ASSESSING GEO-TYPICAL TECHNIQUES FOR MODELING BUILDINGS USING THERMAL SIMULATIONS

Dimitri Bulatov[a,*], Benedikt Kottler[a], Eva Strauss[a], Gisela Häufel[a], Marie May[b], Petra Helmholz[c], Francesco Mancini[b]

[a] Fraunhofer IOSB, Gutleuthausstrasse 1, 76275 Ettlingen, Germany,
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
[b] School of Design and the Built Environment, Curtin University, Australia
[c] School of Earth and Planetary Sciences, Faculty of Science and Engineering, Curtin University, Australia

KEY WORDS: Buildings, Digital Twin, Landcover Map, Modeling, Thermal Simulation, Urban

ABSTRACT:

Building modeling from remote sensing data is essential for creating accurate 3D and 4D digital twins, especially for temperature modeling. In order to represent buildings as gap-free, visually appealing, and rich in details models, geo-typical prototypes should be represented in the scene. The sensor data and freely available OSM data are supposed to provide guidelines for best-possible matching. In this paper, the default similarity function based on intersection over union is extended by terms reflecting the similarity of elevation values, orientation towards the road, and trees in the vicinity. The goodness of fit has been evaluated by architecture experts as well as thermal simulations with a thermal image as ground truth and error measures based on mean average error, root mean square and mutual information. It could be concluded that while intersection over union measure still seems to be most preferred by architects, slightly better thermal simulation results are yielded by taking into account all similarity functions.

1. INTRODUCTION

1.1 Motivation

Buildings represent a crucial component of the urban scene, and their accurate modeling from airborne 2D and 3D data has many applications. For example, only a decade ago, creating visually appealing models was considered the main aim for applications in real estate management, augmented and virtual reality, and training and education for quick response tasks (Bulatov et al., 2014). Nowadays, the focus has changed into the direction of Digital Twins, which means that accurate numerical output is expected from the interaction process with these buildings and surrounding terrain (Lu et al., 2020).

For example, in city planning, Urban Heat Islands (UHIs) became an issue of concern. The goal is to construct and to design buildings in a way to minimize the effects of trapping and radiation and to ensure that the hot air enclosed between high buildings and paved roads in urban street canyons can escape and does not keep heating up the surface even during the night (Voogt and Oke, 2003).

One may wonder whether the digital twin of the scene is necessary for the identification and future mitigation of UHIs. One may argue that direct temperature measurements in the vicinity of the building in question and analysis of peak or average temperatures should be sufficient. However, since the surroundings of buildings differ, we cannot scale up the strategy of the direct measurements without simultaneously increasing the costs for data acquisition and processing. Even if we were interested in one specific building, we could hardly predict the development of temperatures in the course of the years. Just imagine that a tree on the same roadside of the building has grown and now provides more shadow while the tree across the road has faded and lost its leaves. On a more ample scale, what if the

dominant direction of winds has altered due to global climate change, which can either exacerbate or mitigate the local microclimate concerning UHIs? Digital twins of urban scenes are increasingly able not only to assess existing or not yet existing designs, such as sustainable roofing materials in development, but also to simulate their developments according to different scenarios affecting local micro-climate. Because of this temporal dimension and, since the temperature is always a function of its previous state, we refer to the 4D digital twin of the scene.

1.2 Related works

The simulation of surface temperatures is based on the heat balance equation with terms reflecting convection, radiation, and conduction (Kottler et al., 2019). The 3D digital twin results from the semantic model instantiation of combined optical (images) and distance (LiDAR) airborne measurements. The reason is that the airborne sensor data represent the most common way to model three-dimensional scenes of large extensions while the state-of-the-art simulators (Kottler et al., 2019, Guo et al., 2018) are able to compute temperatures efficiently. While the ground (Elmqvist et al., 2001, Mousa et al., 2017) and the trees (Dai et al., 2018, Koch et al., 2014) can be modeled at a sufficient detail level, accurate modeling of buildings is tricky because of their varying appearances. The number of related approaches is vast (Bulatov et al., 2014, Meidow et al., 2016, Xiong et al., 2014, Xiong et al., 2015, Verdie et al., 2015, Verma et al., 2006, Lafarge et al., 2010, Sohn et al., 2012, Hu et al., 2018, Shajahan et al., 2019, Yu et al., 2021). This by no means exhaustive list of related works may serve as a hint that the research in this field is not yet completed. A scrutinized glance into the appropriate methods reveals that they either have very special requirements on the underlying sensor data (such as an airborne laser scan cloud of a high point density), or they are limited to rather simple building types. In contrast, for the highly detailed buildings with complex roof forms, the quality of the reconstruction results is not sufficiently high for thermal simulation. The reason is that the models should be gap-free

---

* dimitri.bulatov@iosb.fraunhofer.de

since otherwise, the heat will escape from inside the building to the environment leading to the implausible heat exchange.

One very well-known concept to measure the granularity of building models is level of details (LOD, (Albert et al., 2003)), which is being continuously refined and improved (Biljecki et al., 2016). For example, LOD 1 denotes prismatic models with flat, non-necessarily horizontal roofs while LOD 3 already possesses roof forms, including roof overhangs and façade elements (Verdie et al., 2015). The work of (Tang et al., 2020), where among other, windows and doors are neglected for shadow model computation, is a good example of application-based handling the LOD concept. Another taxonomy, mentioned in (McAlinden et al., 2015, Bulatov et al., 2020) and adopted for this paper as well, is to differentiate between geo-typical and geo-specific buildings. The appearance of a geo-specific building model reflects that of the prototypes. It can be reconstructed in an automatic way by one of the methods mentioned above, corrected interactively (Xiong et al., 2014), or even modeled manually to a significant part (Döllner et al., 2005, Bulatov et al., 2021). A set of reconstructed buildings is referred to as a library. The geo-typical modeling means to find for every building in the dataset the most similar library entry whereby the degree of resemblance is quantified using some similarity function and is based on 2D and 3D sensor data. By reusing the reconstruction, the building in the dataset will look almost the same way as a building appearing somewhere else; however, it will slightly differ in its real appearance.

1.3 Contribution

The idea of modeling whole cities using building libraries is not new. For example, (Döllner et al., 2005) define a top-down hierarchy of parameters for procedural modeling buildings. Those buildings mostly have rectangular footprints and simple roof structures. In (Tang et al., 2020), on the contrary, the meshes could be adjusted for specific simulations, such as shadow computation; however only computing time was assessed, but not the correctness of the simulation. The first contribution of our work is to fit challenging from the architectural point of view building models into the already existing scene. The scene is derived from the sensor data, and taking this same sensor data into account for building assessing results in a uniform, gap-free model. The second contribution is to simulate temperatures of large scenes, together with environmental data and physical properties of the underlying materials. Using thermal airborne imagery, we provide the numerical output of the 4D digital thermal twin with geo-typical building models. It will become clear that the proposed method possesses the necessary trade-off between accuracy, i.e. ability to simulate buildings in a realistic way, and efficiency, that is, keeping the model complexity low in order for the simulation to run fast and for the urban planner to play with numerous scenarios for scene design.

1.4 Organization

The next section 2 starts at providing the necessary overview of the available data and the intermediate results, which are relevant for geo-typical building modeling and simulation as well as the problem statement. Then, we provide solutions for geo-typical building modeling. The functionality of our simulator and the modeling results are described in Section 3. Conclusions and ideas for future research are given in Section 4.

2. METHODOLOGY

2.1 Overview of available sensor data and intermediate results

We are given an aerial image with five channels (red, green, blue, red-edge, and near-infrared) and DSM (Digital Surface Model), from which the ground model (or Digital Terrain Model, DTM), height over ground (Normalized DSM or NDSM), land cover classification map, as well as materials and colors of building roofs can be retrieved (Ilehag et al., 2018, Bulatov et al., 2020). The GSD (Ground Sampling Distance) of 0.5m means that the buildings are clearly distinguishable. While the 3D planar segments corresponding to larger roof details can be determined with a procedure called J-Linkage (Toldo and Fusiello, 2008), this is not the case for the smaller ones. Those are recognizable in 2D using the nadir high-resolution images. Their GSD is below 0.1, and they can be automatically geo-referenced into the coordinate system of the DSM and aerial image; however, there is no 3D data of such a fine GSD. The building outlines are available from a Geographic Information System cadastral map. There is also a similar vector data for main roads, stored in an OpenStreetMap (OSM) shapefile. Finally, 20 typical buildings in the area of interest were selected and interactively reconstructed to form our library of 20 building models for the upcoming geo-typical modeling. To create visually appealing building models, the 2D high-resolution image cropped around the building together with the information on the outlines from the GIS data and on dominant planes drawn in are processed by the Revit software. Then it is saved into the readable file as a triangle mesh with five classes: Floor, roof, wall, window, and door. These models make up our library for the upcoming geo-typical modeling.

2.2 Principles of geo-typical modeling

We denote by $\mathcal{L}$ the set of library elements and $\mathcal{B}$ is the set of buildings in the dataset. We need to find for each $B \in \mathcal{B}$ a $L \in \mathcal{L}$ with the highest similarity. The similarity measure should keep in mind the main application, that is, simulation and is, therefore, induced from the previously described sensor data. For example, Intersection over Union (IoU) over the ground planes of B and L is one of the simplest similarity measures since it ignores the elevation values and roof forms. The following approach has been implemented: First, the outline of every building (in $\mathcal{B}$ and in $\mathcal{L}$) is rotated in the way that the longest side corresponds to the y axis and then cropped. Doing so in the case of matching piece-wise rectangular buildings means fixing the orientation and thus, reducing the number of degrees of freedom. The resulting binary mask with values 0 and 1 for pixels outside and inside of the polygon is called bitplane o, and the transformation described by a 3×3 homography matrix into the original coordinate system is denoted as h, that is, h_B and h_L. For comparing two bitplanes o_L and o_B, they must have the same size, requiring a scaling (in x and y directions) transformation h_s of L. We augment the number of models in the library by rotating the bitplanes by multiples of $90°$ and flipping them around the x-axis, resulting in transformations h_r and h_f, respectively, as well as the composed index $\tilde{L} = \{L, r, f\} \in \tilde{\mathcal{L}}$ denoting augmentation of $\mathcal{L}$. For every parameter set, the comparison using IoU

$$\text{IoU}(\tilde{L}, B) = \frac{|o_{\tilde{L}} \cap o_B|}{|o_{\tilde{L}} \cup o_B|} \qquad (1)$$

is carried out and the parameters yielding the highest score

$$\tilde{L}^* = \arg\max_{\tilde{L}} \text{IoU}(\tilde{L}, B) \qquad (2)$$

are stored. Now, the 2D transformation of the library building is given by:

$$h = h_B \cdot h_{f^*} \cdot h_{s^*} \cdot h_{r^*} \cdot h_{L^*}^{-1} \qquad (3)$$

and in the direction of z axis, the base position g and elevation e of both buildings can be retrieved either from the sensor data or from the shapefile and used to compute the offset and scale in the direction of z-axis:

$$z_B = s_z (z_L - g_L) + g_B, \text{ where } s_z = e_B/e_L. \qquad (4)$$

In the case of a high deviation between s^* and s_z, which can potentially lead to implausible sizes of doors and windows, our application outputs a warning.

In the example presented in Figure 1, we can see that the roof structure of two buildings in (a) and (b), or, equivalently (f) and (g) are quite similar, however, not that much as the very high score, exceeding 0.96, may suggest. This is not surprising because only the 2D bitplanes are compared. In the next sections, we present and assess further ways to assess similarity.

2.3 Generalized similarity function for geo-typical modeling

In the previous section, geo-typical modeling could be accomplished without sensor data, strictly speaking, because the building footprints, its base level and elevation can be retrieved from the shapefile. The aim of this section is to extend the similarity function from (1) to be able to take into account the sensor data since the building orientation towards the road and the trees too close to the outlines may provide important clues for exchangeability of B and L.

Once again, the best match L^* is defined by

$$L^* = \arg\max_{L \in \mathcal{L}} M(L, B), \qquad (5)$$

whereby this time, we generalize L^* omitting augmentation parameters f and r from (2). Furthermore,

$$M(L, B) = \text{IoU}(L, B) + \alpha_z M_z(L, B) \\ + \alpha_{\mathcal{R}} M_{\mathcal{R}}(L, B) + \alpha_{\mathcal{T}} M_{\mathcal{T}}(L, B), \qquad (6)$$

whereby M_z measures similarities with respect to the height information while $M_{\mathcal{R}}$ and $M_{\mathcal{T}}$ denote similarity terms induced by the road class and tree class, respectively. The corresponding weights α_z, $\alpha_{\mathcal{R}}$, and $\alpha_{\mathcal{T}}$ can be determined by experts, see next section. For example, in order to penalize collisions with surrounding trees that may negatively affect the simulation, $\alpha_{\mathcal{T}}$ must be chosen very high. Note that setting one or more weight to infinity would simply mean to ignore the IoU based term.

The term M_z extends the similarity function from (1) by the 3rd dimension. Since we scale with respect to z in (4), M_z must be invariant with respect to multiplication of NDSM values of the overlapping masks by a constant. Thus, let $Z_B : \mathbb{R}^2 \to \mathbb{R}$ the function that assigns the NDSM value to each point for building B and $Z_B\{o_L \cap o_B\}$ the restriction to the overlapping

part between o_L and o_B, then

$$M_z(L, B) = 1 - \|N_{Z,B}(Z_B) - N_{Z,B}(Z_L)\|, \qquad (7)$$

where

$$N_{Z,B}(Z_B) = \frac{Z_B\{o_L \cap o_B\}}{\max_{o_L \cap o_B} Z_B \cdot |o_L \cap o_B|}^{1/2}. \qquad (8)$$

The reason to prefer the L^2 norm is that we want to consider the whole roof structure and not only single characteristics of its distribution.

The measure $M_{\mathcal{R}}$ considers the orientation of the building towards the road. For this purpose, we calculate the Euclidean distance transformation R (road distance map) of the road class, which indicates the distance from each point to the nearest road. Hereby the road class is the rasterized by means of the Bresenham algorithm shapefile from the freely available OSM data. The alternative, namely to take the binary image induced from the land cover classification result, appeared less promising since the algorithm outputs general asphalt surface and not the main roads. Then R_B is the cutout of the building B in the road class map. The distance data is truncated and concatenated into a vector, which is then normalized to have the unit norm. To find out which building in the library has a similar street orientation, the cosine metric induced by the L_2 norm is used:

$$M_{\mathcal{R}} = v(R_B)^T v(R_L), \|R_B\| = \|R_L\| = 1. \qquad (9)$$

The last similarity measure indicates whether the geo-typical building has a collision with a tree in the vicinity. The tree class T is identified from the classification result and adjusted with morphological operators to cut off overhanging tree tops. Let $T_{B,L}$ be cutout of the building B in the tree class map joint with o_L. The result is a function $T_{B,L} : o_B \cup o_L \to \{0, 1\}$, which assigns to each point whether there is a tree (1) or not (0). Then, $M_{\mathcal{T}}$ is defined as the proportion of tree area in the joint area of o_L and o_B, thus

$$M_{\mathcal{T}} = \frac{|T_{B,L}^{-1}(0)|}{|o_L \cup o_B|}. \qquad (10)$$

3. RESULTS

We used the Melville dataset from the eponymous district of Perth, Australia, which was captured within a measurement campaign in 2016. The data and the pre-processing steps are described in Section 2.1. We replaced 72 buildings to the north of a park with the geo-typical models resulting from four sets of parameters α_z, $\alpha_{\mathcal{R}}$, and $\alpha_{\mathcal{T}}$ from (2):

1. $\alpha_z = \alpha_{\mathcal{R}} = \alpha_{\mathcal{T}} = 0$ (the default IoU configuration)

2. $\alpha_z = \infty$ (only uniformity of elevations matter)

3. $\alpha_z = \alpha_{\mathcal{T}} = 1$ and $\alpha_{\mathcal{R}} = 0$ (roads omitted)

4. $\alpha_z = \alpha_{\mathcal{R}} = \alpha_{\mathcal{T}} = 1$ (everything weighted equally).

A thermal infrared image, needed to evaluate the temperatures, was acquired with the sensor FLIR A615 during two nights. The weather server provided the weather data for the day of the measurement campaign.
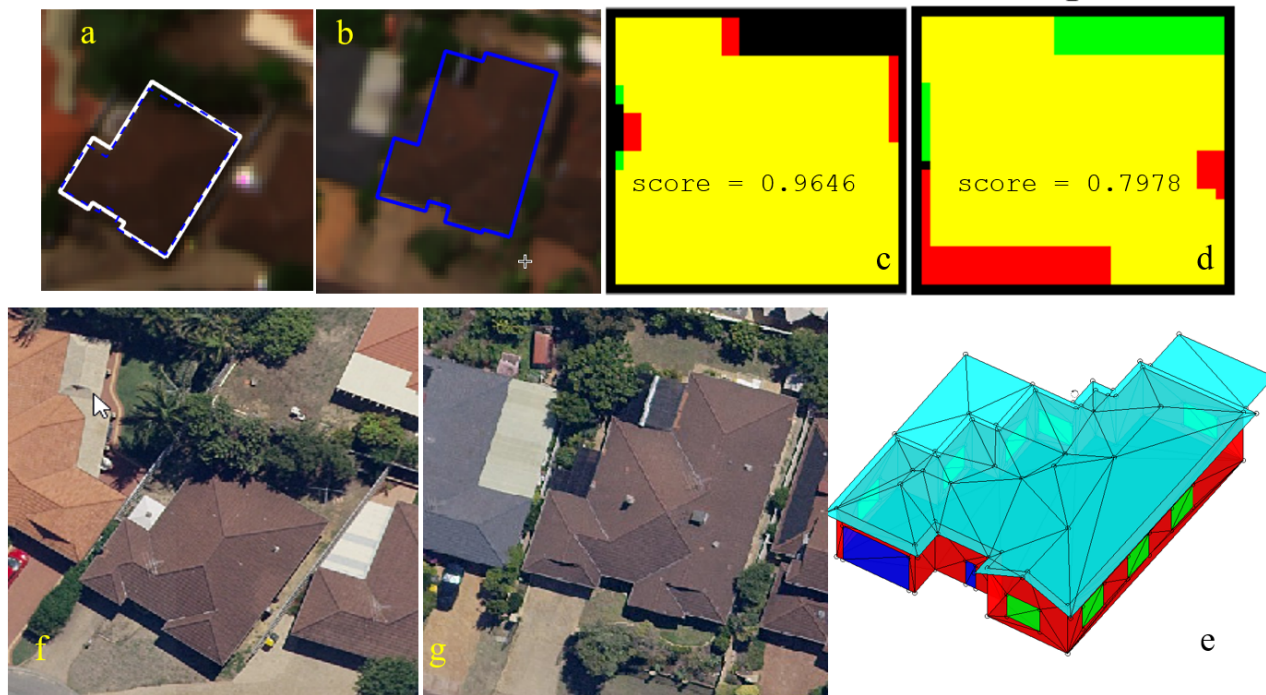
Figure 1. Illustration of geo-typical building modeling. Image (a): A building in the multi-spectral image outlined by white polygon stemming from the GIS data, (b) shows a library building having the highest similarity. Its blue outline polygon is projected into (a) using homography h, yielding the dashed blue polygon. Image (c) shows the superposition of bitplanes yielding the highest score. Here, yellow means the intersection of two rasterized images while red and green denote parts of the library building not present in the source building and, respectively, vice-versa. Image (d) shows the superposition of bitplanes induced by a different homography and yielding a considerably lower score. Image (e) shows the 3D model of the library building to be reproduced. Images (f) and (g) are cutouts from the high-resolution images around buildings in (a) and (b), respectively.

3.1 Interactive assessment of geo-typical modeling

In the section, we asked two independent experts with the architectural background to assess the aforementioned configurations. Both experts were given high-resolution images with cutouts around buildings $B \in \mathcal{B}$ with 3D polygons driven in four semi-transparent colors for four configurations. The experts were shown one configuration for one building after another and asked how expensive they would consider replacing the building B by L regarding the similarity of the outline and the roof. Moreover, the same question was asked about the outlines only. The results are recorded in Table 1, and it is important to note that both experts had different criteria in mind resulting in a non-uniform scale of entries, which were later converted in percents. Both experts clearly preferred configurations 1 and 3 over 2 and 4, while they seem to disagree upon the worst choice of parameters. Overall, it can be concluded that considering intersection over union only yields the most stable results.

In Figure 2, we illustrate some qualitative results. In the top row, we display two examples in which we are able to recognize the important patterns in outlines and roof structures. The bottom row shows another two examples of buildings, for which the scarcity of the library leads to suboptimal results since none of the proposed configurations can even fill out the footprint properly. The observation of experts that well-matching outlines are necessary but not sufficient condition for overall good matching seems to be confirmed.

	C1	C2	C3	C4
Expert 1	75.7/100	70.0/93.1	78.6/100	70.0 / 82.4
Expert 2	77.1/95.6	41.7/45.8	66.7/91.7	56.3/62.5

Table 1. Evaluation of configurations (C) by experts. The numbers refer to scores converted into percents. The first number always indicates the overall fitting score and second number evaluation of the outline.

3.2 Assessment of geo-typical modeling based on thermal simulation

The thermal simulation was carried out based on the workflow of (Kottler et al., 2019) for our four 3D digital twins. In this workflow, the scene is subdivided into triangle mesh, after which the temperature calculation for each triangle and discrete time step takes place by solving the heat balance equation with terms for thermal radiation, convection and conduction. Starting at an initial value, a differential equation is solved using Euler forward integration method. An excerpt of the results of the 4D digital twins is shown in Figure 3. For each configuration, the same viewpoint is displayed to highlight the differences in geo-typical modeling. The temperature ranges comply with each other. However, differences in surface temperatures are more profound on roofs, where deviations in structure occur, and on the ground, where building outlines change between configurations.
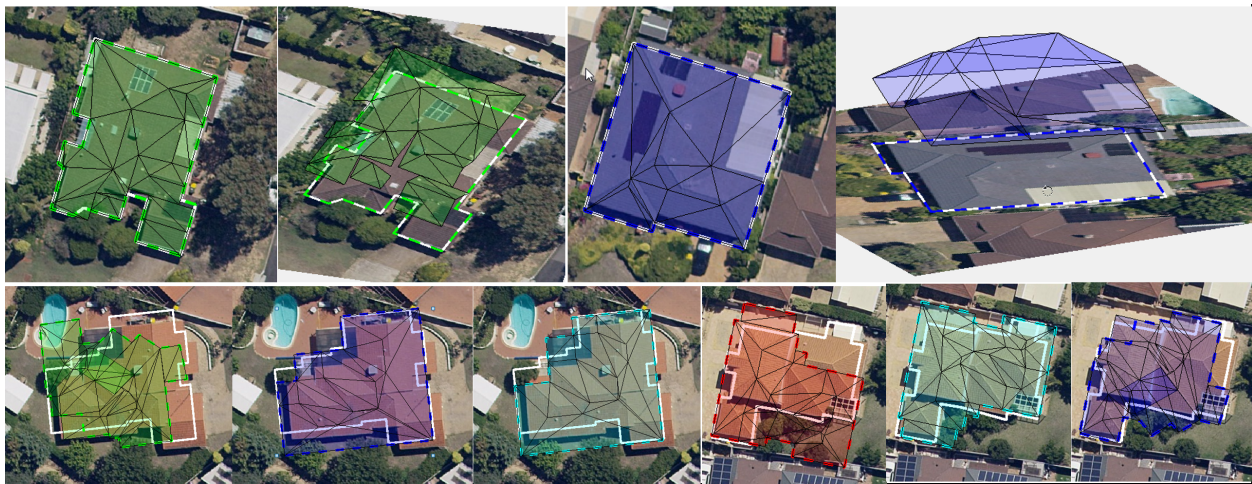
Figure 2. Examples of successful (top row, 2D and 3D views) and not successful (bottom row) results of geo-typical modeling.
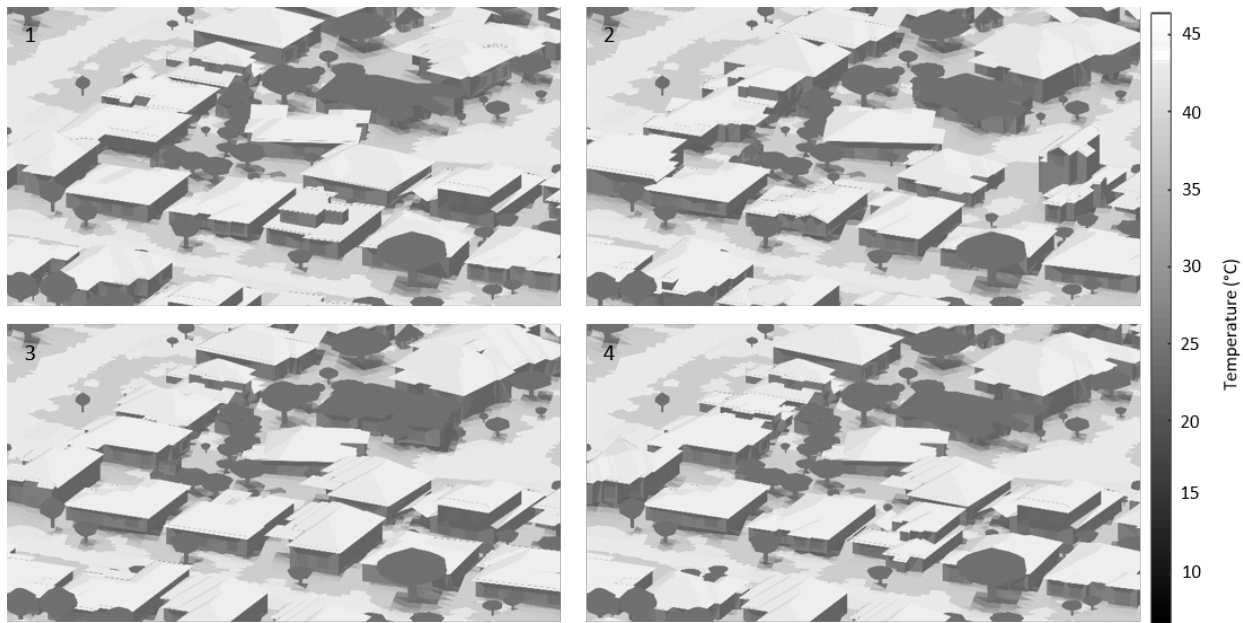

Figure 3. 4D Digital Twins of configurations 1 to 4. The displayed temperatures correspond to 2pm on a typical, i.e. hot and dry summer day in the City of Melville.

For a quantitative evaluation, the orthographic thermal image of the area in consideration was used. In order to compare the simulation results with this image, an orthogonal projection of the three-dimensional results was performed, yielding the simulated orthophoto. Since all given data is geo-referenced, the measured thermal image could easily be cropped to the area in consideration. Figure 4 shows the cropped measured thermal image together with the simulated orthophoto for each of the four building configurations.

Measured and simulated thermal orthophotos were compared pixel-wise based on the following indicators: mean temperature difference, i.e., Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mutual Information (MI). The latter is a similarity metric arising from information theory, quantifying the statistical correlation of two datasets by the two marginal and the joint image probability function. A high MI value indicates a high similarity between temperature distributions.

Table 2. Thermal simulation results for each configuration.

Configuration	MAE	RMSE	MI
1	-0.220	3.013	0.687
2	-0.216	3.033	0.675
3	-0.261	3.008	0.689
4	-0.197	3.002	0.697

As Table 2 shows, the configurations give an average MAE of -0.224, however, a mean RMSE of 3.006, indicating an unbalanced distribution of temperature differences in measurement and simulation. This is caused by deviations in building outlines, misclassified pixels, geo-referencing errors, or deviations in tree shapes, which aligns with our expectations. The results further indicate that configuration 4, in which building outline, roof elevation, orientation towards road, and tree overlay are equally weighted, yields slightly better agreement with the
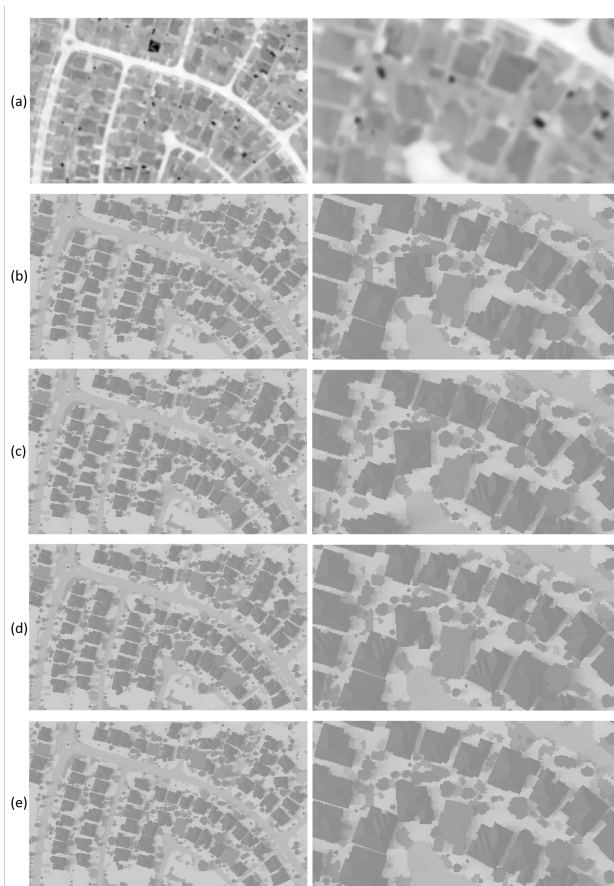
Figure 4. Measured thermal image (a) and simulated thermal images for configuration 1 (b), configuration 2 (c), configuration 3 (d) and configuration 4 (e), showing the full modelled area (left) and zoomed in (right). Color representation is *white-hot*.

measures temperatures both considering MAE and RMSE as well as general similarity quantified by MI.

Since our geo-typical modeling is focused on buildings and roofs, respectively, we attempted to evaluate the thermal simulation results by focusing on roof structures as well. Figure 5 displays the temperature distribution of a single geo-typical building that has been selected differently by the four similarity functions. Only configurations 2 to 4 can approximate the temperature profile. For quantification, a straight-forward method like binary masks creation for buildings in the orthophotos is not appropriate since this method would penalize in an unjustifiably hard way the building outlines. Therefore, we indirectly measure the effect of our geo-typical modeling method by evaluating the overall temperature gradients. As a result, the MAEs yield a value of -0.001 with no significant difference between the configurations. The RMSE varies between 0.590 and 0.596, with the highest, i.e., worst value for configuration 4. However, as before, a high RMSE is attributed to deviations in building outlines. Considering the MI, configuration 4 yields the best (highest) MI value of 0.0706, directly followed by configuration 1 with 0.0699. Based on these findings, we observe a tendency that the inclusion of building elevation, orientation, and trees does not have a significant impact on the realism of the 4D digital twin. However, a crucial parameter is the quality of the measured thermal image. With 0.5m per pixel of resolution, averaging of temperature values between two subsequent nights, and interpolation artifacts, such as blending of temperat-

ures, leading to the absence of clear edges, this baseline for the evaluation should be considered with care. Consistently, an in-depth validation based on higher resolution data will be needed in the future.

4. CONCLUSIONS

In this work, a generalized similarity function has been presented for a better specification of rich-in-details building prototypes, allowing the efficient representation of hundreds and thousands of buildings of a dataset. The previously applied IoU similarity function is very convenient because the building cadastre shapefile alone is sufficient: Only the polygonal footprints are needed to create a binary mask, while it is possible to extend the mask-based transformation to 3D using the shapefile attributes base level and height. However, we hoped for a more accurate matching with the additional information contained in sensor data, such as relative elevation and land cover classification map. Using the former one, we assessed to what extent the elevations in the overlapping parts are mutually proportional, while for the second one, we assessed the building's orientation towards the road as well as the presence of trees within the building outline.

Two strategies have been pursued to evaluate the success of four very distinct configurations of proposed measures. On the one hand, experts have evaluated the four configurations and concluded that the IoU is the most suitable for identifying similarities. On the other hand, the experiments with thermal simulations have shown that a slight improvement of simulated temperature distributions is given by an equally weighted inclusion of measures based on the IoU, elevation similarity, building orientation towards the road, and tree collision.

Despite encouraging findings, we must cover two main issues reflecting the shortcomings of our results pending work in the presented research direction. First of all, the database of existing building models is yet very scarce. With 20 models, four rotations, and an optional flipping, there are merely 160 configurations in total, and this is not much. Since the time needed for interactive modeling can reach up to 30 minutes, either much additional human workforce or advanced generative techniques allowing stretch and forge, add and remove building parts are necessary for data augmentation. Once library size is boosted up by several orders of magnitude, intelligent search algorithms must be identified since an exhaustive trial (and error) of candidates would cause unbearable computing times. Secondly, only four sets of buildings were evaluated in the Results section. However, the number of choices should ideally be exponential in that of the free parameters, so that with four terms needed to form the similarity function and three free parameters in (2), we would need at least a dozen of configurations. We focused on the proof of principle in the current work, but with a larger database, at least for the interactive assessment, a different, better concept shall be developed. In the long term, our purpose is to assess the temperatures of many digital twins and to provide the architects, city planners, and local councils helpful guidelines about efficient designs of building roofs in terms of UHIs.
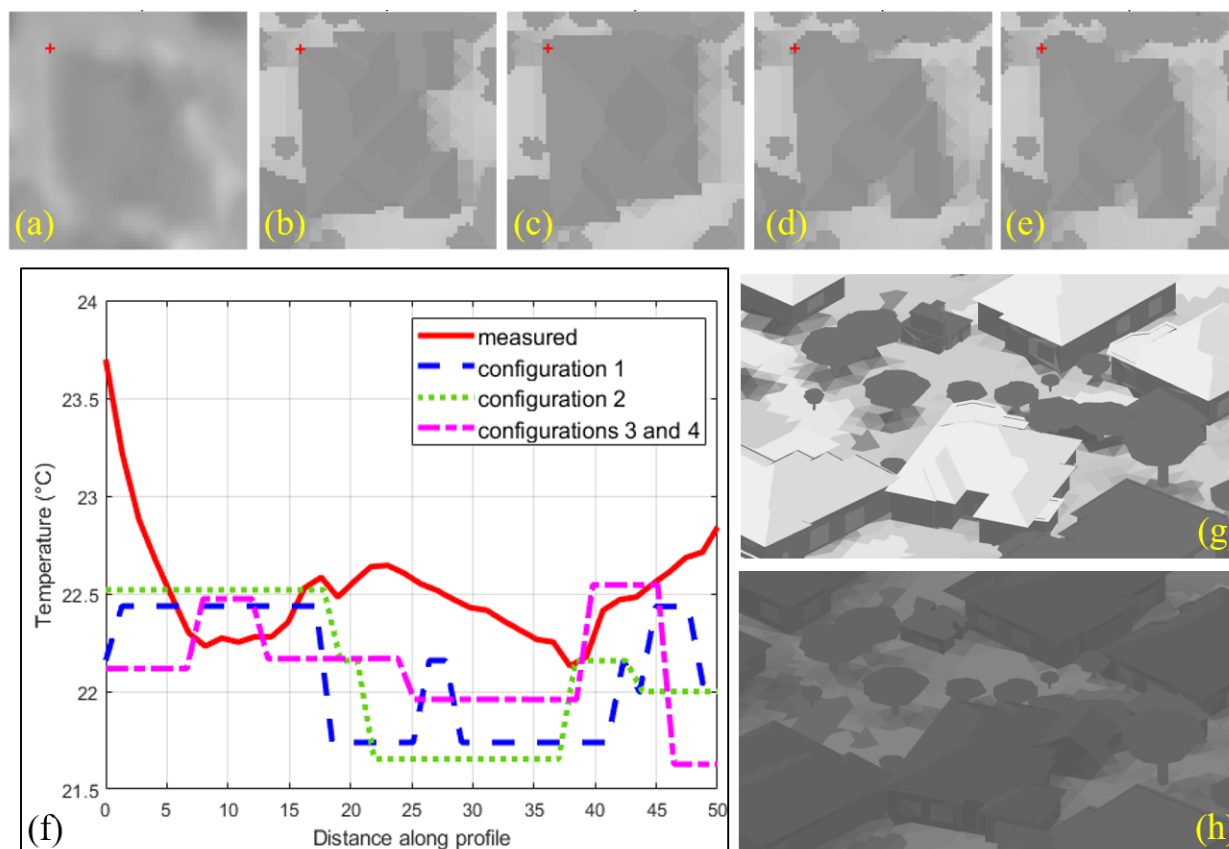
ACKNOWLEDGEMENTS

Figure 5. Temperature distribution in the (a) thermal image around a single building: Images (b) to (e) show the corresponding distributions resulting from configurations 1 to 4, respectively. Image (f) shows the corresponding temperature profiles along the line connecting the left upper and right lower corner of the building in the measured thermal image and the four configurations, respectively. The building itself (configuration 1, at 2 pm and 8 pm) is depicted in images (g) and (h). Note that the output for configurations 3 and 4 results in the same building structure.

REFERENCES

Albert, J., Bachmann, M., Hellmeier, A., 2003. Zielgruppen und Anwendungen für digitale Stadtmodelle und digitale Geländemodelle. *Erhebungen im Rahmen der SIG 3D der GDI NRW.*

Biljecki, F., Ledoux, H., Stoter, J., 2016. An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59, 25–37.

Bulatov, D., Burkard, E., Ilehag, R., Kottler, B., Helmholz, P., 2020. From multi-sensor aerial data to thermal and infrared simulation of semantic 3D models: Towards identification of urban heat islands. *Infrared Physics & Technology*, 105, 103233.

Bulatov, D., Häufel, G., Meidow, J., Pohl, M., Solbrig, P., Wernerus, P., 2014. Context-based automatic reconstruction and texturing of 3D urban terrain for quick-response tasks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 157–170.

Bulatov, D., May, M., Strau, E., Mancini, F., Kottler, B., Helmholz, P., 2021. Increasing level of detail of buildings for improved simulation of 4D urban digital twin. *To appear in: International Seminar on Urban Form (ISUF2021): Urban Form and the Sustainable and Prosperous Cities.*

Dai, W., Yang, B., Dong, Z., Shaker, A., 2018. A new method for 3D individual tree extraction using multispectral airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, 400–411.

Döllner, J., Buchholz, H., Brodersen, F., Glander, T., Jütterschenke, S., Klimetschek, A., 2005. Smart Buildings – A concept for ad-hoc creation and refinement of 3D building models. *Proceedings of the 1st International Workshop on Next Generation 3D City Models*, 1number 3.3.

Elmqvist, M., Jungert, E., Lantz, F., Persson, A., Söderman, U., 2001. Terrain modelling and analysis using laser scanner data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/W4), 219–226.

Guo, S., Xiong, X., Liu, Z., Bai, X., Zhou, F., 2018. Infrared simulation of large-scale urban scene through LOD. *Optics Express*, 26(18), 23980–24002.

Hu, P., Yang, B., Dong, Z., Yuan, P., Huang, R., Fan, H., Sun, X., 2018. Towards reconstructing 3D buildings from ALS data based on gestalt laws. *Remote Sensing*, 10(7), 1127.

Ilehag, R., Bulatov, D., Helmholz, P., Belton, D., 2018. Classification and representation of commonly used roofing material using multisensorial aerial tata. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1, 217–224.

Koch, B., Kattenborn, T., Straub, C., Vauhkonen, J., 2014. Segmentation of forest to tree objects. *Forestry Applications of Airborne Laser Scanning*, 89–112.

Kottler, B., Burkard, E., Bulatov, D., Haraké, L., 2019. Physically-based thermal simulation of large scenes for infrared imaging. *VISIGRAPP (1: GRAPP)*, 53–64.

Lafarge, F., Descombes, X., Zerubia, J., Pierrot-Deseilligny, M., 2010. Structural approach for building reconstruction from a single DSM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1), 135–147.

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., Schooling, J., 2020. Developing a digital twin at building and city levels: case study of West Cambridge campus. *Journal of Management in Engineering*, 36(3), 05020004.

McAlinden, R., Suma, E., Grechkin, T., Enloe, M., 2015. Procedural reconstruction of simulation terrain using drones. *Proc. of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, 1–12.

Meidow, J., Hammer, H., Pohl, M., Bulatov, D., 2016. Enhancement of generic building models by recognition and enforcement of geometric constraints. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, Copernicus GmbH, 333–338.

Mousa, A.-k., Helmholz, P., Belton, D., 2017. New DTM Extraction Approach From Airborne Images Derived DSM. *ISPRS Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42-47.

Shajahan, D. A., Nayel, V., Muthuganapathy, R., 2019. Roof classification from 3-D LiDAR point clouds using multiview CNN with self-attention. *IEEE Geoscience and Remote Sensing Letters*, 17(8), 1465–1469.

Sohn, G., Jwa, Y., Kim, H. B., Jung, J., 2012. An implicit regularization for 3D building rooftop modeling using airborne LIDAR data. *Proc. of ISPRS-Congress, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-3, 305–310.

Tang, L., Ying, S., Li, L., Biljecki, F., Zhu, H., Zhu, Y., Yang, F., Su, F., 2020. An application-driven LOD modeling paradigm for 3D building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 161, 194–207.

Toldo, R., Fusiello, A., 2008. Robust multiple structures estimation with j-linkage. *European Conference on Computer Vision*, Springer, 537–547.

Verdie, Y., Lafarge, F., Alliez, P., 2015. LOD generation for urban scenes. *ACM Transactions on Graphics*, 34(ARTICLE), 30.

Verma, V., Kumar, R., Hsu, S., 2006. 3D building detection and modeling from aerial lidar data. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2, IEEE, 2213–2220.

Voogt, J. A., Oke, T. R., 2003. Thermal remote sensing of urban climates. *Remote Sensing of Environment*, 86(3), 370–384.

Xiong, B., Elberink, S. O., Vosselman, G., 2014. A graph edit dictionary for correcting errors in roof topology graphs reconstructed from point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 227–242.

Xiong, B., Jancosek, M., Elberink, S. O., Vosselman, G., 2015. Flexible building primitives for 3D building modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, 275–290.

Yu, D., Ji, S., Liu, J., Wei, S., 2021. Automatic 3D building reconstruction from multi-view aerial images with deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 155–170.