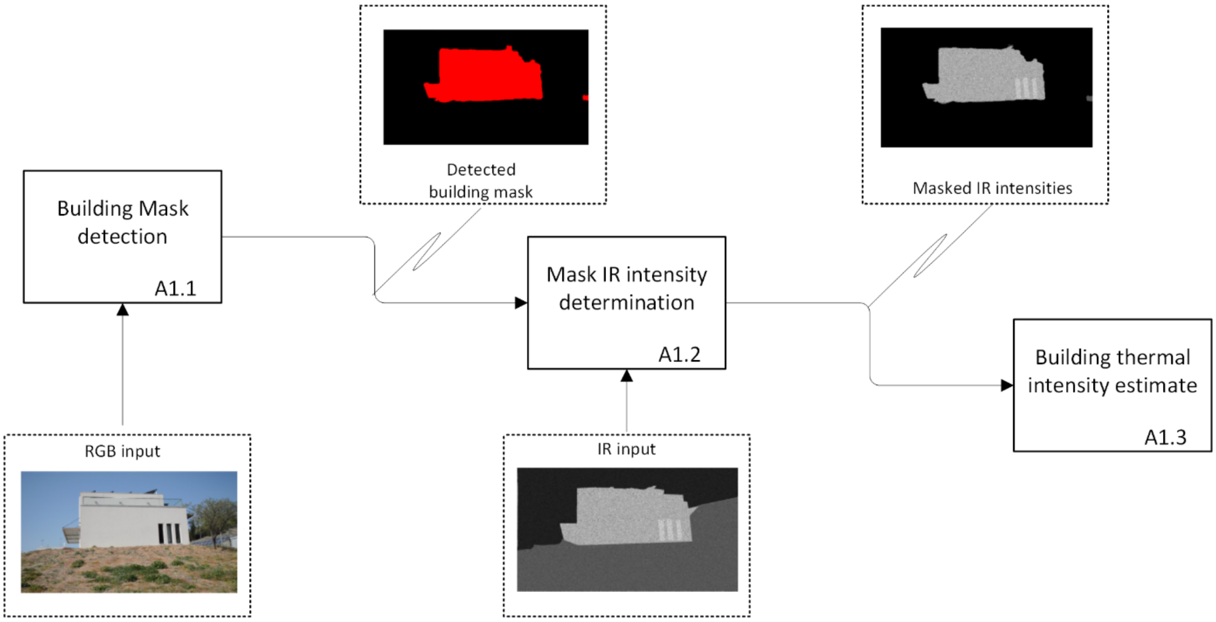

Figure 1: IDEF0 process diagram for the proposed image processing pipeline. The RGB image is used to detect the Region of Interest (target building). The IR image is then masked accordingly and thermal intensity statistics are computed over relevant IR pixels. Without the proposed RGB-based detection component, statistics over the IR image take into account irrelevant pixels, thereby leading to an erroneous thermal signature estimate.

2.1 Region of interest detection

Detection of the region of interest, or in other words localisation of the target building is performed solely over the visible-spectrum/RGB input. Our motivation in using an assumed RGB input for this task is that the latter is semantically much richer than the corresponding IR input; detection of an object is a task related to semantics, as opposed to thermal determination which does not involve image understanding at all, but low-level processing of the IR channel. To the end of target localization, we have used DeepLab v3+ [5], a state-of-the-art neural network-based model for semantic image segmentation. Instead of using this model to directly segment the RGB input, we have used it as a feature extractor, providing so-called Deep Features [6,7]. In this work, we use the activations of the last convolutional layer as our 'deep' features. This is performed without any further training over a pre-trained model[1], by a simple feed-forward pass of the network. We then perform k-means [8] over the extracted deep features, after having reduced the set of the 256-dimensional deep features to 8-dimensional vectors with Principal Component Analysis (PCA, [8]). We use 3 clusters for k-means, which we have found to roughly correspond to the target building, the sky, and other

1 The model backbone we used is Xception, pre-trained on the ADE20K dataset [5].

areas/objects (ground, etc.). Clusters are initialised with k-means++. The advantage of using DeepLab directly over this approach, is that the learning task ‘becomes’ unsupervised, despite that the base model (the neural network) had been pre-trained with a supervised learning process. Consequently, no annotated data is necessary to perform ROI localisation. The cluster with the least average per-pixel distance to the centre image pixel is chosen as the cluster that is related to the target building. The rest are classified as background.

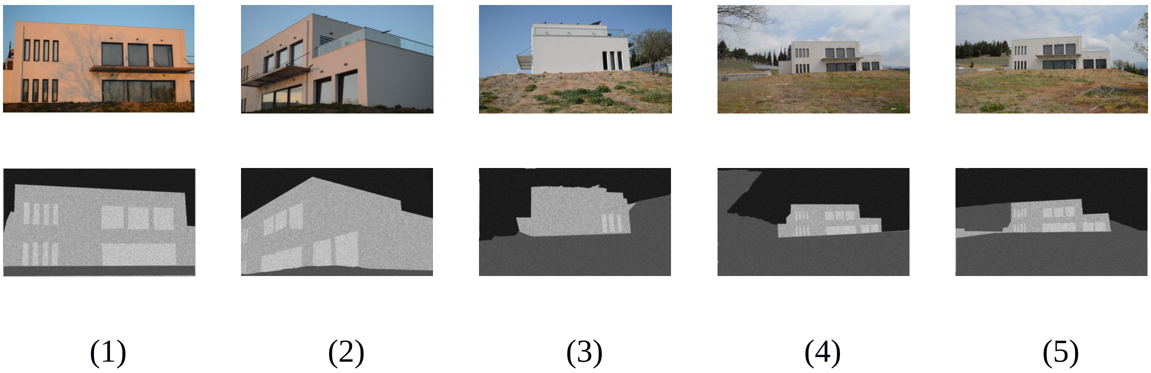

Figure 2: Visible-spectrum/RGB images used to test the proposed pipeline (top row) and corresponding synthetic thermal IR images (bottom row), created using real IR image statistics (see text for details).

Figure 3: IR image used to obtain pixel-level intensity statistics, used to create the synthetic IR images in fig. 2.

2.2 Thermal output determination

Once the area corresponding to the target structure is localized in the previous step, we simply compute mean and standard deviation of the recorded IR intensities. Note that if the previous step was missing, we can expect that IR statistics should be a bad estimate of the target building thermal signature, since it would include pixels not corresponding to the region of interest. This is indeed corroborated by our experiments (cf. following section).

2.3 Infrared image synthesis

While this is not part of the main proposed pipeline per se, this step is necessary to evaluate the algorithm, in order to create aligned RGB/IR pairs where the IR image is as close as possible to a real IR image, w.r.t. to thermal output statistics. We have used a real IR image of the target building (fig. 3) to obtain statistics of IR intensities. In particular, we have computed mean and standard deviation values for pixels in four semantic groups: sky, walls, windows and ground/other. To this end, the real IR image has been manually annotated w.r.t. to the aforementioned semantic classes, and intensity statistics were computed per class. Subsequently, we performed manual annotations on our RGB images w.r.t. to the same classes. For each of the semantic classes and manually delineated areas per image, we drew normally distributed intensity samples that follow the statistics of the real IR image corresponding class statistics. In the context of the current problem, reassuring that the synthesized images follow real image statistics is important, as thermal intensity statistics is what we want to measure.

Table 1: Quantitative comparison of proposed pipeline versus baseline method that does not incorporate visible-spectrum based salient building detection and an Otsu-based localizer. Average absolute values over test images are presented. Lower absolute values are better.

	Percentage offset from ground truth					
	Image 1	Image 2	Image 3	Image 4	Image 5	Mean +- St.dev.
Baseline method	25.7%	29.8%	17.4%	25.1%	24.8%	24.5 +- 4
Otsu-based method	0.9%	38.3%	50.1%	69.6%	46.1%	40.9 +- 22.5
Deep feature-based method (Proposed)	0.3%	2.9%	0.7%	21.9%	21.4%	9.4 +- 10.0

3. Results and Discussion

In order to test the proposed thermal determination method, we have evaluated it quantitatively versus a 'baseline' method and another method where an alternative ROI detection scheme is used ('Otsu-based'). The baseline method involves simply computing statistics of IR image raw intensity values over the whole IR image, as if the full image were of interest. The Otsu-based method employs Otsu [8], a standard binarization algorithm to segment into building and background. Again, the class that is closest to the centre is tagged as building. The proposed method gives the lowest divergence w.r.t. to the true IR mean thermal intensity, i.e. the best result. For 3 out of 5 images, this divergence is minimal, corresponding to an extremely accurate delineation of the target structure boundary. On the other hand, the other methods are consistently far from the real thermal signature; perhaps surprisingly, Otsu results in a worse estimate compared to the baseline, i.e. versus not using ROI localisation at all. This should be attributed to the low quality of ROI localisation attained by Otsu. Results can be examined in table 1, where quantitative results over each shot individually (following the numbering of figure 1), and average offset values are reported.

4. Conclusion

We have shown how to use a neural network-based system to compute a significantly improved estimate of the thermal signature of a target structure. This estimate is useful to applications using heat loss estimates as input, such as Demand Response programs. While our result is to be understood on the premise that we use only synthetic infrared data in this work, we believe that our conclusion can be in all probability extended to real-world data, as we have taken care to use statistics of real IR footage to construct our dataset. Hence, we look forward to testing the proposed pipeline on data that fully include real RGB+infrared image pairs. Future experiments on non-artificial IR data will also determine whether a combination of IR+RGB is useful towards salient object detect, and compare versus using the IR input only for the full processing pipeline.

Acknowledgments

The work presented was carried out as part of the eDREAM project which is co-funded by the EU's Horizon 2020 framework programme for research and innovation under grant agreement No 774478. The authors wish to acknowledge the European Commission for their support, the efforts of the project partners, and the contributions of all those involved in eDREAM.

References

1. BEIS (2014) Standard Assessment Procedure Guidance on how buildings will be SAP energy assessed under the Green Deal and on recent changes to incentivise low carbon developments. Available at: https://www.gov.uk/guidance/standard-assessment-procedure Accessed on: 16th May 2019
2. Paris A. Fokaides, Soteris A. Kalogirou (2011), Application of infrared thermography for the determination of the overall heat transfer coefficient (U-Value) in building envelopes, Applied Energy, Volume 88, Issue 12, 2011, Pages 4358-4365, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2011.05.014.
3. D. González-Aguilera, S. Lagüela, P. Rodríguez-Gonzálvez, D. Hernández-López (2013), Image-based thermographic modeling for assessing energy efficiency of buildings façades, Energy and Buildings, Volume 65, 2013, Pages 29-36, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2013.05.040.
4. Y. Ham & M. Golparvar-Fard (2012) Rapid 3D Energy Modeling for Retrofit Analysis of Existing Buildings Using Thermal and Digital Imagery. LiDAR Magazine, Vol. 2, No. 4
5. L.C. Chen, Y.Zhu, G.Papandreou, F. Schroff, A. Hartwig (2018), Encoder-Decoder with Atrous Separable convolution for semantic image segmentation, the European Conference on Computer Vision (ECCV), pp. 801-818
6. G.Sfikas, G.Retsinas, B.Gatos, Zoning Aggregated Hypercolumns for Keyword Spotting (2016), International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China
7. G.Retsinas, G.Louloudis, N.Stamatopoulos, G.Sfikas, B.Gatos, An Alternative Deep Feature Approach to Line Level Keyword Spotting (2019), IEEE Computer Vision and Pattern recognition (CVPR), Long Beach (CA), USA
8. C.M.Bishop (2006), Pattern Recognition and Machine Learning, Springer.