Classification of photo-realistic 3D window views in a high-density city: The case of Hong Kong

Maosu Li[1]*, Fan Xue[2], Anthony G.O. Yeh[3], Weisheng Lu[4]

Abstract: Window view is an intimate medium between occupants and nature, especially in high-density cities like Hong Kong; and thus belongs to the quality of a house or apartment. In literature, researchers found that window views of nature are vital to the occupants' physical and psychological health and productivity improvement. Understanding the view situation at the urban level can facilitate urban environment optimization, urban planning and development policies, and smart city management. Currently, views of nature have been quantitatively studied in satellite images and cars' cameras at a macro or micro level, respectively. However, as an essential supplement to the greenery view information hub at a mesoscale, few studies on efficient visualization and classification of window views at the urban level seem available. This paper presents an automatic approach that captures and classifies photo-realistic views at the windows in a 3D photogrammetric city model. First, by triangulating the window geometries from geo-matched 3D photogrammetric and 2D digital maps, the rich window semantics are registered to the 3D models. Then, the similar window views are visualized in batch with an appropriate focal length and field of view. Finally, the view at each window is analyzed and classified through transfer learning automatically. We applied the proposed approach to the 3D model of Hong Kong Island and found satisfactory results for identifying nature scenes or urban scenes. Once massively adopted, the presented approach can offer novel geographic indicators for billions of urban inhabitants and the Architecture, Engineering, Construction, and Operation (AECO) industry.

Keywords: Window view, View visualization, View classification, View quality, 3D models, Urban.

[1]* Maosu Li, PhD student
Corresponding author, Department of Urban Planning and Design, University of Hong Kong, Hong Kong SAR, China
Email: maosulee@connect.hku.hk

[2] Fan Xue, Assistant Professor
Department of Real Estate and Construction, University of Hong Kong, Hong Kong SAR, China
Email: xuef@hku.hk

[3] Anthony G.O. Yeh, Chair Professor
Department of Urban Planning and Design, University of Hong Kong, Hong Kong SAR, China
Email: hdxugoy@hku.hk

[4] Weisheng Lu, Professor
Department of Real Estate and Construction, University of Hong Kong, Hong Kong SAR, China
Email: wilsonlu@hku.hk

1 Introduction

Windows physically bridge the interior indoor space and exterior environment. A window view is a significant medium for people's access to nature, especially for high-density urban areas such as Hong Kong. A number of studies identified that a good window view of nature plays an essential part in human mental health such as stress relief[1,2], physical health such as sleep improvement[3], disease remission[4] and working status such as productivity improvement[5]. However, crowded 3D cityscapes and high-rise buildings in high-density development scatter nature views and disparate inhabitants from getting close to nature equally. Window view quality became especially significant during the Coronavirus disease 2019 (COVID-19) pandemic because home isolation and social distancing measures kept billions of people from staying outside.

Besides, nature views from windows could also supplement the urban greenery information hub as the last piece of the puzzle at the mesoscale. In literature, views of nature as an important urban greenery indicator have been widely studied in urban landscape planning and management. For representing an overhead view at a macro scale, remote sensing has been commonly used[6,7] because of its advantages such as repeatability, synoptic view, and area coverage[8]. For capturing street view at a micro-scale, viewer tools based on handheld or cars' cameras were introduced to visualize the profile view of urban greenery at the street level[8,9]. However, for a 3D real world, massive window views visualized and classified to achieve greenery vertically at a mesoscale are still indispensable because of the uneasy access, immature technology and insufficient data. Thus, how to capture and classify the 3D window views in high-density cities at the urban level has become a significant and valuable research question, which can facilitate urban sustainability and optimization, including urban health[5,10], urban environment optimization[11], planning and development policy adjustment[12,13] and smart cities management[14,15].

Several studies have been conducted to visualize views of windows in different ways, which mainly include manual drawing, computer simulation modeling, and camera capture. Hellinga and Hordijk[16] proposed an approach using a 180-degree equidistant projection to achieve the view of the window, which could be realized by manual drawing. Turan and Reinhart developed frameworks employing vector raytracing in 3D models to evaluate views quality[17]. For the window design, Li and Samuelson[11] do some manual work such as window location information inputs to visualize the window views based on 3D models of Google Earth, in which the images captured owned more realism due to the fine resources. Abd-Alhamid et al. [18] utilized a fish-eye camera to evaluate window views at different viewing locations in a manually built virtual environment. Some studies relied on questionnaires and surveys[19,20] to take real photographs from the already-built windows. Although these methods have achieved relatively good results to evaluate the window view quality for different aims, they are all oriented to the single house or few estates with low automation and do not focus on the window view visualization at a larger scale. As a result, they are not applicable and efficient to help capture window views at the urban level due to increasing time-consuming manual works, limited data support, and expensive visualization costs.

For the classification of window views, most studies have been conducted in three methods, including subjective judgement, view feature proportion calculation and their hybrid. Studies on window view quality and impacts mainly use survey, interview or questionnaire methods to assess the views[20-22]. The views from the real windows or in the form of images are classified according to human subjective judgements such as a four-point scale[20], not satisfactory, satisfactory, good, and excellent. Some studies with architectural background calculate view features proportion to help guide the window and layout design using subjective or objective classification criteria[11,23]. Studies in the field of computer vision have developed mature approaches for image segmentation[24-26]. The proportion of each kind of content feature and other advanced features such as spatial and color structure would be calculated to evaluate the image quality[24]. However, for the evaluation of the view quality at a large scale, classification methods including human judgement are subjective, which cannot reach a common standard. Approaches for quantitative image quality evaluation on computer vision and architectural design are not targeted to the specific window view content classification details directly and thus are not accurate and efficient for urban-level applications. To sum up, few studies available were applicable to window view realistic visualization and classification at the urban level automatically and efficiently.

Oriented to the urban applications, this paper presents a novel approach to capture and classify the views from existing windows of estates automatically, based on the integration of a 3D photogrammetric city model and 2D digital map. By matching the 2D footprint data with buildings' 3D models, window view semantics are registered into 3D models in batch. Then views of windows with position and orientation information can be efficiently captured through 3D GIS visualization methods. In the end, the window views are analyzed and classified by transfer learning automatically. The remainder of this paper is organized into the following sections: methods, a pilot case study, discussion and conclusion.

2 Methods

This paper presents a novel approach to collecting and processing 3D window views. As shown in Figure 1, the input to the approach is the actual footprint data of the buildings and their photo-realistic models. In the workflow, for view capture and classification of a single building, automatic data processing mainly includes three steps:

1) semantic information registration of windows based on geo-matching of 2D and 3D data at the urban level,
2) window view visualization in batch with appropriate camera settings, and
3) automatic window view classification by transfer learning and judgement criteria.

The outputs of the approach include categorized views of windows in three major types.

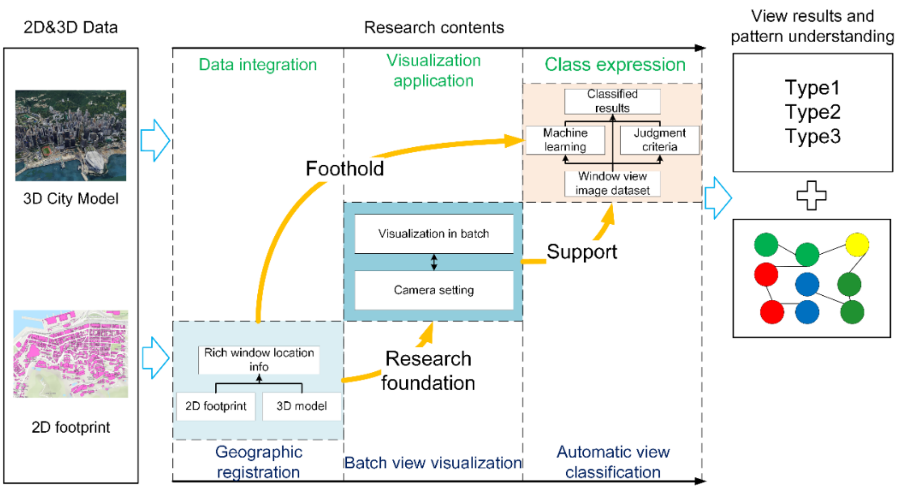


Figure 1. The research design of the proposed view visualization and classification approach

2.1 Window geographic registration

For a target building, the first step is window geographic registration. 3D photo-realistic models depict the actual appearance of buildings and their vivid built environment settings, in which the window location and height information can be easily extracted. 2D footprint provides accurate layout geometry information as well as rich static text attributes, in which the window orientation information can be achieved. The integration of them as an information hub of buildings can provide solid data support for view capture. As both 3D photo-realistic models and 2D footprint data have accurate geographic references, buildings can be easily matched. Then semantic information of windows is extracted and registered to 3D models based on the data hub. In this approach, windows within one façade are considered to have the same orientation as their façade. In this way, by turning to the elevation view of each façade, windows can be determined correctly to reduce the visual deviation.

2.2 Window view batch visualization

The second step is to visualize window views in batch. In this paper, the center of the window is considered the location of the digital virtual camera. Based on the windows' location and orientation information, the camera can be placed on the targeted window automatically. The focal length of the camera is set appropriately to control the involved view features details. The field of view of the camera is adjusted to control the view size. In this way, a similar visualized result can be achieved to stimulate the real window view. By repeating this process automatically, views of windows in each façade of the target building are captured in batch.

2.3 Automatic window view classification

The final step of the proposed approach is the automatic window view classification. In this paper, a deep transfer learning model is used to classify the window views automatically. Views of windows are simply categorized into three types according to the proportion of nature scenes, which regards vegetation view, sky view, sea view or their hybrid as nature view, house view, street view, road view or their hybrid as urban view, and the hybrid of nature view and urban view as hybrid view. The *a-priori* rules are applied to classify the window views. To achieve a complete nature view, we assume that window views with more than 95% nature elements are regarded as nature views. Window views with less than 5% nature elements are urban views while if the proportion of window views' nature elements ranges from 5% to 95%, they are hybrid views. Besides, to extract more urban views from hybrid views, if the percentage of green elements of natural features such as vegetation is less than 5% and the figure for the sky and sea elements is less than 50%, the hybrid view is also considered as the urban view in this article.

3 A Pilot Study

3.1 Window geographic registration

Lap-Chee College on Lung Wah Street is a high-rise student hostel of the University of Hong Kong, as shown in Figure 2. We selected the Lap-Chee College to validate the proposed approach. The scene of the building and its neighborhood environment from the photogrammetric 3D city model is shown in Figure 2 (a). The 2D footprint data of it in the GeoJSON format were transferred from shapefile data (iB1000) purchased by the University of Hong Kong. The photogrammetric 3D city model was collected from the Government of Hong Kong[27]. The matched results were shown in Figure 2 (b) visualized by Cesium (version 1.73, https://cesium.com/).

Figure 2. Geographic registration of the case building. (a) Target building and its environment, (b) geographic registration

Although the boundary line of the building is given by the 2D geometry data, some construction points which are not vertices still exist. Thus, an approach by calculating the angle difference of the line between adjacent points was proposed to determine the vertices and orientation of the building's each façade. The detailed steps are as follows.

1) A 2D building footprint A in the GeoJSON format contains a list of n points $A = \{p_1, p_2, \ldots, p_n\}$.
2) Let $p_{n+1} = p_1$, the direction a_i can be computed for the i-th edge (p_i, p_{i+1}), $1 \leq i \leq n$.
3) If $a_{i+1} - a_i$ $(1 \leq i \leq n)$ is larger than a threshold β, the point p_{i+1} is considered as a vertex of the geometry and recorded.
4) After all the vertices are recorded, the orientation of each facade edge can be calculated. In Cesium, orientation represents the rotation from the local north direction where a positive angle is increasing eastward.

For the case building, the windows can be recognized from the building's front and back views, as shown in Figure 3 (a) and 3 (b). Then by flying to each façade in the elevation view, the center of the window was selected, as shown in figure 3 (c). In the meanwhile, the location and orientation information of windows was registered to 3D models and then collected in the database.

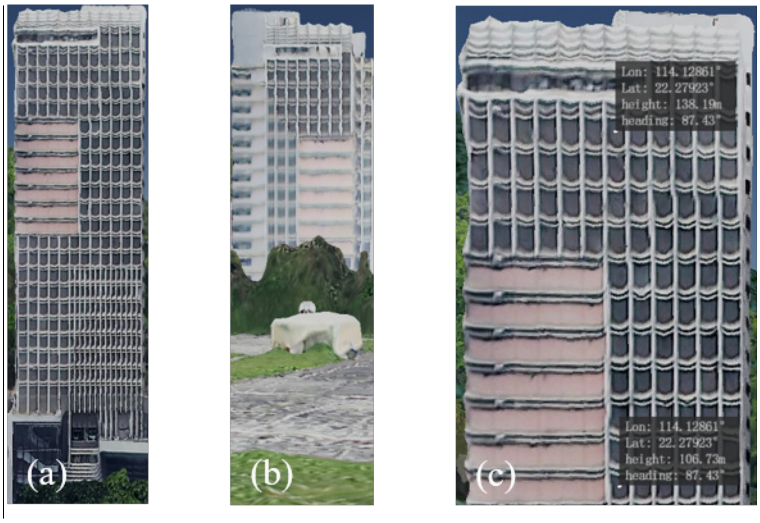


Figure 3. Location information extraction of windows. (a) Front, (b) back, (c) location information collection of windows

3.2 Window view batch visualization

The data visualization and procession were based on an open-source platform Cesium. Cesium is a fast, simple, end-to-end platform for tiling, visualizing, and analyzing 3D geospatial data, which was first released in 2011 and has been used as a tool for 3D visualizations that are accurate, performant, and time-dynamic on the web. Camera settings, including camera destination, camera heading, camera tilt, and camera focal length, can be customized using this tool, which makes it possible for users to set the camera at a proposed position and obtain the view shots efficiently by computer programs.

To visualize the window views in batch, a virtual camera mounts on the target window at the 3D coordinate (longitude, latitude, height) with known orientation. And then, by adjusting the camera's focal length and field of view, a similar view of the window is achieved. In Cesium, we used the "zoom" function to control the virtual camera's focal length, and the canvas size was adjusted to match the window's real size. Four test windows were selected to make a comparison, as shown in Figure 4. Photographs (1a, 2a, 3a, and 4a) for the windows were taken using iPhone X, while the virtual camera captured views of windows (1b, 2b, 3b, and 4b). Based on the comparison results, Cesium's view visualization approach is considered realistic and accurate enough to satisfy the intended purpose. The average time spent on one view image visualization and procession was 3s.

Figure 4. Comparison of real photos and window view results. (a) Ground truth photographs, (b) our view results

3.3 Automatic window view classification

In this paper, the DeepLabv3, a deep learning model trained on the cityscapes dataset, was used to analyze the visualized results at a pixel level[28]. By deep transfer learning, each pixel of a view image could be segmented into one of twenty classes including vegetation, building, sky, terrain, etc. By calculating their percentages in the photo, classifying the classes into nature or urban groups, and implementing a quantitative comparison according to the *a-priori* rules, the type of view could be determined, as shown in Figure 5. For instance, the proportion of natural elements of window view #1 is 0.98, which is more than 95%, and thus it is regarded as a nature view.

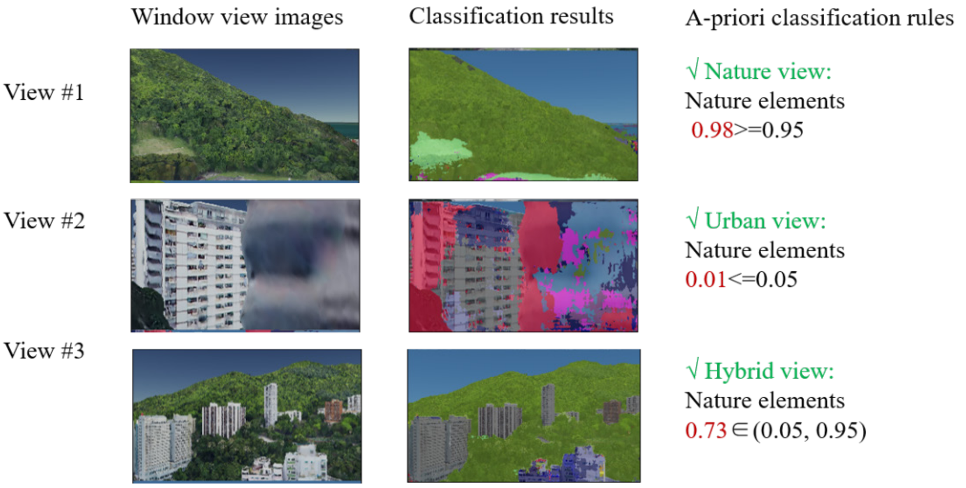
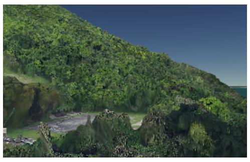
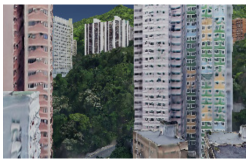
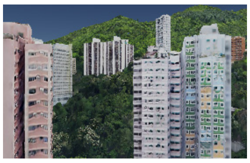
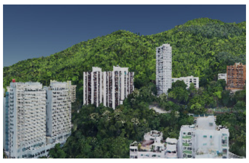
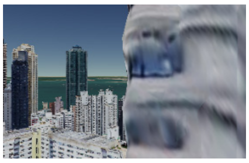


Figure 5. Window view classification process

The deep transfer learning was conducted on the workstation with two Intel XEON E5-2690 v4 CPUs (2.6GHz, 28 cores), 64 GB memory, an Nvidia Quadro P5000 GPU and Windows 10 Enterprise 64-bit operating system. The average procession time of one view image was 1.96s. Based on this way, view images were captured and classified in batch. Examples of three typical kinds of views on floors with different heights are shown in Table 1. For the nature views, higher flats can see more pixels of the sky while the lower flats can only see the greenery. The differences were also noticeable for urban and hybrid views, despite the composition of more types of pixels. A clear trend is that higher-level flats offer views with higher-level natural sights, landscape distance, and perceived spaces. In a high-density city like Hong

Kong, this access to nature could be preferred. Example applications such as view scoring and analysis can be used for real estate valuation and built environment optimization in the AECO industry.

Table 1. Examples of three kinds of views on floors with different heights

View type	Lower-rise flat	Middle-rise flat	Higher-rise flat
Nature			
Hybrid			
Urban			

4 Discussion

View quality on the extent of getting access to nature was difficult to investigate, obtain, and analyze qualitatively or quantitatively in the past, especially in high-density areas. However, with the rising 3D reconstruction technologies, such as oblique photogrammetry, laser scanning, building information modeling, and digital twin cities[29], the proposed approach can provide a novel way to compute the view semantics at the urban level. Compared with existing view visualization approaches that require modeling the outdoor context in 3D or manipulation of extensive data[11,16-19], the proposed view visualization approach reduces the intensity of human-computer interaction and is more applicable for the urban investigation and analysis. By implementing a comparison test with the real windows views, this approach's views of windows captured and visualized, including full context information, are accurate and realistic enough, providing a refined dataset of realistic-looking results for the urban analysis. Compared to existing window view classification methods [11, 20-23], this approach utilizes a deep transfer learning model and *a-priori* rules to categorize window views quantitatively and efficiently, which are more applicable for automatic large-scale view evaluation. To our best knowledge, it is the first method to automate the process from visualization to classification of window views for urban-level view understanding. Related works mainly focus on the view assessment of a single window or building. Applying them to the large-scale window view visualization and classification would face intensive human-computer interactions, expensive production costs and low-level automation.

This pilot study preliminarily confirms the technological feasibility of the proposed approach. Nevertheless, limitations still exist, and future goals for the proposed workflow are described below. They include the workflow automation improvement, construction of view content quality index framework, and correlation analysis between view index and other urban features. The current workflow requires higher automation. For instance, the window centers were determined by batch selection. Future work could develop a more integrated workflow to further improve the automation of the whole process. A nature view or urban view was classified in this paper, which was rough to make use of views information. Future work was planned to focus on constructing the view content quality index framework and corresponded view classification and scoring by machine learning. Based on view index data, more spatial and correlation analyses could indicate urban phenomenon and patterns.

5 Conclusion

Views of windows that measure the extent to which occupants can access nature are vital to various urban topics on health, sustainability, and optimization. However, the detailed and realistic views data are not well prepared to enable such studies. This paper proposes a novel workflow to help visualize and classify building windows at the urban level. By integrating 2D and 3D data of buildings, geometry information of windows could be registered efficiently. Then, by placing a virtual camera in the center of the windows, similar realistic views are visualized in batch. In the end, window views are classified as nature or artifacts automatically.

Compared with existing view computing approaches, it is the first method to automate the whole process of window view visualization and classification for urban-level view understanding, which lowers human-computer interactions, generates accurate and detailed results. A pilot study confirmed that the effective window view capture and visualization approach could provide a tool for view visualization and analysis at the urban level, which can offer a new data hub for smart city management. Future work focuses on higher workflow automation, construction of view index framework, and correlation analysis between view index and other urban features.

Acknowledgements

The work presented in this paper was funded by the Research Grants Council (RGC) of Hong Kong SAR (No. 17200218) and in part by the University Research Committee (URC) of the University of Hong Kong (No. 102009741).

Reference

[1] Kaplan, R. (2001). *The nature of the view from home: Psychological benefits*, Environment and behavior, 33(4), 507-542.

[2] Jiang, B., Chang, C.-Y. and Sullivan, W. C. (2014). *A dose of nature: Tree cover, stress reduction, and gender differences*, Landscape and Urban Planning, 132, 26-36.

[3] Yang, L., Ho, J. Y., Wong, F. K., Chang, K. K., Chan, K. L., Wong, M. S., Ho, H. C., Yuen, J. W., Huang, J. and Siu, J. Y. (2020). *Neighbourhood green space, perceived stress and sleep quality in an urban population*, Urban Forestry & Urban Greening, 54, 126763.

[4] Hartig, T., Astell-Burt, T., Zara, B., Amcoff, J., Mitchell, R. and Feng, X. (2020). *Associations between greenspace and mortality vary across contexts of community change: a longitudinal ecological study*, Journal of Epidemiology & Community Health, 74(6), 534-540.

[5] Aries, M. B., Veitch, J. A. and Newsham, G. R. (2010). *Windows, view, and office characteristics predict physical and psychological discomfort*, Journal of Environmental Psychology, 30(4), 533-541.

[6] Faryadi, S. and Taheri, S. (2009). *Interconnections of urban green spaces and environmental quality of Tehran*, International Journal of Environmental Research, 3(2), 199–208.

[7] Zhu, G., Bian, F. and Zhang, M. (2003). *A flexible method for urban vegetation cover measurement based on remote sensing images*, In: ISPRS WG I/5 Workshop, High Resolution Mapping from Space.

[8] Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q. and Zhang, W. (2015). *Assessing street-level urban greenery using Google Street View and a modified green view index*, Urban Forestry & Urban Greening, 14(3), 675-685.
[9] Xue, F., Li, X., Lu, W., Webster, C.J., Chen, Z. and Lin L. (2020). *Big data-driven pedestrian analytics: Unsupervised clustering and relational query based on deep transfer learning*, Computers, Environment and Urban Systems (under review).
[10] Elsadek, M., Liu, B. and Xie, J. (2020). *Window view and relaxation: viewing green space from a high-rise estate improves urban dwellers' wellbeing*, Urban Forestry & Urban Greening, 55, 126846.
[11] Li, W. and Samuelson, H. (2020). *A New Method for Visualizing and Evaluating Views in Architectural Design*, Developments in the Built Environment, 1, 100005.
[12] Kim, J.-J. and Wineman, J. (2005). *Are windows and views really better? A quantitative analysis of the economic and psychological value of views*, The University of Michigan.
[13] Yeh, A.G.O. and Chen, Z. (2020). *From cities to super mega city regions in China in a new wave of urbanisation and economic transition: Issues and challenges*, Urban Studies, 57(3), 636-654.
[14] Yeh, A.G., Yue, Y., Zhou, X. and Gao, Q.L. (2020). *Big data, urban analytics and the planning of smart cities, In Handbook of Planning Support Science*, Edward Elgar Publishing.
[15] Xue, F., Lu, W., Tan, T. and Chen, K. (2019). *Semantic enrichment of city information models with LiDAR-based rooftop albedo*, In Sustainable Buildings and Structures: Building a Sustainable Tomorrow: Proceedings of the 2nd International Conference in Sustainable Buildings and Structures (ICSBS 2019), 207.
[16] Hellinga, H. and Hordijk, T. (2014). *The D&V analysis method: A method for the analysis of daylight access and view quality, Building and Environment*, 79, 101-114.
[17] Turan, I., Reinhart, C. and Kocher, M. (2019). *Evaluating spatially-distributed views in open plan work spaces*, Proceedings of the IBPSA International Building Simulation Conference.
[18] Abd-Alhamid, F., Kent, M., Calautit, J. and Wu, Y. (2020). *Evaluating the impact of viewing location on view perception using a virtual environment*, Building and Environment, 180, 106932.
[19] Fontenelle, M. and Bastos, L. (2014). *The multicriteria approach in the architecture conception: Defining windows for an office building in Rio de Janeiro*, Building and Environment, 74, 96-105.
[20] Matusiak, B.S. and Klöckner, C.A. (2016). *How we evaluate the view out through the window*, Architectural Science Review, 59 (3), 203-211.
[21] Kent, M. and Schiavon, S. (2020). *Evaluation of the effect of landscape distance seen in window views on visual satisfaction*, Building and Environment, 183, 107160.
[22] Masoudinejad, S. and Hartig, T. (2020). *Window view to the sky as a restorative resource for residents in densely populated cities*, Environment and Behavior, 52 (4), 401-436.
[23] Ko, W.H., Kent, M.G., Schiavon, S., Levitt, B. and Betti, G. (2020). *A window view quality assessment framework*, arXiv preprint arXiv:2010.07025.
[24] Luo, W., Wang, X. and Tang, X. (2011). *Content-based photo quality assessment*, In 2011 International Conference on Computer Vision, 2206-2213.
[25] Kong, S., Shen, X., Lin, Z., Mech, R. and Fowlkes, C. (2016). *Photo aesthetics ranking network with attributes and content adaptation*, In European Conference on Computer Vision, Springer, Cham, 662-679.
[26] Zhang, P., Lu, W., Wang, H., Lei, Y. and Lu, H. (2019). *Deep gated attention networks for large-scale street-level scene segmentation*, Pattern Recognition, 88, 702-714.
[27] PlanD (2019). 3D Photo-realistic Model, Planning Department, Goverment of Hong Kong SAR, Hong Kong.
[28] Chen, Y.H., Chen, W.Y., Chen, Y.T., Tsai, B.C., Frank Wang, Y.C. and Sun, M. (2017). *No more*

discrimination: Cross city adaptation of road scene segmenters, In Proceedings of the IEEE International Conference on Computer Vision, 1992-2001.

[29] Xue, F., Lu, W., Chen, Z. and Webster, C.J. (2020). *From LiDAR point cloud towards digital twin city: Clustering city objects based on Gestalt principles*, ISPRS Journal of Photogrammetry and Remote Sensing, 167, 418-431.